

**PROPOSAL PENELITIAN ISPO**

**C-ATP**

**(Computational Analysis of *Aloe vera* for Transgenic Insulin Production)**

**Disusun Oleh :**

**Aisyah Rifa Fadhilah**

**Muhamad Nabil Alhanif**

**Bidang Lomba Penelitian :**

**BIOLOGI (BIOTEKNOLOGI)**

**Nama Sekolah :**

**SMA Kharisma Bangsa**

**Kota, Provinsi :**

**Tangerang Selatan, Banten**

## **Lembar Pengesahan**

**Judul : C-ATP (Computational Analysis of *Aloe vera* for Transgenic Insulin Production)**

**Nama Penulis : -Aisyah Rifa Fadhilah  
-Muhamad Nabil Alhanif**

**Mengetahui,**

**Kepala Sekolah Kharisma Bangsa,**

**Imam Husnan Nugroho, S.T,M.Pd.**

**Pembimbing,**

**Dr. Satya Nugroho.  
Dessy Norma Juita, M.Pd.**

## **Kata Pengantar**

Puji dan syukur kami panjatkan kepada Allah SWT atas berkat dan rahmat-Nya kami dapat menyelesaikan rancangan penelitian yang berjudul C-ATP (Computational Analysis of *Aloe vera* for Transgenic Proinsulin Production).

Proposal rancangan penelitian ini dibuat dalam rangka mengikuti perlombaan ISPO (Indonesia Science Project Olympiad) 2021.

Proposal ini kami susun dengan maksimal, dan mendapatkan bantuan dari berbagai pihak sehingga dapat memperlancar pembuatan makalah ini. Untuk itu, kami mengucapkan banyak terima kasih pada seluruh pihak yang telah berkontribusi dalam penulisan proposal ini.

Ucapan terima kasih kami berikan kepada:

- Bapak Dr. Satya Nugroho (Mentor pembimbing)
- Bapak Aris Setyawan, S.Si (Koordinator tim karya ilmiah Kharisma Bangsa)
- Ibu Dassy Norma Juita, M.Pd (Guru pembimbing) atas dukungan, bimbingan, serta bantuannya.
- Orang tua, guru dan teman teman atas dukungan, doa, dorongan, serta bantuannya selama ini sehingga kami dapat menyelesaikan proposal ini.

Akhir kata, kami menyadari bahwa proposal dan penelitian ini masih sangat jauh dari sempurna, oleh karena itu, kami mengharapkan kritik serta saran yang membangun agar penelitian ini dapat diselesaikan, direalisasikan, serta dapat bermanfaat bagi masyarakat dan lingkungan.

## Daftar Isi

<b>Kata Pengantar</b>	<b>3</b>
<b>Daftar Isi</b>	<b>4</b>
<b>Abstrak</b>	<b>9</b>
<b>Bab 1</b>	<b>10</b>
<b>Pendahuluan</b>	<b>10</b>
1.1 Latar Belakang	10
1.2 Rumusan Masalah	13
1.3 Batasan Masalah	14
1.4 Tujuan Penelitian	14
1.5 Kebaruan	15
1.6 Hipotesis	15
1.7 Manfaat Penelitian	15
<b>BAB II</b>	<b>16</b>
<b>Tinjauan Pustaka</b>	<b>16</b>
2.1 Metode Penelitian	16
2.2 Diabetes	17
2.2.1 Tipe Diabetes	17
1. Diabetes Tipe 1	17
2. Diabetes Tipe 2	18
3. Gestational diabetes (GDM)	18
2.2.2 Jumlah Penderita Diabetes	18
2.3 Insulin	20
2.3.1 Sintesis Insulin	22
2.3.2 Jalur Sekresi Insulin detail	23
2.3.3 Produksi Insulin Konvensional	24
2.3.4 Produksi Insulin Rekombinan	25
Escherichia coli	25
Ragi	27
Tanaman Transgenik	28
2.3.5 Perbandingan Persentase Produksi Insulin Rekombinan	29
2.3.6 Ekstraksi Protein	29
Trichloroacetic acid (TCA)–acetone Precipitation Method	31
Phenol Extraction with methanol/ammonium acetate precipitation Method	32
2.4 Biopharming	35
2.4.1 Perbandingan Sistem Ekspresi pada Tanaman dan E. coli	36
Pada E. coli:	36
Keuntungan:	37

Kelemahan:	37
Pada tanaman transgenik:	37
2.5 Sistem/Platform Ekspresi	38
Tanaman sebagai biofaktori untuk ekspresi protein heterolog	40
2.6. Tipe Ekspresi	40
2.6.1 Ekspresi Kloroplas	41
Keuntungan dari Transformasi Kloroplas	42
Organisasi Genom Kloroplas : Konsep dari Transformasi Kloroplas	42
Regulasi dari Ekspresi Gen Kloroplas	44
2.7 Metode Transformasi	44
2.7.1 Metode Transformasi Tidak Langsung	45
Transformasi yang dimediasi Agrobacterium tumefaciens	45
2.7.2 Metode Transformasi Langsung	45
Transformasi yang dimediasi Particle Bombardment	45
Aplikasi	46
Keunggulan	47
2.8 Desain Konstruksi Biolistik	48
2.8.1 Promoter	48
2.8.2 Polyadenylation	50
2.8.3 Terminator	51
2.8.4 Marker gene (Gen Penanda)	51
Selectable Marker	51
Hygromycin phosphotransferase (HPT) (hpt)	52
2.9 Memilih Tanaman Model	52
2.9.1 Aloe vera	52
Fisiologi Aloe vera	53
Metabolisme	53
Transformasi dan Rekayasa genetik	53
Kandungan Nutrisi	54
Persebaran Aloe vera	55
Manfaat Aloe vera	56
Pengobatan Tradisional	56
Produk Kesehatan	56
Produk Kosmetik	57
Budidaya Aloe vera	57
Lahan Penanaman	57
Benih, Penanaman, dan Pemeliharaan	57
Pemanenan	57
Hasil Panen	58
2.9.2 Keunggulan Aloe vera	58
2.10 Substansi Genetik Aloe vera	59

2.10.1 Sekuensing dari genome Aloe vera dan transcriptome	61
2.10.2 Genome Kloroplas Aloe vera	62
2.10.3 Aloe vera Menjadi Kandidat yang Cocok untuk Penelitian Ini	63
<b>2.11 Enzim Restriksi dan Ligase</b>	<b>64</b>
<b>2.11.1 Enzim Restriksi</b>	<b>64</b>
Situs Restriksi	64
Tipe	65
Tipe I	66
Tipe II	66
Tipe III	67
Tipe IV	68
<b>2.11.2 Enzim Ligase</b>	<b>69</b>
Mekanisme Enzimatik	69
Tipe	70
E. coli	70
T4	70
Mamalia	71
Thermostable	71
<b>2.12 RT PCR (Reverse Transcription Polymerase Chain Reaction)</b>	<b>72</b>
Prinsip kerja:	74
Aplikasi:	74
<b>2.13 PCR</b>	<b>74</b>
<b>2.13.1 Primer</b>	<b>75</b>
Melting Temperature	76
Annealing Temperature	76
Panjang Primer	77
Kandungan GC Primer	77
Komplementasi Forward dan Reverse Primer	77
<b>2.13.2 Langkah Langkah PCR</b>	<b>78</b>
1. Denaturasi	78
2. Annealing	78
3. Elongasi	78
<b>2.13.3 Primer3</b>	<b>79</b>
Cara Penggunaan Primer3	79
<b>2.14 Sekuensing</b>	<b>81</b>
<b>2.14.1 Maxam-Gilbert Sequencing</b>	<b>81</b>
<b>2.14.2 Sanger Sequencing</b>	<b>82</b>
<b>2.15 Elektroforesis</b>	<b>83</b>
<b>2.16 Western Blot</b>	<b>84</b>
<b>2.17 ELISA</b>	<b>85</b>
<b>2.17.1 Insulin ELISA kit</b>	<b>86</b>

<b>BAB 3</b>	<b>88</b>
<b>Metodologi</b>	<b>88</b>
3.1 In Silico	88
3.2 Data dan Metode Analisis	88
3.3 Prosedur	88
3.4 Mengidentifikasi dan Mengumpulkan informasi	89
3.5 Mengobservasi dan Mengumpulkan Data Konstruksi	90
3.5.1 Mengumpulkan dan Memilih Sequence	90
Menentukan Promoter	90
Menentukan Terminator	90
3.5.2 Memilih Vektor Ekspresi	91
3.6 Langkah-Langkah In silico	91
3.6.1 Identifikasi situs restriksi	92
3.6.2 Optimasi Kodon Proinsulin dengan dnachisel	93
3.6.3 Optimasi PCR	96
3.6.4 Konstruksi transgen menggunakan snapgene	98
Tahap 1:	98
Tahap 2:	99
Tahap 3:	100
3.6.5 Visualisasi Protein	101
<b>Bab 4</b>	<b>102</b>
<b>Hasil dan Pembahasan</b>	<b>102</b>
4.1 Konstruksi Transgen	102
4.1.1 Optimasi Kodon	102
4.1.2 Virtual Digest dan Virtual Electrophoresis	104
4.1.3 Plasmid Transgenik	104
4.2 Visualisasi	106
4.2.1 Model 1	107
4.2.2 Model 2	109
4.2.3 Model 3	112
4.2.4 Model 4	114
4.3 Rencana Penelitian In-Silico Lanjutan	116
4.4 Rancangan Eksperimen In-Vivo	116
4.4.1 Kultur dan Preparasi	117
4.4.2 Persiapan DNA Coated Microcarrier	118
Persiapan Transgen	118
Persiapan Gold Particle dan Microcarrier	118
4.4.3 Particle Bombardment	119
Transformasi Aloe vera	120

4.4.4 Seleksi Embrio, Germinasi, dan Regenerasi	121
4.4.5 Ekstraksi, purifikasi, dan Analisis	121
Ekstraksi dan Purifikasi	121
Analisis Insulin	122
<b>Bab 5</b>	<b>123</b>
<b>Kesimpulan dan Saran</b>	<b>123</b>
5.1 Kesimpulan	123
5.2 Saran	123
<b>Referensi</b>	<b>125</b>

## **Abstrak**

Diabetes merupakan penyakit berbahaya yang berpotensi tinggi menjadi pembunuhan. Pada tahun 2030, diprediksikan akan terjadi peningkatan signifikan dalam jumlah penderita diabetes, dibandingkan pada tahun 2010. Seiring dengan meningkatnya jumlah penderita diabetes, maka permintaan atas insulin akan ikut meningkat. Sehingga, diprediksikan akan terjadi peningkatan pesat permintaan insulin dan kebutuhan untuk menemukan cara produksi insulin baru yang lebih efisien dan lebih terjangkau.

Keterbatasan jumlah produksi, ditambah dengan permintaan tinggi konsumen tentunya akan menaikkan harga insulin pasaran. Selain itu, WHO menyatakan bahwa prevalensi diabetes pada negara dengan pendapatan menengah ke bawah relatif lebih tinggi daripada negara dengan pendapatan tinggi. Hal ini tentunya akan membatasi dan mengurangi akses insulin bagi mereka yang membutuhkan.

Oleh karena itu, kami mengembangkan teknik produksi insulin baru yang lebih efisien dan hemat biaya, dengan menggunakan *Aloe vera* transgenik sebagai pabrik produksi insulin. Dengan menggunakan teknik ini, genome preproinsulin manusia akan disisipkan ke dalam kloroplas *Aloe vera* dengan menggunakan teknik particle bombardment, yang diharapkan akan menghasilkan hasil yang signifikan pada *Aloe vera* transgenik. Keberadaan insulin pada *Aloe vera* akan diperiksa dengan menggunakan teknik ELISA, western blot, dan lain lain.

Namun, dikarenakan batasan waktu dan pandemi, penelitian ini hanya akan berfokus pada simulasi in silico, yang berupa komputasi, pengolahan, dan analisis data, serta rancangan eksperimen in vivo yang diharapkan dapat digunakan sebagai pertimbangan sebelum dan perbandingan sesudah eksperimen in vivo.

Kata Kunci: In silico, computing, diabetes, insulin, *Aloe vera*, kloroplas, *particle bombardment*.

# Bab 1

## Pendahuluan

### 1.1 Latar Belakang

Diabetes merujuk kepada penyakit kronis akibat gangguan metabolisme gula yang mengakibatkan penderitanya memiliki kadar gula darah tinggi. Sekitar 9,3% persen populasi orang dewasa di seluruh dunia menderita diabetes. Angka ini diprediksi akan meningkat melebihi 11% pada 2045.

Diabetes dapat berujung pada berbagai komplikasi kesehatan serius seperti penyakit kardiovaskular, penyakit ginjal kronis, stroke, dan kematian. Bahkan, saat ini diabetes menempati urutan 10 besar penyebab kematian terbanyak secara global, dan merupakan penyebab kematian terbesar ketiga di Indonesia (WHO 2017) [1].

Sampai dengan saat ini, terapi insulin merupakan satu-satunya pengobatan yang efektif untuk diabetes tipe 1, dan umumnya diabetes tipe 2 apabila penyakit berlanjut. Terapi insulin memerlukan pemantauan gula darah teratur, dan injeksi insulin untuk mencegah komplikasi sekunder.

Namun, harga insulin yang tinggi, dan ketidaknyamanan pasien saat injeksi, menyebabkan beberapa pasien tidak patuh pada terapi insulin. Oleh karena itu, peneliti sedang mengembangkan teknologi teknologi baru untuk mengurangi ketidaknyamanan pasien saat terapi insulin. Contohnya adalah melalui paru-paru, mulut dan, hidung (Modi et al, 2002; Goldberg Dan Gomez-Orellana, 2003; Cefalu, 2004).

Insulin oral menggunakan kapsul untuk mencegah rusaknya insulin oleh sistem pencernaan. Adapun kapsul yang digunakan berupa alginate/chitosan, solid lipid nanoparticle, dan lain-lain. (Sarmento, Ribeiro, dkk, 2007; Sarmento, Martins, dan Souto, 2019).

Namun oral insulin, dan beberapa teknologi baru lainnya membutuhkan dosis tinggi untuk mencapai kinerja yang sama dengan insulin suntik. Masalah ini, ditambah dengan meningkatnya penderita diabetes, diprediksi akan menurunkan akses terhadap insulin. Hal lain yang ikut berperan dalam penurunan akses

terhadap insulin adalah pola hidup tidak sehat, seperti terlalu banyak mengkonsumsi junk food (Lancet Diabetes and Endocrinology, 2018).

Jumlah insulin yang dibutuhkan untuk mengobati diabetes tipe 2 diproyeksikan meningkat lebih dari 20 persen di seluruh dunia selama 12 tahun ke depan dari 406 juta botol 1000 unit saat ini menjadi 511 juta botol 1000 unit pada 2030. Pada saat yang sama, penggunaan insulin global juga diproyeksikan meningkat menjadi 634 juta botol 1000 unit pada tahun 2030 dari 526 juta botol 1000 unit saat ini. Pada 2030 mendatang, diprediksikan hanya setengah dari total jumlah penderita diabetes yang mendapatkan akses insulin, dengan kata lain insulin akan mengalami kelangkaan.

Selain itu, perawatan diabetes memakan biaya yang besar. Diperkirakan, pada tahun 2007, biaya pengobatan diabetes di Amerika mencapai \$174 miliar. \$116 miliar diantaranya adalah untuk pengeluaran medis berlebih, yang terdiri dari \$27 miliar perawatan langsung, \$58 miliar mengobati komplikasi diabetes, dan \$31 miliar pengobatan medis umum. Diperkirakan \$1 dari setiap \$10 digunakan untuk perawatan terkait diabetes (American Diabetes Association, 2008). Biaya ini diproyeksikan akan mencapai \$216 sampai \$396 pada tahun 2025 (Giannini et al., 2009).

Banyak negara berkembang yang menghabiskan lebih dari 50% biaya tahunannya untuk insulin dan pengobatan penyakit terkait diabetes, dan biaya insulin dapat mencapai lebih dari \$100 per bulannya (Raab et al., 2004).

Keterbatasan ekonomi dan kapasitas produksi tentunya akan menyebabkan masalah besar apabila tidak segera ditangani. Permintaan saat ini untuk produksi insulin sudah melampaui kapasitas industri. Industri sudah diprediksikan akan perlu meningkatkan kapasitasnya untuk mengatasinya perkiraan permintaan yang banyak di masa depan. Namun, fasilitas produksi untuk insulin mahal dan biasanya membutuhkan waktu lama untuk membuatnya. Oleh karena itu, ada kebutuhan akan metode produksi lebih murah dan dapat mengurangi waktu yang dibutuhkan untuk meningkatkan produksi.

Beberapa metode produksi insulin dari bakteri dan hewan telah digunakan. Namun, hal ini seringkali menimbulkan risiko kesehatan akibat dari penyakit. Risiko tersebut dapat timbul dari kontaminasi silang dengan penyakit yang dapat

mempengaruhi hewan dan pengguna akhir. Dengan demikian, ada kebutuhan metode produksi yang dapat mengeliminasi kemungkinan kontaminasi silang antara produksi oleh organisme dan pengguna akhir. Selain itu, metode konvensional menghasilkan insulin dari E.coli diprediksikan tidak mencukupi peningkatan permintaan akan insulin di masa depan, sehingga diperlukan metode tambahan sebagai salah satu solusi.

Selain itu, banyak metode produksi saat ini memerlukan pemrosesan ekstensif untuk mengekstrak protein terapeutik dari hewan atau organisme lainnya di mana ia berada yang dapat diproduksi dan membuat senyawa tersebut berada dalam kondisi yang dapat dimanfaatkan oleh pasien. Setelah pemurnian, protein dapat digabungkan dengan bahan pembantu atau bahan pembawa lainnya untuk menstabilkan protein sehingga dapat digunakan oleh pasien. Namun, proses ekstraksi, pemurnian, resuspensi antara lain terlibat dengan proses pengolahannya protein yang rumit, tidak praktis, dan mungkin tidak kondusif untuk digunakan di negara-negara terbelakang yang membutuhkannya secara umum. Oleh karena itu, ada kebutuhan akan metode produksi yang mudah dapat mengurangi proses ekstraksi dari protein yang diinginkan.

Tanaman lidah buaya transgenik dapat memberikan alternatif yang layak secara ekonomi untuk produksi protein yang diinginkan, contohnya insulin, kosmetik, dan lain-lain. Selain itu, protein yang diinginkan mungkin di lokalisasi dan dipekatkan di dalam gel daun Aloe vera. Lokalisasi protein yang diinginkan ini dapat menyederhanakan proses ekstraksi ataupun menghilangkan protein yang tidak diinginkan. Dengan demikian, penelitian ini dapat memberikan informasi mengenai Aloe vera transgenik, dimana protein yang diinginkan umumnya lebih banyak dan mudah diakses dibandingkan tanaman transgenik lainnya, seperti tembakau dan jagung. Selanjutnya, penemuan ini dapat memberikan metode yang efisien untuk isolasi protein.

Tanaman lidah buaya dapat menawarkan berbagai keunggulan dibandingkan sistem bakteri dan tanaman konvensional lainnya untuk menghasilkan protein yang diinginkan. Manfaat tanaman Aloe vera termasuk kemampuan mengolah protein dengan cara yang sederhana. Terkadang pada bakteri dan ragi ditemukan ketidaksesuaian dengan protein yang diinginkan, karena kurangnya sistem

modifikasi posttranslational, dimana sistem ini terdapat pada tanaman. Proses ini dapat mencakup modifikasi kimiawi, seperti dengan glikosilasi, dan pelipatan beberapa protein. Selanjutnya, dibandingkan dengan metode produksi protein lainnya, pada sel hewan, produksi tanaman lidah buaya dapat menawarkan keuntungan biaya yang signifikan, keuntungan Skalabilitas, dan pengurangan risiko kontaminasi yang mungkin berbahaya bagi manusia. Keuntungan lainnya menggunakan *Aloe vera* yaitu kemudahan dalam ekstraksi, karena terdapat lokalisasi protein yang diinginkan ke dalam gel *Aloe vera*, sehingga lebih ekonomis dibandingkan menggunakan tumbuhan lain.

*Aloe* dianggap sebagai kandidat pertama untuk penelitian di abad baru ini, karena berpotensi secara ekonomi dan sebagai obat<sup>[2]</sup>. Sebelumnya, kebanyakan penelitian *Aloe* berfokus pada analisis biochemical<sup>[3]</sup> dan obat-obatan atau aplikasi kosmetik dari metabolit sekunder secara efektif<sup>[4]</sup>. Selain itu, *Aloe vera* juga menjadi topik penelitian tanaman transgenik.

Pada penelitian ini, kami akan melakukan pendekatan in-silico, yaitu analisis komputasi data dan desain eksperimen digital untuk mengembangkan *Aloe vera* transgenik sebagai platform ekspresi dan kloroplas sebagai sistem ekskresi dengan menggunakan teknik transformasi *particle bombardment* yang diharapkan akan menghasilkan produk signifikan pada *Aloe vera* transgenik. Kehadiran insulin (protein) akan dianalisis menggunakan ELISA, western blot, dan lain lain. Namun pada penelitian ini, kami akan berfokus pada simulasi in-silico berupa computing, pengolahan dan analisis data, serta rancangan penelitian in-vivo yang diharapkan nantinya dapat menjadi bahan pertimbangan sebelum dan perbandingan setelah eksperimen in vivo.

## 1.2 Rumusan Masalah

Dengan mempertimbangkan latar belakang yang telah dikemukakan maka beberapa masalah yang dapat dirumuskan dari penelitian ini adalah :

1. Bagaimana potensi *Aloe vera* sebagai alat produksi insulin?
2. Bagaimana sistem produksi insulin di tanaman *Aloe vera* dapat dikembangkan?
3. Bagaimana cara mengkonstruksi desain eksperimen *Aloe vera* transgenik?

4. Bagaimana purifikasi dan pematangan insulin hasil produksi dari tanaman *Aloe vera*?
5. Apakah penelitian ini memungkinkan untuk dipelajari secara analisis komputasi?
6. Bagaimana prediksi produksi insulin di *Aloe vera*?

### **1.3 Batasan Masalah**

Adapun batasan masalah pada penelitian ini adalah :

1. Penelitian ini berupa desain *in silico* (komputasi) dan belum diadakan uji coba secara *in vitro*.
2. Penelitian ini hanya mengembangkan teknik produksi insulin, dan belum melakukan pengujian untuk keefektifan pada manusia, baik secara *in silico*, maupun *in vivo*.

### **1.4 Tujuan Penelitian**

1. Memberikan alternatif lain produksi insulin dengan menggunakan tanaman/*Aloe vera*.
2. Memberikan saran konstruksi plasmid binary untuk produksi insulin di *Aloe vera*.
3. Menghasilkan data secara komputasi yang dapat dianalisis sebagai bahan pertimbangan dalam eksperimen digital.
4. Mendesain eksperimen secara digital yang efektif, efisien, ekonomis, aman, dan dapat diterima.
5. Menghasilkan data analisis komputasi dan desain eksperimental digital sebagai acuan dalam melakukan penelitian lanjutan baik secara *in vivo* maupun *in vitro* yang dapat mengembangkan *Aloe vera* transgenic yang dapat menghasilkan insulin.
6. Melakukan penelitian secara *in vivo* dan *in vitro* dengan mempertimbangkan hasil analisis *in silico*.
7. Diharapkan dapat menjadi salah satu solusi dalam memenuhi peningkatan permintaan insulin yang lebih aman dan ekonomis.
8. Setelah mampu memenuhi permintaan insulin, diharapkan terobosan baru dalam pengiriman insulin seperti insulin oral dapat menjadi salah satu

metode terapi insulin selain injeksi, sehingga akan menghasilkan kepatuhan dalam terapi yang dapat menghindari komplikasi sekunder dari diabetes.

## 1.5 Kebaruan

Beberapa ilmuwan sudah meneliti produksi proinsulin pada tanaman transgenik sebelumnya, contohnya pada tembakau yang dimediasi *Agrobacterium tumefaciens*. Namun, pada penelitian ini kami meneliti desain *in silico* produksi proinsulin dari *Aloe vera* transgenik dengan target ekspresi kloroplas menggunakan metode particle bombardment (*biolistics*). Penggunaan *Aloe vera* dan sistem transformasi ini berdasarkan berbagai pertimbangan.

## 1.6 Hipotesis

1. *Aloe vera* transgenik mampu menghasilkan insulin.
2. Hasil yang diperoleh didapat dapat digunakan untuk mendesain eksperimen secara *in vivo* ataupun *in vitro* dalam menghasilkan insulin dari *Aloe vera* transgenik.
3. Desain eksperimen dapat diaplikasikan di dunia nyata dan dapat menghasilkan insulin yang diharapkan dapat mendukung produksi insulin sebagai bagian dari pemenuhan kebutuhan insulin yang kian meningkat.

## 1.7 Manfaat Penelitian

Dapat memberikan metode alternatif seperti desain konstruksi plasmid untuk produksi insulin dengan menggunakan tanaman/*Aloe vera* sehingga dapat berdampak di berbagai bidang.

1. Dari segi kesehatan : - Insulin yang dihasilkan diharapkan mampu memenuhi kebutuhan penderita diabetes dan memiliki efektifitas yang sama atau bahkan lebih baik dari cara produksi konvensional.
2. Dari segi ekonomi : - Insulin dari *Aloe vera* transgenik diharapkan memiliki harga yang lebih ekonomis dibandingkan harga insulin konvensional.
3. Dari segi IPTEK. : - Terobosan terobosan baru dalam metode pemberian insulin, seperti oral insulin, membutuhkan insulin dalam dosis tinggi. Sehingga produksi insulin transgenik ini diharapkan dapat memenuhi meningkatnya permintaan insulin.

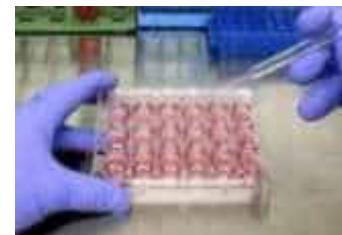
## BAB II

### Tinjauan Pustaka

#### 2.1 Metode Penelitian

Jenis eksperimen :

1. In Vivo : (Di dalam makhluk hidup). Eksperimen dilakukan di dalam tubuh makhluk hidup.
2. In vitro : (Di dalam glass). Eksperimen meliputi sel atau biomolekul yang ada dan dikontrol oleh lingkungan luar.
3. In silico : (In silico = chips) Eksperimen yang dilakukan dengan bantuan model dan simulasi komputer<sup>[24]</sup>



Gambar a

Gambar b

Gambar c

Gambar 2.1.1 Perbandingan (a)In vivo, (b)In vitro, dan (c)In silico

Setelah melakukan berbagai pertimbangan, kami memutuskan untuk melakukan simulasi in silico. Beberapa hal yang kami pertimbangkan meliputi:

1. In silico digunakan sebagai landasan percobaan in vivo dan in vitro, serta untuk menguji hipotesis sebelum melakukan eksperimen langsung di dunia nyata.
2. Waktu yang dibutuhkan pada in vivo dan in vitro lebih panjang daripada in silico, karena in silico berupa simulasi, sedangkan in vivo dan in vitro adalah eksperimen langsung

- Karena hanya simulasi, maka biaya yang dibutuhkan pada eksperimen *in silico* relatif lebih sedikit

Tabel 2.1.1: Perbandingan *in-silico*, *in-vitro* dan *in-vivo* dalam penelitian sunscreen

	<i>in vivo</i>	<i>in vitro</i>	<i>in silico</i>
Purpose of use	For special claims	Development, evaluation of sunscreen formulations	Development, evaluation of sunscreen formulations
Principle	Action spectrum of PPD	Transmission measurement, possibly action spectra of MED and PPD	Transmission calculation with skin roughness model action spectra of MED and PPD
Reliability	medium	low	medium
Time	high	medium	low
Costs	high	medium	low
Diffusion	medium	high	medium
Limitations	No consideration of interaction with UV-B (as the other <i>in vivo</i> or <i>in vitro</i> methods).	Uncertainties introduced by transmission measurement (no biological substrate)	Sunscreen formulations may not behave like simulation. Not available for all types of products
Conclusion	Expensive and time consuming method. Recommended for labelling of products and information of consumers.	Medium cost method. Usefulness for industrial development and labelling of products depending on the type of <i>in vitro</i> method.	Fast and low cost method. Useful for industrial development of sunscreen products.

Comparison of *in vivo*, *in vitro* and *in silico* methods for UV-A assessment.

## 2.2 Diabetes

Diabetes adalah penyakit kronis yang terjadi ketika tubuh tidak dapat memetabolisme gula, sehingga menyebabkan kadar gula dalam darah meningkat. Kondisi ini dapat disebabkan oleh dua hal, yang pertama pankreas tidak lagi mampu memproduksi insulin, dan yang kedua adalah ketika tubuh tidak mampu memanfaatkan insulin yang dihasilkannya dengan baik. (International Diabetes Federation).

### 2.2.1 Tipe Diabetes

#### 1. Diabetes Tipe 1

Dapat berkembang pada usia berapa pun, tetapi paling sering terjadi pada anak-anak dan remaja. Ketika Anda menderita diabetes tipe 1, tubuh Anda memproduksi sangat sedikit atau tidak ada insulin, yang berarti bahwa Anda

memerlukan suntikan insulin setiap hari untuk menjaga kadar glukosa darah tetap terkendali.

## **2. Diabetes Tipe 2**

Lebih umum pada orang dewasa dan menyumbang sekitar 90% dari semua kasus diabetes. Ketika Anda menderita diabetes tipe 2, tubuh Anda tidak menggunakan insulin yang dihasilkannya dengan baik. Landasan pengobatan diabetes tipe 2 adalah gaya hidup sehat, termasuk peningkatan aktivitas fisik dan diet sehat. Namun, seiring waktu sebagian besar orang dengan diabetes tipe 2 akan membutuhkan obat-obatan oral dan / atau insulin untuk menjaga kadar glukosa darah mereka terkendali.

## **3. Gestational diabetes (GDM)**

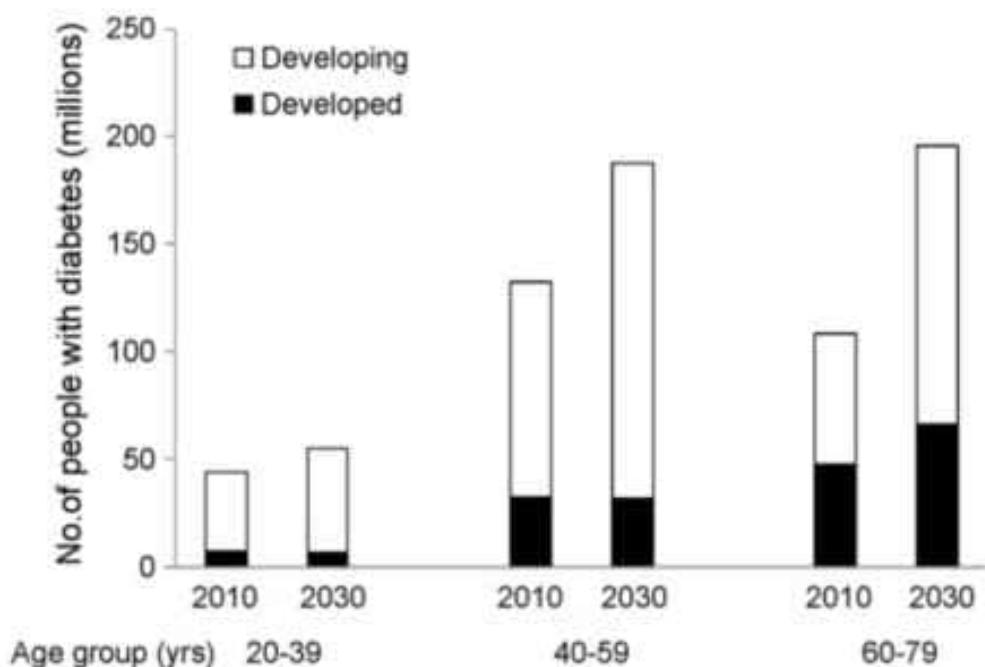
Jenis diabetes yang terdiri dari glukosa darah tinggi selama kehamilan dan dikaitkan dengan komplikasi pada ibu dan anak. GDM biasanya menghilang setelah kehamilan, tetapi wanita yang terkena dan anak-anak mereka berisiko lebih tinggi terkena diabetes tipe 2 di kemudian hari. (IDF International Diabetes).

### **2.2.2 Jumlah Penderita Diabetes**

Secara global, diperkirakan 422 juta orang dewasa hidup dengan diabetes pada tahun 2014, dibandingkan dengan 108 juta pada tahun 1980. Prevalensi diabetes di dunia (dengan usia yang distandarisasi) telah meningkat hampir dua kali lipat sejak tahun 1980, meningkat dari 4,7% menjadi 8,5% pada populasi orang dewasa. Hal ini mencerminkan peningkatan faktor risiko terkait seperti kelebihan berat badan atau obesitas. Selama beberapa dekade terakhir, prevalensi diabetes meningkat lebih cepat di negara berpenghasilan rendah dan menengah daripada di negara berpenghasilan tinggi.

Diabetes menyebabkan 1,5 juta kematian pada tahun 2012. Gula darah yang lebih tinggi dari batas maksimum mengakibatkan tambahan 2,2 juta kematian, dengan meningkatkan risiko penyakit kardiovaskular dan lainnya. Empat puluh tiga persen (43%) dari 3,7 juta kematian ini terjadi sebelum usia 70 tahun. Persentase kematian yang disebabkan oleh diabetes yang terjadi sebelum usia 70 tahun lebih tinggi di negara-negara berpenghasilan rendah

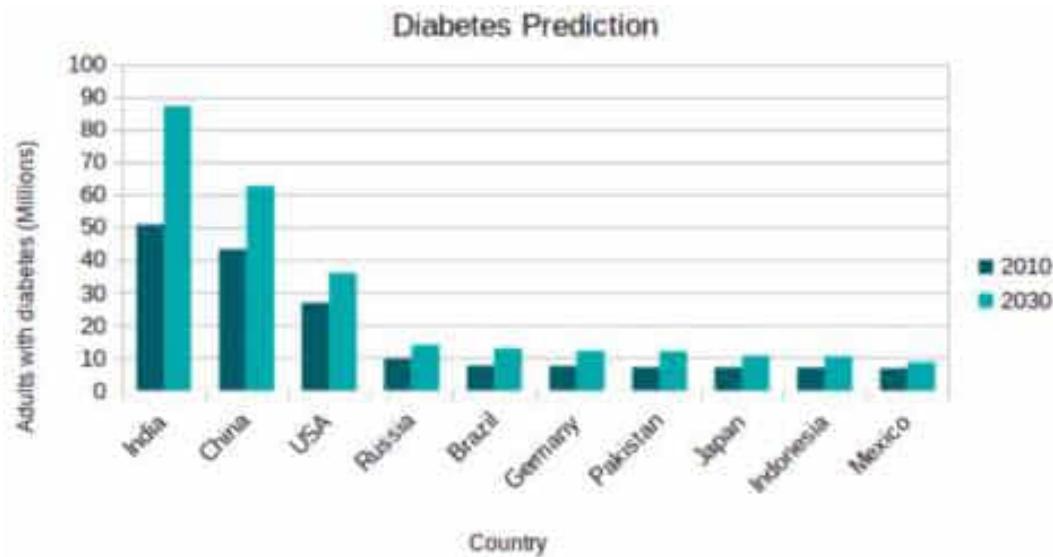
dan menengah daripada di negara-negara berpenghasilan tinggi. (WHO Global Report, 2016).



Grafik 2.2.1. Penderita diabetes pada negara maju dan berkembang pada tahun 2010 dan estimasi penderitanya di tahun 2030. (Sumber : Diabetes Research and Clinical Practice 87 (2010) 4-14)

Tabel 2.2.1: 10 negara teratas dengan penderita diabetes berusia 20-79 pada tahun 2010 dan perkiraan pada 2030.

Table 2.2.1 – Top 10 countries for numbers of people aged 20–79 years with diabetes in 2010 and 2030.				
	2010		2030	
	Country	No. of adults with diabetes (millions)	Country	No. of adults with diabetes (millions)
1	India	59.8	India	87.0
2	China	43.2	China	62.6
3	USA	26.8	USA	36.0
4	Russian Federation	9.6	Pakistan	11.8
5	Brazil	7.6	Brazil	12.7
6	Germany	7.5	Indonesia	12.0
7	Pakistan	7.1	Mexico	11.9
8	Japan	7.1	Bangladesh	10.4
9	Indonesia	7.0	Russian Federation	10.3
10	Mexico	6.8	Egypt	8.6



Grafik 2.2.2: Perbandingan 10 Negara dengan penderita diabetes terbanyak pada 2010 dan 2030.

(Sumber : Diabetes Research and Clinical Practice 87 (2010) 4-14.)

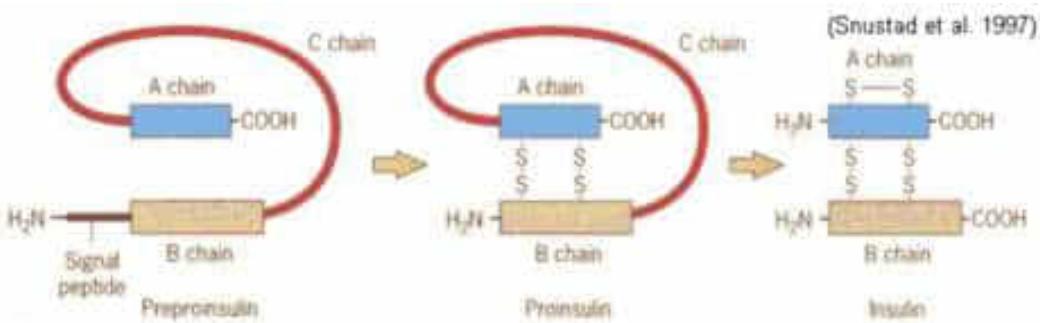
Terlihat bahwa Indonesia menduduki peringkat ke 9 dengan penderita diabetes umur 20-79 terbanyak pada 2010 dan diperkirakan Indonesia akan menduduki peringkat ketiga 3 pada tahun 2030. Selain itu, juga terdapat estimasi peningkatan kasus diabetes di setiap negara.

### 2.3 Insulin

Insulin merupakan hormon dari protein yang berfungsi menurunkan kadar gula darah dengan cara merubah kandungan gula dalam darah menjadi glikogen yang disimpan di hati.

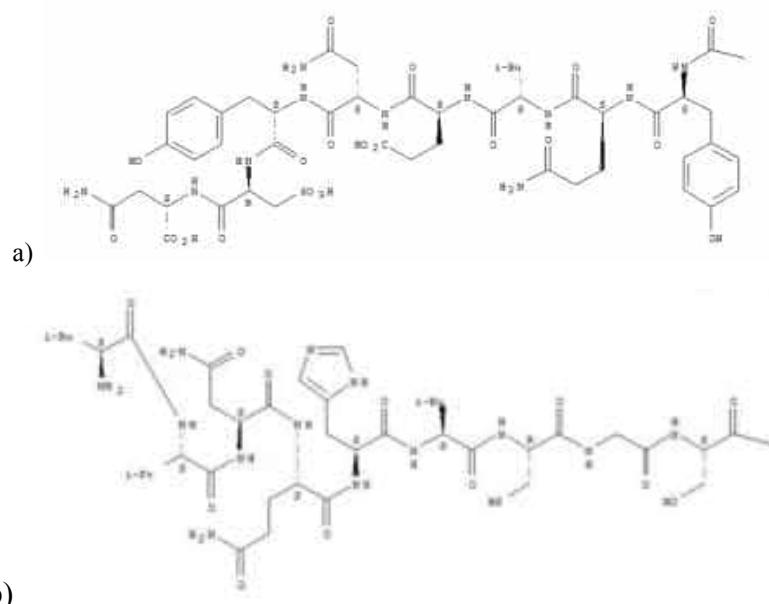
Insulin dihasilkan oleh sel beta pulau langerhans di pankreas, dan disekresikan saat kadar gula darah lebih tinggi dari biasanya. Insulin bekerja berlawanan dengan glukagon, hormon yang berfungsi untuk menaikkan gula darah dengan merubah glikogen dalam hati menjadi glukosa dan diedarkan dalam darah. Insulin pertama kali diisolasi oleh Frederick Banting dan Charles H. Best dari pankreas anjing pada 1921 (wikipedia).

Insulin pada awalnya terbentuk sebagai preproinsulin, rantai panjang yang terdiri dari 110 asam amino. Penghilangan signal peptida dari ujung preproinsulin menghasilkan proinsulin. Pembentukan ikatan disulfida antar rantai A dan B, serta dihilangkannya rantai C membentuk insulin yang aktif secara biologis. Insulin aktif terdiri dari 2 rantai asam amino, rantai A (21 asam amino), dan rantai B (30 asam amino) yang dihubungkan oleh ikatan peptida.



Gambar 2.3.1 Modifikasi *Post-translational* insulin

PAGE 1-A

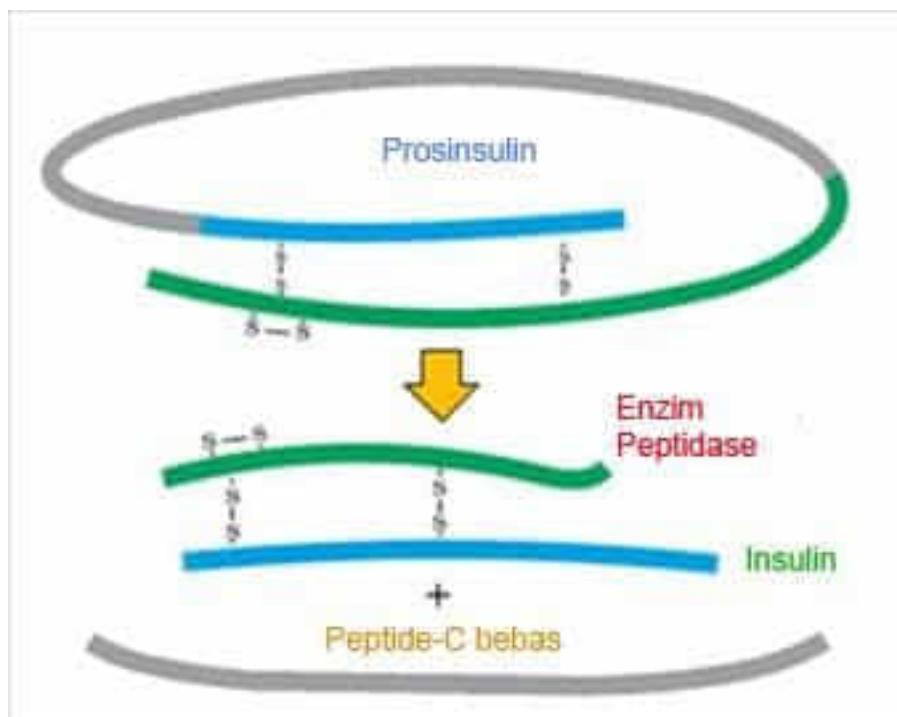


Gambar 2.3.2 a) Insulin chain A oxidized b) Insulin chain B oxidized, lookchem.com



Gambar 2.3.3 Struktur asam amino insulin<sup>[53]</sup>

### 2.3.1 Sintesis Insulin

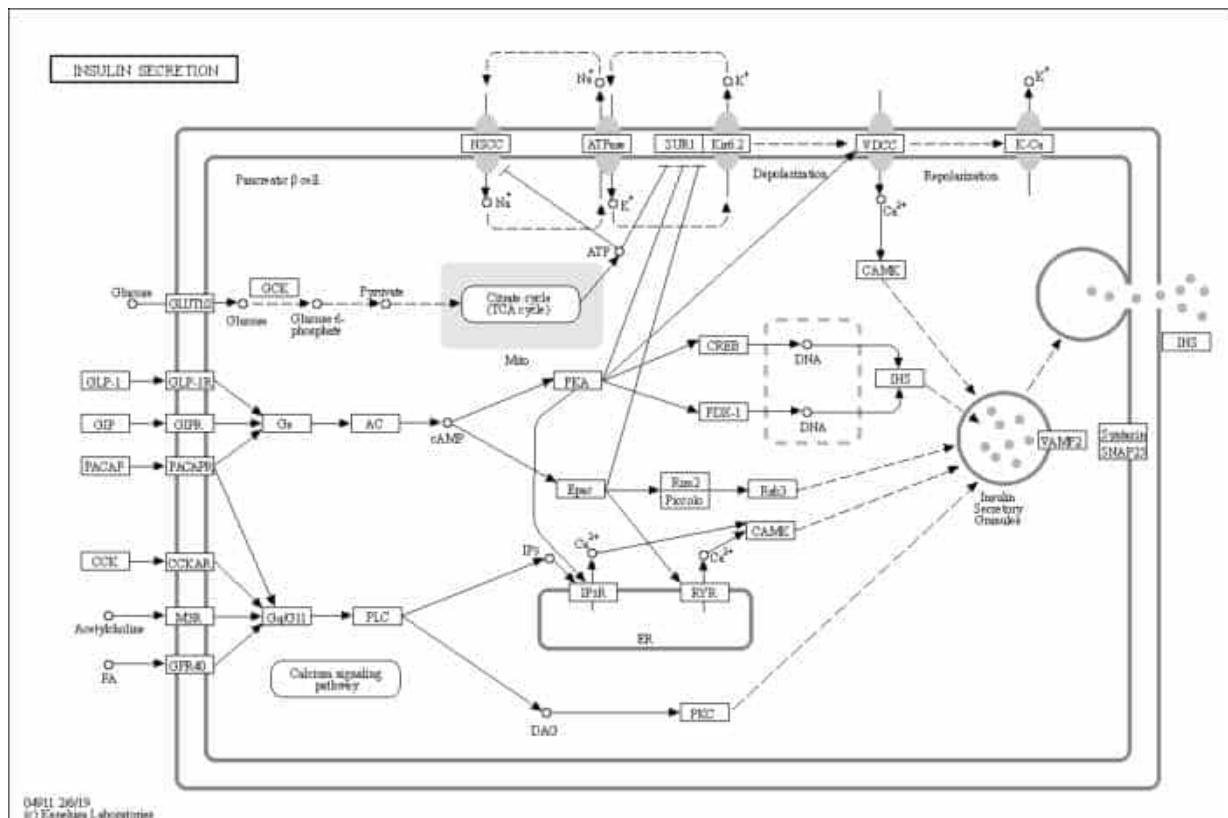


(Gambar 2.3.4 Sintesis Insulin)

Tahapan sintesis insulin adalah sebagai berikut :

1. Sintesis insulin dimulai dalam bentuk preproinsulin (precursor hormon insulin) pada retikulum endoplasma sel beta.
2. Dengan bantuan enzim peptidase, preproinsulin mengalami pemecahan sehingga terbentuk proinsulin, yang kemudian dihimpun dalam gelembung gelembung dalam sel tersebut.
3. Proinsulin kemudian diurai kembali oleh enzim peptidase menjadi insulin dan peptide-C yang keduanya sudah siap disekresikan secara bersamaan melalui membran sel.
4. Fungsi insulin sangat dibutuhkan dalam proses utilisasi glukosa yang ada dalam darah. Kadar glukosa darah yang meningkat merupakan komponen utama yang memberi rangsangan terhadap sel beta dalam memproduksi insulin<sup>[191, 192]</sup>.

### 2.3.2 Jalur Sekresi Insulin detail



(Gambar 2.3.5 Jalur sekresi insulin) Sumber : KEGG (Kyoto Encyclopedia of Gene and Genome)

Sel beta pankreas adalah sel endokrin khusus yang terus-menerus mengatur kadar gula darah dan bahan bakar lainnya, dan sebagai respon, mengeluarkan insulin untuk mempertahankan homeostasis bahan bakar normal. Sekresi insulin yang diinduksi glukosa dan potensinya merupakan mekanisme utama pelepasan insulin. Glukosa diangkut oleh transporter glukosa (GLUT) ke dalam sel beta pankreas. Metabolisme glukosa menghasilkan ATP, yang menghambat saluran K<sup>+</sup> yang sensitif terhadap ATP dan menyebabkan masuknya Ca<sup>2+</sup> yang bergantung pada tegangan. Peningkatan [Ca<sup>2+</sup>] i memicu pelepasan butiran insulin secara eksositosis.

Sekresi insulin diatur lebih lanjut oleh beberapa hormon dan neurotransmitter. Hormon peptida, seperti glucagon-like peptide 1 (GLP-1), meningkatkan level cAMP dan dengan demikian memicu sekresi insulin melalui aksi gabungan PKA dan Epac2. Acetylcholine (ACh), neurotransmitter parasimpatis utama, mengikat reseptor Gq-coupled dan mengaktifkan fosfolipase C- (PLC-), dan efek stimulasi melibatkan aktivasi protein kinase C

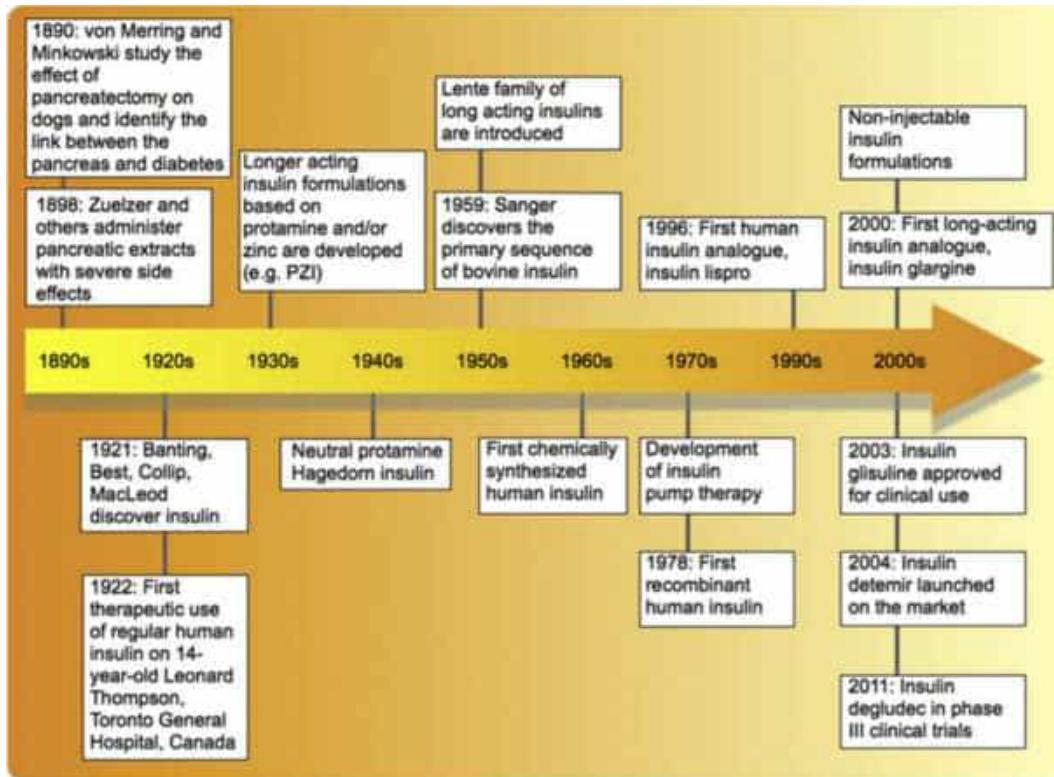
(PKC), yang merangsang eksositosis. Selain itu, ACh memobilisasi Ca<sup>2+</sup> + intraseluler dengan aktivasi reseptor IP<sub>3</sub>. (KEGG (Kyoto Encyclopedia of Gene and Genome)). Insulin juga merupakan senyawa protein yang bersifat mudah dicerna dan rusak saat terkena asam lambung.

### 2.3.3 Produksi Insulin Konvensional

Sebelum 1980-an, sebagian besar produk insulin di pasaran berasal dari pankreas babi(porcine) atau sapi(bovine). Produk insulin ini sangat tidak murni dan dapat menyebabkan komplikasi imunologis seperti alergi insulin, immune-mediated lipodystrophy di lokasi injeksi, dan resistensi insulin yang dimediasi antibodi. Produk insulin ini juga menyebabkan variasi signifikan pada pharmakokinetik dan pharmakodinamik<sup>[54]</sup>.

Seiring berjalannya waktu, perkembangan teknologi insulin menghasilkan insulin dengan kualitas yang lebih baik dan kinerja yang lebih konsisten. Selain itu, long-acting insulin mulai dikembangkan untuk memperpanjang waktu kerja insulin dan mengurangi jumlah injeksi harian. Long-acting insulin tersebut dibentuk dengan menggabungkan zinc dan/atau protamin menjadi protamine insulin dan protamine zinc insulin (PZI) dan mulai dikembangkan pada 1930-an<sup>[54-56]</sup>. Selain itu, NPH (Isophane Neutral Protamine Hagedorn) juga mulai diproduksi pada 1940-an<sup>[57]</sup>, dan masih digunakan sampai sekarang baik digunakan sendiri maupun dicampurkan dengan produk insulin lain<sup>[59]</sup>.

Pada 1950, insulin trilogi “lente” mulai diintroduksikan<sup>[58]</sup> dan Sanger berhasil mendapatkan primary structure insulin bovine(sapi). Seluruh kemajuan ini berujung pada lahirnya insulin hewan murni berkualitas tinggi (monokomponen maupun single-peak insulin) pada 1970<sup>[61]</sup>.



(Gambar 2.3.6 Sejarah insulin<sup>[54]</sup>)

### 2.3.4 Produksi Insulin Rekombinan

Pada 1970-an, teknologi DNA rekombinan mengalami perkembangan pesat yang berujung pada lahirnya insulin manusia rekombinan pada 1978<sup>[62]</sup>.

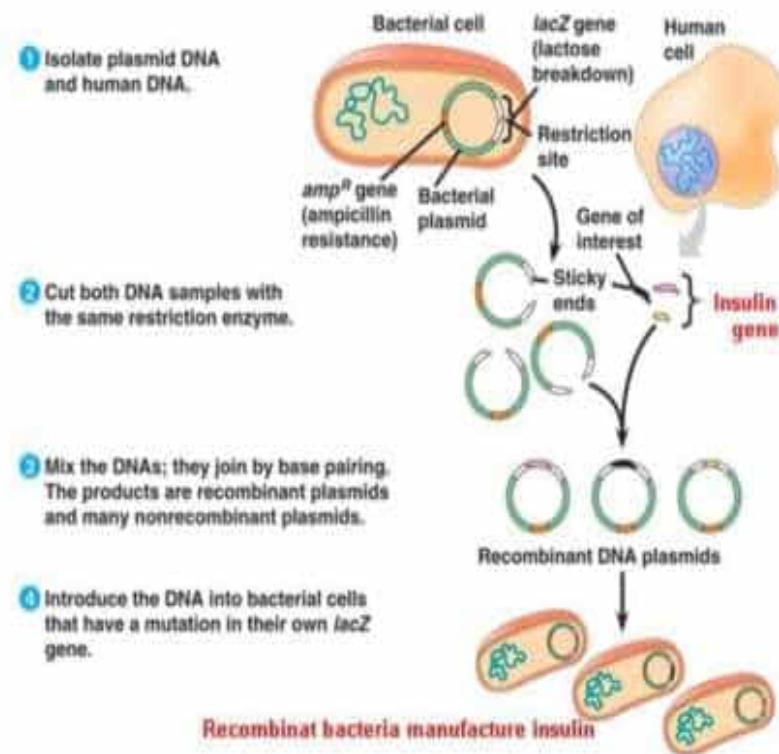
#### *Escherichia coli*

Diantara prokariota, *Escherichia coli* lebih diminati untuk produksi insulin rekombinan karena keuntungan yang ditawarkan. Beberapa keuntungan tersebut termasuk tingkat pertumbuhan tinggi, persyaratan media sederhana, kemudahan dalam penanganan, hasil produksi tinggi, dan sangat hemat biaya.

Bioengineering *E. coli* dimulai pada tahun 1978, dan dilakukan melalui penggunaan plasmid, pBR322 (Goeddel et al., 1979). Plasmid adalah bagian kecil dari DNA yang melayang di sitoplasma, dan tidak penting bagi sel (Madigan et al., 2015). Karena mereka tidak esensial dan sangat kecil, mereka sering dilepas dan diambil oleh prokariota lain melalui proses transformasi, yang memungkinkan mikroba untuk beradaptasi dengan

mengekspresikan gen yang berbeda. Fenomena ini merupakan cara bagi beberapa bakteri untuk mengembangkan resistensi genetik.

Fenomena ini membuka peluang besar bagi peneliti untuk memodifikasi bakteri menjadi pabrik produksi berbagai macam protein dan zat kimia lainnya. Salah satunya adalah tahun 1978 ketika peneliti mampu merubah strain *E. coli* untuk memproduksi insulin manusia. Dengan menggunakan plasmid pBR322, peneliti berhasil memotong urutan spesifik pasangan basa DNA insulin manusia dan plasmid vector menggunakan enzim restriksi (EcoRI), serta menyambungkan gen insulin manusia sintetis ke dalam plasmid. Ketika plasmid ditempatkan ke dalam *E. coli*, bakteri mengambil plasmid; ini memungkinkan bakteri untuk mereplikasi plasmid dan mengkode insulin manusia. Melalui model ini, bakteri terus menerus membuat plasmid baru yang mengandung gen insulin. Skema umum tentang bagaimana transformasi ini terjadi dapat dilihat pada gambar 2.6<sup>[6]</sup>.



(Gambar 2.3.7 Transformasi bakteri transgenik)

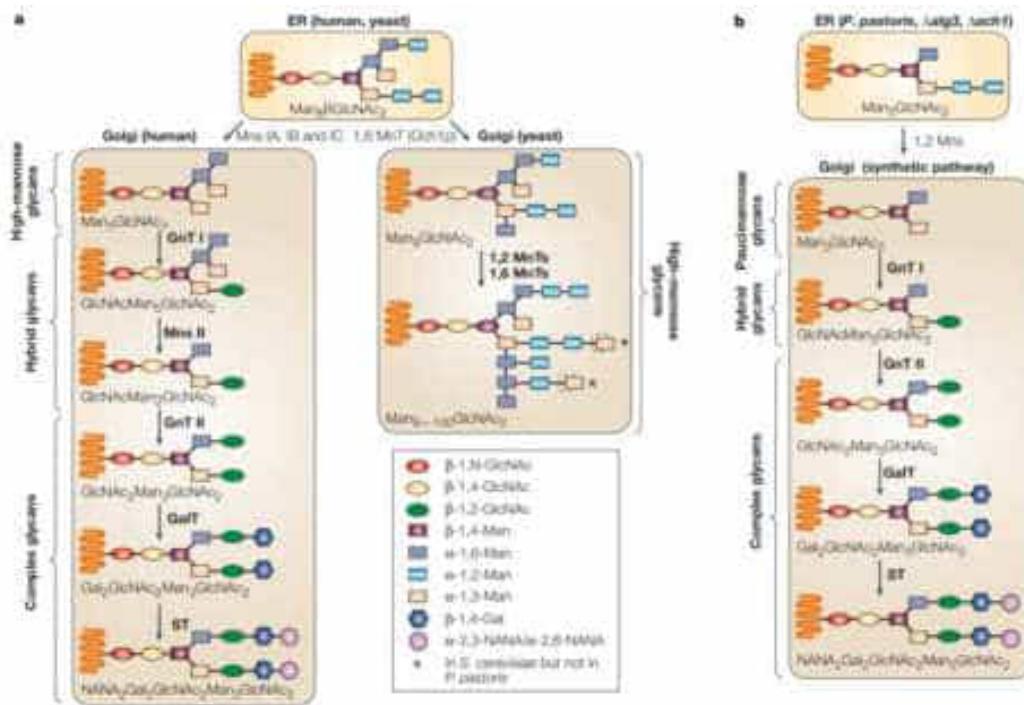
Terlepas dari segala kelebihan yang ditawarkan *E. Coli*, sistem ekspresi *E. Coli* juga memiliki kelemahan. Contohnya adalah hilangnya sifat plasmid, penginduksi yang tidak diinginkan untuk ekspresi gen, akumulasi

protein intraseluler heterolog sebagai badan inklusi, refolding protein yang tidak tepat, kurangnya modifikasi pasca translasi (termasuk pembentukan ikatan disulfida), beban metabolisme dan stress yang dimediasi protein, kontaminasi endotoksin, sekresi yang buruk, pencernaan proteolitik, serta kompleksitas pada proses hilir<sup>[26-27]</sup>.

### Ragi

Ragi adalah inang yang disukai untuk ekspresi berbagai macam protein heterolog yang memerlukan modifikasi pasca-translasi untuk aktivitas biologisnya. Ragi memiliki kemampuan untuk melakukan banyak modifikasi pasca-translasi seperti fosforilasi, glikosilasi terkait-O, glikosilasi terkait-N, asetilasi dan asilasi. Protein rekombinan diekspresikan dalam bentuk larut dan terlipat dengan baik dalam bentuk aktif secara fungsional. Produksi biofarmasi menggunakan sistem ekspresi ragi juga sangat hemat biaya dan dapat ditingkatkan dengan menggunakan bioreaktor besar.

Namun, satu perhatian utama dalam produksi glikoprotein terapeutik untuk manusia adalah jalur N-glikosilasi ragi merupakan tipe manosa tinggi, yang memberikan waktu paruh pendek in vivo dan hiper imunogenisitas yang mengakibatkan glikoprotein terapeutik kurang efektif. Berbagai upaya telah dilakukan untuk merubah jalur N-glikosilasi ragi untuk menghasilkan glikoprotein terapeutik dengan struktur glikosilasi N yang sesuai dengan manusia (humanisasi)<sup>[29, 35]</sup>. Contoh ragi yang digunakan adalah *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Pichia pastoris* dan sering juga digunakan untuk menghasilkan protein rekombinan<sup>[93-96]</sup>.



(Gambar 2.3.8: Jalur utama N-glycosylation pada ragi dan manusia. **a)** Perbandingan jalur N-glycosylation manusia (kiri) dengan ragi (kanan). **b)** Jalur sintetis pada ragi yang dibuat menggunakan assembly oligosakarida mutan.)

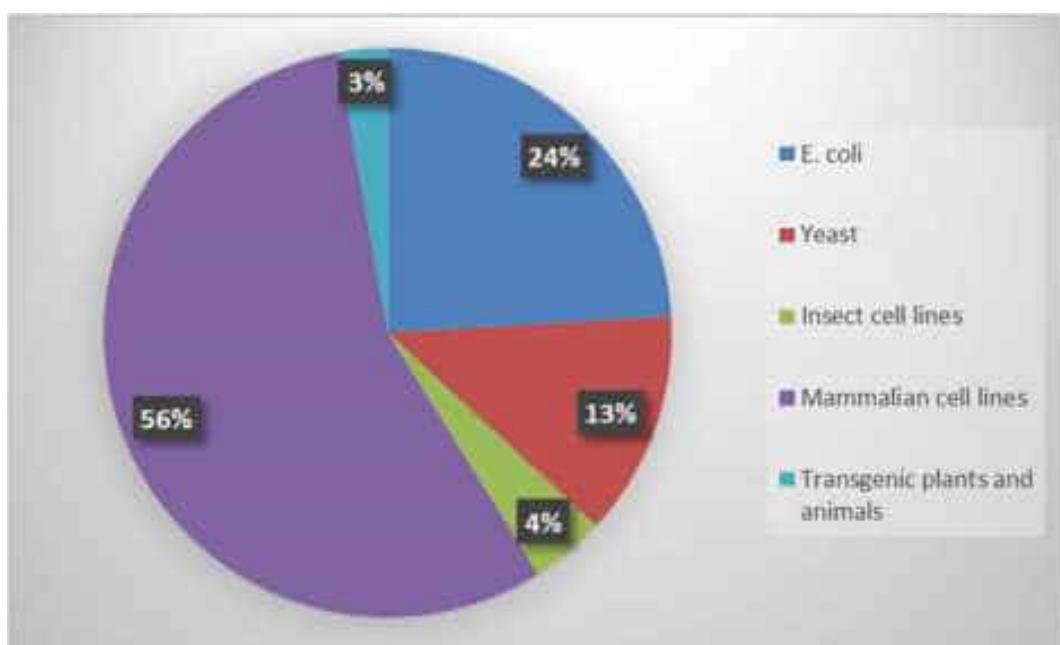
## Tanaman Transgenik

Tanaman transgenik telah lama digunakan sebagai alat produksi protein rekombinan, karena kelebihan kelebihan yang mereka tawarkan. Beberapa kelebihan ini meliputi efektivitas biaya, pemrosesan protein berkualitas tinggi, ketiadaan patogen manusia, kemudahan produksi, serta adanya sistem eukaryotik untuk modifikasi post translasi. Contoh protein yang berhasil diproduksi dengan tanaman transgenik adalah: Growth hormone<sup>[30]</sup>; Virus hepatitis-B surface antigen; antibodi; protein industri, serta beberapa protein susu.

Insulin manusia rekombinan telah berhasil diekspresikan dan diproduksi dari oil body *Arabidopsis thaliana*. Oil Body merupakan organel yang terdapat pada oilseed dan terdiri dari inti triacylglycerol hidrophobik yang dibungkus oleh membran fosfolipid dan dinding luar protein bernama oleosin. Selain itu, insulin rekombinan juga sudah berhasil diekspresikan dan dipanen dari tembakau dan selada.

Secara keseluruhan, tanaman transgenik menyimpan potensi tinggi untuk produksi insulin berkapasitas tinggi dan hemat biaya. Tingkat ekspresi proinsulin tinggi dalam daun dengan stabilitas jangka panjang menawarkan teknologi berbiaya rendah untuk injeksi maupun oral insulin. Selain itu, sel tanaman transgenik dapat digunakan sebagai tempat penyimpanan alami bagi insulin rekombinan<sup>[36]</sup>.

### 2.3.5 Perbandingan Persentase Produksi Insulin Rekombinan



(Gambar 2.3.9: Perbandingan sumber produksi insulin rekombinan.)

### 2.3.6 Ekstraksi Protein

Insulin merupakan hormon yang terbuat dari protein. Analisa berskala besar terhadap protein disebut sebagai analisa proteomik dan biasanya digunakan untuk mengidentifikasi protein<sup>[116, 117]</sup>. Analisa proteomik pertama kali dikembangkan pada 1975 setelah ditemukannya gel dua dimensi, dan dipetakannya protein E. coli. Beberapa contoh analisis proteomik akan dijelaskan di akhir bab 2.

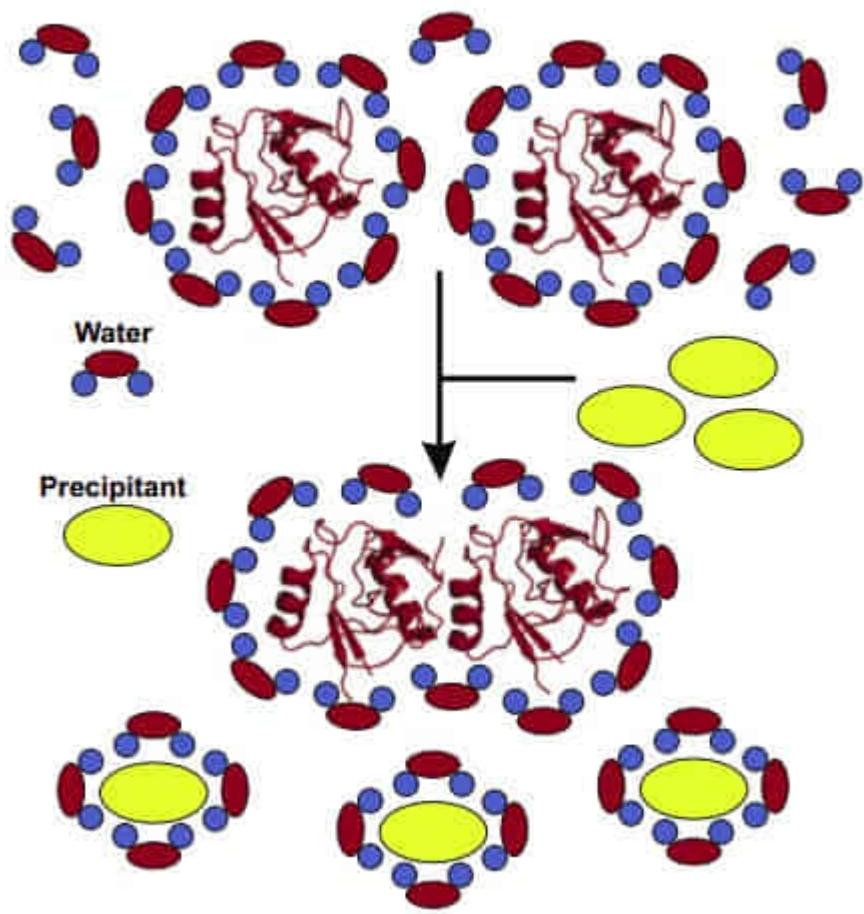
Analisis proteomik yang sukses bergantung pada sampel protein yang bagus<sup>[108]</sup>. Terdapat beberapa hal yang perlu diperhatikan saat mempersiapkan sampel protein untuk analisis proteomik. Dua diantaranya adalah kadar kontaminan dan kadar degradasi protein. Walaupun kontaminan dalam jumlah kecil dapat ditoleransi, beberapa protein tidak stabil, dan kontaminasi dapat

mengganggu analisa proteomik<sup>[108]</sup>. Sedangkan degradasi protein adalah rusaknya struktur primer protein, dan menyebabkan protein terpecah menjadi satuan asam amino, atau rantai peptida yang lebih kecil.

Perlu diingat bahwa denaturasi protein tidak sama dengan degradasi protein. Denaturasi protein hanya merusak struktur sekunder, tersier, dan kuarter protein. Hal ini hanya merubah struktur tiga dimensi protein dan menjadikannya rantai peptida linear. Denaturasi tentunya menghilangkan sifat biologis protein, sehingga protein tidak dapat digunakan untuk analisis yang memerlukan sifat biologis aktifnya<sup>[115]</sup>.

Di sisi lain, degradasi protein adalah rusaknya struktur primer protein akibat terputusnya ikatan peptida antar asam amino. Karena degradasi protein memecah protein menjadi rantai peptida yang lebih kecil, analisis proteomik tidak bisa dilakukan pada protein yang telah terdenaturasi<sup>[115]</sup>.

Pada beberapa kasus, protein harus dikonsentrasi terlebih dahulu sebelum diadakannya analisis. Terdapat beberapa teknik yang tersedia untuk melakukan hal ini, seperti ultrafiltrasi, kromatografi afinitas, dan presipitasi. Dibandingkan teknik lain, presipitasi dipercaya merupakan teknik dengan ekstraksi terbanyak dan paling sesuai untuk volume dengan kadar protein rendah. Hal ini terjadi karena protein dapat menempel pada membran ultrafiltrasi, dan fase stasioner yang digunakan dalam kromatografi afinitas<sup>[118]</sup>.



(Gambar 2.3.10 Prinsip dasar presipitasi protein adalah merubah potensial pelarutan zat pelarut, dan mengurangi kelarutan zat terlarut dengan bantuan reagen yang disebut presipitan<sup>[108]</sup>)

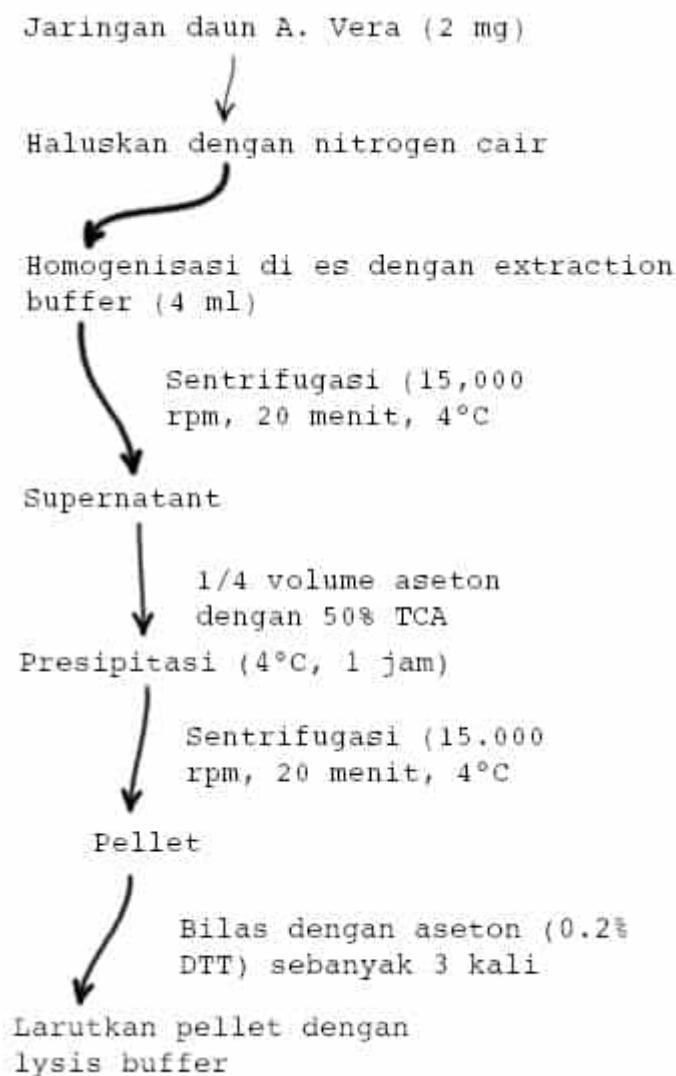
### **Trichloroacetic acid (TCA)-acetone Precipitation Method**

Metode presipitasi trichloroacetic acid (TCA)-acetone merupakan metode yang dikembangkan oleh Damerval<sup>[106]</sup> dan biasanya digunakan untuk menyiapkan sampel analisa proteomik. Metode ini diyakini dapat meminimalisir degradasi protein dan kadar kontaminan<sup>[107]</sup>.

Asam trikloroasetat (TCA) merupakan analog dari asam asetat dengan ketiga atom hidrogen dari gugus metil diganti dengan atom klorin (wikipedia). Penambahan TCA dalam larutan cair mengganggu ikatan hidrogen antar molekul air, sehingga memudahkan penggumpalan protein dan pengumpulan dengan sentrifugasi. Namun, karena TCA memecah ikatan hidrogen, maka TCA dapat merusak struktur sekunder protein dan menyebabkan protein terdenaturasi<sup>[109]</sup>. Meski begitu, TCA tidak cukup kuat

untuk memecah ikatan peptida antar asam amino penyusun protein, sehingga kerusakan struktur primer dapat diminimalisir<sup>[107]</sup>.

Skema TCA/Acetone precipitation method adalah sebagai berikut:



Skema 2.3.1 TCA/Acetone Precipitation Method<sup>[109, 110]</sup>

### Phenol Extraction with methanol/ammonium acetate precipitation Method

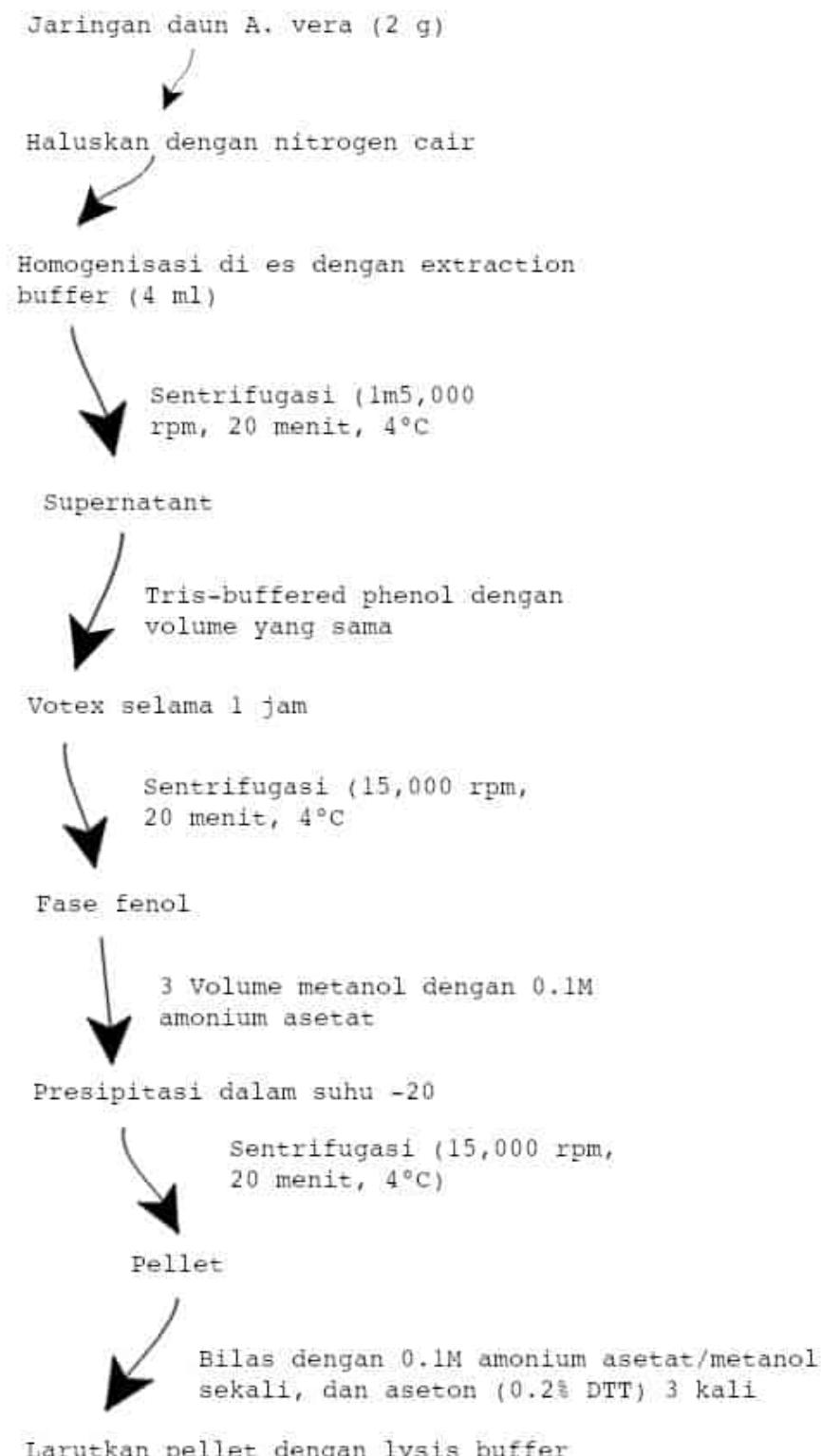
Ekstraksi fenol dengan presipitasi metanol/amonium asetat merupakan cara lain yang dapat digunakan untuk ekstraksi protein. Ekstraksi ini dideskripsikan oleh Hurkman dan Tanaka untuk penyiapan sampel uji proteomik<sup>[111]</sup>. Ekstraksi fenol pertama kali dikembangkan untuk purifikasi

karbohidrat dan asam nukleat, dan kini biasa digunakan untuk menghilangkan protein dari larutan asam nukleat<sup>[112]</sup>.

Fenol merupakan alkohol aromatik paling sederhana dengan rumus kimia C<sub>6</sub>H<sub>5</sub>OH, dan terdiri dari gugus fenil yang berikatan dengan gugus hidroksil. Fenol memiliki sifat asam lemah, korosif, dan beracun. Fenol berinteraksi dengan ikatan hidrogen protein dan memecahnya, menyebabkan protein terdenaturasi<sup>[112]</sup>.

Perbandingan antara ekstraksi fenol dan TCA/Aseton menunjukkan bahwa keduanya merupakan metode ekstraksi yang efisien<sup>[113, 114]</sup>, dan dapat meminimalisir degradasi protein. Namun, ekstraksi fenol lebih efisien dalam menghilangkan kontaminan dan menghasilkan gel dengan kualitas terbaik untuk elektroforesis. Ekstraksi fenol juga menghasilkan glikoprotein dalam jumlah yang lebih besar dibandingkan TCA<sup>[114]</sup>.

Skema phenol extraction adalah sebagai berikut:



Skema 2.3.2 Phenol Extraction with methanol/ammonium acetate precipitation Method

## **2.4 Biopharming**

Biopharming adalah usaha tani (farming) yang memanfaatkan tanaman atau hewan ternak untuk menghasilkan protein atau senyawa metabolit tertentu yang bernilai kesehatan atau pengobatan<sup>[64]</sup>. Biopharming tidak terlepas dari produk rekayasa genetika yang diterapkan pada hewan atau tumbuhan untuk menyisipkan gen dari organisme lain, sehingga hewan atau tumbuhan itu mengekspresikan senyawa tertentu, yang kemudian dapat dipanen hasilnya<sup>[65,66]</sup>.

Produk yang dihasilkan biopharming adalah protein rekombinan. Biopharming menawarkan kelebihan berupa rendahnya biaya karena diintegrasikannya usaha tani yang dapat menghasilkan tanpa memerlukan infrastruktur yang mahal dan kapasitas produksinya dapat diatur dengan lebih fleksibel<sup>[67]</sup>. Kemajuan pada pendekatan plant bio/molecular pharming dalam dekade terakhir telah menjadikan tanaman sebagai sistem manufaktur yang menarik, bahkan dapat mencapai tingkat produksi yang relevan secara komersial dalam waktu singkat<sup>[69-71]</sup>. Adapun pentingnya dan prospek dari ekspresi tanaman untuk produksi protein rekombinan yang hemat biaya. Kandidat vaksin potensial, antibodi monoklonal, dan juga enzim industri yang diekspresikan dalam tumbuhan<sup>[68]</sup>.

Praktik penggunaan tumbuhan untuk produksi protein rekombinan yang bernilai tinggi mulai dari terapi farmasi untuk produk non farmasi seperti antibodi, antigen vaksin, enzim, faktor pertumbuhan, penelitian atau reagen diagnostik, dan bahan kosmetik<sup>[72]</sup> telah meningkat dari waktu ke waktu dan berkembang secara signifikan dalam beberapa dekade terakhir, yang menyebabkan paradigma utama pergeseran di sektor farmasi.

Keuntungan utama dari semua plant-based system adalah budidaya yang mudah, biaya rendah, beban patogen rendah atau tidak ada, produksi massal cepat, protein rekombinan, dan kemampuan tumbuhan untuk menyusun protein kompleks yang mirip eukariotik modifikasi pasca-translasi (PTM)<sup>[73]</sup>.

Pelipatan protein sangat penting untuk mempertahankan aktivitas biologis dari recombinant therapeutic proteins. Karena kurangnya kompleks pemrosesan protein dan kapasitas terbatas untuk PTM, protein yang tepat melipat tidak dapat dicapai dalam sistem ekspresi prokariotik<sup>[75]</sup>. Tanaman memiliki kapasitas untuk merakit dan melakukan PTM protein multimerik besar yang diperlukan untuk aktivitas fungsi biologisnya. Namun, tanaman tidak memiliki mekanisme

pemrosesan N-glikosilasi manusia yang asli telah diatasi oleh pendekatan glycoengineering menuju sintesis manusia yang ditargetkan dan struktur non-manusia untuk meningkatkan homogenitas, kualitas, dan kuantitas produk<sup>[76,77]</sup>.

Tabel 2.4.1 : Contoh biofarmasi yang menggunakan sistem ekspresi tumbuhan skala besar yang diproduksi oleh berbagai perusahaan selama beberapa dekade terakhir. (Stoger et al., 2014 yang diadaptasi dari Wang, Eu Sheng 2016)

Company	Plant species	Plant organ/cells	Expression technology	Lead product
Protalix Biotherapeutics	Carrot	Suspension cells	Stable nuclear	ELELYSO® (taliglucerase alfa)
Medicago	<i>Nicotiana benthamiana</i> Alfalfa	Leaves	Proficia (transient), Stable nuclear	Vaccine for pandemic/ Seasonal flu
Kentucky BioProcessing	<i>Nicotiana benthamiana</i>	Leaves	Transient, agroviral launch vector	Subunit vaccine for influenza
Fraunhofer IME	Tobacco	Leaves	Stable nuclear	Antibody for HIV (microbicide)
Ventria Biosciences	Rice	Seeds	Stable Nuclear	VEN100 (lactoferrin)
Synthon/Biolex Therapeutics	Duckweed	Leafy biomass	Stable nuclear LEX system	Antibody for non-Hodgkin's lymphoma
Planet Biotechnology	Tobacco	Leaves	Stable nuclear	CaroRx (cavity inhibiting treatment)
Icon Genetics	<i>Nicotiana benthamiana</i>	Leaves	MagnICON (transient)	Vaccine for non-Hodgkin's lymphoma
VAXX	Potato	Tubers	Stable nuclear	Oral subunit vaccine for Norwalk virus (NoroVAXX)
SAFC	Tobacco	Leaves	Stable nuclear	Contract manufacturing
NAIST	Strawberry	Fruits	Stable nuclear	Canine interferon alpha
Mapp Biopharmaceutical/ LeafBio	<i>Nicotiana benthamiana</i>	Leaves	MagnICON (transient)	Multiple vaccines and microbicides

Singkatan : Fraunhofer CMB, Fraunhofer Centre for Molecular Biotechnology; Fraunhofer IME, Fraunhofer Institute for Molecular Biology and Applied Ecology; NAIST, National Institute of Advanced Industrial Science and Technology; SAFC, Sigma-Aldrich Fine Chemicals.

#### 2.4.1 Perbandingan Sistem Ekspresi pada Tanaman dan *E. coli*

##### Pada *E. coli*:

Tabel 2.4.2 Permasalahan pada E.Coli sebagai sistem ekspresi dan strateginya

Issue Addressed	Engineering Strategy	Example Reference
Lack of posttranslational modifications in <i>E. coli</i>	Glycosylation	Expression without the glycosylation (only if the biological activity is not impaired) IL-2 [3], IFN- $\beta$ -1b [4]
		Mutation of residues on the surface to create more soluble protein EPO [5]
		Mutation of glycosylation sites to cysteines to allow subsequent glycosylation <i>in vitro</i> EPO [6]
	Proteolytic maturation	Expression in two separate strains or cleavage of the precursor <i>in vitro</i> insulin [7]
	Disulfide bridge formation	Expression in the periplasmic space hGIF [8], proinsulin [9]
	Protein stability	Decreasing number of free cysteine residues by mutation to alanine IL-2 [10], IFN- $\beta$ -1b [11]
Modulation of protein activity		Deletion of the hydrophobic region KGF [12]
		Design of the rapid-acting or long-acting protein version insulin [13-17]
		Enhanced activity by improved affinity to the target molecule DNaseI [18]
		Design of protein consisting of a consensus sequence IFN- $\alpha$ -con [19, 20]

Sumber : J. Shanmugaraj, B.; Malla, A.; Phoolcharoen, W. Emergence of novel coronavirus 2019-nCoV: Need for rapid vaccine and biologics development. Pathogens 2020, 9, 148. [CrossRef]

### Keuntungan:

Tingkat pertumbuhan tinggi, persyaratan media sederhana, kemudahan dalam penanganan, hasil produksi tinggi, dan sangat hemat biaya.

### Kelemahan:

Hilangnya sifat plasmid, penginduksi yang tidak diinginkan untuk ekspresi gen, akumulasi protein intraseluler heterolog sebagai badan inklusi, refolding protein yang tidak tepat, kurangnya modifikasi pasca translasi (termasuk glikosilasi, maturasi proteolitik, pembentukan ikatan disulfida), beban metabolisme dan stress yang dimediasi protein, kontaminasi endotoksin, sekresi yang buruk, pencernaan proteolitik, serta kompleksitas pada proses hilir<sup>[26-27]</sup>.

### Pada tanaman transgenik:

Pada E.coli (prokariotik) tidak terdapat post translational modification (glikosilasi, maturasi proteolitik, dan pembentukan jembatan disulfida) sehingga diperlukan strategi dengan pendekatan yang berbeda untuk menghasilkan insulin yang fungsional. Sedangkan pada eukariotik memiliki mekanisme post translational modification yang mana akan meningkatkan

efisiensi dari produksi insulin dibandingkan dengan menggunakan E.Coli sebagai platform ekspresi.

Alasan dalam penelitian ini menggunakan tanaman sebagai platform ekspresi karena tumbuhan menawarkan berbagai keuntungan diantaranya lebih ekonomis, skalabilitas yang tinggi, skala produksi yang sudah mendunia, dan tingginya hasil protein, dengan kualitas produk yang tinggi. Selain itu, tumbuhan menawarkan akurasi pelipatan protein yang tinggi dibandingkan bakteri, terdapat modifikasi post translational seperti glikosilasi, jembatan disulfida dan lain lain, dimana pada bakteri tidak dapat mekanisme tersebut. Sedangkan post translational modification sangat dibutuhkan dalam pembentukan protein, terutama pada struktur tersier yang mempengaruhi fungsionalitas dari insulin tersebut.

## 2.5 Sistem/Platform Ekspresi

Terdapat berbagai sistem/platform transformasi diantaranya bakteri, sel mamalia, ragi, sel serangga, dan tumbuhan yang memiliki keuntungan masing-masing. Untuk lebih jelasnya pada tabel di bawah ini.

Tabel 2.5.1. Keuntungan dan kerugian berbagai platform ekspresi [78]

Expression System	Advantages	Disadvantages
Bacteria	Easy to manipulate Low cost High expression Ease of scale up Short turnaround time Established regulatory procedures and approval	Improper folding Lack of post-translational modifications which may affect the protein function Endotoxin accumulation
Mammalian Cells	Proper folding and authentic post-translational modifications Existing regulatory approval	High production cost Expensive media and culture condition requirements
Yeast	Rapid growth and scalable Easy to manipulate Simple and inexpensive media requirements and culture conditions Post-translational modifications of recombinant proteins	Difficulty in cell disruption due to the thick and hard cell walls Hyperglycosylation of proteins Limited glycosylation capacity
Insect cells	High expression levels Ability to produce complex proteins including secreted, membrane, and intracellular proteins Proper folding and post-translational modifications	High cost and time consuming Expensive media and culture condition requirements
Plant	Rapid and affordable Optimized growth conditions Free from pathogen and bacterial toxin contaminants Economical Post-translational modification somewhat similar like mammalian system	Regulatory compliance Limited glycosylation capacity

Tabel 2.5.2. Perbandingan lain dari berbagai sistem ekspresi untuk protein farmasi rekombinan (diadaptasi dari Schillberg et al., 2003., Ma et al. 2003)

Comparisons	Transgenic Plant	Plant Cell Culture	Bacteria	Yeast	Mammalian Cell Culture	Transgenic Animals
Overall cost	Very low	Medium	Low	Medium	High	High
Scale-up capacity	Very high	Medium	High	High	Very low	Low
Production scale	Worldwide	Limited	Limited	Limited	Limited	Limited
Protein yield	High	High	Medium	High	Medium-High	High
Protein folding accuracy	High	High	Low	Medium	High	High
Glycosylation	Minor differences	Minor differences	None	Incorrect	Correct	Correct
Product quality	High	High	Low	Medium	High	High
Contamination risks	Low	Low	Endotoxins	Low	Virus, Prions, oncogenic DNA	Virus, Prions, oncogenic DNA
Safety	High	Non-specific	Low	Unknown	Medium	High
Storage cost	Inexpensive	Moderate	Moderate	Moderate	Expensive	Expensive

Meskipun sistem ekspresi berbasis sel mamalia memiliki keuntungan dalam memproduksi antigen vaksin yang secara umum identik dengan inang alaminya, kultur sel mamalia membutuhkan biaya yang besar sehingga menghambat produksi skala besar.

Sistem ekspresi berbasis bakteri atau ragi memungkinkan produksi protein transgenik berskala besar, namun menghasilkan produk yang mana berbeda dari bentuk aslinya (Daniell et al., 2001). Contohnya, protein yang biasanya mengalami glikosilasi pada hewan tidak mengalami proses tersebut pada bakteri dan kemungkinan akan mengalami hiperglikosilasi pada jamur, yang akan menghasilkan ekspresi protein yang sedikit pada kedua sistem (Harashima, 1994).

Sistem ekspresi berbasis tumbuhan menawarkan beberapa keuntungan, diantaranya adalah skalabilitas cepat tanpa membutuhkan fermentor rumit atau bioreaktor, relatif lebih ekonomis, ekspresi protein yang tinggi, risiko minimal kontaminasi dari patogen mamalia atau racun bakteri, heat stable, serta, teknologi yang tersedia untuk panen dan pengolahan tanaman dan produk tanaman dalam skala besar (Daniell et al., 2001) dibandingkan dengan produksi konvensional. Oleh karena itu, maka kami memilih tumbuhan sebagai platform yang ideal untuk ekspresi protein transgenik.

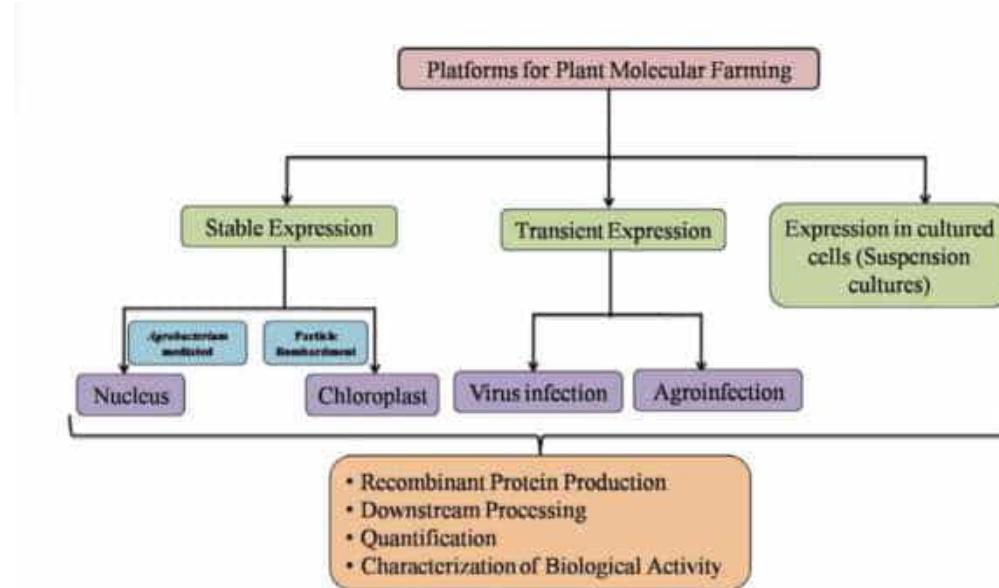
Secara umum, metode transformasi yang efisien harus hemat biaya, tidak memerlukan penggunaan bahan kimia berbahaya atau karsinogenik, memiliki prosedur langsung, memungkinkan pengenalan transgen yang tepat ke dalam host yang dituju, memungkinkan integrasi transgen pada jumlah salinan rendah; dan memungkinkan regenerasi tanaman yang nyaman pasca-transformasi.

Metode transformasi dibagi menjadi langsung atau kategori tidak langsung tergantung pada mode tindakan mereka. Metode tidak langsung biasanya membutuhkan agen biologis seperti bakteri untuk pengiriman transgen ke dalam Tanaman inang sedangkan metode langsung biasanya bersifat fisik, melibatkan langsung penetrasi DNA ke dalam sel inang yang dituju.

### Tanaman sebagai biofaktori untuk ekspresi protein heterolog

Kemajuan signifikan dalam dekade terakhir yang mengarah ke produksi protein yang diturunkan dari biofarmasi tumbuhan, seperti vaksin, antibodi, regulator pertumbuhan manusia, produksi hormon dan darah yang mengarah pada munculnya molecular pharming. (Fischer and Emans, 2000; Giddings, 2001).

### 2.6. Tipe Ekspresi



(Gambar 2.6.1. Skema strategi pendekatan tanaman transgenik untuk memproduksi protein pada tumbuhan<sup>[79]</sup>.)

Kunci utama dalam menghasilkan level transkripsi yang tinggi adalah regulatory genetic elements, terutama promoter dan sisi poliadenilasi. Level ekspresi bergantung pada tingginya level translasi, pelipatan yang benar, target yang tepat, dan stabilitas protein.

Tabel 2.6.1. perbandingan tipe ekspresi manipulasi genetik<sup>[80]</sup>.

Ekspresi Transient	Ekspresi Nuclear	Ekspresi Kloroplas
Manipulasi yang mudah pada vektor virus dan bakteri	Banyak metode manipulasi genetik yang sesuai untuk sistem ini.	Manipulasi yang mudah, resisten terhadap silencing gen
Hasil ekspresi yang tinggi	Membutuhkan waktu yang relatif lama untuk manipulasi genetik	Hasil ekspresi yang tinggi karena banyaknya kloroplas per sel
Risiko kontaminasi yang tinggi dari transfer gen ke lingkungan		Risiko kontaminasi yang rendah dari transfer gen ke lingkungan
Ekspresi yang cepat		Metode transformasi yang terbatas
		Tidak terdapat jalur glikosilasi yang tetap

### 2.6.1 Ekspresi Kloroplas

Ekspresi kloroplas melibatkan pengenalan transgen ke dalam genom kloroplas menggunakan gene gun. Sistem transformasi kloroplas memiliki lebih banyak keuntungan dibandingkan sistem ekspresi nuclear. Genom kloroplas lebih mudah untuk dimanipulasi. Jika genom kloroplas telah diurutkan, kaset transgen dapat dibuat untuk memasukkan gen asing ke daerah spacer antara kloroplas fungsional gen, menggunakan dua urutan flanking sequences pada genom kloroplas, melalui rekombinasi homolog (Daniell, Lin, Yu, & Chang, 2016; Danielly, Streatfield, Streatfield, & Wycoff, 2001).

Transformasi genom kloroplas lebih susah daripada transformasi genom nuclear dikarenakan halangan membran ganda pada kloroplas dan kurang diketahuinya virus yang dapat menginfeksi kloroplas. Tetapi, transformasi yang efektif telah didapatkan dengan menggunakan metode gene gun dengan menembak jaringan tumbuhan muda dengan emas dan partikel tungsten yang dilapisi DNA (Verma, Samson, Koya, & Daniell, 2008). Terdapat ribuan kopi pada genom kloroplas di setiap sel daun, telah didapatkan hasil yang tinggi (lebih dari 70% protein yang larut pada daun tanaman) (Daniell et al., 2016). Keuntungan lainnya yaitu resiko yang rendah gen terkontaminasi ke lingkungan karena kloroplas maternally inherited.

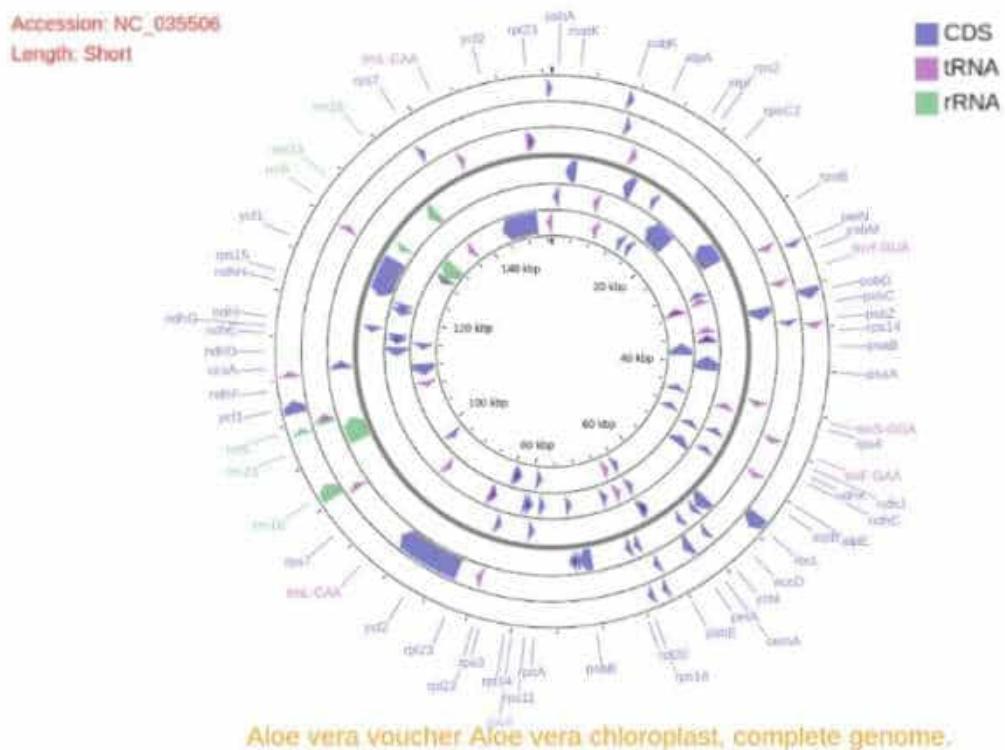
### **Keuntungan dari Transformasi Kloroplas**

Kloroplas mampu mengakumulasi protein transgenik dalam jumlah besar hingga 46% dari total protein pada daun yang diikuti dengan ekspresi dan integrasi transgene yang stabil (De Cosa et al., 2001). Hal ini terutama disebabkan oleh sifat poliploid dari genom kloroplas di mana hingga 100 genom salinan bisa ada dalam satu kloroplas. Sebuah sel tumbuhan yang aktif secara fotosintesis bisa mengandung hingga 100 kloroplas yang berjumlah sekitar 10.000 salinan genom per sel yang fungsional (Verma dan Daniell, 2007).

Selain itu, transgen ke dalam genom kloroplas terjadi di situs secara spesifik melalui dua peristiwa rekombinasi homolog. Berbeda dengan transformasi inti, pada transformasi kloroplas terdapat jumlah salinan yang tinggi karena tidak terdapat mekanisme silencing gene. Keuntungan lainnya dari sistem ekspresi berbasis kloroplas adalah kapasitas dari ekspresi polisistronik yang menyebabkan vaksin atau antibodi multivalen yang akan diekspresikan dari transkrip mRNA tunggal (Quesada-Vargas et al., 2005). Modifikasi post translational seperti pembentukan ikatan disulfida (Staub et al., 2000; Ruhlman et al., 2007), modifikasi lipid (Glenz et al., 2006) juga terjadi di kloroplas yang mengakibatkan pelipatan yang benar dan dari ekspresi protein kloroplas.

### **Organisasi Genom Kloroplas : Konsep dari Transformasi Kloroplas**

Genom kloroplas tumbuhan tingkat tinggi terdiri dari unit dasarDNA untai ganda sekitar 120 hingga 220kb dan dapat ditemukan di beberapa konfigurasi yang berbeda seperti bentuk multimerik, monomerik atau linier (Palmer, 1985; Lilly et al., 2001).



(Gambar 2.6.2. Visualisasi genome kloroplas *Aloe vera* dengan CGView)

DNA asing dapat dimasukkan ke dalam kloroplas dengan cara pengiriman melalui bombardemen partikel (Daniell et al., 1990; Sanford et al., 1993) atau melalui permeasi PEG (Golds et al., 1993; O'Neill et al., 1993). DNA asing itu kemudian diintegrasikan ke dalam genom kloroplas melalui two site-directed homologous recombination diantara homologous flanking sequences pada vektor transformasi dan sekuens endogenous pada genom. Proses ini difasilitasi oleh mekanisme RecA-type yang sama dengan mekanisme E.coli DNA repair (Cerutti et al., 1992). Oleh karena itu, flanking homologous recombination pada vektor transformasi kloroplas harus didesain secara spesifik sesuai dengan urutan genom yang dipilih sebagai transgene insertion (gambar di atas desain nya). Selain itu, sekuens 1kb lebih sering dipilih agar lebih spesifik dan mencegah penggabungan dari transgen di situs non-spesifik (Verma and Daniell, 2007).

Transformasi dicapai ketika transgen berhasil dimasukkan ke dalam beberapa salinan genom kloroplas, diikuti oleh 25 hingga 30 siklus pembelahan sel di bawah tekanan seleksi antibiotik untuk menghilangkan yang tidak ditransformasi kloroplas. Akibatnya, populasi kloroplas menjadi

homogen dengan genom yang ditransformasikan tercapai, disebut sebagai populasi homoplastomik.

Dalam percobaan transformasi kloroplas awal, transgen ditargetkan ke daerah spacer transcriptionally silent untuk mencegah segala kemungkinan untuk gangguan gen kloroplas endogen (Svab dan Maliga, 1993). Meskipun ekspresi transgen berhasil mencapai integrasi transgen ke dalam daerah spacer transkripsi aktif terbukti lebih efektif dan menawarkan sejumlah keunggulan unik seperti mengurangi

ketergantungan pada elemen regulasi (UTR) 5' dan 3' yang tidak diterjemahkan atau promotor untuk regulasi ekspresi transgen (Verma dan Daniell, 2007). Sejauh ini, wilayah spacer transkripsi aktif antara gen trnI dan trnA di dalamnya operon rrn yang terletak di kedua ujung wilayah IR yang paling dekat dengan SSC dari genom kloroplas adalah situs yang paling umum digunakan untuk integrasi dari transgen. Transgen yang terintegrasi dan diekspresikan dari situs ini telah mencatat tingkat ekspresi tertinggi dibandingkan dengan situs integrasi lainnya pada genom kloroplas tembakau (De Cosa et al., 2001). Sehingga, berdasarkan hasil studi literatur dapat diambil kesimpulan bahwa integrasi transgen ke dalam daerah spacer daerah transkripsi aktif mungkin terbukti sangat bermanfaat untuk tingkat ekspresi transgen yang tinggi.

### **Regulasi dari Ekspresi Gen Kloroplas**

Aspek fundamental dari ekspresi gen kloroplas meliputi transkripsi gen, post-transcriptional mRNA processing, modulasi stabilitas mRNA, dan translasi mRNA telah ditinjau secara ekstensif oleh Gillham et al., 1994; Mayfield et al., 1995; Rochaix, 1996; Sugita dan Sugiura, 1996; Barkan dan Goldschmidt-Clermont, 2000; Monde dkk., 2000; Zerges, 2000.

## **2.7 Metode Transformasi**

Semua metode transformasi terutama bertujuan untuk mencapai satu tujuan bersama yaitu, transfer molekul asam deoksiribonukleat (DNA) ke dalam sel target.

## 2.7.1 Metode Transformasi Tidak Langsung

### Transformasi yang dimediasi *Agrobacterium tumefaciens*

Metode transformasi tidak langsung yang sering digunakan melibatkan penggunaan *Agrobacterium tumefaciens* dan *Agrobacterium rhizogenes* untuk transfer transgen dalam bentuk plasmid ke dalam genom nuklear tanaman.

*Agrobacterium tumefaciens* dan *Agrobacterium rhizogenes* berasal dari keluarga rhizobiaceae. Rhizobiaceae merupakan keluarga yang dihuni oleh berbagai macam bakteri pemfiksasi nitrogen yang bersimbiosis mutualisme dengan kacang-kacangan. Berbeda dengan anggota keluarga rhizobiaceae lainnya (wikipedia). *A. tumefaciens* merupakan jenis bakteri fitopatogen yang dapat menyebabkan *crown gall disease*. *Crown gall disease* disebabkan TI-plasmid bakteri yang membawa T-DNA menuju tanaman melalui bacterial type IV secretion system (T4SS).

Ukuran plasmid yang digunakan untuk transformasi agrobacterium umumnya antara 5 dan 12kb termasuk backbone plasmid yang sangat penting untuk replikasi plasmid. Plasmid untuk transformasi agrobacterium terutama didasarkan pada tumor-inducing (Ti) plasmid yang menyebabkan pembentukan tumor dan sintesis opine melalui transfer gen yang terlibat dalam sintesis auksin dan sitokin dan sintase nopolin ke dalam genom inti tanaman yang terinfeksi (Gelvin, 1990).

Transformasi gen yang dimediasi Agrobacterium pengiriman gen biasanya dikaitkan dengan jumlah salinan yang rendah dan integrasi transgen yang stabil ke dalam genom inti inang (Newell, 2000). Kerugian umum termasuk efisiensi transformasi yang rendah, integrasi transgen yang acak, ketergantungan spesies dan persyaratan pengoptimalan yang ketat untuk protokol transformasi.

## 2.7.2 Metode Transformasi Langsung

### Transformasi yang dimediasi Particle Bombardment

Dalam rekayasa genetika, gene gun atau sistem pengiriman partikel biolistic/particle bombardment adalah alat yang digunakan untuk mengirimkan DNA eksogen (transgen), RNA, atau protein ke sel. Dengan

melapisi partikel logam berat dengan gen yang diinginkan dan menembakkan mikro-proyektil ini ke dalam sel menggunakan kekuatan mekanis, integrasi informasi genetik yang diinginkan dapat diinduksi ke dalam sel. Teknik yang terlibat dengan pengiriman mikro-proyektil DNA sering disebut sebagai biolistik<sup>[177]</sup>.

Particle Bombardment juga dikenal sebagai gene gun, dikembangkan untuk memungkinkan penetrasi dinding sel sehingga materi genetik yang mengandung gen yang diinginkan dapat ditransfer ke dalam sel. Sistem ini juga dapat mentransformasi hampir semua tipe sel dan tidak hanya terbatas pada nukleus, namun juga dapat untuk transformasi organel, meliputi plastid (kloroplas) dan mitokondria<sup>[233]</sup>.

Penembakan partikel (Particle bombardment atau biolistik) adalah metode yang paling banyak digunakan pada transformasi tanaman langsung. Materi DNA yang bisa berupa DNA, dalam bentuk plasmid atau dalam vektor virus atau bakteri yang terikat pada mikron inert hingga sub microcarrier berukuran mikron dari emas atau tungsten dan dibombardir ke dalam jaringan tanaman (Stanford, 1988; Klein et al., 1988).

Metode bombardir partikel dimulai dengan melapisi partikel tungsten atau emas (mikro proyeksi) dengan DNA plasmid. Partikel yang dilapisi dilapisi pada proyektil makro, yang dipercepat dengan tekanan udara dan ditembakkan ke jaringan tanaman pada cawan petri (Gan, 1989), seperti yang ditunjukkan pada gambar 2.7.1. Pelat berlubang digunakan untuk menghentikan proyektil makro, sambil membiarkan mikro proyeksi melewati sel di sisi lain. Saat mikro proyeksi memasuki sel, transgen dilepaskan dari permukaan partikel dan dapat bergabung ke dalam DNA kromosom sel. Selectable marker digunakan untuk mengidentifikasi sel yang mengambil transgen. Sel tumbuhan yang telah diubah kemudian beregenerasi menjadi tumbuhan utuh menggunakan kultur jaringan<sup>[181]</sup>.

## Aplikasi

Teknik particle Bombardment kebanyakan digunakan pada tanaman. Namun, juga dapat diaplikasikan pada manusia dan hewan. Target dari gene gun biasanya berupa kalus dari sel tumbuhan. Setelah partikel emas

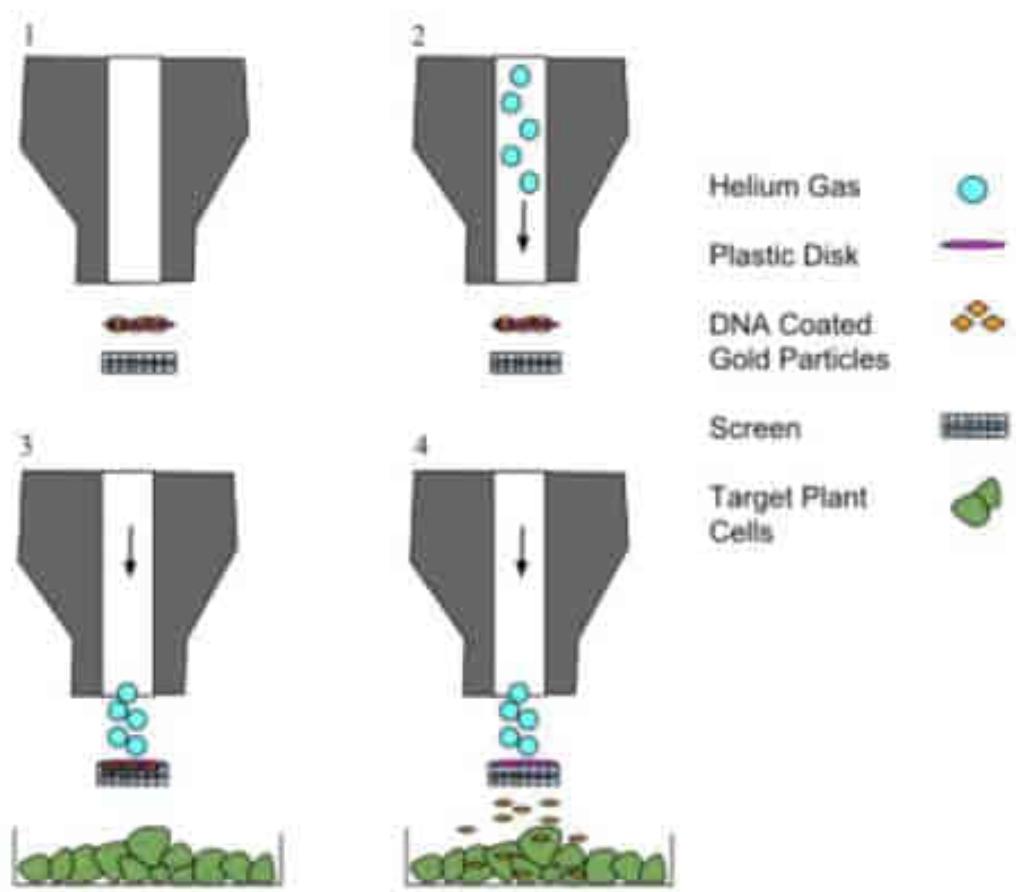
berlapis DNA dikirim ke sel, DNA digunakan sebagai template untuk transkripsi (ekspresi transien) dan terkadang terintegrasi ke dalam kromosom tumbuhan (transformasi 'stabil')<sup>[178]</sup>.

Partikel logam yang sudah ditempel gen target dimasukkan ke dalam senjata gen (Gene Gun) dan kemudian ditembakkan pada sel tanaman. Partikel akan mempenetrasikan dinding sel tanaman sehingga masuk ke dalam sitoplasma. Partikel akan tercuci oleh cairan sel sehingga gen target yang menempel akan lepas dan kemudian masuk ke dalam nukleus dan menyisip pada kromosom sel tanaman<sup>[178]</sup>.

### **Keunggulan**

Biolistik telah terbukti menjadi metode modifikasi genetik yang serbaguna dan banyak digunakan pada rekayasa genetika tanaman<sup>[179]</sup>. Transformasi plastid dimediasi particle bombardment mengalami kesuksesan besar dibandingkan dengan teknik lain, seperti transformasi yang dimediasi Agrobacterium, yang mengalami kesulitan menargetkan vektor dan mengekspresikan secara stabil di kloroplas<sup>[178,179]</sup>. Selain itu, kloroplas tidak membungkam transgen yang dimasukkan dengan gene gun<sup>[181]</sup>. Dan, hanya dengan satu tembakan, seorang teknisi yang terampil dapat menghasilkan dua organisme transgenik<sup>[180]</sup>.

Keuntungan particle bombardment termasuk species independency, kemampuan untuk meregenerasi garis transgenik yang stabil, transformasi yang cepat dan kemungkinan transformasi pada nuclear dan genom plastida. Kekurangan dari metode ini termasuk kecenderungan menghasilkan jumlah salinan yang tinggi, kemungkinan kerusakan pada sel, dan biaya yang relatif tinggi dibandingkan dengan transformasi yang dimediasi Agrobacterium.



Gambar 2.7.1 Gene gun. (Sumber : AP\_Biology\_Bapst\_2015)

## 2.8 Desain Konstruksi Biolistik

Transformasi biolistik melibatkan integrasi fragmen fungsional DNA yang dikenal sebagai konstruksi DNA ke dalam sel target. Sebuah konstruksi gen adalah kaset DNA yang berisi semua elemen pengaturan yang diperlukan untuk ekspresi yang tepat dalam organisme target<sup>[179]</sup>. Sementara konstruksi gen dapat bervariasi dalam desainnya tergantung pada hasil yang diinginkan dari prosedur transformasi, semua konstruksi biasanya mengandung kombinasi urutan promotor, urutan terminator, gen yang diinginkan, dan gen reporter.

### 2.8.1 Promoter

Promoter mengontrol lokasi dan magnitude dari ekspresi gen dan berfungsi sebagai “roda pedal dan pedal gas” dari sebuah gen<sup>[179]</sup>. Promoter adalah faktor utama yang mempengaruhi level dan stabilitas dari ekspresi transgenik (Curtis et al. 1994). Promotor mendahului gen yang diinginkan dalam konstruksi DNA

dan dapat diubah melalui desain laboratorium untuk menyempurnakan ekspresi transgen.

Promoter yang sering digunakan pada tanaman dikotil adalah CaMV 35S dari Cauliflower mosaic virus (Ma et al., 2003), promoter konstitutif kuat yang dapat ditingkatkan dengan duplikasi enhancer (Kay, Chan, Daly, & McPherson, 1987).

Pada penelitian ini kami menggunakan 2 promoter :

- **psbA Promoter (Sebagai promoter untuk gene of interest ke dalam Kloroplas)**

Promotor psbA kloroplas telah ditemukan sebagai promotor terkuat untuk ekspresi gen asing dalam kloroplas (Daniell dan McFadden, 1987; Daniell, 1993). Ekspresi gen asing biasanya diperiksa pada *E. coli* sebelum melanjutkan transformasi kloroplas karena mesin transkripsi dan translasi pada *E. coli* mirip dengan kloroplas tanaman (Daniell, 1993).

Promotor psbA efektif dalam transkripsi gen asing pada *E. coli* dan tidak memerlukan induksi IPTG. Telah dibuktikan bahwa promotor kacang psbA mengandung urutan 10 dan 35 bp hulu dari situs inisiasi transkripsi yang homolog dengan urutan promotor prokariotik dan bahwa transkrip yang dihasilkan secara *in vivo*, *in vitro*, dan di *E. coli* serupa (Boyer dan Mullet, 1986). Promotor ini juga berisi sekuen konsensus kotak CAAT dan TATA eukariotik yang diperlukan untuk inisiasi transkripsi dalam inti eukariotik (Cornelissen dan Vandewiele, 1989).

Kebanyakan promotor plastida yang dikenali oleh RNA polimerase (PEP) mirip bakteri mirip dengan promotor tipe *E. coli*  $\sigma$ 70 yang terdiri dari elemen “-35” dan “-10”. Di antara mereka, promotor psbA unik dalam membawa elemen tambahan antara elemen -35 dan -10 yang dikonservasi. Aktivitas promotor psbA dipertahankan secara berbeda di kloroplas dewasa dimana aktivitas sebagian besar promotor PEP menurun.

Sudah terdapat penelitian dua jenis kegiatan PEP pada bahan gandum [Satoh et al. (1999) Tanaman J. 18: 407]; PEP yang ada dalam kloroplas matang dari ujung daun (PEP tipe ujung) dapat memulai transkripsi dari promotor psbA yang rusak -35, tetapi elemen -35 penting untuk transkripsi

oleh PEP yang ada dalam kloroplas yang belum matang dari pangkal daun (basis- ketik PEP). Unsur promotor -35 penting untuk inisiasi transkripsi pada promotor psbA pada semua jenis plastida, termasuk kloroplas pada daun dewasa, leukoplas pada akar, etioplas pada kotiledon yang mengalami etiolasi<sup>[149]</sup>.

- **CamV 35s promoter (*Sebagai promoter selectable marker*)**

Promotor 35S, yang berasal dari virus mosaik kembang kol (CaMV), adalah komponen konstruksi transgenik lebih dari 80% tumbuhan hasil rekayasa genetika (GM). Promotor RNA 35S adalah promotor konstitutif yang sangat kuat yang bertanggung jawab atas transkripsi seluruh genom CaMV. Ini terkenal karena penggunaannya dalam transformasi tanaman. Studi terbaru menunjukkan bahwa promoter CaMV 35S juga berfungsi pada beberapa sel hewan, meskipun elemen promotor yang digunakan berbeda dengan yang ada pada tumbuhan.

Struktur :

CaMV mengandung molekul DNA untai ganda melingkar sekitar 8,0 kilobase, terganggu oleh torehan yang dihasilkan dari tindakan RNase H selama transkripsi balik. Torehan ini berasal dari Met-tRNA, dan dua primer RNA yang digunakan dalam transkripsi terbalik. Setelah memasuki sel inang, "torehan" untai tunggal dalam DNA virus ini diperbaiki, membentuk molekul superkoil yang mengikat histon.

## 2.8.2 Polyadenylation

Poliadenilasi adalah modifikasi pasca transkripsional dari Ribonucleic Acid (RNA) yang ditemukan di semua sel dan di dalam organel<sup>[193]</sup>. Poliadenilasi dari ujung 3' terjadi sebelum messenger-RNA (mRNA) meninggalkan nukleus. Ekor polyadenylate ini, panjangnya sekitar 100-200 nukleotida, yang melindungi mRNA dari aksi degradasi fosfatase dan nuklease<sup>[194]</sup>. Nantinya pre-mRNA dibelah pada situs poliadenilasi (A) dan ditambahkan ekor poliadenilasi (A), yang merupakan proses yang diperlukan untuk pembentukan mRNA normal. Gen dengan beberapa situs poli (A) dapat menjalani alternative polyadenylation (APA), menghasilkan isoform mRNA yang berbeda dengan 3'

daerah yang tidak diterjemahkan (3' UTR) dan juga dalam beberapa kasus wilayah pengkodean berbeda<sup>[195]</sup>.

### 2.8.3 Terminator

Urutan terminator diperlukan untuk ekspresi gen yang tepat dan ditempatkan setelah wilayah pengkodean gen yang diinginkan dalam konstruksi DNA. Terminator yang umum digunakan dalam transformasi biolistik adalah terminator NOS yang diturunkan dari *Agrobacterium tumefaciens*.

Metode ini memiliki keunggulan kemudahan dan kecepatan dalam penyusunan konstruk, karena gen yang diinginkan dapat dimasukkan ke dalam vektor biner yang sudah berisi promotor dan domain inverted nos dalam satu langkah kloning, dan tidak memerlukan pengetahuan apapun tentang urutan DNA<sup>[97]</sup>.

### 2.8.4 Marker gene (Gen Penanda)

Dalam biologi nuklear dan biologi molekuler, gen penanda adalah gen yang digunakan untuk menentukan apakah urutan asam nukleat telah berhasil dimasukkan ke dalam DNA suatu organisme. Secara khusus, ada dua subtipe dari gen penanda ini: *selectable marker* dan *screenable marker*.

#### *Selectable Marker*

*Selectable marker* memungkinkan pemilihan sel yang telah mengalami transformasi. Umumnya, penanda ini memberikan resistensi terhadap senyawa fototoksik seperti antibiotik dan herbisida. Merupakan gen dominan yang stabil dan merupakan bagian integral dari vektor transformasi. Kriterianya adalah seleksi positif untuk gen resisten yang tanpanya sel yang tidak ditransformasi akan mati saat agen seleksi diterapkan. Idealnya, agen seleksi tidak boleh mempengaruhi sel yang ditransformasikan dan harus efisien bahkan pada konsentrasi yang lebih rendah. Contoh selectable marker adalah cat, nptII (resisten terhadap kanamycin, hpt (resisten terhadap hygromycin), bar, manA, dan lain-lain<sup>[98]</sup>.

### **Hygromycin phosphotransferase (HPT) (hpt)**

hpt (atau aph IV) adalah gen dari Escherichia coli yang mengkode enzim HPT (E.C. 2.7.1.119) dan resisten terhadap antibiotik hygromycin B (Waldron Et al., 1985). Ketika higromisin menempati tempat pengikatan ribosomal dari faktor elongasi 2 (EF-2) dalam sel prokariotik, akibatnya elongasi rantai polipeptida terhambat dan sintesis protein terputus, menyebabkan gejala yang sama yang dijelaskan untuk antibiotik aminoglikosida lainnya.

Dalam sel tumbuhan, antibiotik ini mengerahkan pengaruhnya pada mitokondria dan kloroplas, bekerja dengan cara yang sama dengan merusak sintesis protein. Organel ini memiliki ribosom yang mirip dengan yang ditemukan pada bakteri dan juga rentan terhadap antibiotik aminoglikosida. Oleh karena itu, dengan adanya antibiotik, jaringan tanaman akan menunjukkan klorosis, yang disebabkan oleh kurangnya sintesis klorofil dan terhambatnya pertumbuhan.

## **2.9 Memilih Tanaman Model**

Tabel 2.9.1 Perbandingan beberapa tanaman model (diadaptasi dari Spok and Karner 2008, European Communities and other sources).

Features/Crop	Organ	Yield	Storage/Protein stability	Transformation	Production Costs	Specialty
Tobacco	Leaf	High	Limited	Well established	Good	Human/food/feed
Alfalfa	Leaf	High	Limited	Established	Good	Heterogenous N glycosylation, use atmospheric N <sub>2</sub>
Wheat	Seed	Good	Optimal	Inefficient	Optimal	
Maize	Seed	High	Optimal	Established	Optimal	
Pea	Seed	Good	Optimal	Limited	Good	
Rapeseed	Seed	Good	Optimal	Established	Optimal	Fusion with oleosin for easy purification
Potato	Tuber	Good	Good	Well established	Good	
Banana	Fruit	Good	Good	Inefficient	Good	Can be eaten raw
Aloe vera	Leaves	High	Good	Efficient	Optimal	As a food, medicine, cosmetic plant

### **2.9.1 Aloe vera**

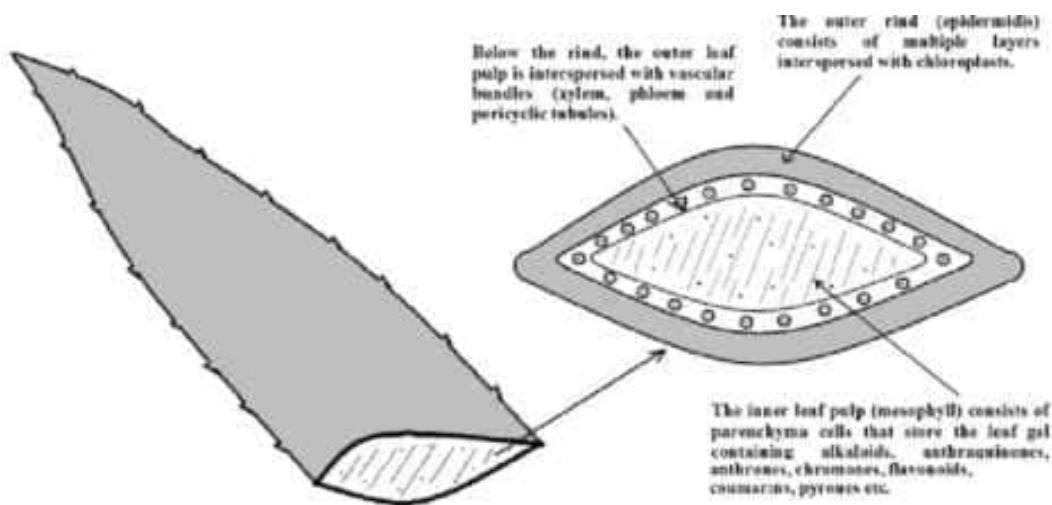
Aloe vera adalah tanaman sukulen yang berasal dari Jazirah Arab. Saat ini, Aloe vera telah menyebar ke berbagai belahan dunia. Menurut Dowling (1985), hanya 3 jenis lidah buaya yang dibudidayakan secara komersial di dunia, yakni: Curacao aloe (*Aloe barbadensis* Miller), Cape aloe (*Aloe ferox* Miller), dan Socotrine (*Aloe perryi* baker).

Jenis yang paling banyak dibudidayakan adalah *Aloe barbadensis* Miller, sedangkan di Indonesia yang paling banyak dikembangkan adalah *Aloe chinensis* Baker. Jenis ini dikembangkan di Kalimantan dan lebih dikenal dengan nama Lidah Buaya Pontianak. Ciri-ciri tanaman ini adalah bunga berwarna oranye, pelepas berwarna hijau muda, pelepas bagian atas agak cekung dan mempunyai totol putih di pelepasnya ketika tanaman masih muda<sup>[46]</sup>.

### Fisiologi *Aloe vera*

#### Metabolisme

Dari segi fisiologi, *Aloe vera* termasuk dalam kelompok CAM (Crassulacean Acid Metabolism). Tanaman CAM hanya membuka stomata pada malam hari untuk menyerap CO<sub>2</sub> dan air, berbeda dengan tumbuhan yang membuka stomatanya di siang dan malam hari. Hal ini memberikan keuntungan bagi tumbuhan CAM, karena ia dapat menutup rapat stomatanya pada siang hari dan mengurangi penguapan air. Oleh karena itu, tumbuhan CAM dapat hidup di daerah yang kering. Contoh tumbuhan CAM selain *Aloe vera* adalah kaktus dan nanas.



(Gambar 2.9.1 Skema daun *Aloe vera* (Cock 2015))

#### Transformasi dan Rekayasa genetik

Semenjak akhir abad 20, *Aloe vera* mulai mendapat perhatian untuk rekayasa genetika. Hal ini disebabkan oleh berbagai macam keuntungan

yang ditawarkan *Aloe vera*. Komponen *Aloe vera* dilaporkan memiliki sifat antijamur, antiseptik, antivirus, antibakteri, anti-inflamasi, antioksidan, dan penyembuhan luka<sup>[7]</sup>. Namun, terdapat beberapa halangan dalam kultur *Aloe vera* transgenik. Halangan tersebut berupa:

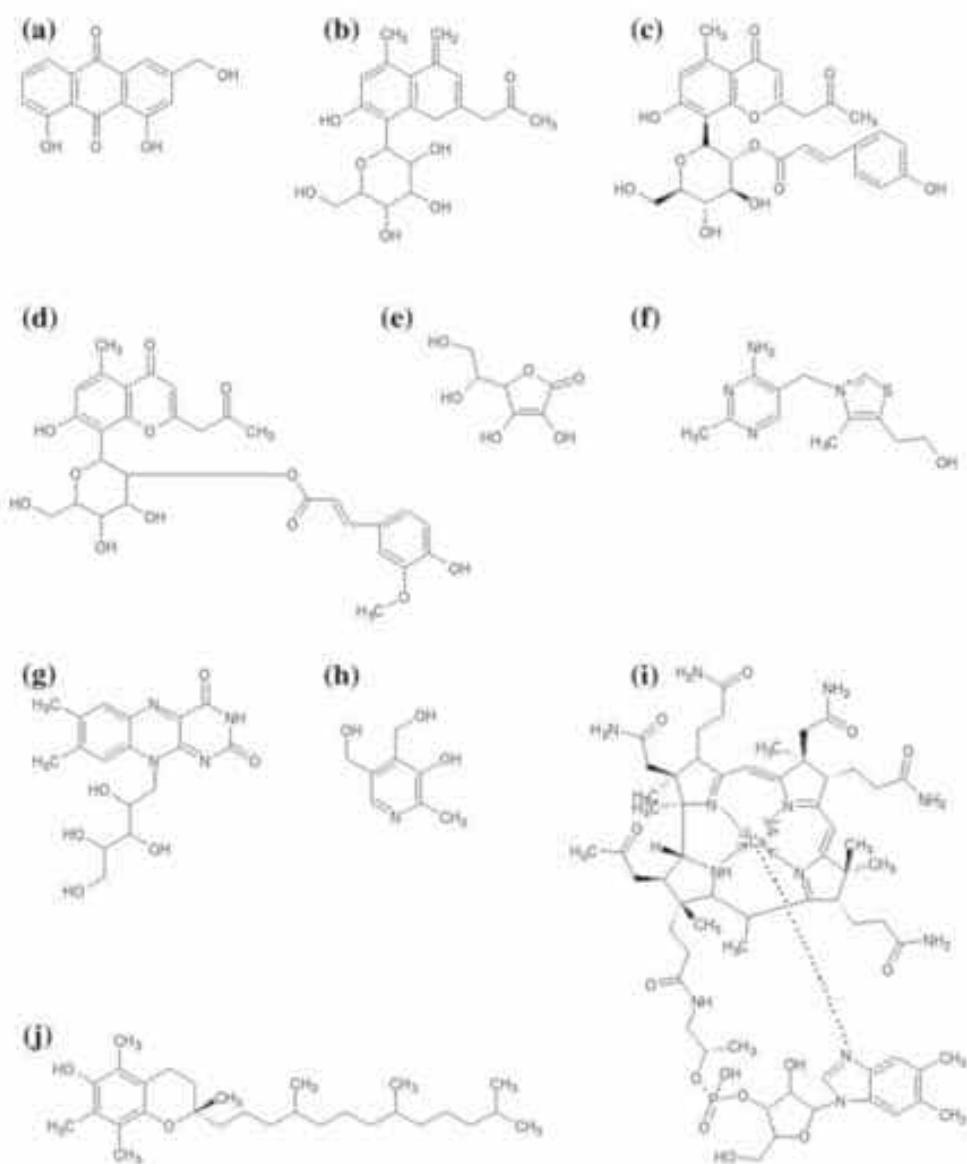
1. browning dan letusan oksidatif terutama pada kultur bermedium cair sering menyebabkan kematian eksplan *Aloe* (Natali et al, 1990);
2. sitokin, yang biasanya digunakan untuk mempromosikan regenerasi, telah dilaporkan dapat menyebabkan peningkatan sekresi fenol dalam sel *Aloe* (Natali et al. 1990; Meyer dan van Staden 1991; Roy dan Sarkar 1991),;
3. *Agrobacterium* yang merupakan faktor tambahan terjadinya nekrosis pada eksplan setelah kultivasi.
4. Ekskresi polyphenol sebelumnya telah dilaporkan menjadi pembatas transformasi tanaman pada spesies lainnya (Lagrimini 1992; Perl et al.1996; Dan 2008)<sup>[63]</sup>.

Pretreatment dengan antioksidan menunjukkan pengurangan letusan oksidatif dan nekrosis pada eksplan (Enriquez-Obregon et al. 1999; Dan 2008). Selain itu, dithiothreitol (DTT) dilaporkan lebih efektif dibandingkan jenis antioksidan lainnya (Apostol et al. 1989)<sup>[63]</sup>.

### Kandungan Nutrisi

Berbagai macam kandungan nutrisi dalam pelepasan lidah buaya antara lain adalah<sup>[46]</sup>:

1. Vitamin, yaitu A, B1, B2, B3, B12, C, E, Choline, Inositol, Folic Acid
2. Mineral, yaitu Kalsium, Magnesium, Potassium, Sodium, Iron, Seng, Chromium
3. Enzim, yaitu Amylase, Catalase, Cellulose, Carboxy Epilase, Carboxy Cellulose, Bradykinase
4. Asam Amino, yaitu Arginin, Asparagin, Aspartat Acid, Alanine, Serine, Glutamate, Threonine, Glycine, Phenylalanine, Histidine, Isoleucine



(Gambar 2.9.2 Struktur kimia molekul dengan antioksidan teridentifikasi pada *Aloe* sp. gel daun:  
 a) Aloe emodin, b) Aloesin, c) 2'-Op-Coumaroyl Aloesin, d) 2'-O-Feruloyl Aloesin, e) Asam askorbat (vitamin C), f) Vitamin B1 (tiamin), g) Vitamin B2 (riboflavin), h) Vitamin B6 (pyridoxal phosphate), i) Vitamin B12, j)  $\alpha$ -Tocopherol (vitamin E) (Cock 2015).)

### Persebaran *Aloe vera*

*Aloe vera* merupakan tanaman asli dari Jazirah Arab, namun, semenjak abad 17, *Aloe vera* mulai dikembang biakan di Cina dan Eropa. Kini *Aloe vera* tersebar luas di daerah beriklim tropis, subtropis, dan kawasan kering di benua Amerika, Asia, dan Australia (wikipedia).



(Gambar 2.9.3 Peta persebaran *Aloe vera* www.discoverlife.org)

## Manfaat Aloe vera

### Pengobatan Tradisional

Aloe vera telah dikenal sebagai tanaman obat sejak puluhan abad silam. Catatan sejarah terawal penggunaan lidah buaya terdapat di Papyrus Ebers dari Mesir 16 abad sebelum masehi<sup>[48]</sup>. Abad ke-1 M, penggunaannya dicatat dalam De Materia Medica karya tabib Yunani Pedanius Dioscorides, dan Naturalis Historia karya penulis Romawi Plinius Tua<sup>[48]</sup>. Di Bizantium abad ke-6 M, penggunaan tanaman ini dicatat dalam Juliana Anicia Codex<sup>[47]</sup>. Dalam pengobatan Ayurveda tumbuhan ini disebut kadhala (sama dengan tumbuhan agave)<sup>[49]</sup>.

### Produk Kesehatan

Terdapat dua zat dari lidah buaya yang sering digunakan, yaitu gel (berwarna bening) dan latex (berwarna kuning). Gel lidah buaya digunakan untuk obat oles untuk berbagai gejala kulit, seperti luka bakar, luka, radang, radang dingin, psoriasis, Herpes labialis, atau kulit terlalu kering. Lateks lidah buaya dijadikan produk (baik bahan itu sendiri maupun digabungkan dengan bahan lain) untuk obat yang ditelan untuk menyembuhkan sembelit<sup>[50,51]</sup>.

## **Produk Kosmetik**

Seperti yang tertera di atas, lidah buaya memiliki banyak khasiat bagi kulit. Oleh sebab itu, banyak penelitian yang dilakukan untuk mencari tahu khasiat lidah buaya dan memproduksinya. Selain itu, sudah ada beberapa penelitian mengenai Aloe vera transgenik untuk produksi protein protein yang dapat dimanfaatkan untuk keperluan kosmetik<sup>[52]</sup>.

## **Budidaya *Aloe vera***

*Aloe vera* sudah mulai dibudidayakan di Indonesia semenjak 1980, terutama di daerah Siantan Hulu, Pontianak. Jenis lidah buaya yang dibudidayakan adalah *Aloe chinensis Baker* atau lebih dikenal dengan sebutan Lidah Buaya Pontianak. Keterlibatan pemerintah dan industri menengah dalam budidaya lidah buaya menjadikannya “ikon” Kota Pontianak<sup>[46]</sup>. Data data yang tertera di bawah berasal dari website Dinas Pangan, Pertanian, dan Perikanan Pontianak serta SOP penanaman lidah buaya direktorat sayuran dan tanaman obat direktorat jenderal hortikultura kementerian pertanian 2019

### **Lahan Penanaman**

Lahan penanaman lidah buaya yang ideal memenuhi persyaratan berikut: Ketinggian 0-10 m dpl; Curah hujan 2500-4000 mm/tahun; Suhu udara 24 - 36°C; pH tanah 3.5-5.5; Jenis tanah podsolik, organosol dengan drainase yang baik; Dataran rendah dengan penyinaran matahari yang baik; Tanah yang subur, gembur, dan mengandung materi organik.

### **Benih, Penanaman, dan Pemeliharaan**

Benih yang digunakan memiliki tinggi 25-30 cm, dengan pelepasan 3-4 helai. Benih ditanam dalam barisan 70-80 cm, dengan jarak 100-150 cm antar barisnya. Pemupukan pra-tanam dilakukan 3-4 hari sebelum penanaman, kemudian dilanjutkan saat tanaman berusia 1-2 bulan, dan pemupukan lanjutan dapat dilakukan 2-3 bulan sekali.

### **Pemanenan**

Panen dapat dilakukan setelah tanaman berusia 8-12 bulan, atau setelah memiliki 15 pelepah. Pemanenan dapat dilakukan sampai

tanaman berusia 2-3 tahun, setelahnya dilakukan peremajaan untuk memperpanjang usia tanaman.

### **Hasil Panen**

No.	Tahun	Luas Tanam (Ha)	Luas Panen (Ha)	Produktivitas (Kg/Ha)	Produksi (Ton)	Keterangan
1.	2007	70	20	51	2.458	1 Tahun = 24 X Panen/Ha
2.	2008	67	45	51	5.530	Produktivitas 1 X Panen
3.	2009	46	43	51	5.284	
4.	2010	37	37	51	4.546	
5.	2011	46	46	51	5.652	
6.	2012	79	46	58	6.359	
7.	2013	84	84	40	7.879	

(Tabel 2.5.1 Data hasil panen lidah buaya Kota Pontianak,  
<https://pertanian.pontianakkota.go.id/produk-unggulan-detil/4-lidah-buaya.html>)

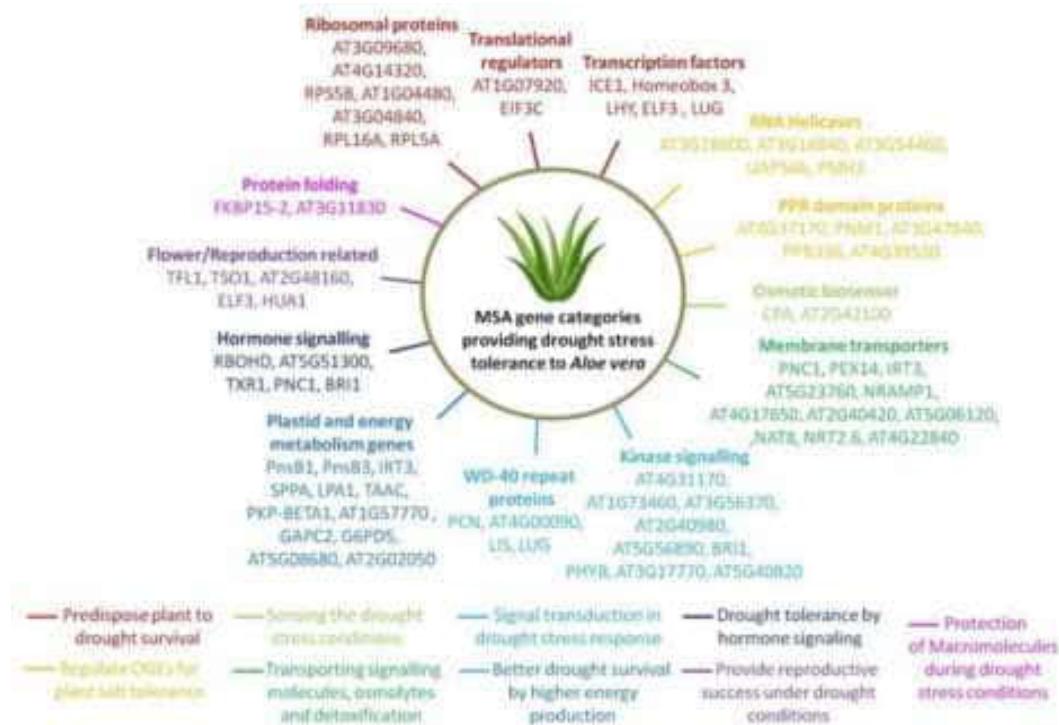
#### **2.9.2 Keunggulan *Aloe vera***

- Menawarkan kemudahan dalam ekstraksi karena protein yang diinginkan terlokalisasi di gel pusat.
- Hasil produksi yang banyak
- Mudah dibudidayakan di iklim tropis seperti Indonesia

## **2.10 Substansi Genetik Aloe vera**

Aloe vera adalah tanaman sukulen dan tahan kekeringan yang termasuk dalam genus family Aloe Asphodelaceae<sup>[81]</sup>. Lebih dari 400 spesies dikenal dalam genus Aloe, empat di antaranya memiliki obat properti dengan lidah buaya sebagai spesies yang paling kuat<sup>[82]</sup>. Lidah buaya adalah tanaman tropis abadi dengan daun sukulen dan memanjang yang terdiri dari jaringan mucilaginous transparan sel parenkim di bagian tengah disebut gel Aloe vera<sup>[83]</sup>. Tanaman ini banyak digunakan sebagai ramuan dalam praktik tradisional di beberapa negara, dan dalam kosmetik dan produk perawatan kulit karena itu sifat farmakologis termasuk anti inflamasi, anti tumor, anti virus, anti bisul, fungisida, dll<sup>[84,85]</sup>. Aloe vera juga mengandung bahan kimia seperti antrakuinon, vitamin, mineral, enzim, sterol, asam amino, asam salisilat, dan karbohidrat<sup>[86,87]</sup>. Properti ini menjadikannya penting secara komersial, dengan pasar global senilai 1,6 miliar<sup>[88]</sup>.

Aloe vera (L.) (*Aloe barbadensis* Miller) adalah xerophyte sukulen abadi, yang mengembangkan jaringan penyimpanan air di daun untuk bertahan dalam kondisi lingkungan kering. Beberapa Gen yang dikenal luas untuk toleransi stres diisolasi dari genom Aloe vera dan terbukti juga mengekspresikan DREB1. HVA22 adalah gen yang dapat diinduksi stres yang diisolasi dari lapisan aleuron *Hordeum vulgare*. Homolog HVA22 telah ditemukan pada 354 spesies tumbuhan. Ekspresi HVA22 pada jaringan vegetatif dapat diinduksi oleh ABA dan tekanan lingkungan, seperti dingin dan kekeringan. Lidah buaya diketahui dapat bertahan hidup tanpa air dalam jangka waktu yang lama. Selain itu, gen penyebab stres dari Aloe vera ini dapat diisolasi. Planlet Aloe vera yang diberi ABA digunakan untuk mengisolasi cDNA pool. Homolog HVA22 diperoleh dengan menggunakan RT-PCR yang diverifikasi dengan sekruensing dan bioinformatika.



(Gambar.2.10.1 Gen MSA pada Aloe vera yang menyebabkan respon terhadap stress kekeringan<sup>[89]</sup>. )

Beberapa penelitian tentang Aloe vera transgenik dan beberapa pencapaian di bidang modifikasi genetik.

Tabel 1.

a)

Plants	Target	Transformation method	Product	References
Aloe vera	Apoplast	Biolistic/gene gun	Active human interferon alpha 2	<a href="https://doi.org/10.1007/s11248-012-9616-0">https://doi.org/10.1007/s11248-012-9616-0</a>
Aloe vera	Chloroplast	Biolistic/gene gun	Polisakarida (karbohidrat) pada AP-2 dan AP-3	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6515206/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6515206/</a>
Aloe vera	Chloroplast	Biolistic or infection with Agrobacterium t.	Some transgenic protein	<a href="https://patents.google.com/patent/US8008546B2/en">https://patents.google.com/patent/US8008546B2/en</a> Paten no : US8.008.546 B2
Aloe vera	Chloroplast	Agrobacterium t.	Introduced TaDREB	Zhao et al. (2009)

			gene isolated from wheat into Aloe by Agrobacterium.	
Aloe vera	Chloroplast	Agrobacterium	Transformed with otsA gene mediated by Agrobacterium t.	Chen et al. (2007)
Aloe vera	Chloroplast	Biolistic	To improve the resistance to cold stress of Aloe plant by introduced otsA gene into its genome via microprojectile bombardment	Chen et al. (2005)

b)

Plants	Achievements in Genetic Engineering	References
Aloe vera	Assessment of genetic stability and instability of tissue culture propagated plantlets	Rathore et al. (2011)
Aloe vera	Assessments of genetic fidelity of micropropagated plants	Gantait et al. (2010)
Aloe vera	Assessments of genetic similarity of micropropagated plantlets with their mother plant	Gantait et al. (2011)

Tabel 1. a) Beberapa penelitian tentang Aloe vera transgenik. b). Beberapa pencapaian pada Aloe vera transgenik.

### 2.10.1 Sekuensing dari genome Aloe vera dan transcriptome

Perkiraan ukuran genom lidah buaya adalah 16,04 Gbp dengan level diploid ploidi yang mengandung 14 (2n) kromosom[2229], dan untuk menutupi ukuran genom yang besar ini, total 506,4 Gbp (~ 32X) data bacaan pendek dan 123,5 Gbp (~ 7,7X) data bacaan panjang dihasilkan menggunakan platform Illumina dan nanopore, masing-masing [90, 59]. Untuk transkriptom, total data

RNA-seq 6,6 Gbp dan 7,3 Gbp dihasilkan masing-masing dari daun dan root. Data transkriptom dari penelitian ini dan RNA-seq yang tersedia untuk data umum dari penelitian sebelumnya [91] digabungkan bersama, menghasilkan total 37,1 Gbp RNA- seq data untuk *Aloe vera*, yang digunakan untuk analisis. Semua genomik dan data pembacaan RNA-seq dipangkas dan disaring menggunakan Trimmomatic, dan hanya data berkualitas tinggi yang digunakan untuk membangun kumpulan genom dan transkriptom akhir.

### **2.10.2 Genome Kloroplas *Aloe vera***

Banyak penelitian menunjukkan bahwa jumlah kromosom untuk sel somatik sebagian besar lidah buaya adalah  $2n = 14$ , dan set genom haploidnya terdiri dari tiga kromosom pendek dan empat yang panjang (Brandham dan Doherty, 1998; Ji et al., 2002; Alam dan Khanam, 2005). Banyak penelitian menunjukkan

bahwa sebagian besar tumbuhan memiliki telomer tipe Arabidopsis terdiri dari banyak salinan berulang dari urutan 5'-TTTAGGG-3 '(Adams et al., 2000a), namun, jenis telomer ini tidak ditemukan di *Aloe*, mirip dengan *Allium*, *Nothoscordum* dan *Tulbaghia*. *Aloe* kekurangan tipe telomerik berulang Arabidopsis, tetapi memiliki telomer seperti vertebrata urutan (T2AG3) dan menurut laporan oleh Weiss dan Scherthan (2002). Tetraploid dan hexaploid *Aloe* bisa diinduksi dari diploid *Aloe* (Ren et al., 2007). Dalam laporan oleh Wang et al. (2001), laju induksi tertinggi meningkat hingga 50% setelah pemaparan dengan 0,06% kolkisin selama 12 jam.

Genom kloroplas memiliki struktur quadripartite tipikal yang berisi daerah salinan tunggal (LSC) besar 83.505bp, wilayah salinan tunggal kecil (SSC) 16.178bp dan sepasang daerah berulang terbalik (IR) dari 26.596 bp. Komposisi nukleotida genom kloroplas secara keseluruhan adalah: 47,185bp A (30,8%), 48,123bp T (31,5%), 29,326bp C (19,2%), 28,241bp G (18,5%) dan kandungan G þ C total 37,7% . Kemudian, 131 gen ditemukan yang mencakup 85 gen penyandi protein (PCG), 38 RNA transfer (tRNA) dan 8 RNA ribosom (rRNA). Analisis filogenetik menunjukkan bahwa *A. vera* berkerabat dekat dengan *A. maculata* dalam

hubungan filogenetik famili Asphodelaceae dengan metode Maximum-Likelihood (ML)<sup>[92]</sup>.

### **2.10.3 *Aloe vera* Menjadi Kandidat yang Cocok untuk Penelitian Ini**

*Aloe* menawarkan sistem produksi protein yang menarik yang memiliki keuntungan menghasilkan protein yang aktif secara biologis serta terkait dengan kosmetik. Protein dapat disekresikan ke dalam pulp yang mudah diakses sehingga meningkatkan efisiensi isolasi dan menawarkan potensi untuk aplikasi topikal langsung dari ekstrak pulp yang diproses secara minimal. Komponen asli dari pulp membuat lidah buaya menjadi sarana pengiriman yang sangat baik yang sangat cocok untuk aplikasi topikal dan penyerapan melalui kulit.

Sistem transformasi yang kami gambarkan juga dapat digunakan untuk mengekspresikan protein nabati untuk meningkatkan atau mengembangkan sifat lidah buaya tertentu. *Aloe vera* adalah tanaman yang kuat dan mudah tumbuh yang dapat diproses hingga empat kali setahun, memasok 3-4 daun dengan berat masing-masing 1 kg (total 12-16 daun per tanaman per tahun). Ini setara dengan kira-kira 6–8 kg pulp per tahun, yang menunjukkan bahwa potensi produksi lidah buaya tinggi.

Aksesibilitas protein yang disekresikan dan biomassa yang melekat membuat *Aloe vera* menjadi platform produksi yang menarik untuk molecular pharming. Lebih jauh lagi, lidah buaya bukanlah tanaman pangan tradisional dan tidak seperti beberapa model tanaman lain yang digunakan untuk modifikasi genetik, tidak ada risiko genetik yang melekat dari penyerbukan silang ke tanaman pangan terkait

Hal ini menjadi salah satu keuntungan menggunakan *Aloe vera*. Daun tanaman aloe transgenik dapat memudahkan ekstraksi karbohidrat, protein, dan air yang terkandung di dalam gel yang terutama berasal dari empulur yang pada umumnya terletak di bagian tengah *Aloe vera*. Berbagai senyawa dalam gel telah terbukti mengobati berbagai penyakit. Selain itu, gel/empulur membantu menstabilkan protein transgenik yang diproduksi oleh *Aloe vera* transgenik dan dilokalisasi dalam gel dan / atau empulur<sup>[150]</sup>.

## **2.11 Enzim Restriksi dan Ligase**

Enzim restriksi dan ligase merupakan dua komponen penting dalam transformasi tanaman. Alasan dari hal tersebut telah dijelaskan pada bab 2.7 Transfer DNA Binary System. Namun untuk mengulas, enzim restriksi dan ligase bertugas untuk memotong dan menyambungkan kembali fragmen fragmen DNA. Hal ini memungkinkan sebagian atau keseluruhan DNA untuk diganti dengan DNA lainnya, dengan syarat keduanya memiliki situs restriksi yang sama.

### **2.11.1 Enzim Restriksi**

Enzim restriksi, disebut juga sebagai restriksi endonuklease atau restriktase, merupakan enzim yang bertugas memotong DNA menjadi fragmen pada situs pengenalan spesifik yang disebut sebagai situs restriksi<sup>[165-167]</sup>. Enzim restriksi merupakan salah satu kelas dari grup enzim lebih besar yang disebut sebagai enzim endonuklease. Enzim restriksi terbagi empat tipe berdasarkan struktur dan lokasi pemotongan. Untuk memotong DNA, enzim restriksi melakukan dua kali pemotongan, sekali pada setiap *sugar-phosphate backbone* DNA untai ganda.

#### **Situs Restriksi**

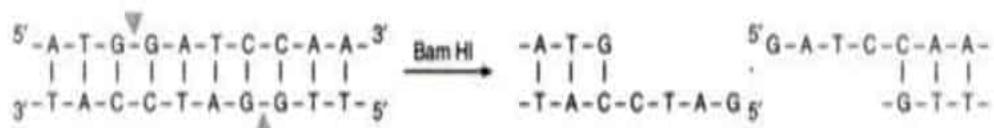
Enzim restriksi mengenali sequence nukleotida spesifik<sup>[166]</sup> dan menghasilkan potongan untai ganda pada DNA. Sequence pengenalan dapat diklasifikasi berdasarkan jumlah basa pada situs restriksi, biasanya antara 4 sampai 8 basa. Selain itu, jumlah basa pada situs restriksi juga ikut menentukan seberapa besar kemungkinan situs restriksi yang sama akan muncul kembali. Sebagai contoh, sequence sebesar 4 bp secara teori akan muncul setiap  $4^4$  bp yaitu 256 bp, sedangkan sequence sebesar 6 bp akan muncul setiap  $4^6$  bp yaitu 4.096 bp, dan sequence sebesar 8 bp akan muncul setiap 48 bp yaitu 65.536 bp.

Banyak sequence pengenalan enzim restriksi merupakan palindrom, yaitu sequence basa yang dapat dibaca sama dari depan maupun belakang<sup>[175]</sup>. Secara teori, terdapat dua tipe palindrom yang dapat muncul pada DNA. Tipe pertama adalah mirror-like palindrome, yaitu sequence yang dibaca sama dari depan maupun belakang dan berada pada satu strand tunggal. Sedangkan yang kedua adalah inverted repeat palindrome, yaitu

segmen yang dibaca sama dari depan maupun belakang namun berada pada strand komplementer DNA untai ganda. Sebagai contoh, GTATAC merupakan sequence komplemen dari CATATG<sup>[176]</sup>. Di antara keduanya, inverted repeat palindrome lebih umum dan memiliki efek biologis yang lebih besar,

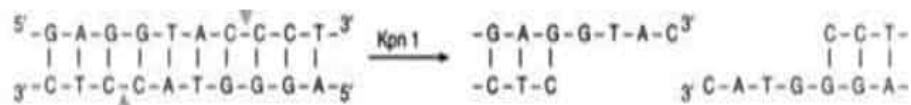
Berdasarkan situs restriksinya, terdapat 3 macam hasil pemotongan DNA, yaitu:

1. 5' sticky-end overhangs



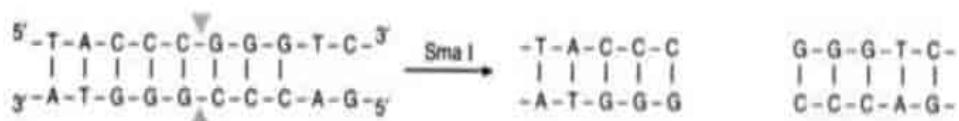
(Gambar 2.11.1 5' overhangs sebagai hasil dari enzim restriksi Bam HI)

2. 3' sticky-end overhangs



(Gambar 2.11.2 3' overhangs sebagai hasil dari enzim restriksi Kpn 1)

3. blunt-ends



(Gambar 2.11.3 Blunt-ends sebagai hasil dari enzim restriksi SMA 1)

### Tipe

Terdapat 4 tipe enzim restriksi yang dapat ditemukan secara alami. Keempat tipe tersebut dibedakan oleh lokasi pemotongannya.

## **Tipe I**

Enzim restriksi tipe 1 merupakan enzim restriksi pertama yang ditemukan dan diidentifikasi pada dua strain (K-12 dan B) *E. coli*<sup>[177]</sup>. Enzim dari tipe ini memotong DNA pada lokasi acak dengan jarak minimal 1000 bp dari situs restriksinya. Pemotongan pada lokasi acak ini mengikuti proses translokasi DNA, yang menunjukkan bahwa enzim ini juga merupakan motor molekuler. Situs restriksinya tidak simetris dan terdiri dari dua bagian spesifik. Bagian pertama mengandung 3-4 nukleotida, dan yang kedua mengandung 4-5 nukleotida, keduanya dipisahkan oleh *non-specific spacer* berukuran sekitar 6-8 nukleotida.

Enzim ini multifungsi dan dapat melakukan pemotongan maupun modifikasi DNA, tergantung status metilasi DNA target. Aktivitas enzim ini membutuhkan kofaktor S-Adenosyl methionine (AdoMet), hydrolyzed adenosine triphosphate (ATP), dan ion magnesium ( $Mg^{2+}$ ). Enzim restriksi tipe I memiliki tiga subunit, yang disebut HsdR, HsdM, dan HsdS; HsdR diperlukan untuk pemotongan; HsdM untuk menambahkan grup metil ke DNA inang (methyltransferase), dan HsdS untuk pengenalan situs penempelan DNA<sup>[171,178]</sup>.

## **Tipe II**

Enzim restriksi tipe II berbeda dengan tipe I dalam beberapa cara. Perbedaan mereka adalah<sup>[175]</sup>:

1. Tipe II membentuk homodimer, dengan situs restriksi yang biasanya tergabung dan terdiri dari 4-8 nukleotida;
2. Mereka mengenali dan memotong DNA pada situs yang sama, dan hanya membutuhkan  $Mg^{2+}$  sebagai kofaktor.

Enzim ini memotong ikatan fosfodiester DNA untai ganda, ia bisa memotong di tengah kedua strand, menghasilkan *blunt-ends* atau pada posisi berbeda dan menghasilkan *overhangs*<sup>[179]</sup>. Pada 1990-an, dan awal 2000-an, ditemukan enzim baru dari keluarga ini yang tidak sesuai dengan ciri ciri umum dari kelas enzim ini. Sehingga, diambil keputusan

untuk membagi kelas ini menjadi beberapa kategori berdasarkan ciri ciri umumnya<sup>[175]</sup>.

Pembagian dari kelas ini adalah<sup>[175]</sup>:

1. Tipe IIB (contoh, Bcgl dan Bpll) adalah multimer, yang mengandung lebih dari satu subunit. Mereka memotong DNA di kedua sisi situs restriksi untuk menghilangkan situs restriksi. Mereka membutuhkan kofaktor AdoMet dan Mg<sup>2+</sup>;
2. Tipe IIE (contoh, Nael) memotong DNA setelah terjadi interaksi antar dua salinan sequence pengenalan. Situs yang pertama merupakan target pemotongan, sedangkan situs yang kedua adalah efektor alosterik yang mempercepat pemotongan;
3. Tipe IIF (contoh, NgoMIV) sama seperti IIE, namun proses kedua pemotongan terjadi bersamaan;
4. Tipe IIG (contoh, Eco57I) memiliki 1 subunit namun memerlukan AdoMet agar bisa aktif;
5. Tipe IIM (contoh, CpnI) dapat mengenali dan memotong DNA termetilasi<sup>[175,180,181]</sup>;
6. Tipe IIS (contoh, fokI) memotong DNA pada jarak yang telah ditentukan dari situs restriksi asimetris non-palindromik. Karakteristik ini banyak digunakan dalam teknik kloning in-vitro seperti *golden gate cloning*, dan enzim ini juga dapat berperan sebagai dimer.
7. Tipe IIT (contoh, Bpu10I dan BsII) terdiri dari dua subunit berbeda. Sebagian mengenali sequence palindromik, sedangkan sebagian sisanya mengenali situs asimetrik.

### Tipe III

Enzim restriksi tipe III (contoh, EcoP15) dua sequence non-palindromik berbeda yang berorientasi terbalik. Mereka memotong DNA sekitar 20-30 bp setelah situs restriksi<sup>[182]</sup>. Enzim ini mengandung lebih dari satu subunit dan membutuhkan kofaktor AdoMet dan ATP<sup>[183]</sup>. Mereka merupakan komponen dari *prokaryotic DNA restriction modification mechanism* yang melindungi organisme dari DNA asing.

Enzim tipe III merupakan protein multifungsi dan hetero-oligomerik yang terdiri dari dua subunit, Res (P08764) dan Mod (P08763). Subunit Mod mengenali sequence DNA spesifik dan merupakan modifikasi methyltransferase. Sedangkan Res diperlukan untuk pemotongan DNA, walaupun ia tidak memiliki aktivitas enzimatis sendiri.

Enzim tipe III mengenali sequence DNA asimetrik pendek berukuran 5-6 bp, dan memotong 25-27 bp ke arah ujung 3' dan meninggalkan tonjolan pendek 5' strand tunggal. Mereka membutuhkan keberadaan dua situs pengenalan berorientasi terbalik yang tidak termetilasi agar pemotongan dapat terjadi. Enzim ini hanya memetilasi satu strand DNA, pada posisi N-6 residu adenosil, sehingga DNA yang baru hanya memiliki satu strand termetilasi, hal ini cukup untuk melindungi DNA dari pemotongan restriksi.

Enzim tipe III termasuk dalam beta-subfamili N6 adenine methyltransferase, dan memiliki 9 motif yang mengkarakteristik keluarga ini, termasuk motif I, Adomet binding pocket (FXGXG), dan motif IV, daerah katalitik (S/D/N (PP) Y/F)<sup>[184,185]</sup>.

#### **Tipe IV**

Enzim tipe IV mengenali DNA yang dimodifikasi, biasanya termetilasi dan dicontohkan oleh sistem McrBC dan Mrr dari *E. coli*.

Enzyme	Source	Recognition Sequence	Cut
<i>EcoRI</i>	<i>Escherichia coli</i>	5'GAAITC 3'CTTAAG	5'—G AATTC—3' 3'—CTTAA G—5'
<i>BamHI</i>	<i>Bacillus amyloliquefaciens</i>	5'GGATCC 3'CCTAGG	5'—G GATCC—3' 3'—CCTAG G—5'
<i>HindIII</i>	<i>Haemophilus influenzae</i>	5'AAGCTT 3'TTCGAA	5'—A AGCTT—3' 3'—TTCGA A—5'
<i>TaqI</i>	<i>Thermus aquaticus</i>	5'TCGA 3'AGCT	5'—T CGA—3' 3'—AGC T—5'
<i>HinfI</i>	<i>Haemophilus influenzae</i>	5'GANTC 3'CTNAG	5'—G ANTC—3' 3'—CTNA G—5'
<i>Sau3A</i>	<i>Staphylococcus aureus</i>	5'GATC 3'CTAG	5'— GATC—3' 3'—CTAG —3'
<i>PovII</i>	<i>Proteus vulgaris</i>	5'CAGCTG 3'GTCGAC	5'—CAG CTG—3' 3'—GTC GAC—5'
<i>HaeIII</i>	<i>Haemophilus eggytus</i>	5'GCC 3'CCGG	5'—GG CC—3' 3'—CC GG—5'
<i>AluI</i>	<i>Arthrobacter luteus</i>	5'AGCT 3'TCGA	5'—AG CT—3' 3'—TC GA—5'
<i>EcoRV</i>	<i>Escherichia coli</i>	5'GATATC 3'CTATAG	5'—GAT ATC—3' 3'—CTA TAG—5'
<i>Sall</i>	<i>Streptomyces albus</i>	5'GTCCGAC 3'CAGCTG	5'—G TCGAC—3' 3'—CAGCT G—5'
<i>Scal</i>	<i>Streptomyces caespiosus</i>	5'AGTACT 3'TCAATGA	5'—AGT ACT—3' 3'—TCA TGA—5'

(Gambar 2.11.4 Contoh enzim restriksi dan sumbernya<sup>[10]</sup>.)

### 2.11.2 Enzim Ligase

Enzim ligase, atau DNA ligase, merupakan enzim yang memfasilitasi penggabungan strand DNA dengan mengkatalisis pembentukan ikatan fosfodiester. DNA ligase berperan dalam perbaikan *single-strand break* pada DNA, namun ada beberapa tipe enzim ligase yang dapat memperbaiki *double-strand break* (kerusakan pada kedua strand komplementer DNA). DNA ligase memperbaiki *single-strand break* menggunakan strand komplementer double helix sebagai template<sup>[186]</sup>. DNA ligase juga berperan dalam replikasi DNA dan pembentukan DNA rekombinan.

#### Mekanisme Enzimatik

Mekanisme DNA ligase adalah membentuk dua ikatan fosfodiester kovalen antara ujung 3' hidroksil satu nukleotida (acceptor) dan ujung 5' fosfat nukleotida lainnya (donor). Dua molekul ATP digunakan tiap

pembentukan ikatan fosfodiester, dan AMP dibutuhkan untuk reaksi ligase, yang dapat dibagi menjadi 4 tahap<sup>[187]</sup>, yaitu:

1. Reorganisasi situs aktivitas seperti fragmen okazaki pada DNA;
2. Adenilasi (penambahan AMP) residu lisin pada pusat aktif enzim, dan melepaskan pirofosfat;
3. Transfer AMP menuju 5' fosfat dan membentuk ikatan pirofosfat;
4. Pembentukan ikatan fosfodiester antara 3' hidroksil dan 5' fosfat.

## Tipe

### E. coli

DNA ligase *E. coli* dikode oleh gen *lig*, dan menggunakan energi yang didapat dari pemotongan nicotinamide adenine dinucleotide (NAD) untuk membentuk ikatan fosfodiester, sama seperti kebanyakan DNA ligase pada prokariota lainnya<sup>[188]</sup>. Ia tidak dapat meligasi *blunt-ends* DNA kecuali dalam kondisi molecular crowding dengan polietilen glikol, dan tidak dapat menggabungkan RNA ke DNA dengan efisien.

Aktivitas DNA ligase *E. coli* dapat ditingkatkan oleh DNA polimerase pada konsentrasi yang tepat. Peningkatan hanya terjadi saat konsentrasi DNA polimerase 1 lebih rendah daripada fragmen DNA yang akan diligasi. Saat konsentrasi DNA polimerase lebih tinggi, ia memiliki efek negatif pada DNA ligase *E. coli*<sup>[189]</sup>.

### T4

DNA ligase dari bakteriofag T4, dan merupakan DNA ligase yang paling sering digunakan dalam laboratorium<sup>[190]</sup>. Ia dapat meligase sticky-ends maupun blunt-ends DNA, oligonukleotida, RNA, dan hibrid RNA-DNA, namun ia tidak meligase asam nukleat rantai tunggal. Tidak seperti *E. coli*, T4 DNA ligase tidak dapat menggunakan NAD dan memiliki kebutuhan absolut terhadap ATP sebagai kofaktor.

Beberapa cara telah dilakukan untuk meningkatkan aktivitas in vitro T4 DNA ligase, salah satunya adalah menguji fusi T4 DNA ligase dengan DNA binding protein lain. Hasil eksperimen ini menunjukkan bahwa konstruksi fusi dari T4 DNA ligase dengan p50 atau NF-kB lebih

efektif 160% pada ligasi blunt-ends dibandingkan dengan T4 DNA ligase liar<sup>[191]</sup>. Reaksi tipikal untuk memasukkan fragmen ke dalam vektor plasmid menggunakan 0.01 unit sticky-ends dan 1 unit blunt-ends. Suhu inkubasi maksimal untuk T4 DNA ligase adalah 16 °C.

T4 DNA ligase mutan memiliki sensitivitas terhadap irradiasi UV<sup>[192,193]</sup> dan methyl methanesulfonate yang lebih tinggi daripada tipe liar<sup>[194]</sup>. Hal ini menunjukkan bahwa DNA ligase berperan dalam perbaikan kerusakan DNA akibat hal hal tadi.

## Mamalia

Terdapat 4 jenis spesifik DNA ligase pada mamalia, yaitu:

1. DNA ligase I: Meligase DNA baru pada lagging strand setelah Ribonuclease H menghilangkan primer RNA dari fragmen okazaki;
2. DNA ligase II: Digunakan dalam perbaikan, dan terbentuk dari splicing alternatif fragmen proteolitik DNA ligase III dan tidak memiliki gen sendiri. Oleh karena itu, ia dianggap identik secara virtual dengan DNA ligase III;
3. DNA ligase III: Kompleks dengan protein perbaikan DNA XRCC1 untuk membantu menyegel DNA dalam proses nucleotide excision repair dan fragmen rekombinan. DNA ligase III juga merupakan satu satunya DNA ligase mamalia yang ditemukan di dalam mitokondria;
4. DNA ligase IV: Kompleks dengan XRCC4 yang mengkatalisis tahap akhir non-homologous end joining DNA double strand break repair pathway. Ia juga dibutuhkan untuk rekombinasi V(D)J, proses yang menghasilkan keanekaragaman lokus imunoglobulin dan T-cell receptor saat perkembangan sistem imun.

DNA ligase pada eukariota dan beberapa mikroorganisme menggunakan ATP daripada NAD<sup>[188]</sup>.

## Thermostable

Didapatkan dari bakteri termofilik, dan merupakan enzim yang stabil pada suhu yang jauh lebih tinggi daripada DNA ligase konvensional. Ia

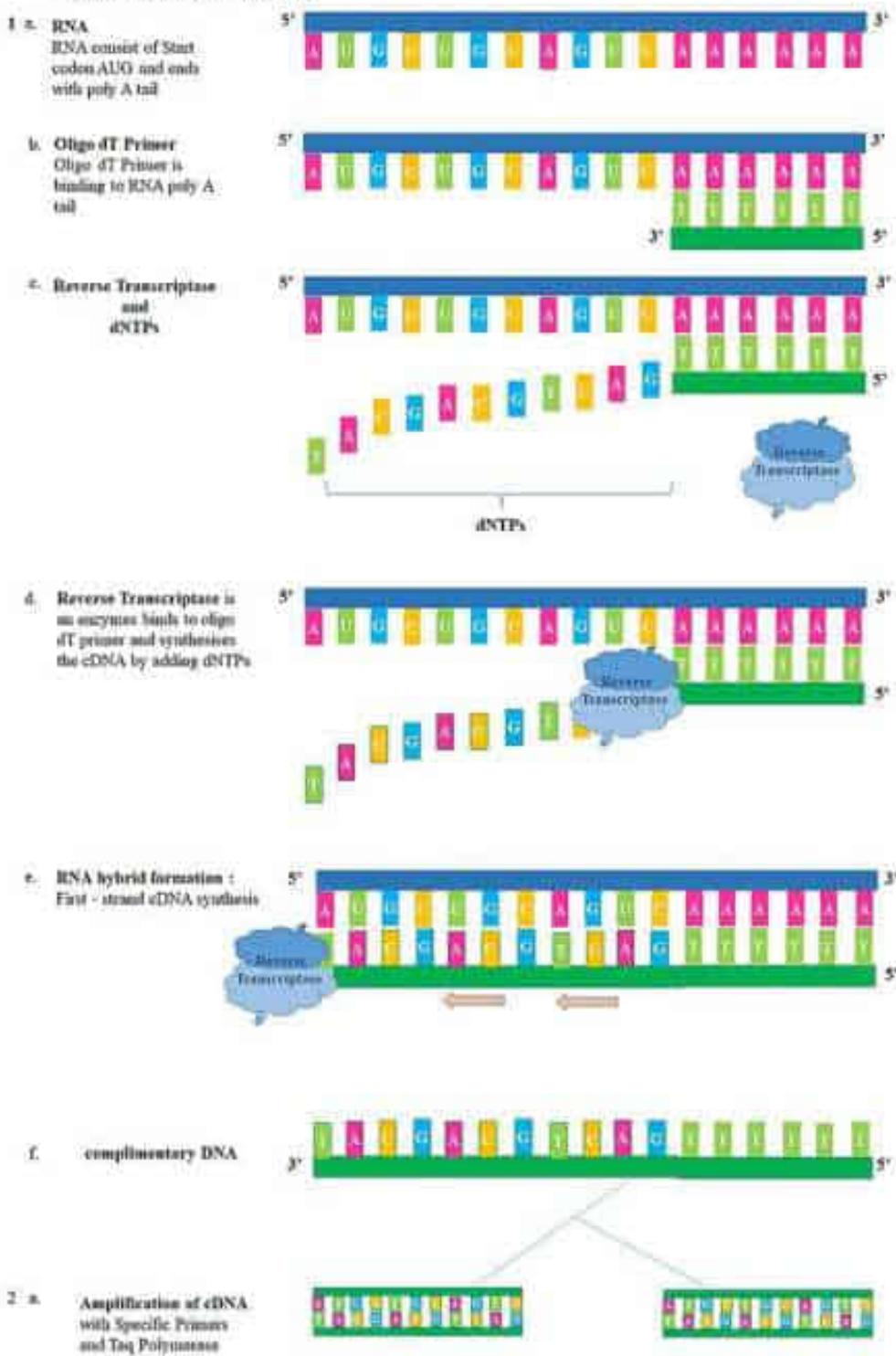
memiliki half-life sepanjang 48 jam pada suhu 65 °C dan lebih dari 1 jam pada suhu 95 °C. Ampligase DNA ligase telah ditunjukkan aktif selama 500 siklus termal (94 °C/80 °C) atau 16 jam siklus<sup>[195]</sup>. Termostabilitas yang luar biasa ini memungkinkan keketatan hibridisasi dan spesifitas ligasi yang sangat tinggi<sup>[195]</sup>.

## **2.12 RT PCR (Reverse Transcription Polymerase Chain Reaction)**

Reverse transcription polymerase chain reaction (RT-PCR) adalah teknik laboratorium yang menggabungkan reverse transcription RNA menjadi DNA (dalam konteks ini disebut complementary DNA atau cDNA) dan amplifikasi target DNA tertentu menggunakan polymerase chain reaction (PCR). [145] Ini terutama digunakan untuk mengukur jumlah RNA tertentu. Ini dicapai dengan memantau reaksi amplifikasi menggunakan fluoresensi, teknik yang disebut PCR waktu nyata atau PCR kuantitatif (qPCR).

## 4.8 Reverse transcription polymerase chain reaction (RT-PCR)

In RT-PCR, the RNA population is converted to cDNA by reverse transcription (RT), and then the cDNA is amplified by the polymerase chain reaction. The cDNA amplification step provides opportunities to further study the original RNA species, even when they are limited in amount or expressed in low abundance. Common applications of RT-PCR include detection of expressed genes, examination of transcript variants, and generation of cDNA samples for cloning and sequencing.



(Gambar 2.12.1 RT-PCR) (sumber : Lokesh Thimmana, under the guidance of Dr. G. Mallikarjuna, Assistant Professor, Molecular Biology, Agri Biotech Foundation.)

### **Prinsip kerja:**

Dalam RT-PCR, template RNA pertama-tama diubah menjadi DNA pelengkap (cDNA) menggunakan reverse transcriptase. cDNA kemudian digunakan sebagai template untuk amplifikasi eksponensial menggunakan PCR. QT-NASBA saat ini merupakan metode paling sensitif untuk mendeteksi RNA yang tersedia. [146] Penggunaan RT-PCR untuk mendeteksi transkrip RNA telah merevolusi studi ekspresi gen dengan cara-cara penting berikut:

1. Secara teoritis memungkinkan untuk mendeteksi transkrip dari hampir semua gen<sup>[417]</sup>;
2. Mengaktifkan amplifikasi sampel dan menghilangkan kebutuhan akan bahan awal yang melimpah yang diperlukan saat menggunakan analisis *northern blot*<sup>[18,19]</sup>;
3. Memberikan toleransi untuk degradasi RNA selama RNA yang mencakup primer masih utuh<sup>[148]</sup>.

### **Aplikasi:**

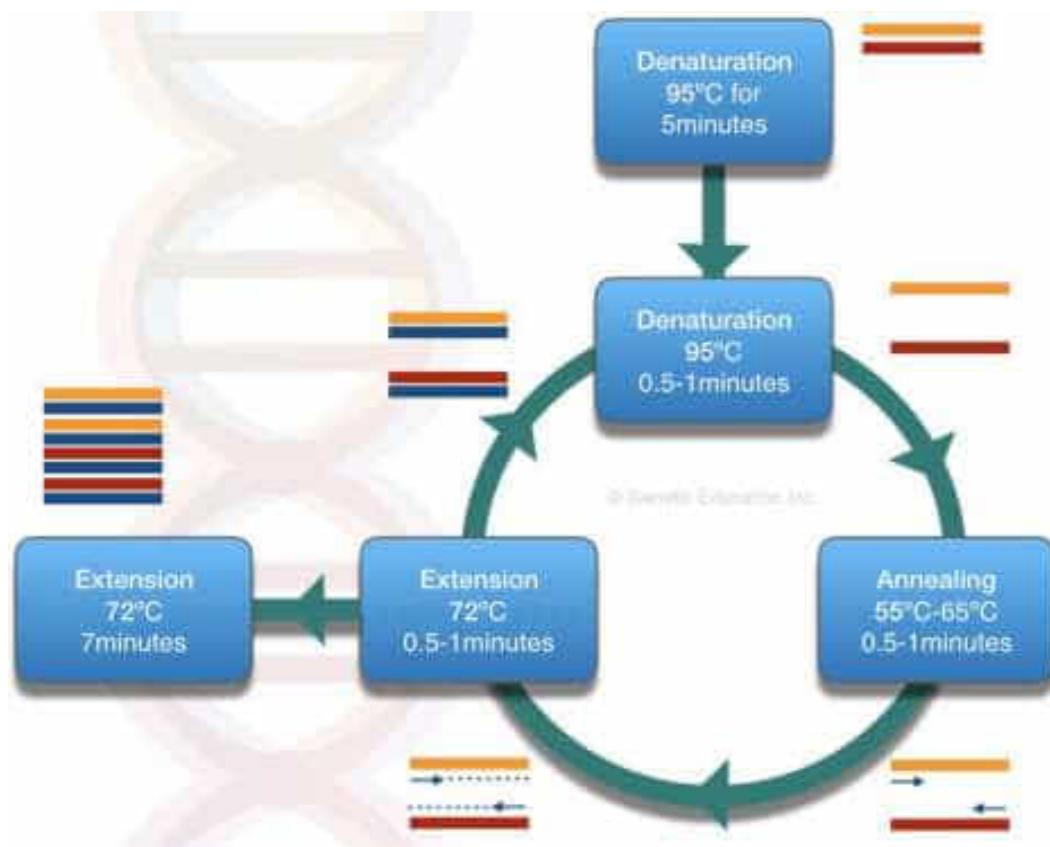
Amplifikasi eksponensial melalui reverse transcription polymerase chain reaction menyediakan teknik yang sangat sensitif di mana jumlah salinan molekul RNA yang sangat rendah dapat dideteksi. RT-PCR banyak digunakan dalam diagnosis penyakit genetik dan, secara semi kuantitatif, dalam penentuan kelimpahan molekul RNA spesifik yang berbeda di dalam sel atau jaringan sebagai ukuran ekspresi gen.

## **2.13 PCR**

Polymerase Chain Reaction (PCR) adalah teknik amplifikasi asam nukleat secara in vitro (Guyer dan Koshland, 1989). PCR memungkinkan peneliti untuk mengamplifikasi urutan DNA tertentu secara eksponensial menjadi ribuan bahkan jutaan salinan. Teknik PCR ditemukan oleh peneliti Amerika, Kary Mullis pada 1984.

Mayoritas metode PCR bergantung pada thermal cycling. Thermal cycling mengekspos reaktan pada siklus pemanasan dan pendinginan berulang. Hal ini

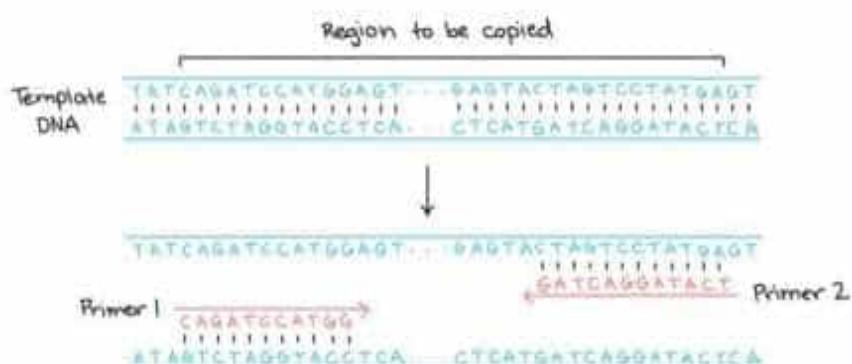
memungkinkan terjadinya reaksi yang bergantung pada suhu. PCR terdiri atas 2 komponen utama, primer (fragmen DNA pendek bernama oligonukleotida dan komplementer terhadap situs DNA target), dan DNA Polymerase.



(Gambar 2.13.1 Polymerase Chain Reaction)

### 2.13.1 Primer

Primer adalah sequence pendek yang komplementer dengan sequence yang akan di amplifikasi. Dalam PCR, digunakan dua jenis primer, forward dan reverse primer.



(Gambar 2.13.2 Forward dan reverse primer)

Pada umumnya, primer PCR adalah DNA. DNA memiliki beberapa keuntungan signifikan dibandingkan RNA untuk PCR. Beberapa keuntungan tersebut adalah:

1. DN lebih stabil dibandingkan RNA, dan tidak terdegradasi dalam suhu tinggi.
2. Proses polimerasi DNA searah pada PCR, sehingga tidak ada kemungkinan hilangnya primer RNA pendek setelah polimerisasi selesai.
3. DNA polimerase 1 membantu menghilangkan RNA primer pendek pada replikasi *in vivo* yang tidak terdapat di PCR.

Setiap enzim membutuhkan kofaktor dan substrat untuk melakukan reaksi, oleh karena itu, Taq DNA polymerase membutuhkan ujung 3'OH bebas untuk memulai polimerisasi. Primer menyediakan ujung 3'OH bebas untuk polimerase dan berfungsi sebagai substrat untuk enzim agar bekerja<sup>[38]</sup>.

### **Melting Temperature**

Suhu dimana setengah template primer terpecah, dipengaruhi oleh kandungan GC dan AT DNA. Alasannya karena primer menempel pada bagian DNA yang memiliki melting temperature yang sama<sup>[38]</sup>.

Formula untuk menghitung melting temperature adalah sebagai berikut:

$$T_m = 4(G + C) + 2(A + T)$$

### **Annealing Temperature**

Suhu yang memungkinkan primer terikat dengan DNA template komplemennya. Sekitar 5°C lebih rendah dari melting temperature, dan berkisar antara 56°C to 65°C. Suhu di luar kisaran tersebut dapat menghambat amplifikasi.

Apabila suhu terlalu rendah, primer dapat menempel dimana saja, sehingga menghasilkan penempelan yang tidak spesifik. Sebaliknya, apabila suhu terlalu tinggi, maka primer tidak dapat menempel dengan DNA template<sup>[38]</sup>.

## Panjang Primer

Primer dengan panjang antara 18-23 bp biasanya memberikan hasil terbaik dalam PCR. Primer yang lebih pendek dari 18 bp dapat menghasilkan penempelan tidak spesifik. Hal ini terjadi karena semakin pendek suatu primer, maka annealing temperature nya akan semakin kecil.

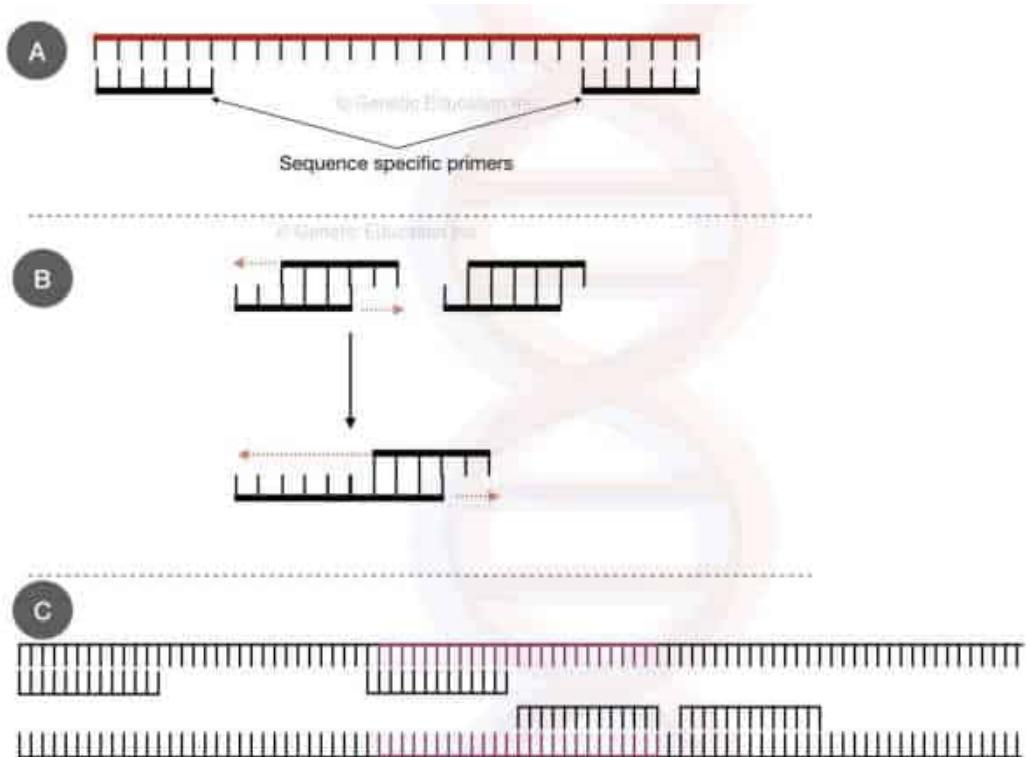
Apabila primer terlalu panjang, annealing temperature akan semakin tinggi, dan dapat mengurangi efisiensi amplifikasi<sup>[38]</sup>.

## Kandungan GC Primer

Kandungan GC disarankan berkisar antara 40%-60%. Daerah kaya GC tidak spesifik, dengan kata lain, kemungkinan terjadinya amplifikasi tidak spesifik di daerah kaya GC lebih tinggi daripada daerah kaya AT<sup>[38]</sup>.

## Komplementasi Forward dan Reverse Primer

Forward dan reverse primer tidak boleh sesuai atau komplementer terhadap satu sama lain. Apabila sesuai, maka kedua primer tersebut akan terikat dan membentuk primer dimer. Primer dimer dapat teramplifikasi dengan mudah dalam PCR karena terbentuk oleh sequence pendek<sup>[38]</sup>.



(Gambar 2.13.3; **a** Penempelan primer spesifik; **b** Pembentukan primer dimer; **c** Penempelan tidak spesifik)

## 2.13.2 Langkah Langkah PCR

PCR terdiri dari 3 langkah, yaitu:

### 1. Denaturasi

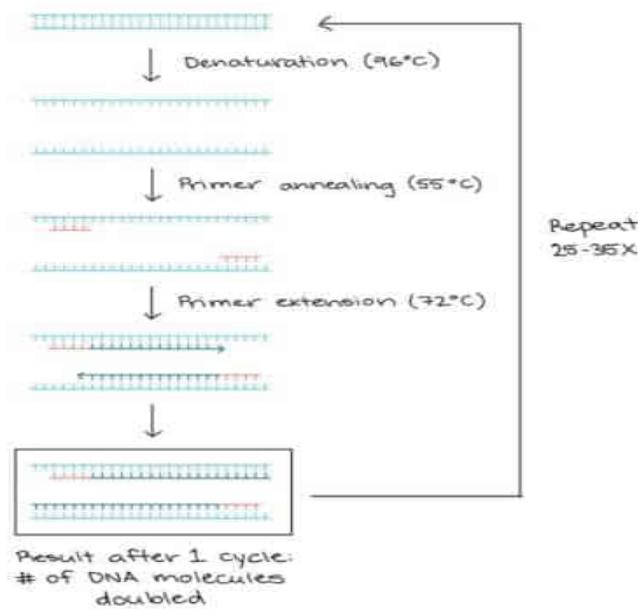
DNA dipanaskan sekitar suhu 95°C, dengan tujuan memecah DNA menjadi dua single-strand molekul DNA dengan memecah ikatan hidrogennya.

### 2. Annealing

Suhu diturunkan menjadi antara 50°C - 60°C, ini memungkinkan annealing primer dengan ujung 3' dari DNA.

### 3. Elongasi

Suhu dinaikkan kembali hingga mencapai antara 70°C-80°C, agar enzim DNA polimerase dapat bekerja. DNA polimerase membentuk strand DNA baru yang komplementer terhadap strand DNA template dari arah 5' ke 3' .



(Gambar 2.13.4 Tahapan PCR)

### 2.13.3 Primer3

Primer3 adalah software online yang digunakan untuk membentuk primer secara otomatis berdasarkan DNA template. Primer 3 dapat diakses di situs <http://bioinfo.ut.ee/primer3/>.

## Cara Penggunaan Primer3

Pertama tama, buka primer3, setelah itu masukkan sequence DNA template.



(Gambar 2.13.5 Primer3)

Lalu, pilih opsi untuk forward dan reverse primer, seperti yang ditunjuk oleh tanda panah pada gambar. Selanjutnya, tekan tombol Pick Primers yang ditunjukkan oleh kotak biru pada gambar.

## Primer3 Output

(Gambar 2.13.6 Hasil primer3)

Hasil primer3 akan nampak seperti gambar diatas. Perhatikan bagian atas, disana tertera spesifikasi forward(left) dan reverse(right) primer. Sekarang perhatikan tulisan bergaris bawah merah, primer ini menghasilkan yield sebesar 250 bp, sehingga saat melakukan PCR berdasarkan kriteria primer ini, hasilnya seharusnya berukuran 250 bp.

Tanda panah (>>>>>>>>>) dan <<<<<<<<<) menunjukkan lokasi annealing primer saat PCR. Software ini juga memberikan kemungkinan primer lain di bagian bawah(lihat gambar dibawah).

```

1201 ACGTCCTGGCAGTGGGGCAGGTGGAGCTGGGGGGGGCCCTGGTGCAAGGCAGCCTGCAGC
1261 CCTTGGCCCTGGAGGGCTCTGCAGAAGCGTGGCATTGTGGAACAATGCTGACCAAGCA
1321 TCTGCTCCCTACCAAGCTGGAGAACTACTGCAACTAGAACGAGCCGAGGCAGGCCA
1381 CACCCGCCCTGCACCGAGAGAGATGGAATAAAGCCCTGAAACCAGC

KEYS (in order of precedence):
>>>> left primer
<<<< right primer

ADDITIONAL OLIGOS:
  start  len      tm      gc%  any_th  3'_th hairpin seq
1 LEFT PRIMER      904   20    58.86  55.00    0.00    0.00    0.00 TCAGGAGATGGGGAAAGATGC
RIGHT PRIMER     1070   20    58.80  60.00    0.00    0.00    0.00 CAGGTTAGAGGGAGGGTCAC
PRODUCT SIZE: 167, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

2 LEFT PRIMER      683   20    58.75  50.00    0.00    0.00    0.00 ACTCGCCCCCTCAAAACAAATG
RIGHT PRIMER     923   20    58.86  55.00    0.00    0.00    0.00 GCATTTCCCCATCTCTGA
PRODUCT SIZE: 241, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

3 LEFT PRIMER     1141   20    58.96  55.00    0.00    0.00    0.00 TGACTGTGTCTCTGTGTC
RIGHT PRIMER    1353   20    59.39  55.00    27.87  27.87    0.00 TGCACTAGTTCTCCAGCTGG
PRODUCT SIZE: 213, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

4 LEFT PRIMER      22   20    59.01  50.00    5.01    0.00    0.00 TCAGAAGAGGCCATCAAGCA
RIGHT PRIMER    239   20    58.35  50.00    0.00    0.00    0.00 TGGCAGAAGGACAGTGATCT
PRODUCT SIZE: 218, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

Statistics
  con  too  in  in  not  no  tm  tm  high  high  high
  sid  many  tar  excl  ok  bad  GC  too  any_th  3'_th hair-  poly  end
  ered  Ns  get  reg  reg  GC% clamp  low  high  compl  compl  pin  X  stab  ok
Left  7787  0  0  0  2791  0  637  2756  0  0  17  58  0  1528
Right 7790  0  0  0  2982  0  719  2574  0  0  19  56  0  1440
Pair Stats:
considered 1585, unacceptable product size 1570, primer in pair overlaps a primer in a better pair 156, ok 7
libprimer3 release 2.4.0

(primer3_results.cgi release 4.1.0)

```

(Gambar 2.13.7 Kemungkinan primer lain)

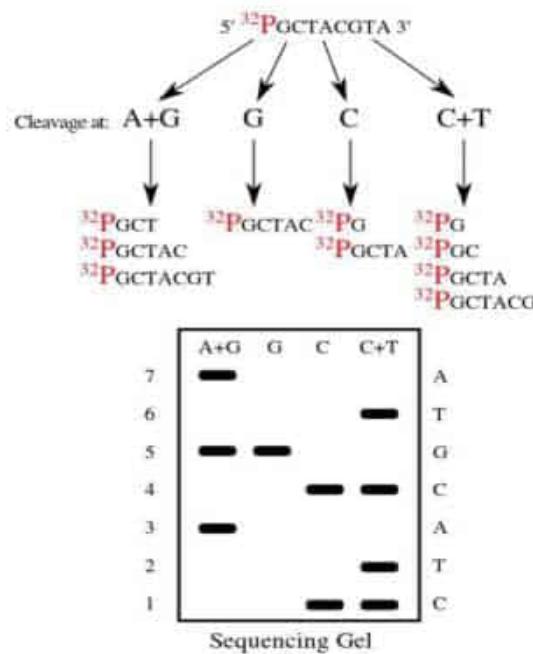
## 2.14 Sekuensing

Sekuensing DNA atau pengurutan DNA adalah proses atau teknik penentuan urutan basa nukleotida pada suatu molekul DNA<sup>[14]</sup>. Sekuensing DNA dapat dimanfaatkan untuk menentukan identitas maupun fungsi gen atau fragmen DNA lainnya dengan cara membandingkan sekuens-nya dengan sekuens DNA lain yang sudah diketahui<sup>[15]</sup>. Teknik ini digunakan dalam riset dasar biologi maupun berbagai bidang terapan seperti kedokteran, bioteknologi<sup>[16]</sup>, dan antropologi.<sup>[17]</sup>

### 2.14.1 Maxam-Gilbert Sequencing

Sekuensing Maxam-Gilbert membutuhkan pelabelan radioaktif pada ujung 5' dari DNA dan pemurnian fragmen DNA untuk diurutkan. Perlakuan kimia kemudian menghasilkan jeda pada proporsi kecil satu atau dua dari empat basa nukleotida di masing-masing dari empat reaksi (G, A + G, C, C + T). Dengan

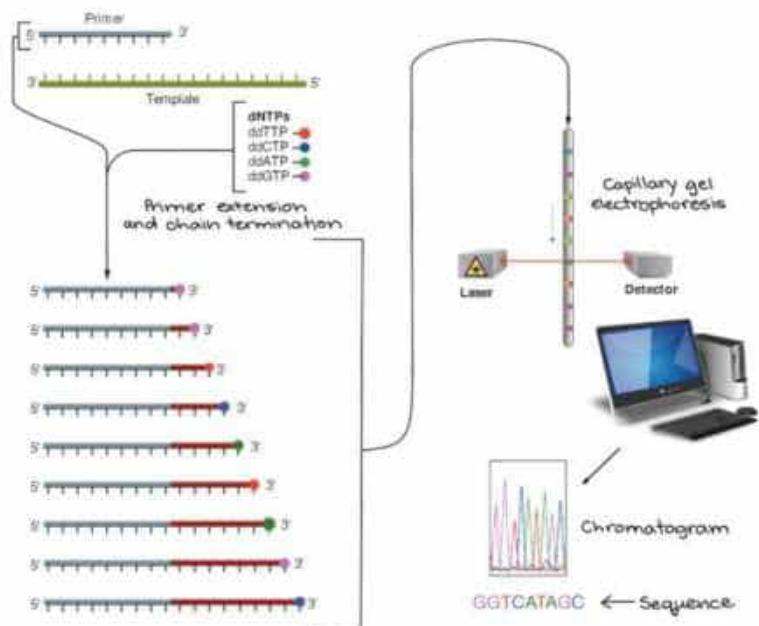
demikian serangkaian fragmen berlabel dihasilkan, dari ujung radiolabeled ke situs “cut” pertama di setiap molekul. Fragmen-fragmen dalam empat reaksi berdampingan secara elektroforesis dalam mendenaturasi gel akrilamid untuk pemisahan ukuran. Untuk memvisualisasikan fragmen, gel dipaparkan ke film sinar-X untuk autoradiografi, menghasilkan serangkaian pita gelap yang masing-masing sesuai dengan fragmen DNA yang diberi label radiolabel, dimana urutan dapat disimpulkan<sup>[18]</sup>.



(gambar 2.14.1 Maxam-Gilbert Sequencing)

## 2.14.2 Sanger Sequencing

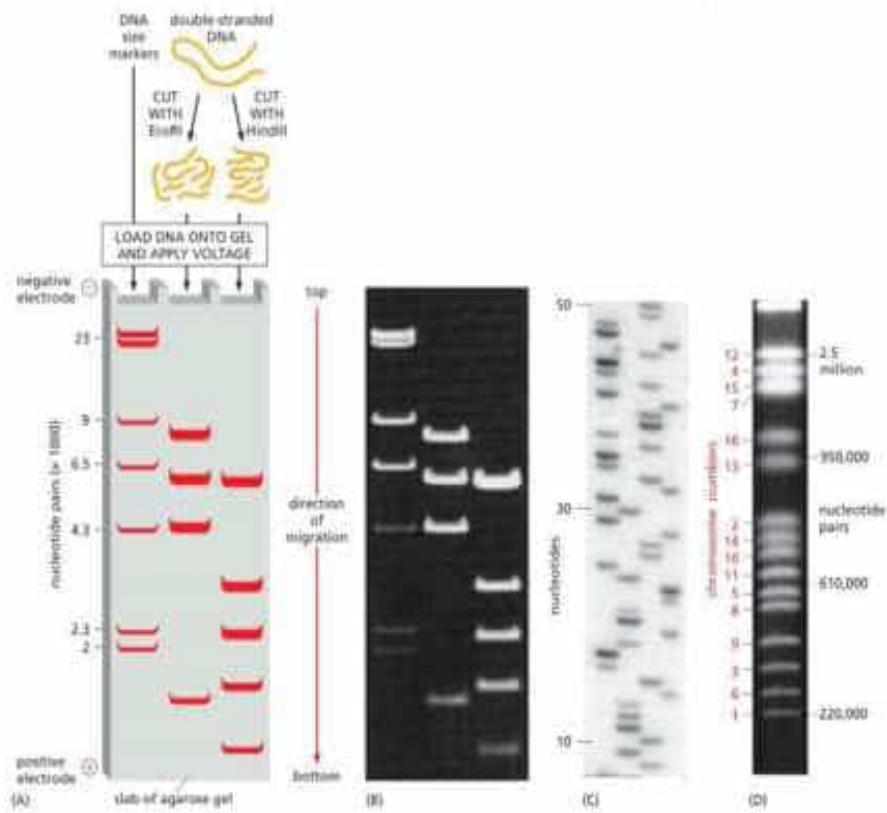
Sanger sequencing adalah metode pengurutan DNA berdasarkan penggabungan selektif dari rantai dideoxynucleotide yang dipecah oleh DNA polimerase selama replikasi DNA in vitro<sup>[19]</sup>.



(Gambar 2.14.2 Sanger Sequencing)

## 2.15 Elektroforesis

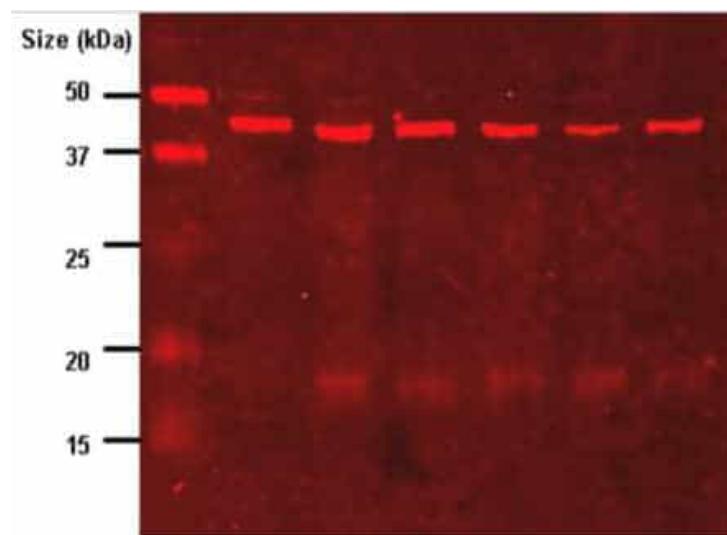
Elektroforesis adalah teknik pemisahan komponen atau molekul bermuatan berdasarkan perbedaan tingkat migrasinya dalam sebuah medan listrik. Teknik ini memanfaatkan muatan listrik pada makromolekul, misalnya DNA yang bermuatan negatif. Jika molekul yang bermuatan negatif dilewatkan melalui suatu medium, kemudian dialiri arus listrik dari suatu kutub ke kutub yang berlawanan muatannya maka molekul tersebut akan bergerak dari kutub negatif ke kutub positif. Kecepatan gerak molekul tersebut tergantung pada nisbah muatan terhadap massanya serta tergantung pula pada bentuk molekulnya<sup>[20]</sup>.



(Gambar 2.15.1 Elektroforesis)

## 2.16 Western Blot

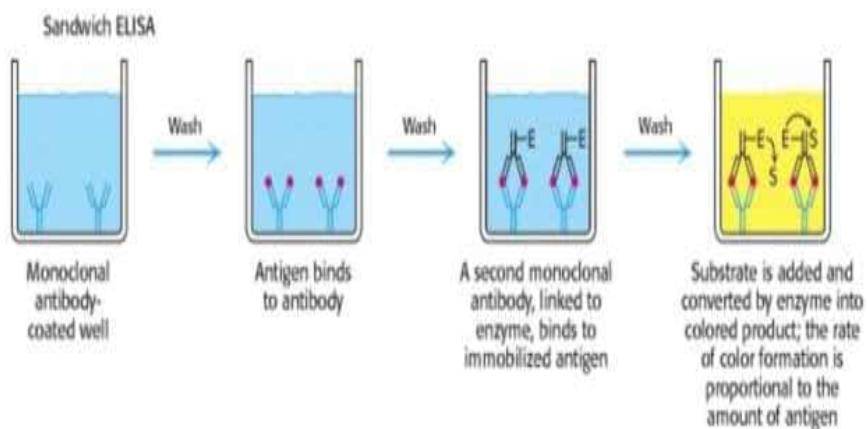
Western blot adalah sebuah metode untuk mendeteksi protein pada sampel jaringan. Western blot menggunakan elektroforesis gel untuk memisahkan protein asli atau perubahan oleh jarak polipeptida atau oleh struktur 3D protein. Protein tersebut dikirim ke membran, di mana mereka dideteksi menggunakan antibodi untuk menargetkan protein<sup>[22]</sup>.



(Gambar 2.16.1. Western blot menggunakan antigen primer asam anti-lipoat dan antigen sekunder berlabel IR dalam ekstrak utama Leishmania. Tim Vickers, 2012)

## 2.17 ELISA

Enzyme-linked immunosorbent assay (ELISA) adalah tes biokimia analitik yang umum digunakan, pertama kali dijelaskan oleh Engvall dan Perlmann pada tahun 1971. Uji ini menggunakan Enzyme Immunoassay (EIA) fase padat untuk mendeteksi keberadaan ligan (umumnya protein) dalam sampel cair menggunakan antibodi yang diarahkan terhadap protein yang akan diukur. ELISA telah digunakan sebagai alat diagnosa dalam kedokteran, patologi tanaman, dan bioteknologi, serta pemeriksaan kualitas di berbagai industri. Mulanya, sampel melekat pada permukaan. Kemudian, antibodi yang cocok diaplikasikan di atas permukaan sehingga dapat mengikat antigen. Antibodi ini terkait dengan enzim, dan pada langkah terakhir, zat yang mengandung substrat enzim ditambahkan. Reaksi selanjutnya menghasilkan sinyal yang dapat dideteksi, misalnya perubahan warna<sup>[23]</sup>.



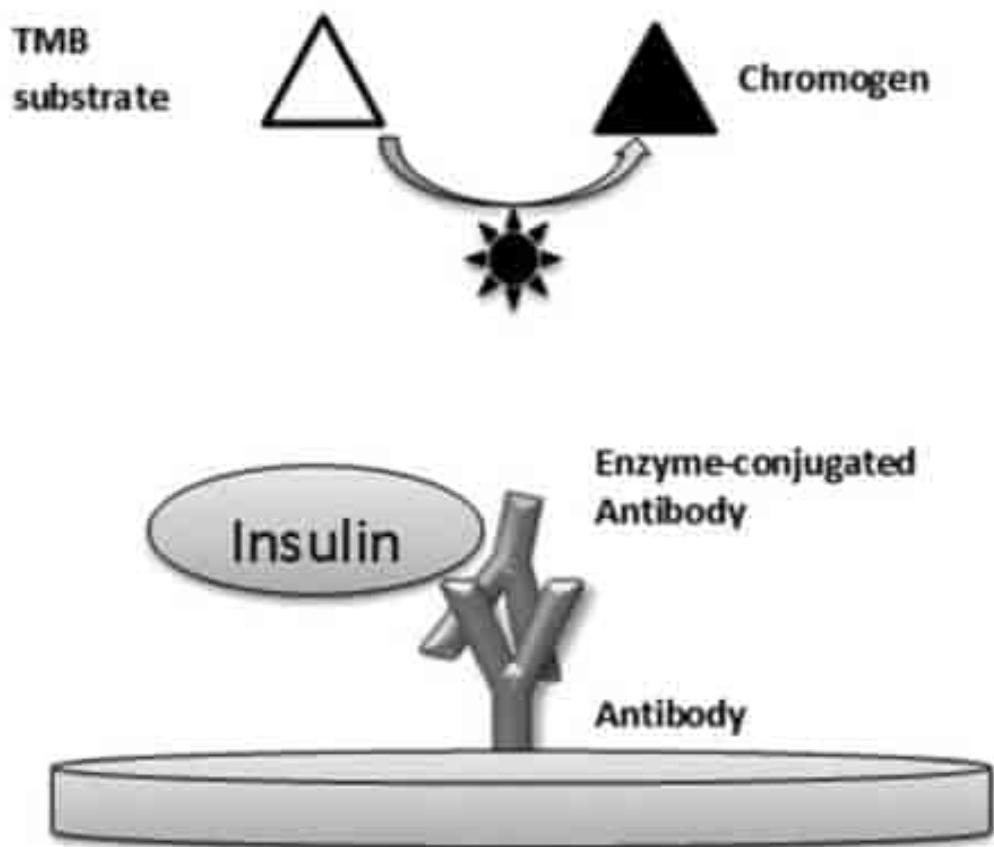
(Gambar 2.17.1 Sandwich ELISA. Pratiksha, Molecular Test, 2015)

### 2.17.1 Insulin ELISA kit

Tujuan dari ELISA adalah untuk mendeteksi apakah antigen target (proinsulin) terdapat pada sampel. Tipe sandwich ELISA adalah yang paling populer untuk mendeteksi keberadaan proinsulin dan mengandung dua antibodi : antibodi penangkap (antibodi anti-insulin) dan pendeteksi antibodi (peroxide labeled anti-insulin antibodies).[151] Selama inkubasi, insulin pada sampel bereaksi dengan antibodi anti insulin yang tidak bisa bergerak di atas piring (Gambar 2.1.1). [152] Antibodi yang tidak terikat dibilas dengan larutan pembilas dan 3,3',5,5'-tetramethylbenzidine (TMB) ditambahkan sebagai substrate chromogenic untuk bereaksi dengan insulin/proinsulin membentuk ikatan konjugasi. [153] Setelah reaksi diberhentikan dengan menambahkan asam, produk yang berwarna diukur pada 450 nm dengan menggunakan spektrofotometer dan mengukur sinyal yang berhubungan dengan konsentrasi analit. [153] kit ELISA yang komersial tersedia untuk mendeteksi proinsulin/insulin menunjukkan batas deteksi (LOD) 3-12pmol/L dan batas atas 600pmol/L. [154] LOD nya  $0.73 \mu\text{IU mL}^{-1}$  dan  $4.9 \mu\text{IU mL}^{-1}$ , and the batas atasnya

$200 \mu\text{IU mL}^{-1}$  dan  $324 \mu\text{IU mL}^{-1}$  (Even et al. 2007 dan Abellan et al. (2009). [155,156] ELISA memiliki sistem seleksi yang bagus terhadap insulin dan meminimalisir interferensi dari protein lain, sehingga cocok untuk tujuan penelitian dimana sampel mengandung larutan fisiologi berupa gara, dan

albumin. [157] Selain itu juga terdapat modifikasi ELISA seperti amplified luminescent proximity homogeneous assay (AlphaLISA) dan homogeneous time-resolved fluorescence (HTRF) telah dikembangkan untuk mendeteksi insulin/proinsulin lebih lanjut.



(Gambar 2.17.2 Direct sandwich ELISA untuk mendeteksi Insulin/Proinsulin<sup>[159]</sup>)

## **BAB 3**

### **Metodologi**

Metodologi yang digunakan dalam penelitian ini adalah teknik *in silico*.

#### **3.1 In Silico**

*In silico* merupakan penelitian yang berbasis pada simulasi dan analisis komputer. Pada penelitian ini kami memakai teknik *In silico* karena beberapa pertimbangan seperti waktu yang dibutuhkan untuk percobaan lebih cepat, lebih ekonomis, dan lain-lain daripada *in vivo* dan *in vitro*.

#### **3.2 Data dan Metode Analisis**

Penelitian ini difokuskan pada analisis, simulasi, dan komputasi data dalam rangka meneliti dan mengembangkan *Aloe vera* transgenik yang mampu menghasilkan proinsulin. Data komputasi dan hasil simulasi akan dikumpulkan sebagai dasar untuk eksperimen di dunia nyata. Kami menggunakan berbagai alat bioinformatika yang kami dapatkan dari internet, dan membangun beberapa lainnya dengan menggunakan biopython. Data data yang kami gunakan kami dapatkan dari berbagai database online seperti GenBank, UCSC Genome Browser, dan lain lain. Perlu diperhatikan bahwa penelitian ini masih dalam tahap uji coba simulasi, sehingga hampir dapat dipastikan bahwa penelitian ini akan mengalami banyak perubahan kedepannya, yang berkaitan dengan hasil dan data dari simulasi yang telah kami lakukan.

#### **3.3 Prosedur**

1. Mengidentifikasi dan mengumpulkan informasi
  - Mengumpulkan informasi mengenai diabetes
  - Menganalisis dari tinjauan pustaka untuk mengkonstruksi sistem dan dasar teori
  - Mengobservasi pendekatan yang paling sesuai untuk mengaplikasikan teknik genetika.
  - Mengumpulkan jurnal dan data transgenik dari database.
2. Mengobservasi dan mengumpulkan data
  - Mengkonstruksi modifikasi transgen untuk meningkatkan produksi insulin pada *Aloe vera*.
  - Mengumpulkan informasi dari jurnal dan database:

- Urutan genom (Sekuens).
- Jalur metabolisme
- Membandingkan platform dan target yang spesifik
- Konstruksi urutan gen
- Menyeleksi enzim restriksi dan ligasi yang digunakan pada penelitian
- Menemukan metode transformasi yang sesuai
- Mencari teori pendukung yang dibutuhkan dalam penelitian
  - Sintesis insulin dan jalur transgenik
  - Mengkonstruksi desain vektor kloning yang sesuai.

### 3. Simulasi *Software*

- Mengidentifikasi sisi restriksi menggunakan Mapper version 3.0;
- Optimasi kodon proinsulin dengan dnachisel dan biopython;
  - Splicing;
  - Codon usage table;
  - Optimasi dengan dnachisel;
  - Pairwise alignment dengan EMBL-EBI<sup>[201]</sup>.
- Optimasi PCR.
- Konstruksi transgen dengan snapgene dan benchling.

### 4. Desain eksperimen (Pendekatan modifikasi transgenik)

- Proses di laboratorium
  - Isolasi promoter
  - Perakitan vektor
  - Kloning
  - Transformasi
  - Identifikasi tumbuhan transgenik
- Tes Laboratorium/Hasil Eksperimen
  - Ekstraksi dan karakterisasi molekuler dari *Aloe vera*
  - Identifikasi gen transgenik
  - Tes konfirmasi dari ekspresi
  - Identifikasi tanaman transgenik
  - Ekstraksi insulin dan *western blotting*
  - Identifikasi level insulin yang dihasilkan

## 3.4 Mengidentifikasi dan Mengumpulkan informasi

Menentukan metode penelitian antara *in vivo*, *in vitro*, dan *in silico*, Menentukan platform dan sistem ekspresi (target transformasi), dan teknik transformasi yang digunakan. Sehingga didapatkan hasil menggunakan pendekatan *in silico*, menggunakan tanaman sebagai platform ekspresi

(spesifiknya *Aloe vera*), menggunakan sistem ekspresi kloroplas, dan teknik transformasi menggunakan particle bombardment.

## 3.5 Mengobservasi dan Mengumpulkan Data Konstruksi

### 3.5.1 Mengumpulkan dan Memilih Sequence

Terdapat tiga sequence yang akan kami gunakan dalam penelitian ini. Ketiga sequence itu adalah promoter, *gene of interest*, dan terminator. Dikarenakan tujuan dari penelitian ini adalah merancang *Aloe vera* transgenik yang dapat menghasilkan proinsulin, maka *gene of interest* yang kami gunakan adalah sequence preproinsulin manusia dari genbank NC\_000011 REGION: complement(2159779..2161209). Sequence insulin tersebut kemudian akan kami optimasi untuk meningkatkan hasil ekspresi.

#### Menentukan Promoter

Seperti yang telah dijelaskan pada bab 2.8.1 promoter merupakan salah satu bagian terpenting dalam ekspresi gen, karena promoter berfungsi sebagai semacam roda pedal bagi ekspresi gen. Oleh karena itu, menentukan promoter yang sesuai dengan kebutuhan dan permintaan eksperimen ini merupakan bagian yang sangat krusial.

Dalam memilih promoter yang akan kami gunakan, kami memiliki beberapa pertimbangan, pertimbangan tersebut adalah:

- Promoter mana yang tidak bersifat toxic bagi manusia
- Promoter mana yang sesuai dengan ekspresi kloroplas

Dan setelah berbagai macam pertimbangan, kami memutuskan untuk menggunakan psbA promoter, sebuah promoter yang terspesifikasi untuk kloroplas. Sequence untuk psbA promoter kami dapatkan dari UniProt P83755.

#### Menentukan Terminator

Selain promoter, terminator juga merupakan bagian yang krusial dari sebuah gen. Terminator berfungsi sebagai penanda batas akhir dari ekspresi. Terminator telah kami jelaskan pada bab 2.8.3. Terminator yang kami gunakan adalah NOS terminator, dengan pertimbangan bahwa terminator ini adalah terminator yang biasa digunakan dalam transformasi yang dimediasi

oleh particle bombardment. Sequence NOS terminator kami dapatkan dari pCAMBIA 1305.1 GenBank AF354045.1

### **3.5.2 Memilih Vektor Ekspresi**

Setelah menentukan sequence yang akan digunakan, berikutnya adalah memilih vektor ekspresi yang sesuai. Setelah melakukan penelitian lebih lanjut pada berbagai jurnal dan forum diskusi online seperti researchgate, kami memutuskan untuk menggunakan vektor dari keluarga pCAMBIA. Tahap berikutnya adalah memutuskan vektor yang kami inginkan dari keluarga pCAMBIA cloning vector. Beberapa hal yang kami pertimbangkan adalah:

- Vektor tidak boleh mengandung zat yang bersifat toxic bagi manusia.
- Vektor tidak boleh mengandung gen yang dapat menghambat ekspresi pada kloroplas tanaman monokotil seperti *Aloe vera*.
- Vektor harus memiliki gen yang mengkode ketahanan terhadap antibiotik.

Kebanyakan vektor pCAMBIA memiliki GUS reporter gene, sayangnya, enzim  $\beta$ -glucuronidase, yang dikode oleh gen GUS, bersifat toxic bagi manusia<sup>[196]</sup>. Oleh karena itu, pada akhirnya kami memilih untuk menggunakan pCAMBIA 1300 cloning vector. pCAMBIA 1300 adalah cloning vector dari keluarga pCAMBIA yang tidak memiliki gen GUS dan memiliki gen yang mengkode ketahanan terhadap kanamisin dan higromisin. pCAMBIA 1300 memiliki promoter CaMV 35S sebelum sequence HygR yang mengkode ketahanan terhadap higromisin.

## **3.6 Langkah-Langkah In silico**

### **• Data dan Metode Analisis**

Projek ini berfokus pada analisis komputasional dan metode konstruksi dalam pengembangan *Aloe vera* transgenik yang menghasilkan proinsulin. Data komputasi dan hasil simulasi didapatkan beberapa sumber serta menggunakan beberapa algoritma dan bahasa pemrograman phyton. Rencana dan metode detail eksperimental ditetapkan berdasarkan hasil in silico, dasar teori, data yang diperoleh, dan juga berpedoman dengan penelitian-penelitian yang telah dilakukan sebelumnya. Selain itu, pada bagian eksperimen

penelitian akan dilakukan selanjutnya untuk membandingkan data antara hasil komputasi dan hasil eksperimen nyata.

- **Software yang digunakan**

1. Genbank, UCSC Genome
  - Database genome
2. KEGG: Kyoto Encyclopedia of Genes and Genomes
  - Sifat molekul, jalur metabolisme *Aloe vera*
3. Restriction Mapper 3.0
  - Pemetaan situs restriksi
4. Snapgene
  - Konstruksi vektor, PCR
5. Benchling
  - Virtual electroforesis
6. Python dnachisel
  - Optimasi kodon
7. Python biopython
  - Modul bioinformatika python
8. Jupyter Notebook
  - Modul python untuk data science

### **3.6.1 Identifikasi situs restriksi**

Kami mendapatkan data mengenai situs restriksi dengan software Restriction Mapper 3 (<http://restrictionmapper.org>). Kami menggunakan vektor pCAMBIA1300.

- Situs restriksi unik MCS pCAMBIA1300:

HindII, SmaI, AccI, ApoI, AvaI, BamHI, BspMI, EcoRI, HindIII, PleI, Sall, XbaI, XhoII, KpnI, NspI, PstI, SacI, SduI, SphI, Sse8387.

Tabel 3.6.1.1 sisi restriksi pada setiap gen (sumber : Restriction Mapper Version 3.0)

Gen	Non Cutter
psbA Promoter	AarI, AatII, AbsI, AccI, AclI, AflII, AflIII, AgsI, AjuI, AlfI, AloI, AlwNI, ApaI, ApaLI, ApoI, ArsI, AscI, AsuII, AvaI, AvaII, AvrII, BaeI, BamHI, BarI, BbvCI, BciVI, BclI, BdaI, BfiI, BglII, BplI, Bpu10I, BsaAI, BsaBI,

	BsaXI, BseRI, BseSI, BsgI, BsmAI, Bsp1407I, BspHI, BspMI, BsrDI, BstXI, BtrI, BtsI, ClaI, CspCI, DraII, DraIII, DrdI, Eam1105I, Eco31I, Eco47III, Eco57I, Eco57MI, EcoNI, EcoRI, EcoRV, Esp3I, Fall, FseI, FspAI, GsuI, HaeIV, HgaI, Hin4I, HindII, HindIII, HpaI, Hpy99I, KpnI, MauBI, MfeI, MluI, NaeI, NheI, NmeAIII, NotI, OliI, PacI, PasI, PfoI, PleI, PmaCI, PmeI, PpiI, PpuMI, PshAI, PsiI, PI-PspI, PspXI, PsrI, PstI, PvuI, PvuII, RsrII, SacI, SalI, SanDI, SapI, Scal, PI-SceI, SexAI, SfiI, SgfI, SgrDI, SmaI, SmII, SnaBI, SpeI, SrfI, Sse8387I, SspI, StuI, SwaI, TaqII, TatI, TspRI, TstI, Tth111I, VspI, XbaI, XcmI, XhoI, Xmnl
Proinsulin	AatII, AbsI, AccI, AcII, AflIII, AflIII, AgeI, AgsI, AjuI, AlfI, AloI, ApaLI, ApoI, ArsI, AscI, AsuII, AvrII, BaeI, Ball, BamHI, BarI, BbvCI, BccI, BcgI, BciVI, BclI, BdaI, BfI, BglII, BplI, Bpu10I, BsaAI, BsaBI, BseMII, BsePI, BsmI, BsmAI, Bsp1407I, BspHI, Bsrl, BsrBI, BsrDI, BstEII, BstXI, BtgZI, CfrI, ClaI, CspCI, DrdI, Eam1105I, EciI, Eco31I, Eco47III, Eco57I, EcoRI, EcoRV, Esp3I, Fall, FseI, FspAI, HaeIV, HgaI, Hin4I, HindII, HindIII, HpaI, Hpy99I, KpnI, MauBI, MfeI, MluI, MmeI, MsII, NcoI, NdeI, NheI, NotI, NruI, NspI, OliI, PacI, PasI, PflMI, PfoI, PleI, PmaCI, PmeI, PpiI, PshAI, PsiI, PI-PspI, PspXI, PsrI, PvuI, RsrII, SacI, SacII, SalI, SanDI, SapI, Scal, PI-SceI, SfaNI, SgfI, SgrAI, SgrDI, SmII, SnaBI, SpeI, SphI, SspI, StuI, StyI, SwaI, TaqII, TatI, TfII, TsoI, Tsp45I, TspDTI, TspGWI, TstI, Tth111I, VspI, XbaI, XcmI, XhoII, Xmnl
NOS	AarI, AatII, AbsI, AccI, AcII, AcyI, AgeI, AjuI, AlfI, AloI, AlwNI, ApaI, ApaLI, ApoI, ArsI, AscI, AsuII, Aval, Aval, AvrII, BaeI, Ball, BamHI, BarI, BbvI, BbvCI, BcgI, BciVI, BclI, BdaI, BfI, BglII, BplI, Bpu10I, BsaAI, BsaXI, BseMII, BseRI, BseSI, BseYI, BsgI, BsmI, BsmAI, Bsp1407I, BspHI, BspMI, Bsrl, BsrBI, BsrDI, BstEII, BstXI, BtrI, BtsI, CfrI, ClaI, CspCI, DraII, DraIII, DrdI, Eam1105I, EciI, Eco31I, Eco47III, Eco57I, Eco57MI, EcoNI, EcoP15I, EcoRI, EcoRII, EcoRV, Esp3I, Fall, FokI, FseI, FspAI, GsuI, HaeII, HgaI, HindII, HindIII, HpaI, HphI, Hpy99I, KpnI, MboII, MfeI, MluI, MmeI, MsII, NaeI, NarI, NcoI, NdeI, NheI, NmeAIII, NotI, NruI, OliI, PacI, PasI, PflMI, PfoI, PmaCI, PmeI, PpiI, PpuMI, PshAI, PsiI, PI-PspI, PspXI, PsrI, PstI, PvuI, PvuII, RsrII, SacI, SacII, SalI, SanDI, SapI, Scal, PI-SceI, SduI, SexAI, SfaNI, SfiI, SgfI, SgrAI, SgrDI, SmaI, SnaBI, SpeI, SphI, SrfI, Sse8387I, SspI, StuI, StyI, SwaI, TaqII, TatI, TauI, TseI, TsoI, Tsp45I, TspDTI, TspGWI, TspRI, TstI, Tth111I, VspI, XbaI, XcmI, XhoII, Xmnl

### 3.6.2 Optimasi Kodon Proinsulin dengan dnachisel

Sequence proinsulin yang kami inginkan berada dalam dua exon berbeda, yang pertama pada basa 239..425, dan yang kedua pada basa 1213..1358, yang dipisahkan oleh intron sepanjang 787 bp. Untuk mendapatkan sequence proinsulin yang tidak terpisah oleh intron, kami memutuskan untuk melakukan langkah-langkah berikut:

1. Transkripsi ke RNA
2. Splicing
3. Reverse Transcriptase PCR ke DNA

Setelah langkah-langkah diatas selesai, akan didapatkan sequence berikut:  
239 at

```
241 ggcctgtgg atgcgcctcc tgcccctgct ggccgtgtcg gccctctggg gacctgaccc
301 agccgcagcc ttgtgaacc aacacctgtg cggctcacac ctgttggaaag ctctctaccc
361 agtgtgcggg gaacgaggct tcttctacac acccaagacc cgccgggagg cagaggacct
421 gcaggtggggcag gtggagctgg gcgggggccc ttgtgcaggc agcctgcagc
1261 cttggccct ggaggggtcc ctgcagaagc gtggcattgt ggaacaatgc tgtaccagca
1321 tctgtccctt ctaccagctg gagaactact gcaactag
```

Warna biru menandakan exon pertama, dan warna merah menandakan exon kedua.

Dengan hasil translasi berikut:

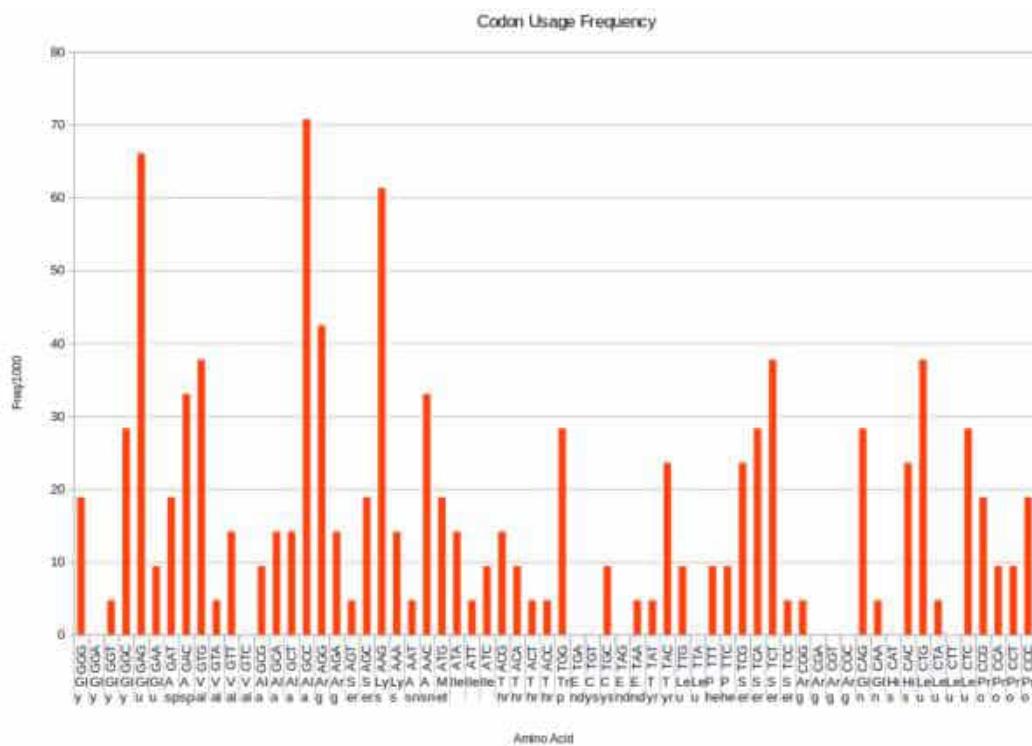
```
MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERG
FFYTPKTRREAEDLQVGQVELGGPGAGSLQPLALE GSLQKRGIVEQC
CTSICSLYQLENYCN*
```

Untuk meningkatkan ekspresi, maka kami melakukan codon optimization menggunakan python dengan modul dnachisel dan biopython. Codon optimisasi kami didasarkan pada index spesies 34199.chloroplast dari kazusa.

Codon usage table yang kami gunakan:

Tabel 3.6.1 Codon kloroplas Aloe vera (gbpln) 1 CDS's (516 kodon), kazusa spesies 34199.chloroplast

chloroplast Aloe vera [gbpln]: 1 CDS's (516 codons)									
fields: [triplet] [amino acid] [frequency] [per thousand] ([number])									
UUU F 0.09 65.9 {	34)	UCU S 0.35 34.9 {	18)	UAU Y 0.91 58.1 {	30)	UGU C 0.08 11.6 {	7)		
UUC F 0.31 29.1 {	15)	UCC S 0.08 7.8 {	4)	UAC Y 0.09 5.8 {	31)	UGC C 0.12 1.9 {	1)		
UUA L 0.33 38.8 {	20)	UCA S 0.31 31.0 {	16)	UAA * 0.00 0.0 {	0)	UGA * 1.00 1.9 {	1)		
UUG L 0.18 21.3 {	11)	UCG S 0.10 9.7 {	5)	UAG * 0.00 0.0 {	0)	UGG W 1.00 11.6 {	6)		
CUU L 0.28 32.9 {	17)	CCU P 0.36 9.7 {	5)	CAU H 0.89 46.5 {	24)	CGU R 0.15 11.6 {	6)		
CUC L 0.05 5.8 {	3)	CCC P 0.14 3.9 {	2)	CAC H 0.11 5.8 {	3)	CGC R 0.05 3.9 {	2)		
CUA L 0.15 17.4 {	9)	CCA P 0.29 7.8 {	4)	CAA Q 0.74 27.1 {	14)	CGA R 0.38 29.1 {	15)		
CUG L 0.02 1.9 {	1)	CCG P 0.21 5.8 {	3)	CAG Q 0.26 9.7 {	5)	CGG R 0.08 5.8 {	3)		
AUU I 0.58 50.4 {	26)	ACU T 0.44 15.5 {	8)	AAU N 0.00 38.8 {	20)	AGU S 0.14 11.6 {	7)		
AUC I 0.16 13.6 {	7)	ACC T 0.22 7.8 {	4)	AAC N 0.28 9.7 {	5)	AGC S 0.02 1.9 {	1)		
AUA I 0.27 23.3 {	12)	ACA T 0.22 7.8 {	4)	AAA K 0.81 40.7 {	21)	AGA R 0.28 21.3 {	11)		
AUG M 1.00 11.6 {	6)	ACG T 0.11 3.9 {	2)	AAG K 0.19 9.7 {	5)	AGG R 0.05 3.9 {	2)		
GUU V 0.48 19.4 {	10)	GCU A 0.50 13.6 {	7)	GAU D 0.07 25.2 {	13)	GGU G 0.45 9.7 {	5)		
GUC V 0.18 3.9 {	2)	GCC A 0.07 1.9 {	1)	GAC D 0.13 3.9 {	2)	GGC G 0.09 1.9 {	1)		
GUA V 0.38 15.5 {	8)	GCA A 0.29 7.8 {	4)	GAA E 0.05 44.6 {	23)	GGA G 0.18 3.9 {	2)		
GUG V 0.05 1.9 {	1)	GCG A 0.14 3.9 {	2)	GAG E 0.15 7.8 {	4)	GGG G 0.27 5.8 {	3)		



Grafik 3.6.1 Frekuensi Codon Usage kazusa spesies 34199.chloroplast

Hasil optimasi berdasarkan codon usage table diatas adalah sebagai berikut:

ATGCCCTGTGGATGAGGCTGCTGCCCTGCTGGCCCTGCTGGCCCT  
GTGGGGCCCGAACCCGCCGCCGCCTTGTGAACCAGCACCTGTGC  
GGCTCTCACCTGGTGGAGGCCCTGTACCTGGTGTGCGGCGAGAGGG  
GCTTCTTCTACACGCCAAGACGAGGGAGGGAGGCCGAGGACCTGCA  
GGTGGGCCAGGTGGAGCTGGCGGCCGGCGCCGGCTCTCTG  
CAGCCCCTGCCCTGGAGGGCTCTGCAGAACAGAGGGCATAGTGG  
AGCAGTGCTGCACGTCTATGCTCTGTACAGCTGGAGAACTAC  
TGCAACTAA

Hasil translasi setelah optimasi:

MALWMRLLPLLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERG  
FFYTPKTRREAEDLQVGQVELGGPGAGSLQPLALEGSLQKRGIVEQC  
CTSICSLYQLENYCN\*

Alignment sequence sebelum dan sesudah optimasi:

Gambar 3.6.12 Pairwise alignment sequence proinsulin sebelum dan sesudah optimasi kodon

Alignment hasil translasi sebelum dan sesudah optimasi:

NP_001278826.	1	MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFY	50
NP_001278826.	1	MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFY	50
NP_001278826.	51	TPKTRREAEDLQVGQVELGGPGAGSLQPLALEGSLQKRKGIVEQCCTSIC	100
NP_001278826.	51	TPKTRREAEDLQVGQVELGGPGAGSLQPLALEGSLQKRKGIVEQCCTSIC	100
NP_001278826.	101	SLYQLENYCN*	111
NP_001278826.	101	SLYQLENYCN*	111

Gambar 3.6.3 Pairwise alignment hasil translasi sebelum dan sesudah optimasi kodon

### 3.6.3 Optimasi PCR

Sistem ini dimaksudkan untuk membantu mengevaluasi kuantitas dari primer dan nukleotida untuk reaksi PCR yang optimal dan menentukan jumlah siklus yang dibutuhkan. Estimasi ini bersifat mendasar. Parameter default berhubungan dengan amplifikasi fragmen 2Kb dari 0.5 µg dari DNA manusia. Sehingga diasumsikan :

1. Kondisi dari reaksi mendekati normal, sehingga tidak perlu khawatir tentang teorinya (Kebanyakan polimerase bisa menyebabkan amplifikasi

yang tidak spesifik, kebanyakan primer bisa menyebabkan primer-dimer);

2. A,T, dan G,C secara ekuivalen dipresentasikan pada hasil PCR;
3. Tidak terdapat primer-dimers;
4. Taq polimerase tidak kehilangan keaktifannya selama reaksi.

Jika:

- Panjang dari produk PCR adalah “L” [kbp];
- Konsentrasi dNTP’s adalah “c” [mM];
- Kuantitas primer adalah “q” [pmol];
- Kuantitas dari Taq polimerase adalah “a” [u];
- Volume reaksi adalah “V” [ $\mu$ l];
- Waktu elongasi adalah “t” [min];
- Kuantitas template adalah “mo”.

Lalu:

1. Hasil yang maksimal merupakan hasil minimum dari dua evaluasi di bawah ini:

Jika semua nukleotida digunakan :

$$mn = 4[\text{nukleotida}] \times c[\text{mmol/l}] \times 324.5 \text{ [g/mol]} \times V [\mu\text{l}] = 1300 cV \text{ [ng]}$$

Jika semua primer digunakan :

$$mp = q[\text{pmol}] \times 2[\text{strands}] \times 324.5[\text{g/mol}] \times L[\text{kbp}] = 650qL \text{ [ng]}$$

2. Kuantitas maksimum dari produk PCR per satu siklus bergantung pada dua faktor :

a. Kecepatan taq polimerase : 2-4 [kbp/min];

b. Aktivitas taq polymerase (1 u adalah jumlah dari enzim, yang berinkorporasi 10 nmol dari keempat dNTPs in 30 min pada 72oC.

$$m_{\text{cycle}} = 10[\text{nmol}] \times 324.5[\text{g/mol}] \times a [\text{u}] \times t[\text{min}] / 30[\text{min}] = 108at \text{ [ng]}$$

3. Jumlah siklus yang berguna untuk sintesis dari “mmax” produk PCR adalah:

$$m_{\text{max}} = 2^n \times mo \Rightarrow n = \ln(m_{\text{max}}/mo)/\ln$$

4. Hubungan antara massa dan kuantitas mol adalah :

$$m[\mu\text{g}] = 649[\text{g/mol}] \times q[\mu\text{mol}] \times L[\text{kbp}] \times 1000$$

### **3.6.4 Konstruksi transgen menggunakan snapgene**

Tahap pertama yang kami lakukan adalah mencari primer dan situs restriksi berdasarkan hasil dari restriction mapper. Dengan mempertimbangkan data dari restriction mapper, kami memutuskan untuk menggunakan 4 enzim restriksi untuk menyambungkan ketiga gen. Keempat enzim restriksi tersebut adalah KpnI, BamHI, SphI dan HindIII, diurutkan dari kiri ke kanan pada vektor akhir. Keempat situs restriksi itu dipilih karena mereka berempat merupakan non cutter bagi ketiga gen. Untuk menyambungkan ketiga gen dengan MCS pCAMBIA1300, kami memutuskan untuk menggunakan teknik insertion cloning dengan enzim ligase T4.

Tahap tahap yang kami lakukan adalah sebagai berikut:

1. Menambahkan restriction site pada kedua ujung gene of interest dengan PCR;
2. Elektroforesis virtual hasil PCR dengan benchling;
3. Melakukan insertion cloning.

#### **Tahap 1:**

Kami melakukan PCR virtual dalam snapgene untuk menghasilkan gen yang diapit oleh enzim restriksi. Primer yang kami gunakan didesain otomatis oleh snapgene, sedangkan situs restriksi kami tambahkan secara manual dengan menggunakan data dari restriction mapper sebagai pertimbangan.

Berikut adalah list primer yang kami gunakan:

1. psbA Promoter
  - a. Forward: 5'-(KpnI)ATGACCGCGATTCTGGAACG-3'
  - b. Reverse: 5'-(BamHI)GCCGTTGGTGCTCGGC-3'
2. Proinsulin
  - a. Forward: 5'-(BamHI)ATGGCCCTGTGGATGAGGC-3'
  - b. Reverse: 5'-(SphI)TTAGTTGCAGTAGTTCTCCAGCTGGT-3'
3. NOS Terminator
  - a. Forward:  
5'-(SphI)GATCGTTCAACATTGGCAATAAAGTTCTTAAGA-3'
  - b. Reverse: 5'-(HindIII)GATCTAGAACATAGATGACACCGCGC-3'

## Tahap 2:

Setelah PCR selesai, kami melakukan elektroforesis dan virtual digest dengan bantuan benchling. Hasil dari virtual digest adalah sebagai berikut:

Enzymes	Cuts	Temp.	1.1	2.1	3.1	4/CS
BamHI	1	37°C	75*	100*	100	100*
KpnI	1	37°C	100	75	10	50*
Start	End	Length	Left Cutter	Left Overhang	Right Cutter	Right Overhang
1	5	5	None	blunt	KpnI	3'
6	1066	1061	KpnI	3'	BamHI	5'
1067	1071	5	BamHI	5'	None	blunt

Gambar 3.6.3 Virtual digest psbA Promoter dengan enzim BamHI dan Kpn I (Benchling)

Enzymes	Cuts	Temp.	1.1	2.1	3.1	4/CS
BamHI	1	37°C	75*	100*	100	100*
SphI	1	37°C	100	100	50	100
Start	End	Length	Left Cutter	Left Overhang	Right Cutter	Right Overhang
1	1	1	None	blunt	BamHI	5'
2	344	343	BamHI	5'	SphI	3'
345	345	1	SphI	3'	None	blunt

Gambar 3.6.4 Virtual digest Proinsulin dengan enzim BamHI dan SphI (Benchling)

Enzymes	Cuts	Temp.	1.1	2.1	3.1	4/CS
HindIII	1	37°C	25	100	50	50
SphI	1	37°C	100	100	50	100
Start	End	Length	Left Cutter	Left Overhang	Right Cutter	Right Overhang
1	5	5	None	blunt	SphI	3'
6	260	255	SphI	3'	HindIII	5'
261	265	5	HindIII	5'	None	blunt

Gambar 3.6.5 Virtual digest NOS Terminator dengan enzim HindIII dan SphI (Benchling)

Hasil dari elektroforesis virtual:

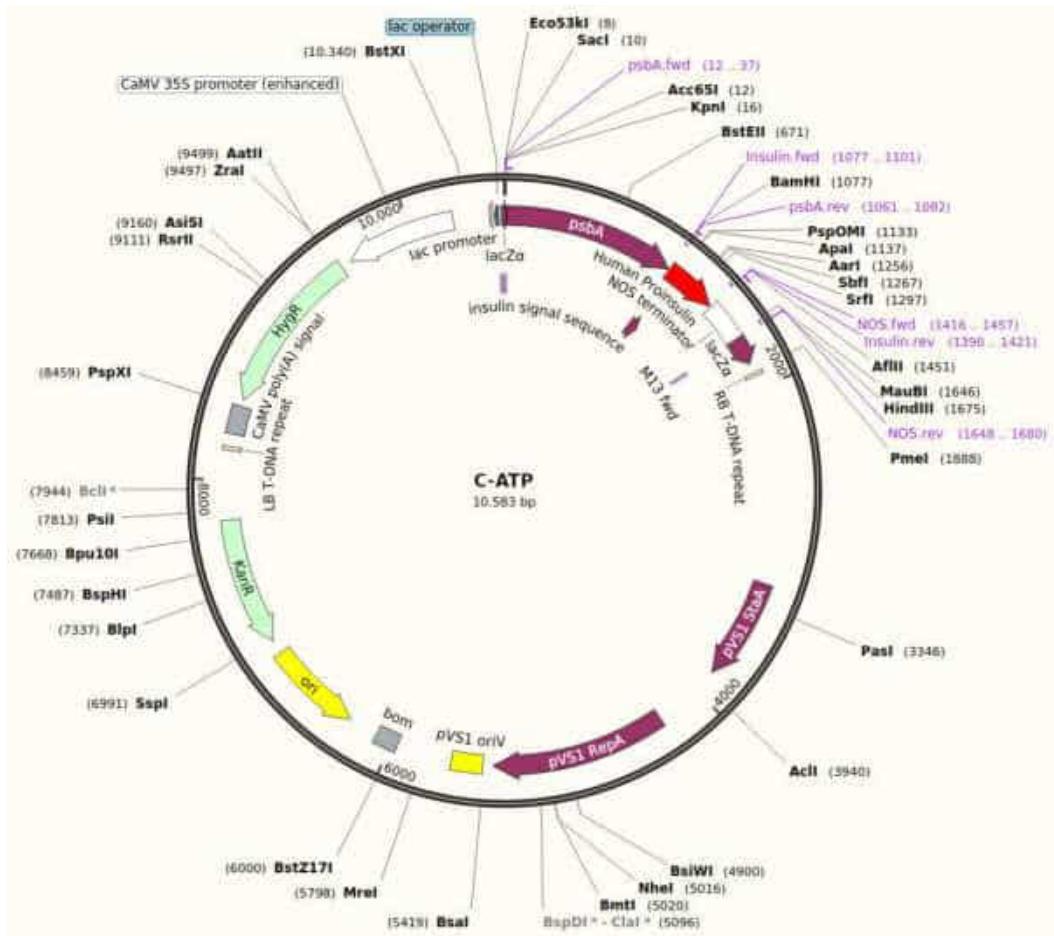
Ladder	Life 1 kb Plus
1	psbA Promoter - BamHI KpnI
2	Optimized Human Insulin - BamHI SphI
3	NOS Terminator - HindIII SphI



Gambar 3.6.6 Elektroforesis virtual promoter psbA, insulin, NOS (benchling)

### Tahap 3:

Setelah virtual digest dan elektroforesis selesai, kami memasukkan ketiga sequence kedalam pCAMBIA menggunakan insertion cloning dengan bantuan software snapgene.



Gambar 3.6.7 Plasmid akhir hasil insertion cloning (Benchling)

### 3.6.5 Visualisasi Protein

Setelah plasmid akhir terbentuk, maka tahap selanjutnya adalah memvisualisasikan protein ke dalam model 3D untuk selanjutnya dianalisa lebih lanjut menggunakan komputer. Untuk visualisasi, kami akan menggunakan software robetta yang tersedia secara online pada website [robetta.bakerlab.org](http://robetta.bakerlab.org). Model 3D lalu akan dianalisa dengan menggunakan PyMol<sup>[205-207]</sup>.

## Bab 4

### Hasil dan Pembahasan

#### 4.1 Konstruksi Transgen

##### 4.1.1 Optimasi Kodon

Optimasi kodon kami lakukan dengan tujuan meningkatkan hasil ekspresi. Kode genetik terdiri dari 64 kodon berbeda yang mengkode 21 jenis asam amino dan beberapa stop kodon, hal ini berarti terdapat beberapa kodon berbeda yang mengkode asam amino yang sama. Oleh karena itu, walaupun hasil asam amino nya sama, setiap spesies memiliki kecenderungan untuk menggunakan suatu kodon tertentu, dibandingkan dengan kodon lainnya. Hal inilah yang menyebabkan adanya Codon Usage Bias, yaitu perbedaan frekuensi kodon sinonim (kodon yang mengkode asam amino yang sama) dalam DNA. Sebagai contoh, asam leucine dikode oleh 6 macam kodon, beberapa diantaranya sangat jarang ditemukan, dengan menyesuaikan penggunaan kodon berdasarkan preferensi spesiesnya, kodon leucine yang jarang digunakan akan diganti dengan kodon leucine yang lebih sering digunakan. Hal ini diperkirakan dapat meningkatkan hasil dari ekspresi protein<sup>[197]</sup>.

Meski begitu, sampai dengan saat ini, masih belum ada metode yang dapat memprediksi hasil ekspresi dengan akurat. Walaupun optimasi kodon dapat meningkatkan hasil ekspresi, hal itu tidak menjamin kesuksesan ekspresi, dan peningkatan hasil ekspresi oleh optimasi kodon akan bervariasi tergantung pada jenis protein dan organisme. Selain itu, ada banyak faktor lain yang mempengaruhi hasil ekspresi seperti kopian tRNA<sup>[198]</sup>, stabilitas mRNA<sup>[199]</sup>, kinetika pelipatan protein<sup>[200]</sup>, stabilitas protein, transpor protein, toksitas protein dalam lingkungan ekspresi sel, dan berbagai faktor lainnya. Oleh karena itu, hasil optimasi kodon perlu diverifikasi dengan eksperimen.

Kami menggunakan DnaChisel untuk melakukan optimasi kodon, dan setelahnya kami melakukan pairwise alignment dengan algoritma Smith-Waterman yang disediakan oleh EMBL-EBI<sup>[201]</sup>.

Alignment sequence sebelum dan sesudah optimasi:

NP_001278826.	1	ATGGCCCTGTGGATGCGCCTCTGCCCTGCTGGCCTGCTGGCCCTCTG        . .   .   . .   . .   .	50
NP_001278826.	1	ATGGCCCTGTGGATGAGGCTGCTGCCCTGCTGGCCCTGCTGGCCCTGTG        . .   .   . .   . .   .	50
NP_001278826.	51	GGGACCTGACCCAGCCGAGCCTTGTGAACCAACACCTGTGCGGCTCAC    . .   .   . .   . .   . .   .	100
NP_001278826.	51	GGGCCGGACCCCGCCGCCCTTGTGAACCAAGCACCTGTGCGGCTCTC        . .   .   . .   .	100
NP_001278826.	101	ACCTGGTGAAGCTCTCACCTAGTGTGCGGGAACGAGGCTTCTTCTAC        . .   .   . .   . .   .	150
NP_001278826.	101	ACCTGGTGGAGGCCCTGTACCTGGTGTGCGCGAGAGGGCTTCTTCTAC        . .   .   . .   .	150
NP_001278826.	151	ACACCCAAGACCCGGGGAGGCAGAGGACCTGCAGGTGGGCAGGTGGA    .   . ..   .   . .   .	200
NP_001278826.	151	ACGCCAAGACGAGGGAGGGAGGCCGAGGACCTGCAGGTGGGCCAGGTGGA        . .   .   .	200
NP_001278826.	201	GCTGGCGGGGGCCCTGGTCAGGCAGCCTGCAGCCCTGGCCCTGGAGG        . .   .   . .   .	250
NP_001278826.	201	GCTGGCGCGGCCGGCCGGCGCCGGCTCTGCAGCCCTGGCCCTGGAGG        . .   .   . .   .	250
NP_001278826.	251	GGTCCTGCAGAACGCTGGCATTTGTGGAAACAATGCTGTACAGCATCTGC  . .   . .   .   . .   .	300
NP_001278826.	251	GCTCTCTGCAGAACGGGCATAGTGGAGCAGTGCTGCACGTCTATATGC        . .   .   .	300
NP_001278826.	301	TCCCTCTACCAGCTGGAGAACACTACTGCAACTAG      .   .	333
NP_001278826.	301	TCTCTGTACCAGCTGGAGAACACTACTGCAACTAA        . .   .	333

Gambar 4.1.1 Pairwise alignment sequence proinsulin sebelum dan sesudah optimasi kodon

Alignment hasil translasi sebelum dan sesudah optimasi:

NP_001278826.	1	MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFY	50
NP_001278826.	1	MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFY	50
NP_001278826.	51	TPKTRREAEDLQVGQVELGGPGAGSLQPLALEGSLQKRKGIVEQCCTSIC	100
NP_001278826.	51	TPKTRREAEDLQVGQVELGGPGAGSLQPLALEGSLQKRKGIVEQCCTSIC	100
NP_001278826.	101	SLYQLENYCN*	111
NP_001278826.	101	SLYQLENYCN*	111

Gambar 4.1.2 Pairwise alignment hasil translasi sebelum dan sesudah optimasi kodon

Berdasarkan gambar diatas, dapat kita lihat adanya perbedaan antara sequence proinsulin sebelum dan sesudah optimasi kodon, namun hasil translasinya tetap sama. Oleh karena itu, kami menyimpulkan bahwa sequence proinsulin tersebut dapat diekspresikan di dalam sel Aloe vera. Meski begitu, dikarenakan tidak adanya metode yang dapat memprediksi hasil optimasi secara akurat, eksperimen lebih lanjut diperlukan untuk memastikan keberhasilan ekspresi.

#### **4.1.2 Virtual Digest dan Virtual Electrophoresis**

Virtual digest dilakukan dalam rangka mensimulasikan pemotongan pada situs restriksi oleh enzim restriksi saat konstruksi transgen. Virtual digest akan memprediksi lokasi pemotongan virtual, serta panjang dari hasil pemotongan. Setelah itu, kami melakukan virtual electrophoresis untuk mendapatkan data ukuran dari masing masing sequence yang dapat digunakan sebagai perbandingan dalam uji laboratorium *in vivo*.

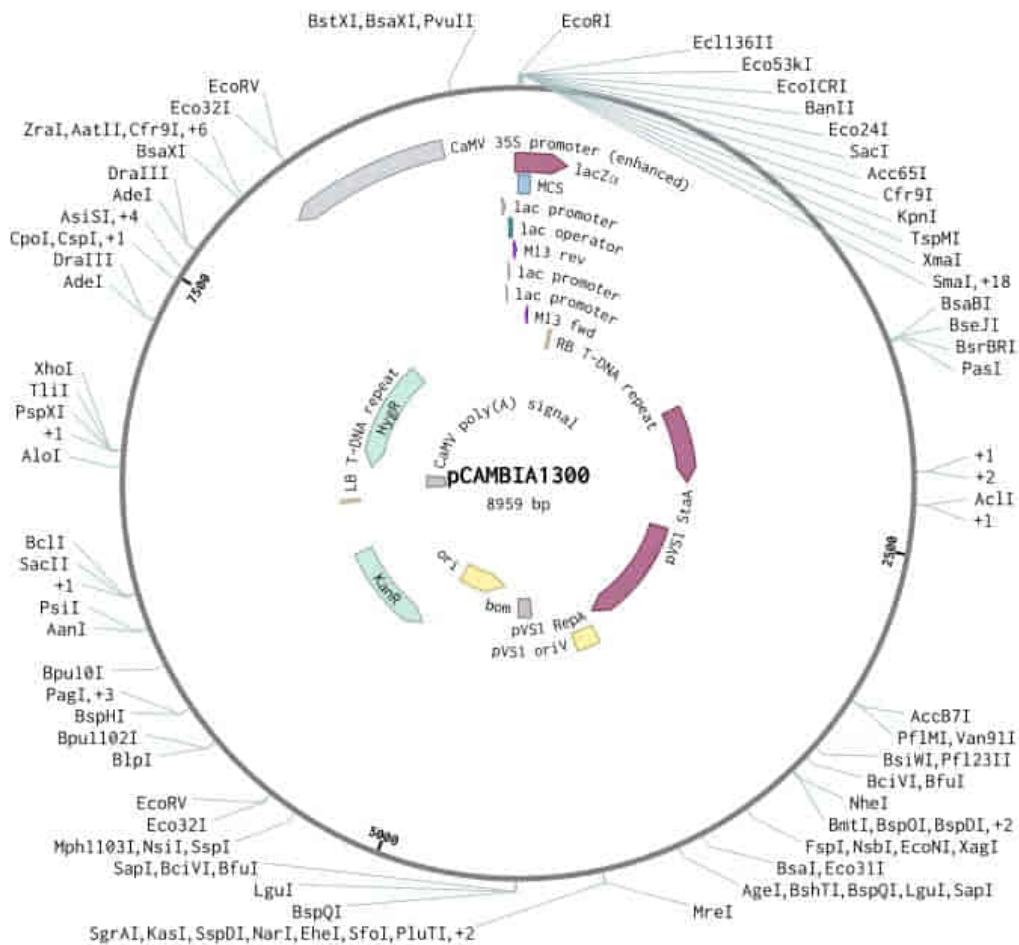
#### **4.1.3 Plasmid Transgenik**

Selanjutnya, kami menggabungkan ketiga sequence yang diinginkan ke dalam plasmid. Plasmid yang kami pilih adalah pCAMBIA 1300 Cloning Vector. Kami memilih pCAMBIA karena ia merupakan salah satu plasmid yang paling sering digunakan, sehingga ia lebih teruji secara laboratorium dibandingkan plasmid plasmid lain. Lalu, kami memilih pCAMBIA 1300 karena ia tidak memiliki GUS reporter gene yang bersifat toxin, dan juga telah memiliki gen hygromycin resistance, dan kanamycin resistance, yang mengkode ketahanan terhadap higromisin dan kanamisin.

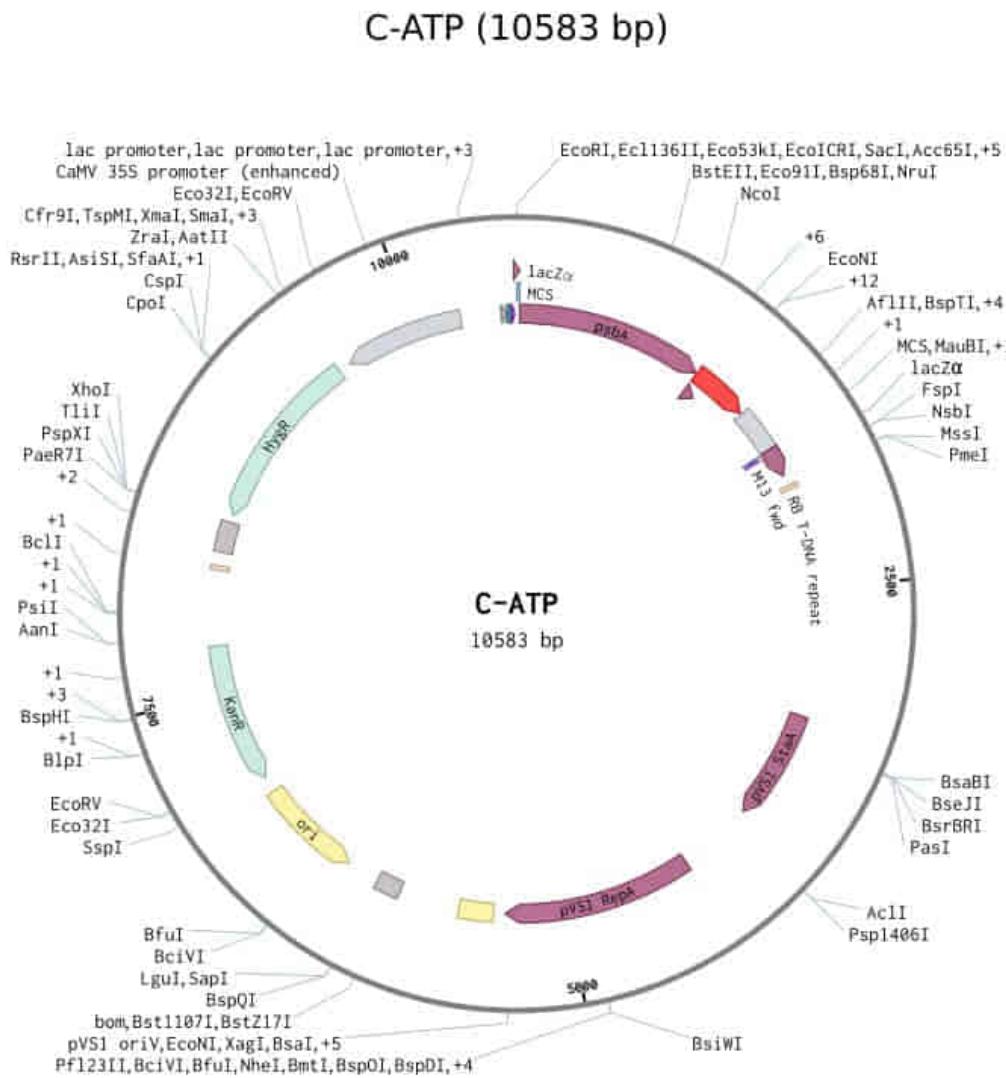
Kemudian kami memasukkan ketiga gene of interest kedalam multiple cloning site (MCS) dari pCAMBIA 1300 searah jarum jam, berurutan dari psbA promoter, proinsulin, dan NOS terminator. Secara keseluruhan, proses insersi menambahkan sebanyak 1659 pasang basa, yang terdiri dari 1061 basa promoter psbA, 343 basa proinsulin, dan 255 basa NOS terminator. Selain itu, insersi juga menghilangkan sebanyak 35 pasang basa dari situs restriksi di dalam MCS yang terpotong saat insersi. Totalnya, terdapat pertambahan sebanyak 1624 pasang basa, dari 8959 pasang basa pCAMBIA 1300 menjadi 10583 pasang basa C-ATP.

Berikut adalah perbandingan pCAMBIA 1300 dan plasmid C-ATP:

## pCAMBIA1300 (8959 bp)



Gambar 4.1.3 pCAMBIA1300 Cloning Vector



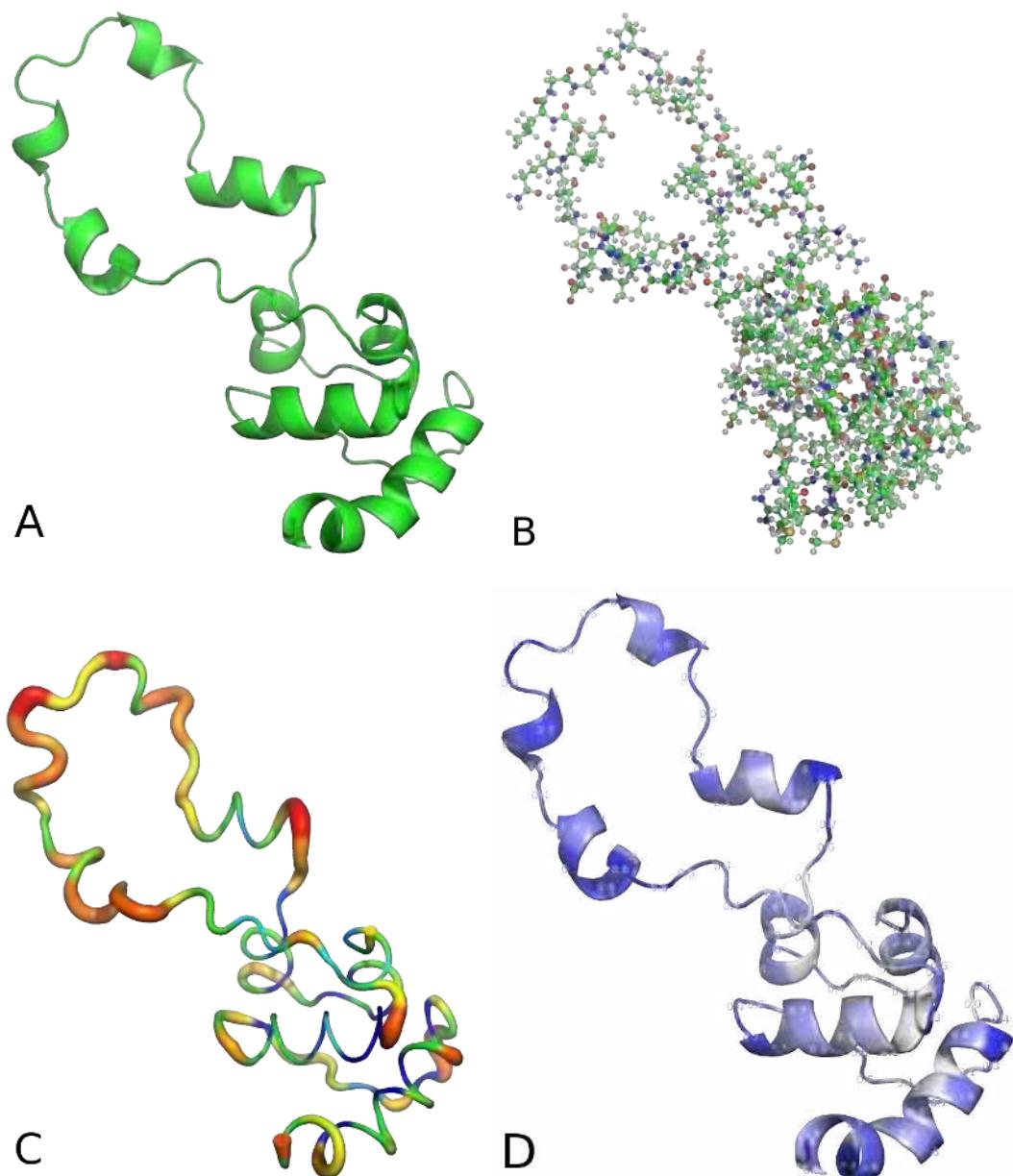
Gambar 4.1.4 Plasmid C-ATP

## 4.2 Visualisasi

Tahapan berikutnya adalah visualisasi dari protein. Kami menggunakan software roBetta untuk menghasilkan model pdb, yang selanjutnya divisualisasikan dengan menggunakan PyMol. Perlu diketahui bahwa belum ada metode yang dapat memvisualisasikan protein hasil percobaan *in silico* secara akurat, software roBetta sendiri menggunakan database protein yang ada dan memprediksikan modelnya berdasarkan dengan kemiripan dengan protein lain. Selain itu, roBetta juga dievaluasi secara terus menerus oleh CAMEO<sup>[202-204]</sup>.

Totalnya, terdapat 4 prediksi model berbeda untuk struktur insulin transgenik kami. Keempat model tersebut kemudian kami analisa dengan menggunakan PyMol<sup>[205-207]</sup>.

#### 4.2.1 Model 1



(Gambar 4.2.1. A). Protein interface B). Ball and Stick model C).  $\beta$ -Factor D). Surface per residue)

Tabel 4.2.1 Data jumlah atom dan berat molekular model 1

Bagian Model	Jumlah Atom	Berat Molekular
Model utuh	1671 atom	11971.7169 u
Polimer	1671 atom	11971.7169 u
Organik	0 atom	0.0000 u

Solvent	0 atom	0.0000 u
Hidrogen Polar	173 atom	174.3736 u
Hidrogen Non-Polar	659 atom	664.2325 u
Donor	148 atom	2096.9040 u
Akseptor	154 atom	2461.9149 u
Atom Permukaan	648 atom	3610.7895 u
C-Alpha	110 atom	1321.1770 u

Tabel 4.2.2 Data muatan model 1

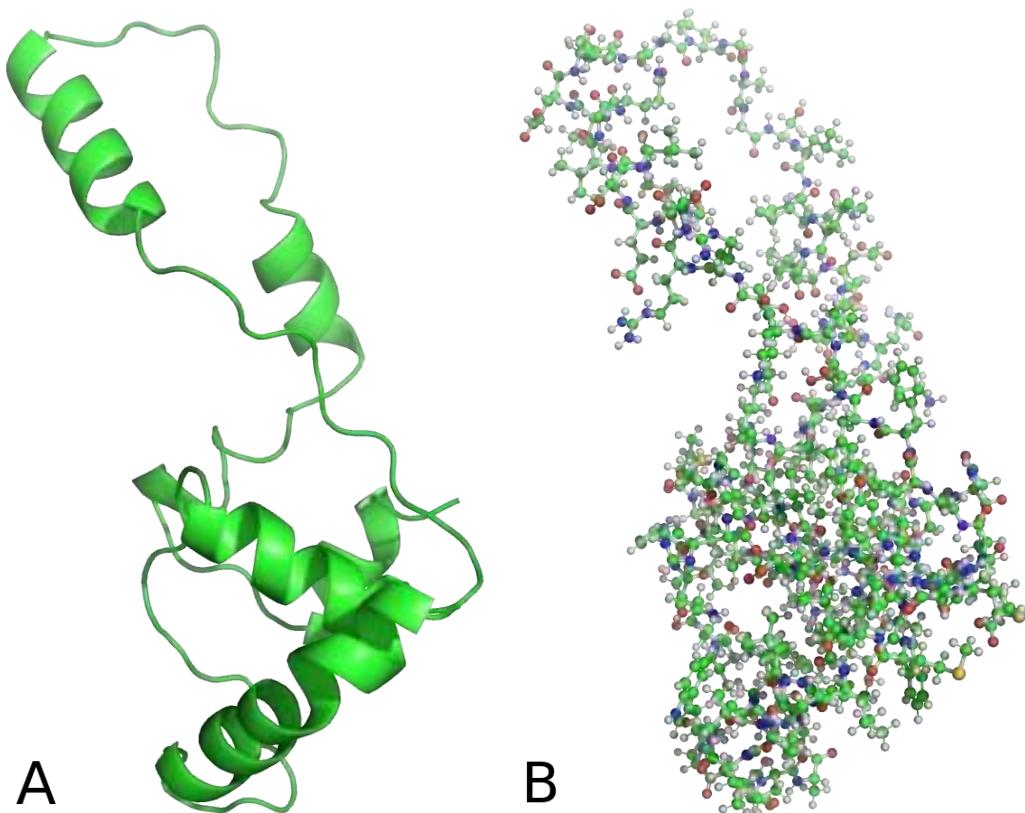
Bagian Model	Muatan Formal	Muatan Partial
Model utuh	-4	0.0000
Polimer	-4	0.0000
Organik	0	0.0000
Solvent	0	0.0000
Hidrogen Polar	0	0.0000
Hidrogen Non-Polar	0	0.0000
Donor	7	0.0000
Akseptor	-11	0.0000
Atom Permukaan	-6	0.0000
C-Alpha	0	0.0000

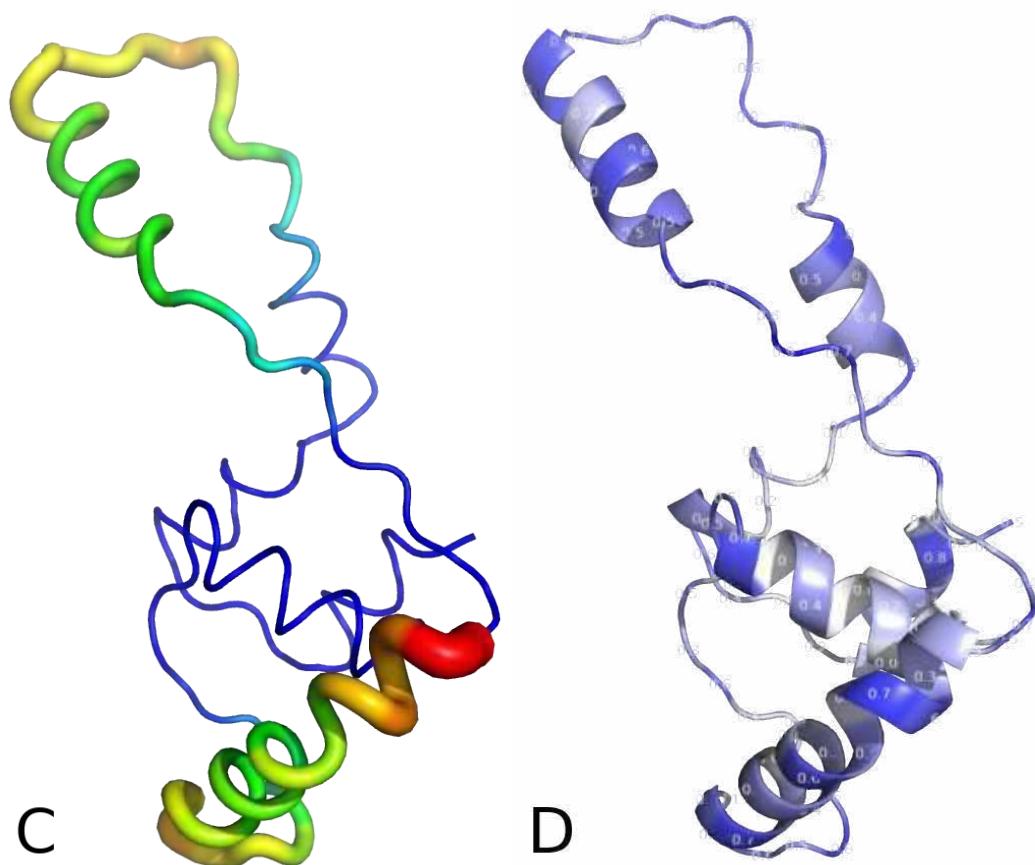
Tabel 4.2.3 Data luas permukaan model 1

Bagian Model	Luas Permukaan Molekular	Luas Permukaan Solvent Accessible
Model utuh	12442.948 Angstroms <sup>2</sup>	8148.517 Angstroms <sup>2</sup>
Polimer	12442.948 Angstroms <sup>2</sup>	8148.517 Angstroms <sup>2</sup>
Organik	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Solvent	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>

Hidrogen Polar	1147.248 Angstroms <sup>2</sup>	1182.374 Angstroms <sup>2</sup>
Hidrogen Non-Polar	5052.217 Angstroms <sup>2</sup>	4292.863 Angstroms <sup>2</sup>
Donor	841.471 Angstroms <sup>2</sup>	385.156 Angstroms <sup>2</sup>
Akseptor	2254.160 Angstroms <sup>2</sup>	1674.963 Angstroms <sup>2</sup>
Atom Permukaan	6046.378 Angstroms <sup>2</sup>	7797.255 Angstroms <sup>2</sup>
C-Alpha	332.974 Angstroms <sup>2</sup>	19.984 Angstroms <sup>2</sup>

#### 4.2.2 Model 2





(Gambar 4.2.2. A). Protein interface B). Ball and Stick model C).  $\beta$ -Factor D). Surface per residue)

Tabel 4.2.4 Data jumlah atom dan berat molekular model 2

Bagian Model	Jumlah Atom	Berat Molekular
Model utuh	1671 atom	11971.7169 u
Polimer	1671 atom	11971.7169 u
Organik	0 atom	0 u
Solvent	0 atom	0 u
Hidrogen Polar	173 atom	174.3736 u
Hidrogen Non-Polar	659 atom	664.2325 u
Donor	148 atom	2096.9040 u
Akseptor	154 atom	2461.9149 u
Atom Permukaan	668 atom	3602.9581 u
C-Alpha	110 atom	1321.1770 u

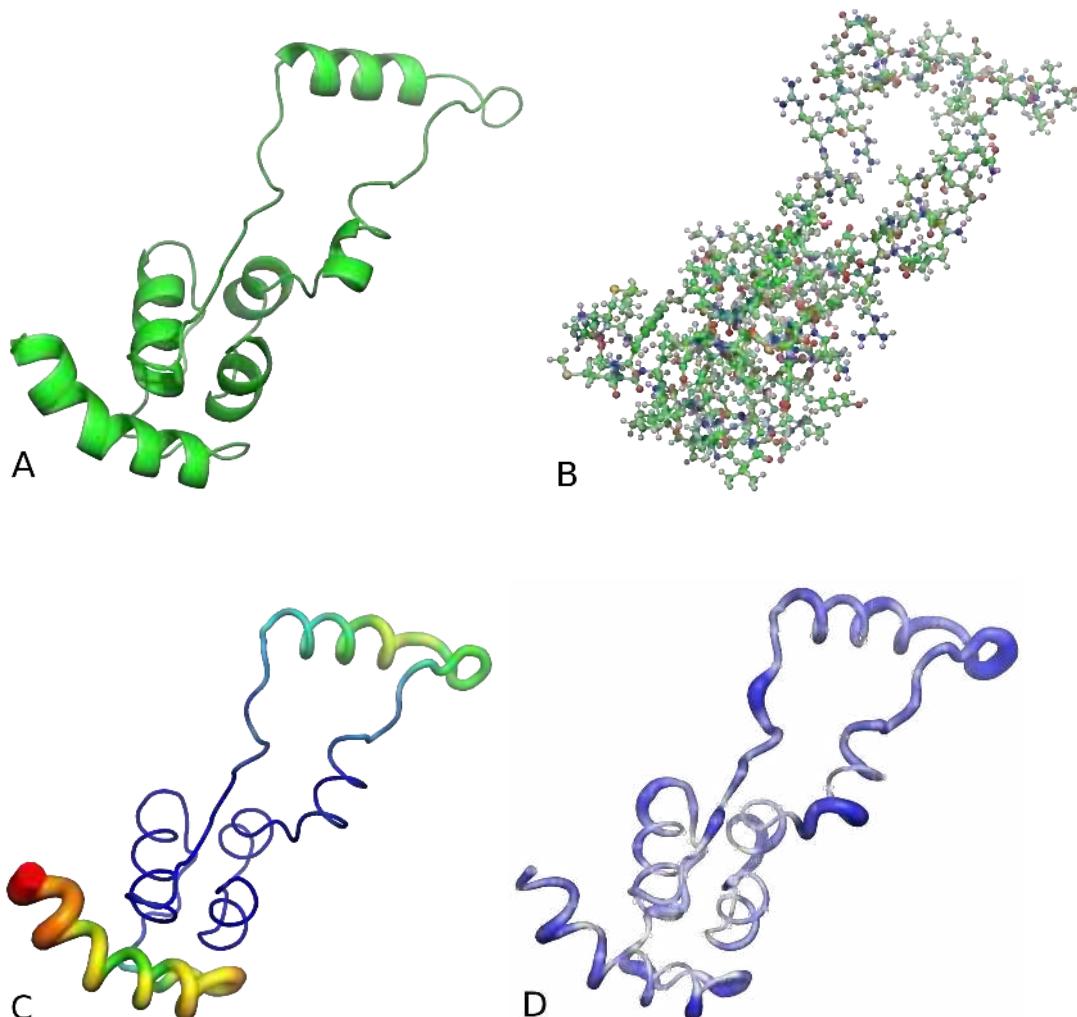
Tabel 4.2.5 Data muatan model 2

Bagian Model	Muatan Formal	Muatan Partial
Model utuh	-4	0.0000
Polimer	-4	0.0000
Organik	0	0.0000
Solvent	0	0.0000
Hidrogen Polar	0	0.0000
Hidrogen Non-Polar	0	0.0000
Donor	7	0.0000
Akseptor	-11	0.0000
Atom Permukaan	-6	0.0000
C-Alpha	0	0.0000

Tabel 4.2.6 Data luas permukaan model 2

Bagian Model	Luas Permukaan Molekular	Luas Permukaan Solvent Accessible
Model utuh	12401.562 Angstroms <sup>2</sup>	8540.121 Angstroms <sup>2</sup>
Polimer	12401.562 Angstroms <sup>2</sup>	8540.121 Angstroms <sup>2</sup>
Organik	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Solvent	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Hidrogen Polar	1159.096 Angstroms <sup>2</sup>	1150.314 Angstroms <sup>2</sup>
Hidrogen Non-Polar	5026.537 Angstroms <sup>2</sup>	4561.397 Angstroms <sup>2</sup>
Donor	825.722 Angstroms <sup>2</sup>	363.157 Angstroms <sup>2</sup>
Akseptor	2249.460 Angstroms <sup>2</sup>	1813.615 Angstroms <sup>2</sup>
Atom Permukaan	6216.088 Angstroms <sup>2</sup>	8228.303 Angstroms <sup>2</sup>
C-Alpha	331.766 Angstroms <sup>2</sup>	18.612 Angstroms <sup>2</sup>

#### 4.2.3 Model 3



(Gambar 4.2.3. A). Protein interface B). Ball and Stick model C).  $\beta$ -Factor D). Surface per residue)

Tabel 4.2.7 Data jumlah atom dan berat molekular model 3

Bagian Model	Jumlah Atom	Berat Molekular
Model utuh	1671 atom	11971.7169 u
Polimer	1671 atom	11971.7169 u
Organik	0 atom	0 u
Solvent	0 atom	0 u
Hidrogen Polar	173 atom	174.3736 u
Hidrogen Non-Polar	659 atom	664.2325 u

Donor	147 atom	2082.8973 u
Akseptor	155 atom	2475.9216 u
Atom Permukaan	642 atom	3534.7366 u
C-Alpha	110 atom	1321.1770 u

Tabel 4.2.8 Data muatan model 3

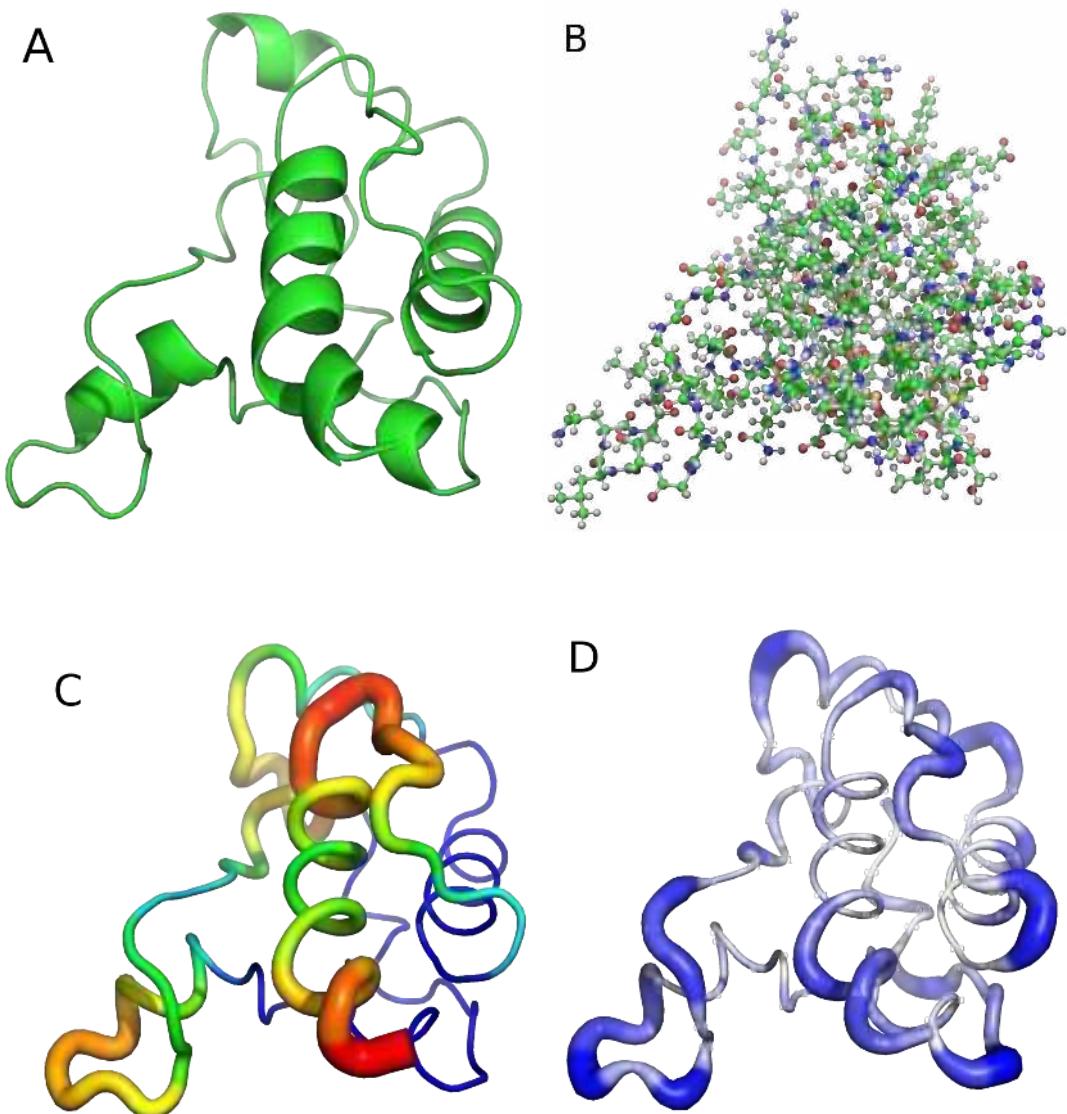
Bagian Model	Muatan Formal	Muatan Partial
Model utuh	-4	0.0000
Polimer	-4	0.0000
Organik	0	0.0000
Solvent	0	0.0000
Hidrogen Polar	0	0.0000
Hidrogen Non-Polar	0	0.0000
Donor	7	0.0000
Akseptor	-11	0.0000
Atom Permukaan	-6	0.0000
C-Alpha	0	0.0000

Tabel 4.2.9 Data luas permukaan model 3

Bagian Model	Luas Permukaan Molekular	Luas Permukaan Solvent Accessible
Model utuh	12379.827 Angstroms <sup>2</sup>	7946.748 Angstroms <sup>2</sup>
Polimer	12379.827 Angstroms <sup>2</sup>	7946.748 Angstroms <sup>2</sup>
Organik	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Solvent	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Hidrogen Polar	1154.202 Angstroms <sup>2</sup>	1148.252 Angstroms <sup>2</sup>
Hidrogen Non-Polar	5025.376 Angstroms <sup>2</sup>	4295.495 Angstroms <sup>2</sup>
Donor	823.042 Angstroms <sup>2</sup>	321.154 Angstroms <sup>2</sup>

Akseptor	2236.014 Angstroms <sup>2</sup>	1606.604 Angstroms <sup>2</sup>
Atom Permukaan	5983.686 Angstroms <sup>2</sup>	7635.942 Angstroms <sup>2</sup>
C-Alpha	330.181 Angstroms <sup>2</sup>	16.935 Angstroms <sup>2</sup>

#### 4.2.4 Model 4



(Gambar 4.2.4. A). Protein interface B). Ball and Stick model C). β-Factor D). Surface per residue)

Tabel 4.2.10 Data jumlah atom dan berat molekular model 4

Bagian Model	Jumlah Atom	Berat Molekular

Model utuh	1675 atom	11975.7486 u
Polimer	1675 atom	11975.7486 u
Organik	0 atom	0 u
Solvent	0 atom	0 u
Hidrogen Polar	177 atom	178.4054 u
Hidrogen Non-Polar	659 atom	664.2325 u
Donor	149 atom	2110.9107 u
Akseptor	153 atom	2447.9082 u
Atom Permukaan	534 atom	3045.8878 u
C-Alpha	110 atom	1321.1770 u

Tabel 4.2.11 Data muatan model 4

Bagian Model	Muatan Formal	Muatan Partial
Model utuh	-4	0.0000
Polimer	-4	0.0000
Organik	0	0.0000
Solvent	0	0.0000
Hidrogen Polar	0	0.0000
Hidrogen Non-Polar	0	0.0000
Donor	7	0.0000
Akseptor	-11	0.0000
Atom Permukaan	-7	0.0000
C-Alpha	0	0.0000

Tabel 4.2.12 Data luas permukaan model 4

Bagian Model	Luas Permukaan Molekular	Luas Permukaan Solvent Accessible
Model utuh	12305.951 Angstroms <sup>2</sup>	6627.765 Angstroms <sup>2</sup>

Polimer	12305.951 Angstroms <sup>2</sup>	6627.765 Angstroms <sup>2</sup>
Organik	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Solvent	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Hidrogen Polar	1135.465 Angstroms <sup>2</sup>	1096.120 Angstroms <sup>2</sup>
Hidrogen Non-Polar	4962.301 Angstroms <sup>2</sup>	3276.135 Angstroms <sup>2</sup>
Donor	831.292 Angstroms <sup>2</sup>	340.515 Angstroms <sup>2</sup>
Akseptor	2215.719 Angstroms <sup>2</sup>	1408.931 Angstroms <sup>2</sup>
Atom Permukaan	5043.382 Angstroms <sup>2</sup>	6346.040 Angstroms <sup>2</sup>
C-Alpha	335.935 Angstroms <sup>2</sup>	10.105 Angstroms <sup>2</sup>

### 4.3 Rencana Penelitian In-Silico Lanjutan

Penelitian in-silico selanjutnya akan lebih berpusat kepada analisis dan simulasi protein ketika berada di dalam tubuh manusia. Salah satu cara yang mungkin bisa dilakukan adalah melakukan docking, yaitu metode yang digunakan untuk memprediksi orientasi suatu molekul dengan molekul lainnya saat mereka berikatan untuk membentuk kompleks stabil<sup>[211]</sup>. Kami belum melakukan docking dikarenakan hal tersebut diluar batasan masalah kami, karena batasan masalah kami hanya mencapai rancangan produksi insulin di *Aloe vera* saja.

### 4.4 Rancangan Eksperimen In-Vivo

Tahap selanjutnya setelah menyelesaikan eksperimen in-silico, adalah melakukan eksperimen in-vivo, karena hasil dari eksperimen in-silico hanyalah prediksi yang dapat digunakan sebagai panduan dan perbandingan untuk eksperimen in-vivo dan in-vitro, sehingga eksperimen in-vivo atau in-vitro diperlukan untuk memastikan hasil dari in-silico. Dikarenakan batasan waktu dan kondisi saat ini yang tidak memungkinkan dilaksanakannya eksperimen in-vivo, maka kami hanya membuat rancangan saja, sebagai panduan untuk di kemudian hari.

Terdapat berbagai macam faktor berbeda yang dapat mempengaruhi hasil akhir eksperimen. Faktor ini terbagi dua, yaitu faktor biologis, dan faktor lingkungan. Faktor biologis yang perlu diperhatikan adalah konstruksi gen pada *cloning*

*vector*, seperti promotor, gene of interest, terminator, reporter gene, serta marker gene, yang masing masingnya penting bagi proses transformasi seperti yang tertera pada bab 2.8. Faktor biologis lainnya meliputi tipe jaringan, ukuran sel, umur kultur, fase mitosis, dan lain sebagainya.

Sedangkan faktor lingkungan meliputi suhu, kelembaban, serta intensitas cahaya, yang masing masingnya memiliki efek langsung kepada fisiologis jaringan kultur. Kelembaban juga penting dalam persiapan dan penembakan *microcarrier*. Kelembaban yang terlalu tinggi dapat menyebabkan *microcarrier* menggumpal, sehingga mengurangi tingkat trasnformasi.

Dalam eksperimen ini, kami akan menyisipkan plasmid yang telah kami rakit, seperti yang tertera pada bab 4.1 ke dalam gen pada kloroplas *Aloe vera*. Sedangkan untuk metode transformasinya, kami akan menggunakan metode biolistik atau gene gun. Perlu kami tekankan kembali bahwa prosedur dibawah ini masih berupa rancangan kasar yang didasari oleh penelitian penelitian menyangkut transformasi *Aloe vera* dengan particle bombardment<sup>[150]</sup>.

Rancangan eksperimen kami adalah sebagai berikut:

- 1) Preparasi dan kultur sel tanaman
- 2) Persiapan DNA *coated microcarrier*
- 3) Particle bombardment
- 4) Seleksi embrio, germinasi, dan regenerasi
- 5) Ekstraksi, purifikasi, serta analisis insulin

#### **4.4.1 Kultur dan Preparasi**

Kalus *Aloe vera* didapatkan dari biji. Benih disterilkan permukaannya dalam etanol 70% (v / v) selama 30 detik diikuti dengan 10 menit dalam pengenceran pemutih untuk konsentrasi efektif 1% (v / v) natrium hipoklorit dan kemudian direndam semalam dalam larutan 0,5% (v / v) media pengawet tanaman. Embrio yang diisolasi dilapisi pada media MS yang dilengkapi dengan gelrite 0,25% (w / v), sukrosa 3% (b / v), PPM 0,01% (v / v), 300 mg / L myo-inositol, 2 mg / L NAA, 0,2 mg / L BA, 0,01 mM AgNO<sub>3</sub> dan 2 mM CaCl<sub>2</sub>. Sesuaikan dengan pH 5,4 pasca autoclave. Embrio yang diisolasi ditumbuhkan pada keadaan gelap pada suhu 25°C. Tunas dan akar yang sedang tumbuh telah dibuang. Kalus dipilih setelah 4 minggu.

Embriogenesis somatik dimulai pada media MS yang mengandung 0,2 mg / L NAA, 0,2 mg / L BA tanpa AgNO<sub>3</sub> dan terkena cahaya 14 jam, siklus gelap 10 jam yang disediakan oleh tabung fluoresen putih dingin. Tunas yang sedang berkembang dipindahkan ke media MS dengan 0,2 mg / L NAA untuk mengembangkan akar di pembuluh magenta, dikeraskan dan dipindahkan ke tanah.

#### **4.4.2 Persiapan DNA *Coated Microcarrier***

##### **Persiapan Transgen**

Transgen disiapkan dengan cara yang sama seperti yang tertera dalam bab 3.6. Materi genetik dapat dipesan dari berbagai macam perusahaan penyuplai untuk bioteknologi, atau juga dapat diisolasi secara mandiri, maupun disintesis.

##### **Persiapan Gold Particle dan *Microcarrier***

Tahapan untuk mempersiapkan DNA coated microcarrier adalah sebagai berikut<sup>[208]</sup>:

- 1) 25 miligram partikel emas ditempatkan didalam 1.5 ml tabung microfuge, yang selanjutnya ditambahkan 50 ml 0.05 M spermidine dan 50 ml DNA (1mg/ml) untuk transfeksi.
- 2) Selanjutnya, tabung tersebut diputar selama 10 detik, yang mana selama pemutaran, 50 ml 1 M CaCl ditambahkan sedikit demi sedikit.
- 3) Suspensi emas/DNA lalu diinkubasi dalam suhu ruang selama 5 menit dan diputar sebentar setiap 30 detik, lalu disentrifugasi pada 3000 rpm selama 10 detik.
- 4) Supernatan lalu dihilangkan dan pellet emas ditangguhkan kembali di dalam 200 ml polyvinylpyrrolidone (PVP;0.075 mg/ml di etanol) dari aliquot 3.5 ml, dan ditransfer ke dalam tabung polypropylene steril 5 ml. Dikarenakan proses transfer yang tidak efisien, prosedur ini diulang beberapa kali dengan aliquot solusi PVP 200 ml hingga seluruh partikel emas (4-6 aliquot). Setelahnya, solusi PVP yang tersisa dimasukkan kedalam tabung.
- 5) Solusi emas/PVP lalu diputar sebentar untuk memastikan suspensi emas yang merata. Setelahnya, solusi tadi dimasukan kedalam

tabung Tefzel 75 cm dengan menggunakan suntikan 10 ml. Tabung tersebut sebelumnya telah dikeringkan di stasiun persiapan tabung dengan nitrogen (0,3 - 0,4 LPM) selama 10 menit.

- 6) Tabung yang telah terisi lalu dimasukkan ke dalam stasiun persiapan tabung dengan suntikan yang masih tersambung, bubur yang terbentuk lalu dipindahkan perlahan lahan dari tabung menggunakan suntikan.
- 7) Suntikan lalu dilepas dari tabung, dan tabung diputar selama 30 - 40 detik.
- 8) Partikel emas/DNA di dalam tabung lalu dikeringkan dengan aliran nitrogen (0,3 - 0,4 LPM) selama 5 sampai 10 menit, dipotong menjadi panjang yang diinginkan, dan disimpan dalam kondisi kering pada suhu 4 °C hingga dibutuhkan.

#### **4.4.3 Particle Bombardment**

Setelah *microcarrier* disiapkan, tahap berikutnya adalah melakukan particle bombardment. Langkah langkah umum untuk melakukan particle bombardment adalah sebagai berikut<sup>[180]</sup>:

Time	Activity (in sequential steps)
<b><i>Week prior to bombardment</i></b>	
(-) 6 d	Sterilize supplies (Whatman and Sharkskin filter papers, funnels, flasks, water, etc.).
(-) 5 d	Prepare media needed for transformation procedure, GM+NOA suspension culture medium. 1/2 MS-HF bombardment medium with osmotica. 1/2 MS-HF medium without osmotica. 1/2 MS-HF selective medium.
(-) 4 d	Subculture or refresh medium of embryogenic cell suspensions.
<b><i>Week of bombardment</i></b>	
(-) 1 d	Set gene gun parameters (distances as described in Fig. 2). Weigh gold particles (microcarriers) and place in an oven overnight. Sterilize macrocarriers, holders and stopping screens. Assemble macrocarriers into holders.
Key d	Bombardment day (suggested day, Tuesday). Examine embryogenic cell suspension for contamination using a microscope. Prepare cells on filter paper for bombardment. Sterilize microcarriers. Coat microcarriers with DNA. Bombard cells. Incubate cells in the dark at $23 \pm 1^\circ\text{C}$ .
(+) 1 d	Transfer cells to medium without osmotica. First transfer approx 16 h after bombardment. Second transfer approx 24 h after bombardment.
(+) 2 d	Transfer cells to selective medium. Analysis of reporter gene (i.e., GUS assay) for transient expression.
(+) 3 d	Examine GUS-positive blue spots per filter paper.
<b><i>Postbombardment weeks</i></b>	
(+) 30 d	Transfer cells to fresh selective medium. Reporter gene assay for transient expression.
(+) 60 d	Check plates for development of embryos. Transfer embryos to germination medium. Transfer remaining cells to fresh selective medium. Reporter gene assay for long-term expression.
(+) 90 d	Items and procedure as in (+) 60 d. Transfer germinated embryos to plant growth medium.

(Gambar 4.4.1 Flowchart transformasi menggunakan particle bombardment secara umum)

### Transformasi *Aloe vera*

Kalus *Aloe vera* dirawat selama 4 jam dalam media osmotik (media MS dilengkapi dengan 36,4 g / L sorbitol dan 36,4 g / L manitol seperti yang

direkomendasikan oleh Frame et al.<sup>[209]</sup>, sebelum ditransformasikan dengan sistem pengiriman partikel Biolistic PD-1000 / He (BioRad, Hercules, CA) menggunakan partikel Au 1 lm, tekanan Helium 1.100 psi dan jarak target 9 cm. Partikel emas dilapisi dengan DNA plasmid dengan protokol yang telah dijelaskan sebelumnya. Vektor ekspresi dibangun dengan mengkloning promoter ubiquitin yang diamplifikasi PCR seperti yang dijelaskan sebelumnya oleh Christensen dan Quail<sup>[210]</sup> ke dalam pCAMBIA1300<sup>[150,160]</sup>.

Sel dikembalikan ke media osmotik (selanjutnya ditambah dengan 200 mg / L glutathione dan 1 mM DTT) selama 1 jam sebelum dipindahkan ke MS yang dilengkapi dengan antioksidan. Dalam waktu 2 minggu sel dipindahkan ke media MS yang dilengkapi dengan 50 mg / L Kan untuk seleksi setelah globules sekunder mulai berkembang. Sel dipindahkan ke media regenerasi dan tunas yang berkembang diisolasi dan diuji untuk ekspresi transgen<sup>[150]</sup>.

#### **4.4.4 Seleksi Embrio, Germinasi, dan Regenerasi**

Setelah transformasi selesai, tahap berikutnya adalah mengidentifikasi tanaman transgenik. Salah satu cara untuk mengidentifikasi tanaman transgenik adalah dengan menggunakan medium yang mengandung antibiotik. Antibiotik tersebut diperkirakan akan membunuh kultur sel yang tumbuh di dalam medium tersebut. Transgen yang ingin dikembangkan biasanya memiliki gen yang mengkode ketahanan atas antibiotik tadi, dalam kasus kami adalah higromisin (posisi 8498-9187 forward) dan kanamisin (posisi 6936-6980 forward).

Tanaman yang telah ditransformasi selanjutnya diaklimatisasi sebagai persiapan untuk pemindahan ke tanah. Setelah aklimatisasi selesai, tanaman selanjutnya dipindahkan ke tanah dan ditumbuhkan hingga siap panen. Insulin hasil pemanenan lalu akan dianalisa dan diuji lebih lanjut.

#### **4.4.5 Ekstraksi, purifikasi, dan Analisis**

##### **Ekstraksi dan Purifikasi**

Ekstraksi protein dapat dilakukan dari total biomassa maupun jaringan tertentu. Ekstraksi dapat menggunakan metode fisik dan/atau metode kimia.

Metode ekstraksi dapat melibatkan ultrasentrifugasi, kromatografi, dan lain lain. Penjelasan lebih lanjut mengenai ekstraksi dan purifikasi protein dapat dilihat pada bab 2.3.6

### **Analisis Insulin**

Beberapa metode yang dapat digunakan untuk menganalisa protein adalah western blot dan ELISA, yang telah kami jelaskan pada bab 2.16 dan 2.17

## Bab 5

### Kesimpulan dan Saran

#### 5.1 Kesimpulan

*Aloe vera* merupakan tumbuhan sukulen yang berpotensi digunakan sebagai bahan penelitian tanaman transgenik. Fisiologis *Aloe vera* yang memungkinkannya untuk tumbuh di berbagai macam iklim, tidak membutuhkan banyak air, dan memiliki waktu produksi yang cepat, ditambah lagi dengan terkenalnya *Aloe vera* di dunia kosmetik, menjadikannya tanaman yang ideal untuk diproduksi secara masal. *Aloe vera* juga menawarkan berbagai keuntungan, diantaranya, kemudahan dalam proses kultivasi dan ekstraksi sehingga menjadi lebih ekonomis, rendahnya risiko kontaminasi oleh substansi toxic, dan membutuhkan waktu relatif singkat dengan hasil produksi yang tinggi, sehingga meningkatkan produktivitas, terutama di daerah beriklim tropis seperti di Indonesia.

Di sisi lain, diabetes merupakan salah satu dari 10 penyebab kematian terbesar di dunia, dan merupakan penyebab kematian terbesar ketiga di Indonesia<sup>[1]</sup>. Saat ini, jumlah penderita diabetes mengalami peningkatan yang signifikan, dan dengan insulin sebagai satu satunya teknik pengobatan yang efektif, diprediksikan bahwa permintaan terhadap insulin akan mengalami peningkatan, dan hal ini dapat menyebabkan kelangkaan insulin.

Oleh karena itu kami meneliti secara *in silico* mengenai *Aloe vera* transgenik dan membuat rancangan untuk *Aloe vera* yang dapat menghasilkan insulin. Prosedur yang kami gunakan telah kami jelaskan dalam bab 3, dan telah kami bahas di bab 4. Berdasarkan hasil yang kami dapat, kami menyimpulkan bahwa *Aloe vera* memiliki kemungkinan untuk digunakan sebagai produsen insulin. Hal ini kami simpulkan dari data yang menunjukkan bahwa hasil translasi insulin *Aloe vera*, sama seperti insulin manusia, seperti yang telah kami jelaskan dalam bab 4.1. Meski begitu, hasil dari eksperimen *in silico* kami masih berupa prediksi yang dapat digunakan sebagai panduan dan perbandingan eksperimen *in vivo*, sehingga masih dibutuhkan penelitian dan eksperimen lebih lanjut di dunia nyata untuk menentukan kemampuan *Aloe vera* sebagai produsen insulin.

## **5.2 Saran**

Penelitian *in silico* berikutnya dapat mulai fokus meneliti kinerja dari insulin transgenik yang diproduksi di dalam *Aloe vera* pada tubuh manusia, sedangkan penelitian *in vivo* berikutnya dapat mulai mengembangkan *Aloe vera* transgenik yang dapat memproduksi insulin.

## Referensi

- [1] Diabetes atlas, international Diabetes Federation, 2019
- [2] Dong YM, Zhu SQ. Shanghai: Aloe Refinery Tech and Its Rules. Shanghai: Shanghai Science-Tech Press, 2002 (in Chinese).
- [3] Gu YX, Zhu SZ. Plant and Application of Aloe. Shanghai: Shanghai Science-Tech Press, 1999 (in Chinese).
- [4] Shi XJ. Discussion on aloe and its exploitation. Southwest Horticulture, 2001, 29: 14–17.
- [5] Sustiprijatno. Jagung Transgenik dan Perkembangan Penelitian di Indonesia. Balai Besar Penelitian dan Pengembangan Bioteknologi dan Sumberdaya Genetik Pertanian,  
[Bogor.\[http://balitsereal.litbang.pertanian.go.id/wp-content/uploads/2016/11/duade\\\_lapan.pdf\]\(http://balitsereal.litbang.pertanian.go.id/wp-content/uploads/2016/11/duade\_lapan.pdf\)](http://balitsereal.litbang.pertanian.go.id/wp-content/uploads/2016/11/duade_lapan.pdf). Diunduh pada 22 April 2020.
- [6] Cells and beyond Austin, 2016.  
<https://cellsandbeyond.wordpress.com/2016/04/28/using-e-coli-for-insulin-production/>.
- [7] Francisco Pouci, 2017. Aloe vera for Tissue Engineering Application. Journal of Functional Biomaterials. DOI: 10.3390/jfb8010006
- [9] Bowen RA, Austgen L, Rouge M. 2002. Biology and Activity of Restriction Endonucleases. [http://www.vivo.colostate.edu/hbooks/genetics/biotech/enzymes/r\\_enzymes.html](http://www.vivo.colostate.edu/hbooks/genetics/biotech/enzymes/r_enzymes.html) [10 Januari 2020].
- [10] Pingoud A, Jeltsch A. 2001. Structure and function of type II restriction endonucleases. Nucleic Acids Res 29(18): 3705-27
- [11] Molecular Biology of the Cells, Bruce Albert, 6<sup>th</sup> edition, 2015
- [12] add gene. <https://www.addgene.org/mol-bio-reference/promoters/>
- [13] Molecular Biology of the Cells, Bruce Albert, 6<sup>th</sup> edition, 2015
- [14] Rogers, K., ed. (2011), New Thinking about Genetics, New York: Britannica Educational Publishing, hlm. 132
- [15] Glick, B.R., Pasternak, J.J., Patten, C.L. (2010). Molecular Biotechnology: Principles and Applications of Recombinant DNA (edisi ke-4). Washington, DC: ASM Press. hlm. 117–118.
- [16] Sweet, K.M., Michaelis, R.C. (2011). The Busy Physician's Guide to Genetics, Genomics and Personalized Medicine. Dordrecht: Springer. hlm. 76.

- [17] Bieber, F.R. (2004), "Science and Technology of Forensic DNA Profiling: Current Use and Future Directions", dalam Lazer, D. (Penyunting), DNA and the Criminal Justice System: the Technology of Justice, Cambridge, MA: MIT Press, hlm. 30.
- [18] Maxam AM, Gilbert W (February 1977). "A new method for sequencing DNA". Proc. Natl. Acad. Sci. USA. 74 (2): 560–64. Bibcode:1977PNAS...74..560M. DOI:10.1073/pnas.74.2.560. PMC 392330. PMID 265521.
- [19] Sanger F; Coulson AR (May 1975). "A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase". J. Mol. Biol. 94 (3): 441–8. DOI:10.1016/0022-2836(75)90213-2. PMID 1100841.)
- [20] Westermeier. 2004. Electrophoresis in Practice: A Guide to Theory and Practice. New Jersey: John Wiley & Sons inc.  
Yuwono T. 2005. Biologi Molekuler. Jakarta: Erlangga.
- [21] U. Albrecht et al., J. Gen. Virol. 75:3353-3363, 1994; C, courtesy of Leander Lauffer and Peter Walter; D, from D. Vollrath and R.W. Davis, Nucleic Acids Res. 15:7865–7876, 1987. With permission from Oxford University Press. Molecular Biology of the cells, bruce Albert,2015)
- [22] Renart J, Reiser J, Stark GR. "Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure", Proc Natl Acad Sci U S A, 1979 Jul;76(7):3116-20. PMID: 91164 abstract\
- [23] Engvall, E (1972-11-22). "Enzyme-linked immunosorbent assay, Elisa". The Journal of Immunology. 109 (1): 129–135. ISSN 0022-1767. PMID 4113792.
- [24][researchgate.net/figure/Comparison-of-in-vivo-in-vitro-and-in-silico-methods-for-UV-A-assessment\\_tbl3\\_280623647](https://www.researchgate.net/figure/Comparison-of-in-vivo-in-vitro-and-in-silico-methods-for-UV-A-assessment_tbl3_280623647)
- [25] Cong Fen He, Jiaxing Zhang, Jie Chen. 2007. Genetic Transformation of *Aloe barbadensis* Miller by *Agrobacterium tumefaciens*. Journal of Genetics and Genomics. College of Chemistry, Beijing. Institute of crops science, Beijing.
- [26] Sahdev S, Khattar SK. Production of active eukaryotic proteins through bacterial expression system: a review of the existing biotechnology strategies. Mol Cell Biochem. 2008;307: 249-264.

- [27] Ferrer-Miralles N, Villaverde A. Bacterial cell factories for recombinant protein production; expanding the catalogue. *Microb Cell Fact*.
- [28] Gellissen G, Hollenberg CP: A comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* - a review. *Gene*. 1997, 190: 87-97. doi:10.1016/S0378-1119(97)00020-6.
- [29] Porro D, Sauer M, Branduardi P, Mattanovich D: Recombinant protein production in yeasts. *Mol Biotechnol*. 2005, 31: 245-259. doi:10.1385/MB:31:3:245.
- [30] Staub JM, Garcia B, Graves J, Hajdukiewicz PT, Hunter P, Nehra N, Paradkar V, Schlittler M, Carroll JA, Spatola L, Ward D, Ye G, Russell DA: High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nat Biotechnol*. 2000, 18: 333-338. doi:10.1038/73796.
- [31] Lower et al. "transgenic aloe plants for production of protein and related methods"
- [32] Jefferson, R. A. (1987). *Assaying chimeric genes in plants: The GUS gene fusion system*. *Plant Molecular Biology Reporter*, 5(4), 387–405. doi:10.1007/bf02667740
- [33] William Lowther • Kevin Lorick • Susan D. Lawrence • Wen-Shuz Yeow Expression of biologically active human interferon-alpha 2 in Aloe vera
- [34] Plotkin, JB, Kudla G. (2011). Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet*, 12:32–42.
- [35] Wildt, S., & Gerngross, T. U. (2005). The humanization of N-glycosylation pathways in yeast. *Nature Reviews Microbiology*, 3(2), 119–128. doi:10.1038/nrmicro1087
- [36] Baeshen, N. A., Baeshen, M. N., Sheikh, A., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A. I., ... Redwan, E. M. (2014). Cell factories for insulin production. *Microbial Cell Factories*, 13(1). doi:10.1186/s12934-014-0141-0
- [37] Boyhan D, Daniell H: Low-cost production of proinsulin in tobacco and lettuce chloroplasts for injectable or oral delivery of functional insulin and C-peptide. *Plant Biotech J* 2011, 9:585–598
- [38] PCR primer design guidelines
- [39] Campbell RK, White JR., Jr Insulin therapy in type 2 diabetes. *J Am Pharm Assoc (Wash)* 2002;42:602–11

- [40] Rolla AR, Rakel RE. Practical approaches to insulin therapy for type 2 diabetes mellitus with premixed insulin analogues. *Clin Ther*. 2005;27:1113–25.
- [41] Yaturu S. Insulin therapies: Current and future trends at dawn. *World J Diabetes* 2013;4:1-7
- [42] Benedict C, Frey WH 2nd, Schiöth HB, Schultes B, Born J, Hallschmid M. Intranasal insulin as a therapeutic option in the treatment of cognitive impairments. *Exp Gerontol* 2011;46:112-5.
- [43] Craft S, Baker LD, Montine TJ, Minoshima S, Watson GS, Claxton A, et al. Intranasal insulin therapy for Alzheimer disease and amnestic mild cognitive impairment: A pilot clinical trial. *Arch Neurol* 2012;69:29-38.
- [44] Kim H.K., Chung H.J., Park T.G. Biodegradable polymeric microspheres with “open/closed” pores for sustained release of human growth hormone. *J Control Release*. 2006;112:167–174.
- [45] Eilleia, S. Y., Soliman, M. E., Mansour, S., & S. Geneidi, A. (2018). Novel technique of insulin loading into porous carriers for oral delivery. *Asian Journal of Pharmaceutical Sciences*, 13(4), 297–309. doi:10.1016/j.ajps.2018.03.003
- [46] Dinas Pangan, Pertanian, dan Perikanan Pontianak - Lidah Buaya.
- [47] Reynolds, Tom (Ed.) (2004) *Aloes: The genus Aloe (Medicinal and Aromatic Plants - Industrial Profiles)*. CRC Press. ISBN 978-0415306720
- [48] Quattrocchi, Umberto (2012) *CRC World Dictionary of Medicinal and Poisonous Plants: Common Names, Scientific Names, Eponyms, Synonyms, and Etymology* (5 Volume Set) CRC Press.ISBN 978-1420080445
- [49] Food Drug Administration, HHS (2002). "Status of certain additional over-the-counter drug category II and III active ingredients. Final rule". *Fed Regist*. 67 (90): 31125–7.PMID 12001972
- [50] National Center for Complementary and Integrative Health, US National Institutes of Health. 1 September 2019. Diakses tanggal 24 August 2019.
- [51] Mayo Clinic. 17 September 2017. Diakses tanggal 21 Januari 2020.
- [52] Transgenic aloe plants for production of proteins and related methods
- [53] Borgoño, C. A., & Zinman, B. (2012). *Insulins: Past, Present, and Future. Endocrinology and Metabolism Clinics of North America*, 41(1), 1–24. doi:10.1016/j.ecl.2012.03.002

- [54] Hagedorn HC, Jensen BN, Krarup NB, et al. Protamine insulinate. *JAMA* 1936; 106:177–80.
- [55] Scott DA, Fisher AM. Studies on insulin with protamine. *J Pharmacol Exp Therapeut* 1936;58:78–92.
- [56] Himssworth HP. Protamine insulin and zinc protamine insulin. *Br Med J* 1937;1: 541–6.
- [57] Krayenbuhl C, Rosenberg T. Crystalline protamine insulin. *Rep Steno Mem Hosp Nord Insulin Lab* 1946;1:60–73.
- [58] Hallas-Moller K. The lente insulins. *Diabetes* 1956;5:7–14.
- [59] Oakley W, Hill D, Oakley N. Combined use of regular and crystalline protamine (NPH) insulins in the treatment of severe diabetes. *Diabetes* 1966;15:219–22.
- [60] Sanger F. Chemistry of insulin; determination of the structure of insulin opens the way to greater understanding of life processes. *Science* 1959;129:1340–4.
- [61] Markussen J, Damgaard U, Jorgensen KH, et al. Human monoclonal insulin. Chemistry and characteristics. *Acta Med Scand Suppl* 1983;671:99–105.
- [62] Keen H, Glynne A, Pickup JC, et al. Human insulin produced by recombinant DNA technology: safety and hypoglycaemic potency in healthy men. *Lancet* 1980;2:398–401.
- [63] Velcheva, M., Faltin, Z., Vardi, A., Hanania, U., Eshdat, Y., Dgani, O., ... Perl, A. (2010). Aloe vera transformation: the role of Amberlite XAD-4 resin and antioxidants during selection and regeneration. *In Vitro Cellular & Developmental Biology - Plant*, 46(6), 477–484. doi:10.1007/s11627-010-z
- [66] ^ Norris, Sonya (4 July 2005). "Molecular farming". Library of Parliament. Parliament of Canada. PRB 05-09E. Diakses tanggal 2008-09-11.
- [67] ^ Kaiser, Jocelyn (25 April 2008). "Is the Drought Over for Pharming?" (PDF). *Science*. 320 (5875): 473–5. doi:10.1126/science.320.5875.473. PMID 18436771.
- [68]. Vieira Gomes, A.M.; Souza Carmo, T.; Silva Carvalho, L.; Mendonca Bahia, F.; Parachin, N.S. Comparison of yeasts as hosts for recombinant protein production. *Microorganisms* 2018, 6, 38. [CrossRef]

- [71]. Shanmugaraj, B.; Ramalingam, S. Plant expression platform [74]. Sahdev, S.; Khattar, S.K.; Saini, K.S. Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies. *Mol. Cell. Biochem.* 2008, 307, 249–264. [CrossRef]
- [75]. Strasser, R.; Stadlmann, J.; Schahs, M.; Stiegler, G.; Quendler, H.; Mach, L.; Glossl, J.; Weterings, K.; Pabst, M.; Steinkellner, H. Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol.*
- [77]. Montero-Morales, L.; Steinkellner, H. Advanced plant-based glycan engineering. *Front. Bioeng. Biotechnol.* 2018, 6, 81. [CrossRef] [PubMed]
- [79]. Menary, J.; Hobbs, M.; Mesquita de Albuquerque, S.; Pacho, A.; Drake, P.M.W.; Prendiville, A.; Ma, J.K.; Fuller, S.S. Shotguns vs Lasers: Identifying barriers and facilitators to scaling-up plant molecular farming for high-value health products. *PLoS ONE* 2020, 15, e0229952. [CrossRef]
- [80]. Floss, D.M.; Sack, M.; Stadlmann, J.; Rademacher, T.; Scheller, J.; Stöger, E.; Fischer, R.; Conrad, U. Biochemical and functional characterization of anti-HIV antibody–ELP fusion proteins from transgenic plants. *Plant Biotechnol. J.* 2008, 6, 379–391. [CrossRef]
- [81]. Silva H, Sagardia S, Seguel O, Torres C, Tapia C, Franck N, Cardemil L: Effect of water availability on growth and water use efficiency for biomass and gel production in Aloe Vera (*Aloe barbadensis* M.). *Industrial Crops and Products* 2010, 31(1):20-27.
- [82]. Grace OM, Buerki S, Symonds MR, Forest F, van Wyk AE, Smith GF, Klopper RR, Bjorå CS, Neale S, Demissew S: Evolutionary history and leaf succulence as explanations for medicinal use in aloes and the global popularity of Aloe vera. *BMC evolutionary biology* 2015, 15(1):29.
- [83]. Reynolds T, Dweck A: Aloe vera leaf gel: a review update. *Journal of ethnopharmacology* 1999, 68(1-3):3-37.
- [84]. Gupta VK, Malhotra S: Pharmacological attribute of Aloe vera: Revalidation through experimental and clinical studies. *Ayu* 2012, 33(2):193.
- [85]. Raksha B, Pooja S, Babu S: Bioactive compounds and medicinal properties of Aloe vera L.: An update. *Journal of Plant Sciences* 2014, 2(3):102-107.
- [86]. Hamman JH: Composition and applications of Aloe vera leaf gel. *Molecules* 2008, 13(8):1599-1616.

- [87]. Joseph B, Raj SJ: Pharmacognostic and phytochemical properties of Aloe vera linn an overview. International journal of pharmaceutical sciences review and research 2010, 4(2):106-110.
- [88]. Choudhri P, Rani M, Sangwan RS, Kumar R, Kumar A, Chhokar V: De novo sequencing, assembly and characterisation of Aloe vera transcriptome and analysis of expression profiles of genes related to saponin and anthraquinone metabolism. BMC genomics 2018, 19(1):427.
- [89]. Zonneveld BJ: Genome size analysis of selected species of Aloe (Aloaceae) reveals the most primitive species and results in some new combinations. Bradleya 2002, 2002(20):5-12.
- [90]. Kadirvelu Sharon1, Sharma Suvarna. 2017. Cloning of HVA22 Homolog from Aloe vera and Preliminary Study of Transgenic Plant Development. ISSN: 2320 – 7051. DOI: <http://dx.doi.org/10.18782/2320-7051.6033>.
- [91]. Jia-Ju Rena, Jian Wang et al. 2020. The complete chloroplast genome of Aloe vera from China as a Chinese herb. aSchool of Nursing, Beijing University of Chinese Medicine, Beijing, China. Taylor & Francis. VOL. 5, NO. 1, 1092–1093 <https://doi.org/10.1080/23802359.2020.1726229>
- [91]. Shubham, Abhisek et al. 2020. The genome sequence of Aloe vera reveals adaptive evolution of drought tolerance  
2 mechanisms. bioRxiv preprint doi: <https://doi.org/10.1101/2020.05.29.122895>.
- [92]. Jia-Ju Ren, Jian wang. 2020. The complete chloroplast genome of Aloe vera from China as a Chinese herb. © 2020 The Author(s). Published by InformaUK Limited, trading as Taylor & FrancisGroup. DOI : VOL. 5, NO. 1, <https://doi.org/10.1080/23802359.2020.1726229>
- [93]. Gellissen G, Hollenberg CP, Application of yeasts in gene expression studies A comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* - a review. Gene. 1997;190:87–97. doi: 10.1016/S0378-1119(97)00020-6. [PubMed] [CrossRef] [Google Scholar] [Ref list]
- [94]. 24. Porro D, Sauer M, Branduardi P, Mattanovich D: Recombinant protein production in yeasts. Mol Biotechnol 2005, 31:245–259.

- [95]. Kjeldsen T, Pettersson AF, Hach M: Secretory expression and characterization of insulin in *Pichia pastoris*. *Biotechnol Appl Biochem* 1999, 29:79–86.
- [96]. Mansur M, Cabello C, Hernandez L, Pais J, Varas L, Valdes J, Terrero Y, Hidalgo A, Plana L, Besada V, Garcia L, Lamazares E, Castellanos L, Martinez E: Multiple gene copy number enhances insulin precursor secretion in the yeast *Pichia pastoris*. *Biotechnol Lett* 2005, 27:339–345.
- [97]. "nopaline synthase terminator: Topics by Science.gov". [www.science.gov](http://www.science.gov). Retrieved 2019-11-20.
- [98] Carrer, Helaine; Hockenberry, Tish Noel; Svab, Zora; Maliga, Pal (October 1993). "Kanamycin resistance as a selectable marker for plastid transformation in tobacco". *MGG Molecular & General Genetics*. 241-241 (1-2): 49–56.
- [99] Vande Broek, A; Lambrecht, M; Vanderleyden, J (1998). "Bacterial chemotactic motility is important for the initiation of wheat root colonization by *Azospirillum brasilense*". *Microbiology*. 144 (9): 2599–606. doi:10.1099/00221287-144-9-2599. PMID 9782509.
- [100] Sundberg, C., Meek, L., Carroll, K., Das, A., & Ream, W. (1996). VirE1 protein mediates export of the single-stranded DNA-binding protein VirE2 from *Agrobacterium tumefaciens* into plant cells. *Journal of Bacteriology*, 178(4), 1207–1212. doi:10.1128/jb.178.4.1207-1212.1996
- [101] Sundberg CD, Ream W (November 1999). "The *Agrobacterium tumefaciens* chaperone-like protein, VirE1, interacts with VirE2 at domains required for single-stranded DNA binding and cooperative interaction". *Journal of Bacteriology*. 181 (21): 6850–5. doi:10.1128/JB.181.21.6850-6855.1999
- [106] Damerval C, Vienne DD, Zivy M, Thiellement H. Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. *Electrophoresis* 1986; 7: 52–4
- [107] Shaw, M. M., & Riederer, B. M. (2003). *Sample preparation for two-dimensional gel electrophoresis*. *PROTEOMICS*, 3(8), 1408–1417. doi:10.1002/pmic.200300471

- [108] Novák, P., & Havlíček, V. (2016). *Protein Extraction and Precipitation. Proteomic Profiling and Analytical Chemistry*, 51–62. doi:10.1016/b978-0-444-63688-1.00004-5
- [109] Koontz, L. (2014). *TCA Precipitation. Laboratory Methods in Enzymology: Protein Part C*, 3–10. doi:10.1016/b978-0-12-420119-4.00001-x
- [110] He, C.-F., & Wang, Y.-M. (2008). *Protein Extraction From Leaves of Aloe vera L., A Succulent and Recalcitrant Plant, for Proteomic Analysis. Plant Molecular Biology Reporter*, 26(4), 292–300. doi:10.1007/s11105-008-0040-9
- [111] Hurkman, W. J., & Tanaka, C. K. (1986). *Solubilization of Plant Membrane Proteins for Analysis by Two-Dimensional Gel Electrophoresis. PLANT PHYSIOLOGY*, 81(3), 802–806. doi:10.1104/pp.81.3.802
- [112] Faurobert, M., Pelpoir, E., & Chaïb, J. (n.d.). Phenol Extraction of Proteins for Proteomic Studies of Recalcitrant Plant Tissues. *Plant Proteomics*, 9–14. doi:10.1385/1-59745-227-0:9
- [113] Carpentier, S. C., Witters, E., Laukens, K., Deckers, P., Swennen, R., & Panis, B. (2005). *Preparation of protein extracts from recalcitrant plant tissues: An evaluation of different methods for two-dimensional gel electrophoresis analysis. PROTEOMICS*, 5(10), 2497–2507. doi:10.1002/pmic.200401222
- [114] Saravanan, R. S., & Rose, J. K. C. (2004). *A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. PROTEOMICS*, 4(9), 2522–2532. doi:10.1002/pmic.200300789
- [115] [https://www.researchgate.net/post/Protein\\_denaturation\\_vs\\_degradation](https://www.researchgate.net/post/Protein_denaturation_vs_degradation)
- [116] Anderson, N. L., & Anderson, N. G. (1998). *Proteome and proteomics: New technologies, new concepts, and new words. Electrophoresis*, 19(11), 1853–1861. doi:10.1002/elps.1150191103
- [117] Blackstock, W. P., & Weir, M. P. (1999). *Proteomics: quantitative and physical mapping of cellular proteins. Trends in Biotechnology*, 17(3), 121–127. doi:10.1016/s0167-7799(98)01245-1
- [118] Komari, T., Takakura, Y., Ueki, J., Kato, N., Ishida, Y., & Hiei, Y. (n.d.). Binary Vectors and Super-binary Vectors. *Agrobacterium Protocols*, 15–42. doi:10.1385/1-59745-130-4:15

- [119] Lee, L.-Y., & Gelvin, S. B. (2007). *T-DNA Binary Vectors and Systems*. *PLANT PHYSIOLOGY*, 146(2), 325–332. doi:10.1104/pp.107.113001
- [120] Hoekema, A., Hirsch, P. R., Hooykaas, P. J. J., & Schilperoort, R. A. (1983). *A binary plant vector strategy based on separation of vir- and T-region of the Agrobacterium tumefaciens Ti-plasmid*. *Nature*, 303(5913), 179–180. doi:10.1038/303179a0
- [121] Fraley, R. T., Rogers, S. G., Horsch, R. B., & Gelvin, S. B. (1986). *Genetic transformation in higher plants*. *Critical Reviews in Plant Sciences*, 4(1), 1–46. doi:10.1080/07352688609382217
- [122] Christensen, A. H., Sharrock, R. A., & Quail, P. H. (1992). *Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation*. *Plant Molecular Biology*, 18(4), 675–689. doi:10.1007/bf00020010
- [123] Zhang, W. (1991). *Analysis of Rice Act1 5[prime] Region Activity in Transgenic Rice Plants*. *THE PLANT CELL ONLINE*, 3(11), 1155–1165. doi:10.1105/tpc.3.11.1155
- [124] GARFINKEL, D. (1981). *Genetic analysis of crown gall: Fine structure map of the T-DNA by site-directed mutagenesis*. *Cell*, 27(1), 143–153. doi:10.1016/0092-8674(81)90368-8
- [125] Zambryski P, Joos H, Genetello C, Leemans J, Montagu MV, Schell J (1983). Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. *EMBO J*. 1983; 2(12):2143-50.
- [126] Fraley, R. T., Rogers, S. G., Horsch, R. B., Eichholtz, D. A., Flick, J. S., Fink, C. L., ... Sanders, P. R. (1985). *The SEV System: A New Disarmed Ti Plasmid Vector System for Plant Transformation*. *Nature Biotechnology*, 3(7), 629–635. doi:10.1038/nbt0785-629
- [127] De Framond, A. J., Barton, K. A., & Chilton, M.-D. (1983). *Mini-Ti: A New Vector Strategy for Plant Genetic Engineering*. *Nature Biotechnology*, 1(3), 262–269. doi:10.1038/nbt0583-262
- [128] Pansegrouw, W., Lanka, E., Barth, P. T., Figurski, D. H., Guiney, D. G., Haas, D., ... Thomas, C. M. (1994). *Complete Nucleotide Sequence of Birmingham*

- IncPa Plasmids. Journal of Molecular Biology*, 239(5), 623–663.  
doi:10.1006/jmbi.1994.1404
- [129] Okumura, M. S., & Kado, C. I. (1992). *The region essential for efficient autonomous replication of pSa in Escherichia coli*. *MGG Molecular & General Genetics*, 235(1), 55–63. doi:10.1007/bf00286181
- [130] Hellens, R. P., Edwards, E. A., Leyland, N. R., Bean, S., & Mullineaux, P. M. (2000). *Plant Molecular Biology*, 42(6), 819–832.  
doi:10.1023/a:1006496308160
- [131] Jouanin, L., Vilaine, F., d' Enfert, C., & Casse-Delbart, F. (1985). *Localization and restriction maps of the replication origin regions of the plasmids of Agrobacterium rhizogenes strain A4*. *MGG Molecular & General Genetics*, 201(3), 370–374. doi:10.1007/bf00331325
- [132] Deblaere, R., Reynaerts, A., Höfte, H., Hernalsteens, J.-P., Leemans, J., & Van Montagu, M. (1987). [16] *Vectors for cloning in plant cells. Recombinant DNA Part D*, 277–292. doi:10.1016/0076-6879(87)53059-2
- [133] Sambrook, J. and Russell, D. W. (2001) *Molecular Cloning, A Laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.  
ISBN: 978-087969577-4
- [134] Lemos, M. L., & Crosa, J. H. (1992). *Highly preferred site of insertion of Tn7 into the chromosome of Vibrio anguillarum*. *Plasmid*, 27(2), 161–163.  
doi:10.1016/0147-619x(92)90016-4
- [135] Pang, S. Z., DeBoer, D. L., Wan, Y., Ye, G., Layton, J. G., Neher, M. K., ... Fromm, M. E. (1996). *An Improved Green Fluorescent Protein Gene as a Vital Marker in Plants*. *Plant Physiology*, 112(3), 893–900. doi:10.1104/pp.112.3.893
- [136] OW, D. W., DE WET, J. R., HELINSKI, D. R., HOWELL, S. H., WOOD, K. V., & DELUCA, M. (1986). *Transient and Stable Expression of the Firefly Luciferase Gene in Plant Cells and Transgenic Plants*. *Science*, 234(4778), 856–859. doi:10.1126/science.234.4778.856
- [137] Bevan, M. (1984). *Binary Agrobacterium vectors for plant transformation*. *Nucleic Acids Research*, 12(22), 8711–8721. doi:10.1093/nar/12.22.8711

- [138] Chung, S.-M., Frankman, E. L., & Tzfira, T. (2005). *A versatile vector system for multiple gene expression in plants*. *Trends in Plant Science*, 10(8), 357–361. doi:10.1016/j.tplants.2005.06.001
- [139] Tzfira, T., Tian, G.-W., Lacroix, B., Vyas, S., Li, J., Leitner-Dagan, Y., ... Citovsky, V. (2005). *pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants*. *Plant Molecular Biology*, 57(4), 503–516. doi:10.1007/s11103-005-0340-5
- [140] Goderis IJWM, De Bolle MFC, Francois IEJA, Wouters PFJ, Broekaert WF, Cammue BPA (2002) A set of modular plant transformation vectors allowing flexible insertion of up to six expression units. *Plant Mol Biol* 50 17–27 doi:10.1023/a:1016052416053
- [141] Citovsky, V., Lee, L.-Y., Vyas, S., Glick, E., Chen, M.-H., Vainstein, A., ... Tzfira, T. (2006). *Subcellular Localization of Interacting Proteins by Bimolecular Fluorescence Complementation in Planta*. *Journal of Molecular Biology*, 362(5), 1120–1131. doi:10.1016/j.jmb.2006.08.017
- [142] Hamilton, C. M. (1997). *A binary-BAC system for plant transformation with high-molecular-weight DNA*. *Gene*, 200(1-2), 107–116. doi:10.1016/s0378-1119(97)00388-0
- [143] Liu, Y.-G., Shirano, Y., Fukaki, H., Yanai, Y., Tasaka, M., Tabata, S., & Shibata, D. (1999). *Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning*. *Proceedings of the National Academy of Sciences*, 96(11), 6535–6540. doi:10.1073/pnas.96.11.6535
- [144] Liu, Y.-G., Nagaki, K., Fujita, M., Kawaura, K., Uozumi, M., & Ogihara, Y. (2000). *Development of an efficient maintenance and screening system for large-insert genomic DNA libraries of hexaploid wheat in a transformation-competent artificial chromosome (TAC) vector*. *The Plant Journal*, 23(5), 687–695. doi:10.1046/j.1365-313x.2000.00827.x
- [145] Freeman WM, Walker SJ, Vrana KE (January 1999). "Quantitative RT-PCR: pitfalls and potential". *BioTechniques*. 26 (1): 112–22, 124–5. doi:10.2144/99261rv01. PMID 9894600.

- [146] Schmittgen TD, Zakrajsek BA, Mills AG, Gorn V, Singer MJ, Reed MW (October 2000). "Quantitative reverse transcription-polymerase chain reaction to study mRNA decay: comparison of endpoint and real-time methods". *Anal. Biochem.* 285 (2): 194–204. doi:10.1006/abio.2000.4753. PMID 11017702. S2CID 258810.
- [147] Deepak S, Kottapalli K, Rakwal R, et al. (June 2007). "Real-Time PCR: Revolutionizing Detection and Expression Analysis of Genes". *Curr. Genomics*. 8 (4): 234–51. doi:10.2174/138920207781386960. PMC 2430684. PMID 18645596.
- [148] Bustin SA (August 2002). "Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems". *J. Mol. Endocrinol.* 29 (1): 23–39. doi:10.1677/jme.0.0290023. PMID 12200227.
- [149] Keiko, Takashi. 2003. A Role of the –35 Element in the Initiation of Transcription at psbA Promoter in Tobacco Plastids
- [150] Lowther, Lorick. 2012. Expression of biologically active human interferon alpha 2 in Aloe vera
- [151] T. Cruz, T. Kanashiro, A. de Castro, C. Baldin, L. Richtzenhain and J. Araujo Jr., Pesqui. *Vet. Bras.*, 2016, 36, 1171–1177.
- [152] T. Csont, Hungary-Romania cross Border co-operation programme, 2013, vol. 2007–2013, pp. 3–4.
- [153] I. A. Darwish, *Int. J. Biomed. Sci.*, 2006, 2, 217–235.
- [154] S. E. Manley, I. M. Stratton, P. M. Clark and S. D. Luzio, *Clin. Chem.*, 2007, 53, 922–932.
- [155] M. S. Even, C. B. Sandusky, N. D. Barnard, J. Mistry and M. K. Sinha, *Clin. Biochem.*, 2007, 40, 98–103.
- [156] R. Abellán, R. Ventura, I. Palma, S. di Carlo, R. di Giovannandrea, M. Bellver, R. Olive, J. A. Pascual, R. Pacifici, J. Segura, P. Zuccaro and S. Pichini, *J. Pharm. Biomed. Anal.*, 2009, 49, 793–799.
- [157] Z. J. Farino, T. J. Morgenstern, J. Vallaghe, N. Gregor.
- [158] Ushimaru, T., Terada, H., Tsuboi, K., Kogou, Y., Sakaguchi, A., Tsuji, G., & Kubo, Y. (2010). *Development of an efficient gene targeting system in*

*Colletotrichum higginsianum* using a non-homologous end-joining mutant and *Agrobacterium tumefaciens*-mediated gene transfer. *Molecular Genetics and Genomics*, 284(5), 357–371. doi:10.1007/s00438-010-0572-1

[159] T. Csont, Hungary-Romania cross Border co-operation programme, 2013, vol. 2007–2013, pp. 3–4.

[160] Zahra, Ravi. 2016. A novel aseptic technique for micropropagation of Aloe vera mill. [herbmed.skums.ac.ir](http://herbmed.skums.ac.ir)

[161] Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, 77(1), 51–59. doi:10.1016/0378-1119(89)90358-2

[162] MICHELSE, C., ARENTSHORST, M., RAM, A., & VANDENHONDEL, C. (2005). -mediated transformation leads to improved gene replacement efficiency in. *Fungal Genetics and Biology*, 42(1), 9–19. doi:10.1016/j.fgb.2004.06.009

[163] Idnurm, A., Giles, S. S., Perfect, J. R., & Heitman, J. (2006). Peroxisome Function Regulates Growth on Glucose in the Basidiomycete Fungus *Cryptococcus neoformans*. *Eukaryotic Cell*, 6(1), 60–72. doi:10.1128/ec.00214-06

[164] Li, L., Shao, Y., Li, Q., Yang, S., & Chen, F. (2010). Identification of *Mga1*, a G-protein α-subunit gene involved in regulating citrinin and pigment production in *Monascus ruber* M7. *FEMS Microbiology Letters*, no–no. doi:10.1111/j.1574-6968.2010.01992.x

[165] Roberts, R. J., & Murray, K. (1976). *Restriction Endonuclease*. CRC Critical Reviews in Biochemistry, 4(2), 123–164. doi:10.3109/10409237609105456

[166] Kessler, C., & Manta, V. (1990). Specificity of restriction endonucleases and DNA modification methyltransferases — a review (edition 3). *Gene*, 92(1-2), 1–240. doi:10.1016/0378-1119(90)90486-b

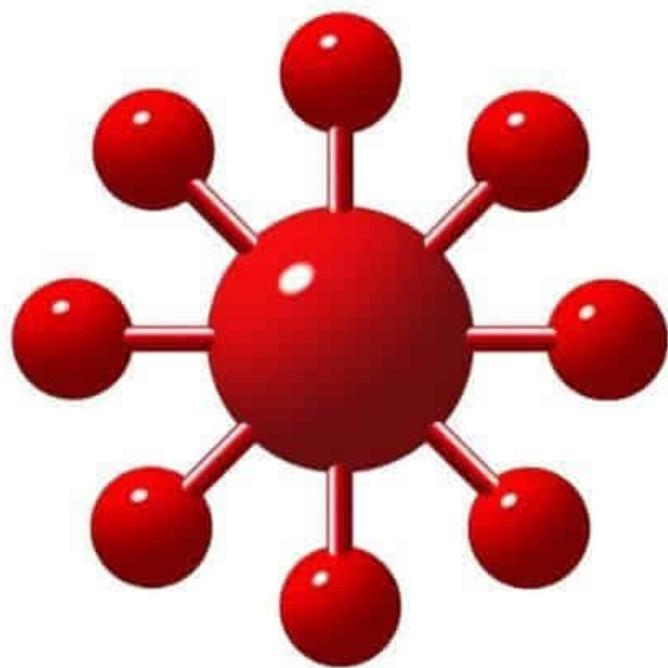
[167] Pingoud A, Alves J, Geiger R (1993). "Chapter 8: Restriction Enzymes". In Burrell M (ed.). *Enzymes of Molecular Biology*. Methods of Molecular Biology. 16. Totowa, NJ: Humana Press. pp. 107–200. ISBN 0-89603-234-5.

- [168] Luria SE, Human ML (October 1952). "A nonhereditary, host-induced variation of bacterial viruses". *Journal of Bacteriology*. **64** (4): 557–69. doi:10.1128/JB.64.4.557-569.1952. PMC 169391. PMID 12999684
- [169] Bertani G, Weigle JJ (February 1953). "Host controlled variation in bacterial viruses". *Journal of Bacteriology*. **65** (2): 113–21. doi:10.1128/JB.65.2.113-121.1953. PMC 169650. PMID 13034700.
- [176] Clark DP (2005). *Molecular biology*. Amsterdam: Elsevier Academic Press. ISBN 0-12-175551-7.
- [177] O'Brien, John A.; Lummis, Sarah CR (2011). "Nano-biolistics: A method of biolistic transfection of cells and tissues using a gene gun with novel nanometer-sized projectiles". *BMC Biotechnology*. **11**: 66. doi:10.1186/1472-6750-11-66. PMC 3144454. PMID 21663596.
- [178] Sanford, John C. (1990). "Biolistic plant transformation". *Physiologia Plantarum*. **79** (1): 206–209. doi:10.1111/j.1399-3054.1990.tb05888.x. ISSN 1399-3054.
- [179] Slater, Adrian; Scott, Nigel; Fowler, Mark (2008). *Plant Biotechnology: the genetic manipulation of plants* (2 ed.). Oxford, New York, USA: Oxford University Press Inc. ISBN 978-0-19-928261-6.
- [180] Kikkert, Julie; Vidal, Jose; Reisch, Bruce (2005). Stable transformation of plant cells by particle bombardment/bioistics. *Methods in Molecular Biology*. 286. pp. 61–78. doi:10.1385/1-59259-827-7:061. ISBN 978-1-59259-827-4. PMID 15310913. S2CID 44395352.
- [181] Gan C. (1989) Gene Gun Accelerates DNA-Coated Particles To Transform Intact Cells. *The Scientist* 3(18):25.
- [182] Baldy MW, Strom B, Bernstein H (March 1971). "Repair of alkylated bacteriophage T4 deoxyribonucleic acid by a mechanism involving polynucleotide ligase". *Journal of Virology*. **7** (3): 407–8. doi:10.1128/JVI.7.3.407-408.1971. PMC 356131. PMID 4927528.
- [183] Baldy, M. W. (1970). *The UV sensitivity of some early-function temperature-sensitive mutants of phage T4*. *Virology*, **40**(2), 272–287. doi:10.1016/0042-6822(70)90403-4

- [184] Baldy MW (1968). "Repair and recombination in phage T4. II. Genes affecting UV sensitivity". *Cold Spring Harbor Symposia on Quantitative Biology*. 33: 333–8.doi:10.1101/sqb.1968.033.01.038. PMID 4891973.
- [185] Wilson, R. H., Morton, S. K., Deiderick, H., Gerth, M. L., Paul, H. A., Gerber, I., ... Patrick, W. M. (2013). *Engineered DNA ligases with improved activities in vitro*. *Protein Engineering Design and Selection*, 26(7), 471–478. doi:10.1093/protein/gzt024
- [186] Pascal, J. M., O'Brien, P. J., Tomkinson, A. E., & Ellenberger, T. (2004). *Human DNA ligase I completely encircles and partially unwinds nicked DNA*. *Nature*, 432(7016), 473–478. doi:10.1038/nature03082
- [187] Lehman, I. R. (1974). *DNA Ligase: Structure, Mechanism, and Function*. *Science*, 186(4166), 790–797. doi:10.1126/science.186.4166.790
- [188] Foster JB, Slonczewski J (2010). *Microbiology: An Evolving Science* (Second ed.). New York: W. W. Norton & Company. ISBN 978-0-393-93447-2.
- [189] Yang, Y., & LiCata, V. J. (2018). *Pol I DNA polymerases stimulate DNA end-joining by Escherichia coli DNA ligase*. *Biochemical and Biophysical Research Communications*, 497(1), 13–18. doi:10.1016/j.bbrc.2018.01.165
- [190] "Ligases" (PDF). *Enzyme Resources Guide*. Promega Corporation. pp. 8–14.
- [191] m-edukasi. Kemdikbud. Pembentukan hormon-hormon Endokrin Pankreas
- [192] Richard. Insulin Synthesis and Secretion. VIVO Pathophysiology.
- [193] Régnier, Philippe; Marujo, Paulo E. (2013). Polyadenylation and Degradation of RNA in Prokaryotes
- [194] "Polyadenylation - an overview | ScienceDirect Topics". www.sciencedirect.com.
- [195] Rehfeld, Anders; Plass, Mireya; Krogh, Anders; Friis-Hansen, Lennart (2013). "Alterations in Polyadenylation and Its Implications for Endocrine Disease". *Frontiers in Endocrinology*
- [196] Beaud, D. (2005). Genetic characterization of the -glucuronidase enzyme from a human intestinal bacterium, *Ruminococcus gnavus*. *Microbiology*, 151(7), 2323–2330. doi:10.1099/mic.0.27712-0

- [197] Plotkin JB, Kudla G. (2011) Synonymous but not the same: the causes and consequences of codon bias.
- [198] Cannarozzi G, Schraudolph NN, et al. (2010) A role for codon order in translation dynamics.
- [199] Wang Y, Liu CL, et al. (2002) Precision and functional specificity in mRNA decay.
- [200] Zhang G, Ignatova Z. (2011) Folding at the birth of the nascent chain: coordinating translation with co-translational folding.
- [201] The EMBL-EBI search and sequence analysis tools APIs in 2019 Fábio Madeira, Young mi Park, Joon Lee, Nicola Buso, Tamer Gur, Nandana Madhusoodanan, Prasad Basutkar, Adrian R N Tivey, Simon C Potter, Robert D Finn, and Rodrigo Lopez
- [202] Haas J., Gumienny R., Barbato A., Ackermann F., Tauriello G., Bertoni M., Studer G., Smolinski A., Schwede T. Introducing "best single template" models as reference baseline for the Continuous Automated Model Evaluation (CAMEO). *Proteins*. (2019), 87, 1378-1387. [DOI: 10.1002/prot.25815]
- [203] Haas J., Barbato A., Behringer D., Studer G., Roth S., Bertoni M., Mostaguir K., Gumienny R., Schwede T. Continuous Automated Model EvaluatiOn (CAMEO) complementing the critical assessment of structure prediction in CASP12. *Proteins*. (2018), 86, 387-398. [DOI: 10.1002/prot.25431]
- [204] Haas J., Roth S., Arnold K., Kiefer F., Schmidt T., Bordoli L. and Schwede T.(2013). The Protein Model Portal - a comprehensive resource for protein structure and model information. Database [PMID: 23624946].
- [205] DeLano, W. L. (2002). PyMOL. DeLano Scientific, San Carlos, CA, 700.
- [206] DeLano, W. L. (2009). The PyMOL Molecular Graphics System; DeLano Scientific: San Carlos, CA, 2002.
- [207] DeLano, W. L. (2002). Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography*, 40, 82-92.
- [208] O'Brien, J., & Lummis, S. C. R. (2002). An improved method of preparing microcarriers for biolistic transfection. *Brain Research Protocols*, 10(1), 12–15. doi:10.1016/s1385-299x(02)00175-7

- [209] Frame, B. R., Zhang, H., Coccilone, S. M., Sidorenko, L. V., Dietrich, C. R., Pegg, S. E., ... Wang, K. (2000). Production of transgenic maize from bombarded type II callus: Effect of gold particle size and callus morphology on transformation efficiency. In *Vitro Cellular & Developmental Biology - Plant*, 36(1), 21–29. doi:10.1007/s11627-000-0007-5
- [210] Christensen, A. H., & Quail, P. H. (1996). Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Research*, 5(3), 213–218. doi:10.1007/bf01969712
- [211] Lengauer T, Rarey M (Jun 1996). "Computational methods for biomolecular docking". *Current Opinion in Structural Biology*. 6(3): 402–6. doi:10.1016/S0959-440X(96)80061-3. PMID 8804827.



## **ISPO RESEARCH PROPOSAL**

**C-ATP**

**(Computational Analysis of *Aloe vera* for Transgenic Insulin Production)**

**Written By:**

**Aisyah Rifa Fadhilah**

**Muhamad Nabil Alhanif**

**Research Field:**

**BIOLOGI (BIOTEKNOLOGI)**

**School :**

**SMA Kharisma Bangsa**

**City, Province :**

**Tangerang Selatan, Banten**

## **Validity Sheet**

**Title : C-ATP (Computational Analysis of *Aloe vera* for Transgenic Insulin Production)**

**Writer : -Aisyah Rifa Fadhilah  
-Muhamad Nabil Alhanif**

**Knowing,**

**Kharisma Bangsa Head Master,**

**Imam Husnan Nugroho, S.T,M.Pd.**

**Supervisor,**

**Dr. Satya Nugroho.  
Dessy Norma Juita, M.Pd.**

## **Foreword**

Praise and gratitude to Allah SWT for His blessings and mercy we were able to complete the research design entitled C-ATP (Computational Analysis of Aloe vera for Transgenic Proinsulin Production).

This research design proposal was made in the context of participating in the 2021 ISPO (Indonesia Science Project Olympiad) competition.

We compiled this proposal optimally, and received assistance from various parties to facilitate the production of this paper. For that, we would like to express our gratitude to all parties who have contributed in writing this proposal.

Our thanks are given to:

- Mr. Dr. Satya Nugroho (Supervisor)
- Mr. Aris Setyawan, S.Si (Kharisma Bangsa Science Team coordinator)
- Mrs. Dessy Norma Juita, M.Pd (Supervisor) for their support, guidance, and assistance.
- Parents, teachers and friends for their support, prayers, encouragement, and assistance so far so that we can complete this proposal.

Finally, we realize that this proposal and research is still very far from perfect, therefore, we expect constructive criticism and suggestions so that this research can be completed, realized, and can benefit society and the environment.

## **Table of Content**

<b>Foreword</b>	<b>3</b>
<b>Table of Content</b>	<b>4</b>
<b>Abstract</b>	<b>9</b>
<b>Chapter 1</b>	<b>10</b>
<b>Preliminary</b>	<b>10</b>
1.1 Background	10
1.2 Formulation of Problem	13
1.3 Problem Limitation	13
1.4 Research Purposes	13
1.5 Novelty	14
1.6 Hypothesis	14
1.7 Research Benefit	14
<b>Chapter II</b>	<b>16</b>
<b>Literature Review</b>	<b>16</b>
2.1 Research Method	16
2.2 Diabetes	17
2.2.1 Type of Diabetes	17
1. Type 1 Diabetes	17
2. Type 2 Diabetes	17
3. Gestational diabetes (GDM)	18
2.2.2 Number of Diabetes Patients	18
2.3 Insulin	20
2.3.1 Insulin Synthesis	22
2.3.2 Insulin Secretion Pathway	23
2.3.3 Conventional Insulin Production	24
2.3.4 Recombinant Insulin Production	25
Escherichia coli	25
Yeast	27
Transgenic Plants	28
2.3.5 Perbandingan Persentase Produksi Insulin Rekombinan	29
2.3.6 Protein Extraction	29
Trichloroacetic acid (TCA)-acetone Precipitation Method	31
Phenol Extraction with methanol/ammonium acetate precipitation Method	32
2.4 Biopharming	35
2.4.1 Comparison of Expression Systems in Plants and E. coli	36
E. coli:	36
Advantages:	37

Disadvantages:	37
In Transgenic Plants:	37
2.5 Expression Platform	38
Plants as a biofactory for heterologous protein expression	40
2.6. Expression Types	40
2.6.1 Chloroplast Expression	41
Advantage	42
Chloroplast Genome Organization: Concept of Chloroplast Transformation	42
Regulation of Chloroplast Gene Expression	44
2.7 Transformation Method	44
2.7.1 Indirect Transformation Method	45
Agrobacterium tumefaciens-mediated transformation	45
2.7.2 Direct Transformation Method	45
Particle Bombardment-mediated transformation	45
Application	46
Advantage	47
2.8 Biolistic Construction Design	48
2.8.1 Promotor	48
2.8.2 Polyadenylation	50
2.8.3 Terminator	51
2.8.4 Marker gene	51
Selectable Marker	51
Hygromycin phosphotransferase (HPT) (hpt)	51
2.9 Choosing a Model Plant	52
2.9.1 Aloe vera	52
Aloe vera Physiology	53
Metabolism	53
Transformation and Genetic Engineering	53
Nutritional Content	54
Distribution of Aloe vera	55
Benefits of Aloe vera	56
Traditional Medicine	56
Health Products	56
Cosmetic Products	56
Cultivation of Aloe vera	57
Planting Land	57
Seed, Planting, and Maintenance	57
Harvesting	57
Yields	58
2.9.2 Aloe vera Excellence	58
2.10 Aloe vera Genetic Substance	58

2.10.1 Sekuensing dari genome Aloe vera dan transcriptome	61
2.10.2 Aloe vera Chloroplast Genome	62
2.10.3 Aloe vera as a suitable candidate for this research	63
<b>2.11 Restriction and Ligases Enzymes</b>	<b>63</b>
<b>2.11.1 Restriction Enzymes</b>	<b>64</b>
Restrictions Site	64
Types	65
Type I	65
Type II	66
Type III	67
Type IV	68
<b>2.11.2 Ligase Enzymes</b>	<b>68</b>
Enzymatic Mechanism	69
Types	69
E. coli	69
T4	69
Mammals	70
Thermostable	71
<b>2.12 RT PCR (Reverse Transcription Polymerase Chain Reaction)</b>	<b>71</b>
Work Principle:	73
Application:	73
<b>2.13 PCR</b>	<b>73</b>
<b>2.13.1 Primer</b>	<b>74</b>
Melting Temperature	75
Annealing Temperature	75
Primer Length	75
Primer GC content	76
Forward and Reverse Primer Complementation	76
<b>2.13.2 PCR Steps</b>	<b>76</b>
1. Denaturation	76
2. Annealing	77
3. Elongation	77
<b>2.13.3 Primer3</b>	<b>77</b>
How to Use Primer3	77
<b>2.14 Sequencing</b>	<b>80</b>
<b>2.14.1 Maxam-Gilbert Sequencing</b>	<b>80</b>
<b>2.14.2 Sanger Sequencing</b>	<b>81</b>
<b>2.15 Electroforesis</b>	<b>82</b>
<b>2.16 Western Blot</b>	<b>83</b>
<b>2.17 ELISA</b>	<b>84</b>
<b>2.17.1 Insulin ELISA kit</b>	<b>84</b>

<b>BAB 3</b>	<b>86</b>
<b>Methodology</b>	<b>86</b>
3.1 In Silico	87
3.2 Data and Analysis Methods	87
3.3 Procedure	87
3.4 Identifying and Gathering Information	88
3.5 Observing and Collecting Construction Data	88
3.5.1 Collecting and Selecting Sequences	88
Determining the Promoter	89
Determining the Terminator	89
3.5.2 Selecting the Expression Vector	89
3.6 In Silico Steps	90
3.6.1 Restriction Sites Identification	91
3.6.2 Optimization of Proinsulin Codons with dnachisel	92
3.6.3 PCR Optimization	95
3.6.4 Transgene Construction Using Snapgene	96
Phase 1:	97
Phase 2:	97
Phase 3:	99
3.6.5 Protein Visualization	100
<b>Bab 4</b>	<b>101</b>
<b>Results and Discussion</b>	<b>101</b>
4.1 Transgene Construction	101
4.1.1 Codon Optimization	101
4.1.2 Virtual Digest and Virtual Electrophoresis	103
4.1.3 Transgenic Plasmids	103
4.2 Visualization	105
4.2.1 Model 1	106
4.2.2 Model 2	108
4.2.3 Model 3	111
4.2.4 Model 4	113
4.3 In-Silico Research Continuation Plan	115
4.4 In-Vivo Experimental Design	115
4.4.1 Culture and Preparation	116
4.4.2 DNA Coated Microcarrier Preparation	117
Transgene Preparation	117
Gold Particle and Microcarrier Preparation	117
4.4.3 Particle Bombardment	118
Aloe vera Transformation	119

4.4.4 Embryo Selection, Germination, and Regeneration	120
4.4.5 Extraction, Purification, and Analysis	120
Extraction and Purification	120
Insulin Analysis	121
<b>Bab 5</b>	<b>121</b>
<b>Conclusions and Recommendations</b>	<b>121</b>
5.1 Conclusions	121
5.2 Recommendations	122
<b>Reference</b>	<b>123</b>

## **Abstract**

Diabetes is a dangerous disease with high potential to be a killer. In 2030, it is predicted that there will be a significant increase in the number of diabetics compared to 2010. As the number of diabetics increases, the demand for insulin will also increase. Thus, it is predicted that there will be a rapid increase in demand for insulin and the need to find new, more efficient and more affordable ways of producing insulin.

The limited amount of production, coupled with high consumer demand, will certainly increase the market price of insulin. Also, the WHO stated that the prevalence of diabetes in middle to lower-income countries is relatively higher than in high-income countries. This of course will limit and reduce insulin access for those who need it.

Therefore, we developed a new, more efficient and cost-effective insulin production technique, using transgenic *Aloe vera* as an insulin production plant. By using this technique, the human preproinsulin genome will be inserted into the *Aloe vera* chloroplast using the particle bombardment technique, which is expected to produce significant results in transgenic *Aloe vera*. The presence of insulin in *Aloe vera* will be checked using ELISA techniques, western blot, and others.

However, due to time constraints and pandemics, this study will only focus on in silico simulations, in the form of computation, processing and data analysis, as well as in vivo experimental designs which are expected to be used as considerations before and after comparisons of in vivo experiments.

*Keyword:* In silico, computing, diabetes, insulin, *Aloe vera*, chloroplast, *particle bombardment*.

# **Chapter 1**

## **Preliminary**

### **1.1 Background**

Diabetes refers to a chronic disease caused by sugar metabolism disorders that result in sufferers having high blood sugar levels. About 9.3% of the adult population worldwide has diabetes. This figure is predicted to increase by more than 11% by 2045.

Diabetes can lead to a variety of serious health complications such as cardiovascular disease, chronic kidney disease, stroke and death. Diabetes currently ranks in the top 10 causes of death globally and is the third-largest cause of death in Indonesia (WHO 2017)<sup>[1]</sup>.

To date, insulin therapy is the only effective treatment for type 1 diabetes, and generally, type 2 diabetes if the disease persists. Insulin therapy requires regular blood sugar monitoring, and insulin injection to prevent secondary complications.

However, high insulin prices, and the discomfort of patients with injection, have led to some patients not adhering to insulin therapy. Therefore, researchers are developing new technological technologies to reduce patient discomfort during insulin therapy. Examples are through the lungs, mouth and nose (Modi et al, 2002; Goldberg and Gomez-Orellana, 2003; Cefalu, 2004).

Oral insulin using capsules to prevent the breakdown of insulin by the digestive system. The capsules used are in the form of alginate/chitosan, solid lipid nanoparticles, and others. (Sarmento, Ribeiro, et al, 2007; Sarmento, Martins, and Souto, 2019).

However, oral insulin and several other newer technologies require high doses to achieve the same performance as injectable insulin. This problem, coupled with the increase in diabetics, is predicted to decrease access to insulin. Another thing that plays a role in decreasing access to insulin is an unhealthy lifestyle, such as consuming too much junk food (Lancet Diabetes and Endocrinology, 2018).

The amount of insulin needed to treat type 2 diabetes is projected to increase by more than 20% worldwide over the next 12 years from 406 million bottles 1000 units currently to 511 million bottles 1000 units by 2030. At the same time,

global insulin use is also projected to increase to 634 million bottles of 1000 units in 2030 from 526 million bottles of 1000 units at present. By 2030, it is predicted that only half of the total number of diabetics will have access to insulin, or in other words, insulin will be scarce.

Besides, diabetes treatment is costly. It is estimated that, in 2007, the cost of diabetes treatment in America was \$ 174 billion. \$ 116 billion is for excess medical expenditures, consisting of \$ 27 billion in direct care, \$ 58 billion treating diabetes complications, and \$ 31 billion for general medical treatment. An estimated \$ 1 out of every \$ 10 is spent on diabetes-related care (American Diabetes Association, 2008). This cost is projected to reach \$ 216 to \$ 396 by 2025 (Giannini et al., 2009).

Many developing countries spend more than 50% of their annual costs on insulin and the treatment of diabetes-related diseases, and insulin costs can reach more than \$ 100 per month (Raab et al., 2004).

Economic limitations and production capacity will certainly cause big problems if they are not addressed immediately. The current demand for insulin production has already exceeded industrial capacity. The industry has predicted that it will need to increase its capacity to cope with the large forecast demand in the future. However, the production facilities for insulin are expensive and usually take a long time to make. Hence, there is a need for cheaper production methods that reduce the time needed to increase production.

Several methods of producing insulin from bacteria and animals have been used. However, this often creates health risks resulting from disease. Such risks can arise from cross-contamination with diseases that can affect animals and end-users. Thus, there is a need for a method of production that eliminates the possibility of cross-contamination between production by organisms and end-users. Besides, the conventional method of producing insulin from *E. coli* is predicted to be insufficient to increase the demand for insulin in the future, so an additional method is needed as a solution.

In addition, many current production methods require extensive processing to extract the therapeutic protein from the animal or other host organism in which it is present which can be produced and put the compound in conditions that can be

utilized by the patient. After purification, proteins can be combined with auxiliary materials or other carriers to stabilize the protein so that it can be used by patients. However, the process of extraction, refining, resuspension is involved, among other things, with the processing of proteins which are complicated, impractical, and may not be conducive to use in underdeveloped countries where general needs are made. Therefore, there is a need for a production method that can easily reduce the extraction process of the desired protein.

Transgenic *Aloe vera* plants can provide an economically viable alternative for the production of desired proteins, for example insulin, cosmetics, and others. Besides, the desired protein may be localized and concentrated in the *Aloe vera* leaf gel. This localization of the desired protein can simplify the extraction process or remove unwanted protein. Thus, this study can provide information about transgenic *Aloe vera*, where the desired protein is generally more abundant and easily accessible than other transgenic crops, such as tobacco and corn. Furthermore, the present invention can provide an efficient method for protein isolation.

The *Aloe vera* plant can offer various advantages over other conventional plant and bacterial systems for producing the desired protein. The benefits of the *Aloe vera* plant include the ability to process protein simply. Sometimes bacteria and yeasts find an incompatibility with the desired protein, due to the lack of a posttranslational modification system, in which these systems exist in plants. This process can include chemical modification, such as by glycosylation, and folding of some proteins. Furthermore, compared to other protein production methods, in animal cells, *Aloe vera* plant production can offer significant cost advantages, scalability advantages, and a reduced risk of contamination that may be harmful to humans. Another advantage of using *Aloe vera* is the ease of extraction, because there is a localization of the desired protein into the *Aloe vera* gel, making it more economical than using other plants.

*Aloe* is considered to be the first candidate for research in this new century, because of its economic and medicinal potential<sup>[2]</sup>. Previously, most of the *Aloe* research focused on effective biochemical analysis<sup>[3]</sup> and drug or cosmetic

application of secondary metabolites<sup>[4]</sup>. In addition, *Aloe vera* is also a research topic for GMO crops.

In this study, we will carry out an in-silico approach, namely computational data analysis and digital experimental design to develop transgenic *Aloe vera* as an expression platform and chloroplasts as an excretion system using particle bombardment transformation techniques which are expected to produce significant products on transgenic *Aloe vera*. The presence of insulin (protein) will be analyzed using ELISA, western blot, and others. However, in this study, we will focus on in-silico simulations in the form of computing, data processing and analysis, as well as in-vivo research designs which are expected to be taken into consideration before and after comparisons after in vivo experiments.

## **1.2 Formulation of Problem**

By considering the background that has been stated, several problems that can be formulated from this study are:

1. What is the potential of *Aloe vera* as a means of insulin production?
2. How can the insulin production system in the *Aloe vera* plant be developed?
3. How to construct a transgenic *Aloe vera* experimental design?
4. How is the purification and maturation of insulin produced from the *Aloe vera* plant?
5. Is this research possible to study computational analysis?
6. How is the prediction of insulin production in *Aloe vera*?

## **1.3 Problem Limitation**

The problem boundaries in this study are:

1. This research is an in silico (computational) design and has not been tested in vitro.
2. This study only develops insulin production techniques and has not tested its effectiveness in humans, either in silico or in vivo.

## **1.4 Research Purposes**

1. Provide another alternative to insulin production by using plants / *Aloe vera*
2. Provides binary plasmid construction advice for insulin production in *Aloe vera*

3. Generate computational data that can be analyzed for consideration in digital experiments.
4. Design digital experiments that are effective, efficient, economical, safe, and acceptable
5. Generate computational analysis data and digital experimental design as a reference in conducting further research both in vivo and in vitro which can develop transgenic *Aloe vera* which can produce insulin.
6. It is expected to be a solution in meeting the increasing demand for insulin which is safer and more economical.
7. After being able to meet the demand for insulin, it is hoped that new breakthroughs in insulin delivery such as oral insulin can become a method of insulin therapy besides injection so that it will result in adherence in therapy that can avoid secondary complications of diabetes.

### **1.5 Novelty**

Several scientists have studied proinsulin production in transgenic plants before, for example in tobacco mediated by *Agrobacterium tumefaciens*. However, in this study, we investigated the in silico design of proinsulin production from transgenic *Aloe vera* with the target of chloroplast expression using the particle bombardment (biolistics) method. The use of *Aloe vera* and this transformation system is based on various considerations.

### **1.6 Hypothesis**

1. Transgenic *Aloe vera* is capable of producing insulin.
2. The results obtained can be used to design experiments in vivo or in vitro to produce insulin from transgenic *Aloe vera*.
3. The experimental design can be applied in the real world and can produce insulin which is expected to support insulin production as part of meeting the increased insulin needs.

### **1.7 Research Benefit**

Can provide alternative methods such as plasmid construction design for insulin production using plants / *Aloe vera* so that it can have an impact in various fields.

1. From a health perspective: - The insulin produced is expected to be able to meet the needs of diabetics and have the same or even better effectiveness than conventional production methods.
2. From an economic point of view: - Insulin from transgenic Aloe vera is expected to have a more economical price than the price of conventional insulin.
3. In terms of science and technology. : -Breakthroughs in insulin delivery methods, such as oral insulin, require high doses of insulin. So that transgenic insulin production is expected to meet the increasing demand for insulin.

## Chapter II

### Literature Review

#### 2.1 Research Method

Experiment type:

1. In Vivo: (Within a living being). Experiments are carried out in the bodies of living things.
2. In vitro: (in glass). Experiments include cells or biomolecules that are present and controlled by the external environment.
3. In silico: (In silico = chips) Experiments carried out with the help of computer models and simulations<sup>[24]</sup>



Gambar a

Gambar b

Gambar c

Gambar 2.1.1 Comparison (a)In vivo, (b)In vitro, dan (c)In silico

After some consideration, we decided to do an in silico simulation. Some of the things we considered include:

1. In silico is used as a basis for in vivo and in vitro experiments, as well as to test hypotheses before conducting direct experiments in the real world.
2. The time needed in in-vivo and in vitro is longer than in-silico because in silico is a simulation, whereas in vivo and in vitro are direct experiments.
3. Because it is only a simulation, the costs required in the in silico experiment are relatively less

Tabel 2.1.1: In-silico, in-vitro and in-vivo comparisons in sunscreen studies

	<i>in vivo</i>	<i>in vitro</i>	<i>in silico</i>
Purpose of use	For special claims	Development, evaluation of sunscreen formulations	Development, evaluation of sunscreen formulations
Principle	Action spectrum of PPD	Transmission measurement, possibly action spectra of MED and PPD	Transmission calculation with skin roughness model action spectra of MED and PPD
Reliability	medium	low	medium
Time	high	medium	low
Costs	high	medium	low
Diffusion	medium	high	medium
Limitations	No consideration of interaction with UV-B (as the other <i>in vivo</i> or <i>in vitro</i> methods).	Uncertainties introduced by transmission measurement (no biological substrate)	Sunscreen formulations may not behave like simulation. Not available for all types of products
Conclusion	Expensive and time consuming method. Recommended for labelling of products and information of consumers.	Medium cost method. Usefulness for industrial development and labelling of products depending on the type of <i>in vitro</i> method.	Fast and low cost method. Useful for industrial development of sunscreen products.

Comparison of *in vivo*, *in vitro* and *in silico* methods for UV-A assessment.

## 2.2 Diabetes

Diabetes is a chronic disease that occurs when the body cannot metabolize sugar, which causes blood sugar levels to increase. This condition can be caused by two things, the first is the pancreas is no longer able to produce insulin, and the second is when the body is unable to properly utilize the insulin it produces. (International Diabetes Federation).

### 2.2.1 Type of Diabetes

#### 1. Type 1 Diabetes

It can develop at any age but is most common in children and adolescents. When you have type 1 diabetes, your body produces very little or no insulin, which means that you need daily injections of insulin to keep blood glucose levels under control.

#### 2. Type 2 Diabetes

It is more common in adults and accounts for about 90% of all diabetes cases. When you have type 2 diabetes, your body doesn't use the insulin it produces properly. The cornerstone of type 2 diabetes treatment is a healthy

lifestyle, including increased physical activity and a healthy diet. However, over time most people with type 2 diabetes will need oral medications and/or insulin to keep their blood glucose levels under control.

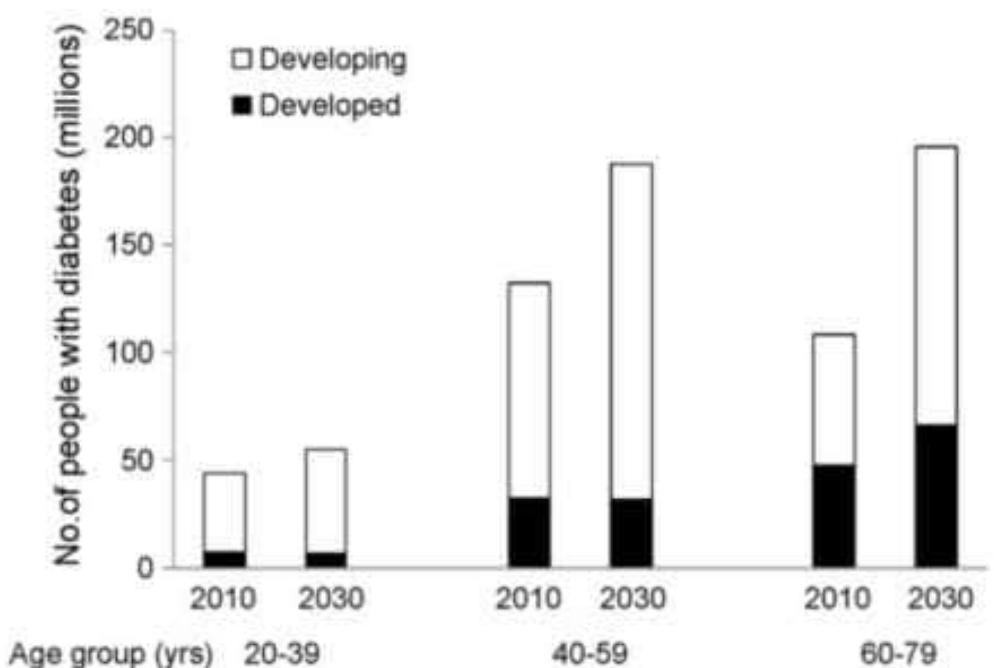
### **3. Gestational diabetes (GDM)**

This type of diabetes consists of high blood glucose during pregnancy and is associated with complications in both mother and child. GDM usually disappears after pregnancy, but affected women and their children are at a higher risk of developing type 2 diabetes later in life. (IDF International Diabetes).

#### **2.2.2 Number of Diabetes Patients**

Globally, an estimated 422 million adults lived with diabetes in 2014, compared with 108 million in 1980. The prevalence of diabetes worldwide (with age-standardized) has nearly doubled since 1980, increasing from 4.7% to 8.5% in the adult population. This reflects an increase in associated risk factors such as being overweight or obese. Over the past several decades, the prevalence of diabetes has increased more rapidly in low- and middle-income countries than in high-income countries.

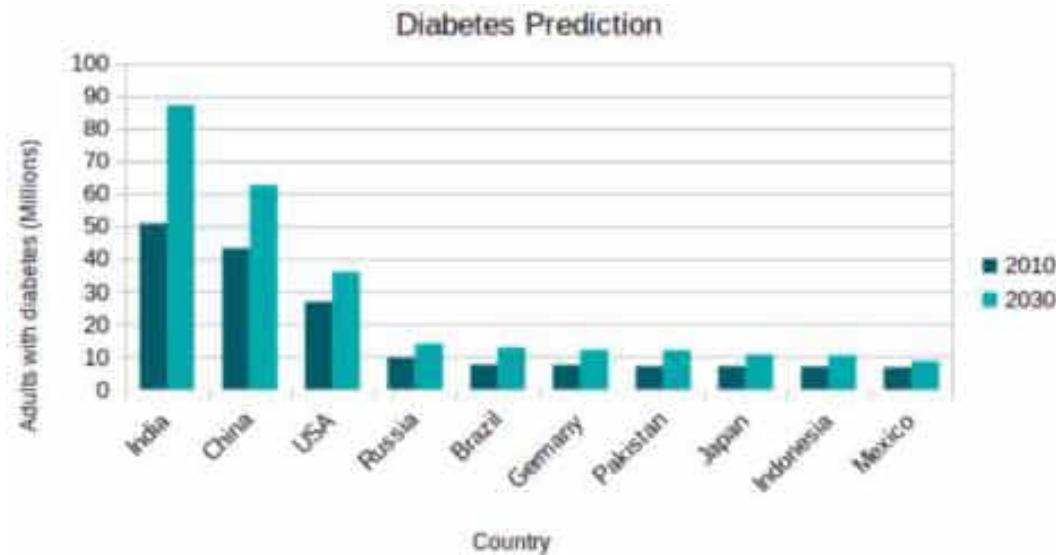
Diabetes caused 1.5 million deaths in 2012. Blood sugar higher than the maximum limit resulted in an additional 2.2 million deaths, increasing the risk of cardiovascular and other diseases. Forty-three percent (43%) of these 3.7 million deaths occurred before the age of 70. The percentage of deaths caused by diabetes occurring before age 70 is higher in low- and middle-income countries than in high-income countries. (WHO Global Report, 2016).



Graph 2.2.1. Diabetics in developed and developing countries in 2010 and estimated sufferers in 2030 (Source: Diabetes Research and Clinical Practice 87 (2010) 4-14)

Table 2.2.1: Top 10 countries with diabetes aged 20-79 in 2010 and forecast in 2030.

Table 2.2.1: Top 10 countries with numbers of people aged 20-79 years with diabetes in 2010 and 2030.				
	2010		2030	
	Country	No. of adults with diabetes (millions)	Country	No. of adults with diabetes (millions)
1	India	50.8	India	87.0
2	China	43.2	China	62.6
3	USA	26.8	USA	36.0
4	Russian Federation	9.6	Pakistan	13.8
5	Brazil	7.6	Brazil	12.7
6	Germany	7.5	Indonesia	12.0
7	Pakistan	7.1	Mexico	11.9
8	Japan	7.1	Bangladesh	10.4
9	Indonesia	7.0	Russian Federation	10.3
10	Mexico	6.8	Egypt	8.6



Graph 2.2.2: Comparison of 10 Countries with the most diabetes sufferers in 2010 and 2030.

(Source: Diabetes Research and Clinical Practice 87 (2010) 4-14.)

It can be seen that Indonesia is in the 9th rank with the most diabetes sufferers aged 20-79 in 2010 and it is estimated that Indonesia will be in the 3rd place in 2030. Also, there is also an estimated increase in diabetes cases in each country.

### 2.3 Insulin

Insulin is a protein hormone that functions to lower blood sugar levels by changing the sugar content in the blood to glycogen which is stored in the liver.

Insulin is produced by the beta islet Langerhans cells in the pancreas and is secreted when blood sugar levels are higher than normal. Insulin works against glucagon, a hormone that increases blood sugar by converting glycogen in the liver into glucose and circulating it in the blood. Insulin was first isolated by Frederick Banting and Charles H. Best from dog pancreas in 1921 (Wikipedia).

Insulin was originally formed as preproinsulin, a long chain consisting of 110 amino acids. Removal of the signal peptide from the preproinsulin tip yields proinsulin. The formation of disulfide bonds between the A and B chains, and the removal of the C chains to form biologically active insulin. Active insulin consists of 2 amino acid chains, an A chain (21 amino acids), and a B chain (30 amino acids) connected by peptide bonds.

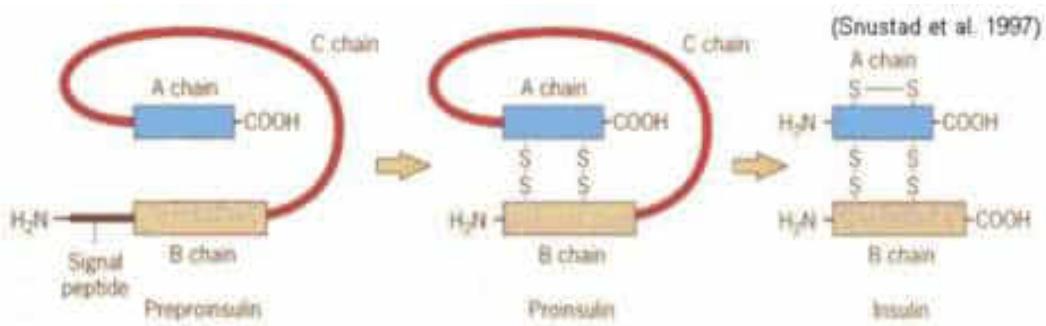


Figure 2.3.1 Post-translational modification of insulin

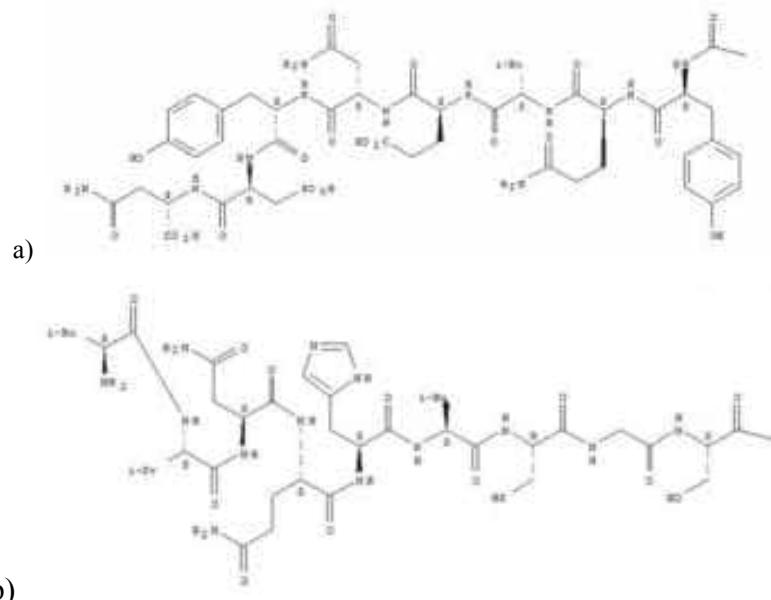
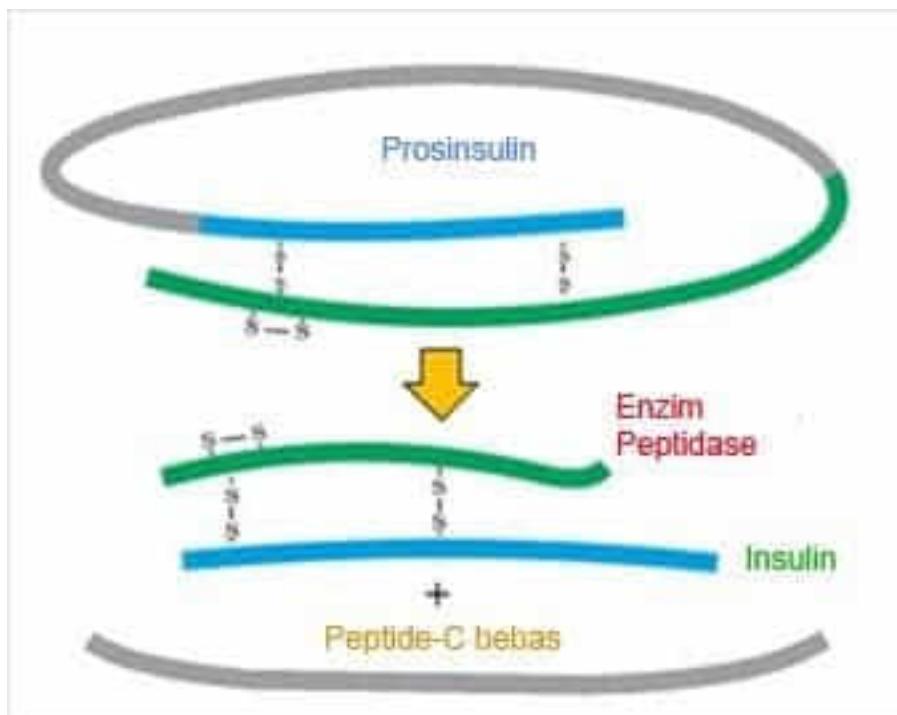


Figure 2.3.2 a) Insulin chain A oxidized b) Insulin chain B oxidized, lookchem.com



Figure 2.3.3 Amino acid structure of insulin<sup>[53]</sup>

### 2.3.1 Insulin Synthesis

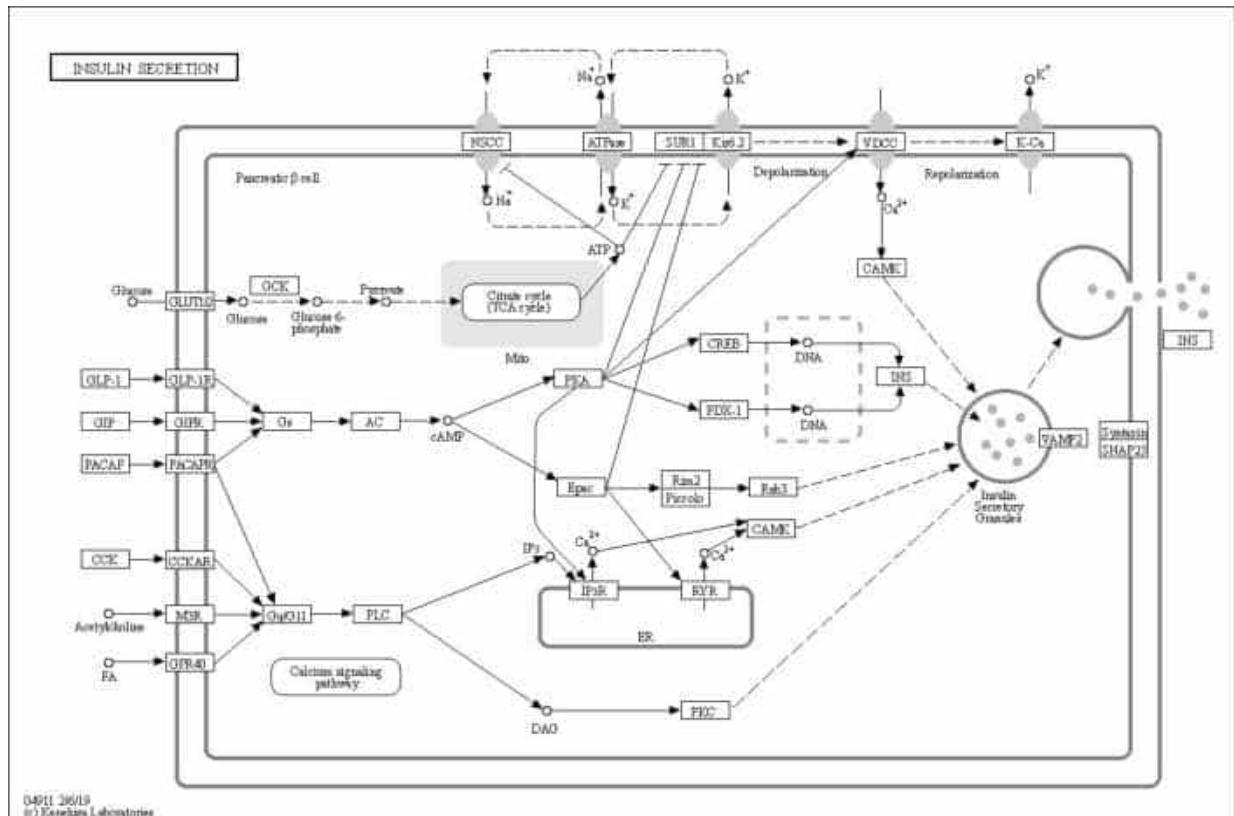


(Figure 2.3.4 Insulin Synthesis)

The stages of insulin synthesis are as follows:

1. Insulin synthesis is initiated in the form of preproinsulin (precursor to the insulin hormone) in the endoplasmic reticulum of beta cells.
2. With the help of the peptidase enzyme, preproinsulin undergoes breakdown to form proinsulin, which is then collected in the bubbles in the cell.
3. Proinsulin is then broken down again by the peptidase enzyme into insulin and peptide-C, both of which are ready to be secreted simultaneously through the cell membrane.
4. The function of insulin is needed in the process of utilizing glucose in the blood. Increased blood glucose levels are the main component that stimulates beta cells to produce insulin<sup>[191, 192]</sup>.

### 2.3.2 Insulin Secretion Pathway



(Figure 2.3.5 Insulin secretion pathways) Source: KEGG (Kyoto Encyclopedia of Gene and Genome)

Pancreatic beta cells are specialized endocrine cells that continuously regulate blood sugar and other fuel levels, and in response, secrete insulin to maintain normal fuel homeostasis. Glucose-induced secretion of insulin and its potential is the main mechanism of insulin release. Glucose is transported by glucose transporter (GLUT) into pancreatic beta cells. Glucose metabolism produces ATP, which blocks ATP sensitive  $K^+$  channels and causes an influx of voltage-dependent  $Ca^{2+}$ . The increase in  $[Ca^{2+}]$  triggers the release of insulin granules by exocytosis.

The secretion of insulin is further regulated by several hormones and neurotransmitters. Peptide hormones, such as glucagon-like peptide 1 (GLP-1), increase cAMP levels and thereby promote insulin secretion through the combined action of PKA and Epac2. Acetylcholine (ACh), the main parasympathetic neurotransmitter, binds to Gq-coupled receptors and activates phospholipase C- (PLC-), and its stimulatory effect involves activation of protein kinase C (PKC), which stimulates exocytosis. In addition, ACh

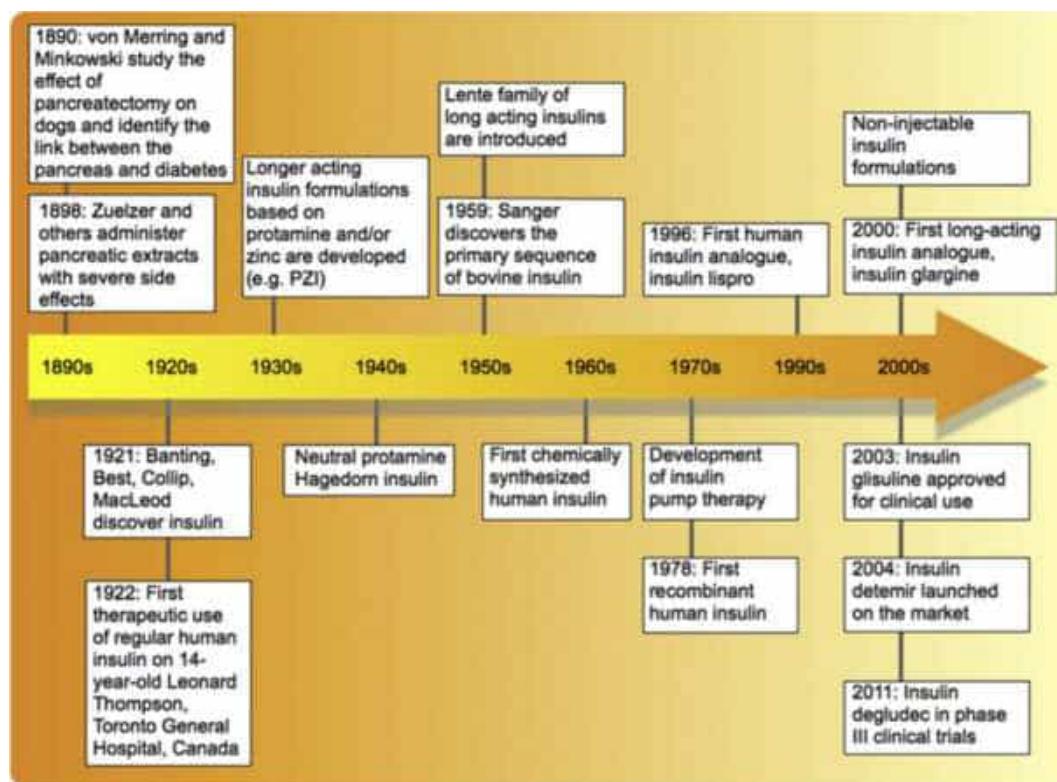
mobilizes intracellular  $\text{Ca}^{2+}$  by activation of IP<sub>3</sub> receptors. (KEGG (Kyoto Encyclopedia of Gene and Genome)). Insulin is also a protein compound that is easily digested and breaks down when exposed to stomach acid.

### 2.3.3 Conventional Insulin Production

Before the 1980s, most insulin products on the market came from porcine (porcine) or bovine (bovine) pancreas. These insulin products are highly impure and can cause immunological complications such as insulin allergy, immune-mediated lipoatrophy at the injection site, and antibody-mediated insulin resistance. These insulin products also cause significant pharmacokinetic and pharmacodynamic variations<sup>[54]</sup>.

Over time, developments in insulin technology have resulted in better quality insulin and more consistent performance. In addition, long-acting insulin was being developed to extend the working time of insulin and reduce the number of daily injections. Long-acting insulin is formed by combining zinc and/or protamine into protamine insulin and protamine zinc insulin (PZI) and began to be developed in the 1930s<sup>[54-56]</sup>. Besides, NPH (Isophane Neutral Protamine Hagedorn) was also produced in the 1940s<sup>[57]</sup>, and is still used today either alone or in combination with other insulin products<sup>[59]</sup>.

In 1950, the insulin trilogy "Lente" was introduced<sup>[58]</sup> and Sanger succeeded in obtaining the primary structure of bovine (bovine) insulin. All these advances led to the birth of high-quality pure animal insulin (both monocomponent and single-peak insulin) in 1970<sup>[61]</sup>.



(Figure 2.3.6 History of insulin<sup>[54]</sup>)

### 2.3.4 Recombinant Insulin Production

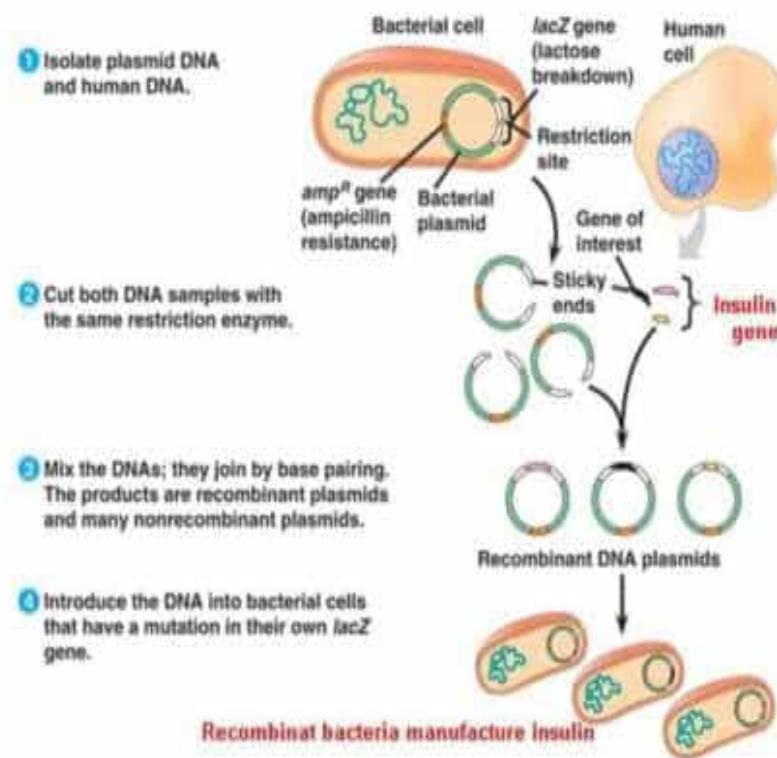
In the 1970s, recombinant DNA technology underwent rapid development which led to the introduction of recombinant human insulin in 1978<sup>[62]</sup>.

#### *Escherichia coli*

Among prokaryotes, *Escherichia coli* is more desirable for the production of recombinant insulin because of the advantages it offers. Some of these advantages include high growth rates, simple media requirements, ease of handling, high yields, and great cost savings.

Bioengineering of *E. coli* began in 1978 and was carried out through the use of a plasmid, pBR322 (Goeddel et al., 1979). Plasmids are small pieces of DNA that float in the cytoplasm and are not important for cells (Madigan et al., 2015). Because they are nonessential and so small, they are often released and taken up by other prokaryotes through a transformation process, which allows the microbes to adapt by expressing different genes. This phenomenon is a way for some bacteria to develop genetic resistance.

This phenomenon opens up great opportunities for researchers to modify bacteria into factories for the production of various kinds of proteins and other chemicals. One of them was in 1978 when researchers were able to change the *E. coli* strain to produce human insulin. Using the plasmid pBR322, the researchers succeeded in cutting the specific sequence of human insulin DNA base pairs and the plasmid vector using restriction enzymes (EcoRI), as well as attaching a synthetic human insulin gene to the plasmid. When plasmids are placed into *E. coli*, the bacteria take up the plasmids; this allows the bacteria to replicate the plasmid and code for human insulin. Through this model, bacteria continuously make new plasmids containing the insulin gene. A general schematic of how this transformation occurs is shown in Figure 2.6<sup>[6]</sup>.



(Figure 2.3.7 Transformation of transgenic bacteria)

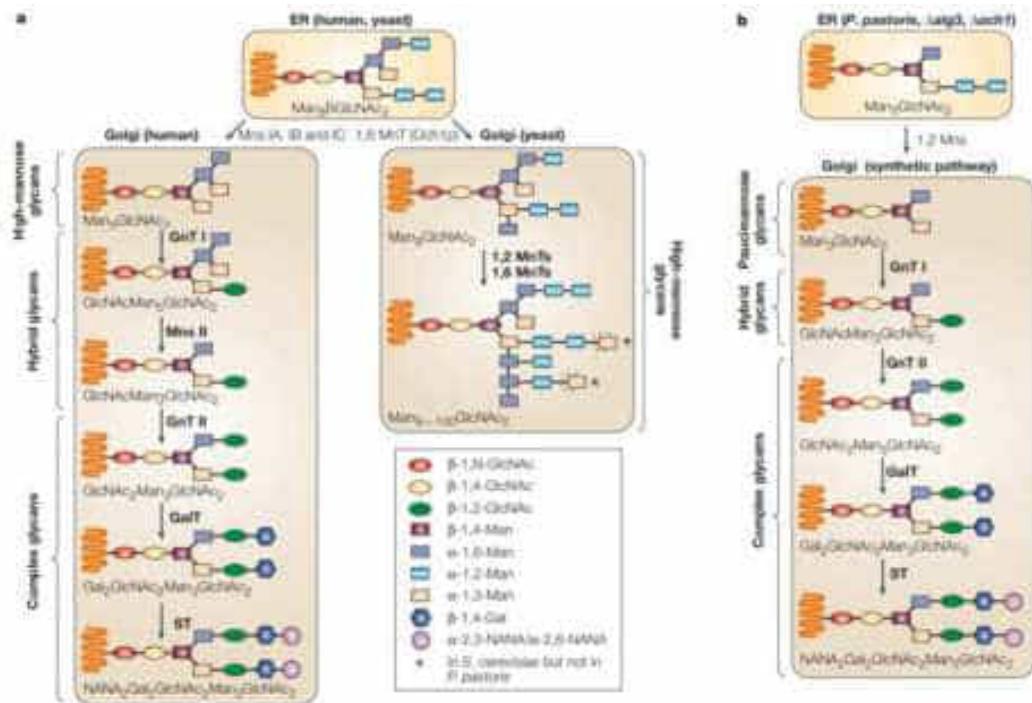
Apart from all the advantages that E. Coli offers, the E. Coli expression system also has drawbacks. Examples are loss of plasmid traits, undesirable inducers for gene expression, accumulation of heterologous intracellular proteins as inclusion bodies, improper refolding of proteins, lack of post-translational modifications (including the formation of disulfide

bonds), metabolic load and protein-mediated stress, endotoxin contamination, poor secretion, proteolytic digestion, as well as complexity in downstream processes<sup>[26-27]</sup>.

## Yeast

Yeast is the preferred host for the expression of a wide variety of heterologous proteins which require post-translational modification for their biological activity. Yeast can carry out many post-translational modifications such as phosphorylation, O-linked glycosylation, N-linked glycosylation, acetylation and acylation. Recombinant protein is expressed in a soluble and well-folded form in a functionally active form. Biopharmaceutical production using yeast expression systems is also very cost-effective and can be increased by using large bioreactors.

However, one major concern in the production of therapeutic glycoproteins for humans is that the yeast N-glycosylation pathway is a high mannose type, which provides a short half-life in vivo and hyper immunogenicity that renders therapeutic glycoproteins less effective. Attempts have been made to alter the yeast's N-glycosylation pathway to produce therapeutic glycoproteins with a humanized (humanized) N-glycosylated structure<sup>[29, 35]</sup>. Examples of yeasts used are *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Pichia pastoris* and are often used to produce recombinant protein<sup>[93-96]</sup>.



(Figure 2.3.8: Main N-glycosylation pathway in yeast and humans. A) Comparison of human N-glycosylation pathway (left) with yeast (right). b) Synthetic pathways in yeast made using mutant oligosaccharide assembly.)

## Transgenic Plants

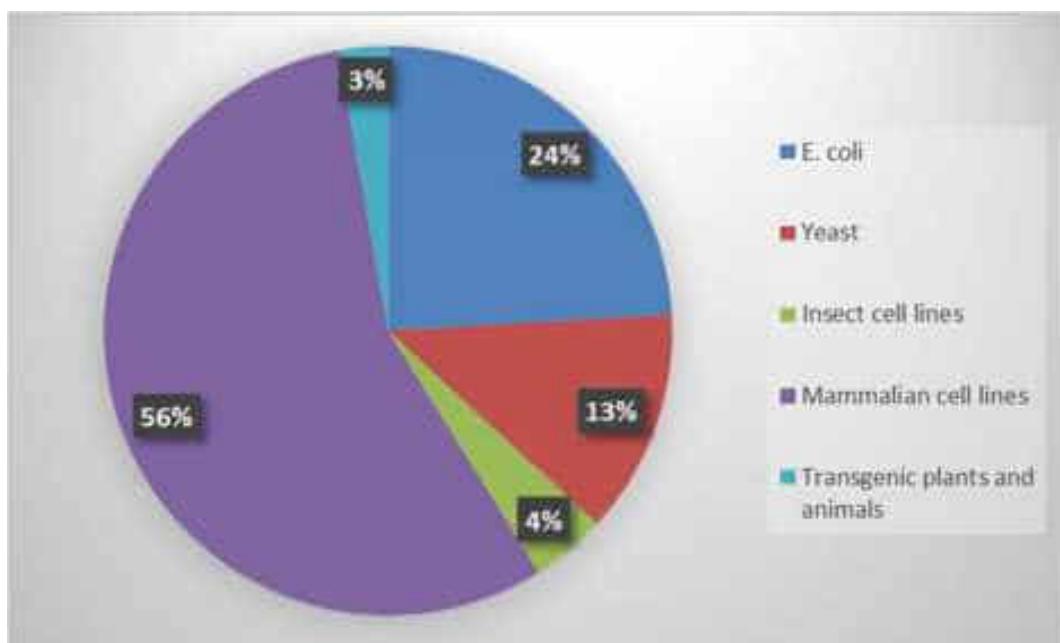
GMO crops have long been used as a means of producing recombinant protein, due to the advantages they offer. Some of these advantages include cost-effectiveness, high-quality protein processing, absence of human pathogens, ease of production, and the existence of a eukaryotic system for post-translational modification. Examples of proteins that are successfully produced with transgenic plants are growth hormone<sup>[30]</sup>; Hepatitis B virus surface antigen; antibody; industrial protein, as well as some milk proteins.

Recombinant human insulin has been successfully expressed and produced from the *Arabidopsis thaliana* oil body. Oil Body is an organelle found in oilseed and consists of a hydrophobic triacylglycerol core which is covered by a phospholipid membrane and an outer wall of a protein called oleosin. In addition, recombinant insulin has also been successfully expressed and harvested from tobacco and lettuce.

Overall, GMO crops hold high potential for high capacity and cost-effective insulin production. The high level of proinsulin expression in

leaves with long term stability offers low-cost technology for both injection and oral insulin. In addition, transgenic plant cells can be used as natural storage sites for recombinant insulin<sup>[36]</sup>.

### 2.3.5 Perbandingan Persentase Produksi Insulin Rekombinan



(Figure 2.3.9: Comparison of sources of recombinant insulin production.)

### 2.3.6 Protein Extraction

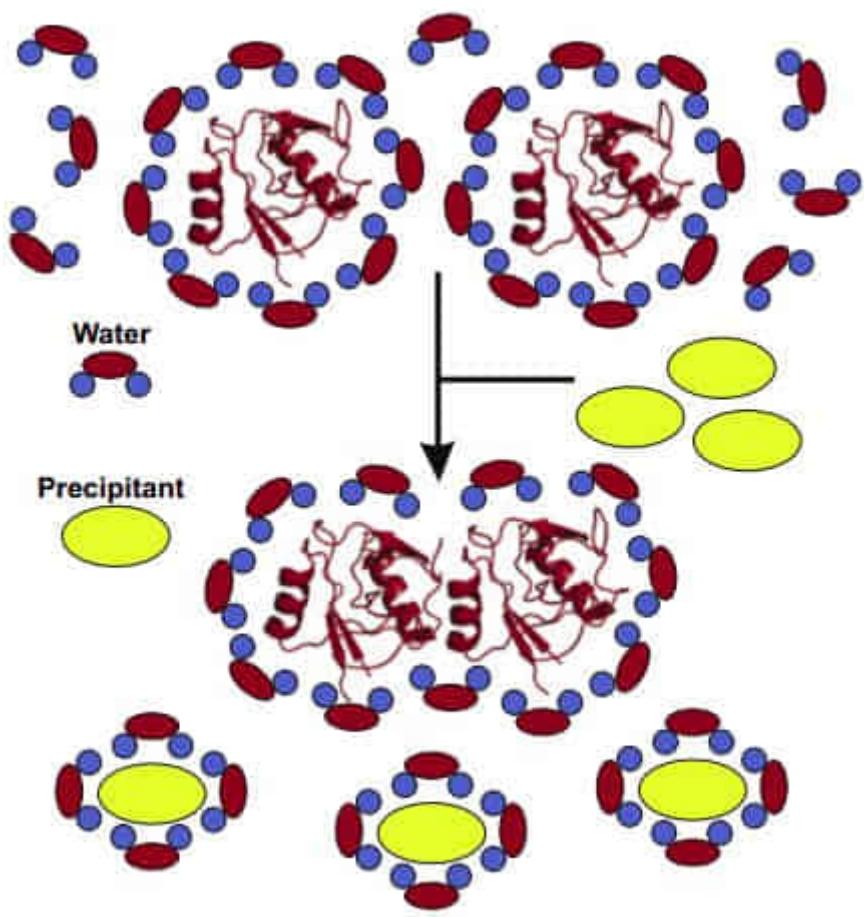
Insulin is a hormone made from protein. Large-scale analysis of proteins is referred to as proteomic analysis and is commonly used to identify proteins<sup>[116, 117]</sup>. Proteomic analysis was first developed in 1975 after the discovery of two-dimensional gels, and mapping of the E. coli protein. Some examples of proteomic analysis will be described at the end of chapter 2.

A successful proteomic analysis relies on a good protein sample<sup>[108]</sup>. There are several points to consider when preparing protein samples for proteomic analysis. Two of them are levels of contaminants and levels of protein degradation. Although small amounts of contaminants are tolerated, some proteins are unstable, and contamination can interfere with proteomic analysis<sup>[108]</sup>. Meanwhile, protein degradation is the destruction of the primary structure of the protein and causes the protein to break down into amino acid units, or smaller peptide chains.

Note that denaturation of protein is not the same as degradation of the protein. Protein denaturation only damages the secondary, tertiary, and quaternary structures of proteins. This simply changes the three-dimensional structure of the protein and makes it a linear peptide chain. Denaturation certainly removes the biological properties of the protein, so that the protein cannot be used for analyses requiring biologically active properties<sup>[115]</sup>.

On the other hand, protein degradation is the destruction of the primary structure of the protein due to the breaking of peptide bonds between amino acids. Since protein degradation breaks the protein into smaller peptide chains, proteomic analysis cannot be performed on denatured proteins<sup>[115]</sup>.

In some cases, the protein must be concentrated prior to carrying out the analysis. There are several techniques available for doing this, such as ultrafiltration, affinity chromatography, and precipitation. Compared to other techniques, precipitation is believed to be the technique with the most extraction and is most suitable for volumes with low protein content. This occurs because the protein can attach to the ultrafiltration membrane, and the stationary phase is used in affinity chromatography<sup>[118]</sup>.



(Figure 2.3.10 The basic principle of protein precipitation is to change the solubility potential of the solvent, and reduce the solubility of the solute with the aid of a reagent called precipitant<sup>[108]</sup>)

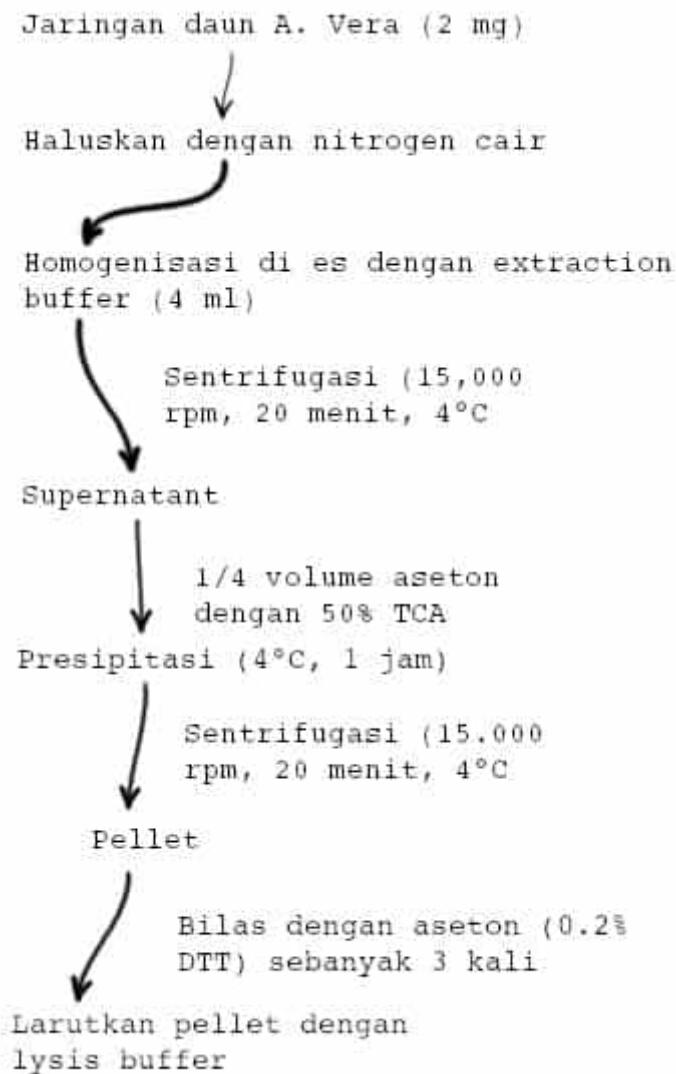
### Trichloroacetic acid (TCA)-acetone Precipitation Method

The trichloroacetic acid (TCA) –acetone precipitation method is a method developed by Damerval<sup>[106]</sup> and is usually used to prepare samples for proteomic analysis. This method is believed to minimize protein degradation and contaminant levels<sup>[107]</sup>.

Trichloroacetic acid (TCA) is an analogue of acetic acid in which the three hydrogen atoms of the methyl group are replaced by chlorine atoms (Wikipedia). The addition of TCA in aqueous solution disrupts the hydrogen bonds between water molecules, making it easier to coagulate and collect by centrifugation. However, because TCA breaks hydrogen bonds, it can damage the secondary structure of proteins and cause denatured proteins<sup>[109]</sup>. Even so, TCA is not strong enough to break peptide bonds

between amino acids making up proteins, so that the primary structure damage can be minimized<sup>[107]</sup>.

The schematic of the TCA / Acetone precipitation method is as follows:



#### 2.3.1 TCA / Acetone Precipitation Method Scheme<sup>[109, 110]</sup>

### Phenol Extraction with methanol/ammonium acetate precipitation Method

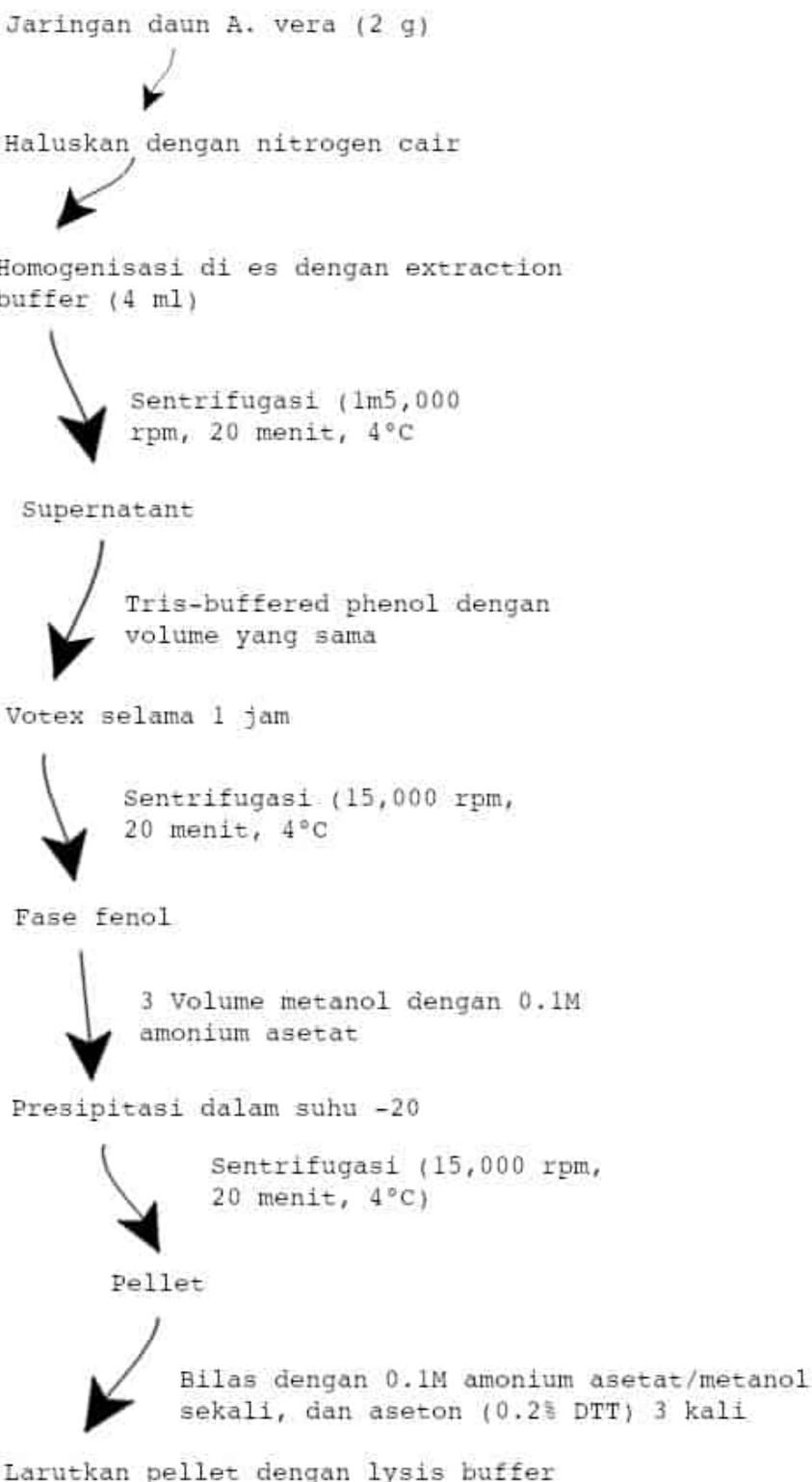
Phenol extraction with methanol/ammonium acetate precipitation is another method that can be used for protein extraction. This extraction was described by Hurkman and Tanaka for the preparation of proteomic assay samples<sup>[111]</sup>. Phenol extraction was first developed for the purification of

carbohydrates and nucleic acids and is now commonly used to remove proteins from nucleic acid solutions<sup>[112]</sup>.

Phenol is the simplest aromatic alcohol with the chemical formula C<sub>6</sub>H<sub>5</sub>OH and consists of a phenyl group that binds to a hydroxyl group. Phenols have weak acidic, corrosive and toxic properties. Phenol interacts with protein hydrogen bonds and breaks them down, causing denatured protein<sup>[112]</sup>.

Comparison between phenol and TCA / Acetone extraction shows that they are both efficient extraction methods<sup>[113, 114]</sup>, and can minimize protein degradation. However, phenol extraction is more efficient in removing contaminants and produces the highest quality gel for electrophoresis. Phenol extraction also produces glycoproteins in greater quantity than TCA<sup>[114]</sup>.

The phenol extraction scheme is as follows:



Scheme 2.3.2 Phenol Extraction with methanol/ammonium acetate precipitation Method

## **2.4 Biopharming**

Biopharming is a farming business (farming) that uses plants or livestock to produce a protein or certain metabolite compounds that have health or medicinal value<sup>[64]</sup>. Biopharming is inseparable from genetically modified products that are applied to animals or plants to insert genes from other organisms so that the animal or plant expresses certain compounds, which can then be harvested as a result<sup>[65,66]</sup>.

The product produced by biopharming is a recombinant protein. Biopharming offers advantages in the form of low costs because it integrates farming businesses that can produce without the need for expensive infrastructure and the production capacity can be managed more flexibly<sup>[67]</sup>. Advances in the plant bio / molecular pharming approach in the last decade have made plants attractive manufacturing systems, even achieving commercially relevant levels of production in a short period<sup>[69-71]</sup>. As for the importance and prospects of plant expression for the cost-effective production of recombinant proteins, potential vaccine candidates, monoclonal antibodies, as well as industrial enzymes expressed in plants<sup>[68]</sup>.

The practise of using plants for the production of high-value recombinant proteins ranging from pharmaceutical therapy to non-pharmaceutical products such as antibodies, vaccine antigens, enzymes, growth factors, research or diagnostic reagents, and cosmetic ingredients<sup>[72]</sup> has increased over time and has grown significantly. in recent decades, which led to a major paradigm shift in the pharmaceutical sector.

The main advantages of any plant-based system are easy cultivation, low cost, low or no pathogen load, rapid mass production, recombinant protein, and the ability of plants to compose post-translational modified eukaryotic-like complex proteins (PTM)<sup>[73]</sup>.

Protein folding is essential for maintaining the biological activity of recombinant therapeutic proteins Due to a lack of protein processing complexes and limited capacity for PTM, proper protein folding cannot be achieved in the prokaryotic expression system<sup>[75]</sup>. Plants can assemble and carry out large multimeric PTM proteins required for their biological function activities. However, plants lacking the original human N-glycosylated processing

mechanisms have been overcome by a glycoengineering approach towards the targeted synthesis of human and non-human structures to increase the homogeneity, quality, and quantity of the product<sup>[76,77]</sup>.

Table 2.4.1: Examples of biopharmaceuticals using large-scale plant expression systems produced by various companies over the past decades. (Stoger et al., 2014 adapted from Wang, Eu Sheng 2016)

Company	Plant species	Plant organ/cells	Expression technology	Lead product
ProtaLix Biotherapeutics	Carrot	Suspension cells	Stable nuclear	ELELYSO® (taliglucerase alfa)
Medicago	<i>Nicotiana benthamiana</i> Alfalfa	Leaves	Proficia (transient), Stable nuclear	Vaccine for pandemic/ Seasonal flu
Kentucky BioProcessing	<i>Nicotiana benthamiana</i>	Leaves	Transient, agroviral launch vector	Subunit vaccine for influenza
Fraunhofer IME	Tobacco	Leaves	Stable nuclear	Antibody for HIV (microbicide)
Ventria Biosciences	Rice	Seeds	Stable Nuclear	VEN100 (lactoferrin)
Synthon/Biolex Therapeutics	Duckweed	Leafy biomass	Stable nuclear LEX system	Antibody for non-Hodgkin's lymphoma
Planet Biotechnology	Tobacco	Leaves	Stable nuclear	CaroRx (cavity inhibiting treatment)
Icon Genetics	<i>Nicotiana benthamiana</i>	Leaves	MagnICON (transient)	Vaccine for non-Hodgkin's lymphoma
VAXX	Potato	Tubers	Stable nuclear	Oral subunit vaccine for Norwalk virus (NoroVAXX)
SAFC	Tobacco	Leaves	Stable nuclear	Contract manufacturing
NAIST	Strawberry	Fruits	Stable nuclear	Canine interferon alpha
Mapp Biopharmaceutical/ LeafBio	<i>Nicotiana benthamiana</i>	Leaves	MagnICON (transient)	Multiple vaccines and microbicides

Abbreviations: Fraunhofer CMB, Fraunhofer Center for Molecular Biotechnology; Fraunhofer IME, Fraunhofer Institute for Molecular Biology and Applied Ecology; NAIST, National Institute of Advanced Industrial Science and Technology; SAFC, Sigma-Aldrich Fine Chemicals.

## 2.4.1 Comparison of Expression Systems in Plants and *E. coli*

### *E. coli*:

Table 2.4.2 Problems in *E. Coli* as an expression system and strategy

Issue Addressed	Engineering Strategy	Example Reference
Lack of posttranslational modifications in <i>E. coli</i>	Glycosylation	Expression without the glycosylation (only if the biological activity is not impaired) IL-2 [3], IFN- $\beta$ -1b [4]
		Mutation of residues on the surface to create more soluble protein EPO [5]
		Mutation of glycosylation sites to cysteines to allow subsequent glycosylation <i>in vitro</i> EPO [6]
	Proteolytic maturation	Expression in two separate strains or cleavage of the precursor <i>in vitro</i> insulin [7]
Disulfide bridge formation		Expression in the periplasmic space hGH [8], proinsulin [9]
Protein stability		Decreasing number of free cysteine residues by mutation to alanine IL-2 [10], IFN- $\beta$ -1b [11]
		Deletion of the hydrophobic region KGF [12]
Modulation of protein activity		Design of the rapid-acting or long-acting protein version insulin [13-17]
		Enhanced activity by improved affinity to the target molecule DNaseI [18]
		Design of protein consisting of a consensus sequence IFN- $\alpha$ -con [19, 20]

Source: J. Shanmugaraj, B .; Malla, A .; Phoolcharoen, W. Emergence of novel coronavirus 2019-nCoV: Need for rapid vaccine and biologics development. Pathogens 2020, 9, 148. [CrossRef]

### Advantages:

high growth rates, simple media requirements, ease of handling, high yields, and very cost-effective.

### Disadvantages:

loss of plasmid traits, undesirable inducers for gene expression, accumulation of heterologous intracellular proteins as inclusion bodies, improper refolding of proteins, lack of post-translational modification (including glycosylation, proteolytic maturation, the formation of disulfide bonds), metabolic load and protein-mediated stress, endotoxin contamination, poor secretion, proteolytic digestion, and complexity of downstream processing<sup>[26-27]</sup>.

### In Transgenic Plants:

In *E. coli* (prokaryotic) there is no post-translational modification (glycosylation, proteolytic maturation, and disulfide bridge formation) so a strategy with a different approach is needed to produce functional insulin. Whereas eukaryotic has a post-translational modification mechanism which will increase the efficiency of insulin production compared to using *E. Coli* as an expression platform.

The reason in this study using plants as an expression platform is because plants offer various advantages including more economical value, high scalability, global production scale, and high protein yield, with high product quality. Plants also offer high protein folding accuracy compared to bacteria, there are post-translational modifications such as glycosylation, disulfide bridges and others, which bacteria cannot. Meanwhile, the post-translational modification is needed in the formation of protein, especially in the tertiary structure which affects the functionality of the insulin.

## 2.5 Expression Platform

There are various transformation systems/platforms including bacteria, mammalian cells, yeast, insect cells, and plants which have their respective advantages. For more details in the table below.

Table 2.5.1. Advantages and disadvantages of various expression platforms<sup>[78]</sup>

Expression System	Advantages	Disadvantages
Bacteria	Easy to manipulate Low cost High expression Ease of scale up Short turnaround time Established regulatory procedures and approval	Improper folding Lack of post-translational modifications which may affect the protein function Endotoxin accumulation
Mammalian Cells	Proper folding and authentic post-translational modifications Existing regulatory approval	High production cost Expensive media and culture condition requirements
Yeast	Rapid growth and scalable Easy to manipulate Simple and inexpensive media requirements and culture conditions Post-translational modifications of recombinant proteins	Difficulty in cell disruption due to the thick and hard cell walls Hyperglycosylation of proteins Limited glycosylation capacity
Insect cells	High expression levels Ability to produce complex proteins including secreted, membrane, and intracellular proteins Proper folding and post-translational modifications	High cost and time consuming Expensive media and culture condition requirements
Plant	Rapid and affordable Optimized growth conditions Free from pathogen and bacterial toxin contaminants Economical Post-translational modification somewhat similar like mammalian system	Regulatory compliance Limited glycosylation capacity

Table 2.5.2. Another comparison of various expression systems for recombinant pharmaceutical proteins (adapted from Schillberg et al., 2003., Ma et al. 2003)

Comparisons	Transgenic Plant	Plant Cell Culture	Bacteria	Yeast	Mammalian Cell Culture	Transgenic Animals
Overall cost	Very low	Medium	Low	Medium	High	High
Scale-up capacity	Very high	Medium	High	High	Very low	Low
Production scale	Worldwide	Limited	Limited	Limited	Limited	Limited
Protein yield	High	High	Medium	High	Medium-High	High
Protein folding accuracy	High	High	Low	Medium	High	High
Glycosylation	Minor differences	Minor differences	None	Incorrect	Correct	Correct
Product quality	High	High	Low	Medium	High	High
Contamination risks	Low	Low	Endotoxins	Low	Virus, Prions, oncogenic DNA	Virus, Prions, oncogenic DNA
Safety	High	Non-specific	Low	Unknown	Medium	High
Storage cost	Inexpensive	Moderate	Moderate	Moderate	Expensive	Expensive

Although mammalian cell-based expression systems have the advantage of producing vaccine antigens that are generally identical to their natural hosts, mammalian cell culture is costly and thus inhibits large-scale production.

Bacterial or yeast-based expression systems allow large-scale production of transgenic proteins but produce products which differ from the original form (Daniell et al., 2001). For example, proteins that normally undergo glycosylation in animals do not undergo this process in bacteria and are likely to undergo hyperglycosylated in fungi, which results in low protein expression in both systems (Harashima, 1994).

Plant-based expression systems offer several advantages, including fast scalability without the need for complicated fermenters or bioreactors, relatively more economical, high protein expression, minimal risk of contamination from mammalian pathogens or bacterial toxins, heat stability, and technology available for harvesting and processing crops and plant products on a large scale (Daniell et al., 2001) compared to conventional production. Therefore, we chose plants as the ideal platforms for transgenic protein expression.

In general, efficient transformation methods should be cost-effective, do not require the use of hazardous or carcinogenic chemicals, have direct procedures, allow the proper introduction of a transgene into the target host, allow integration of transgene at low copy counts; and enables convenient post-transformation of plant regeneration.

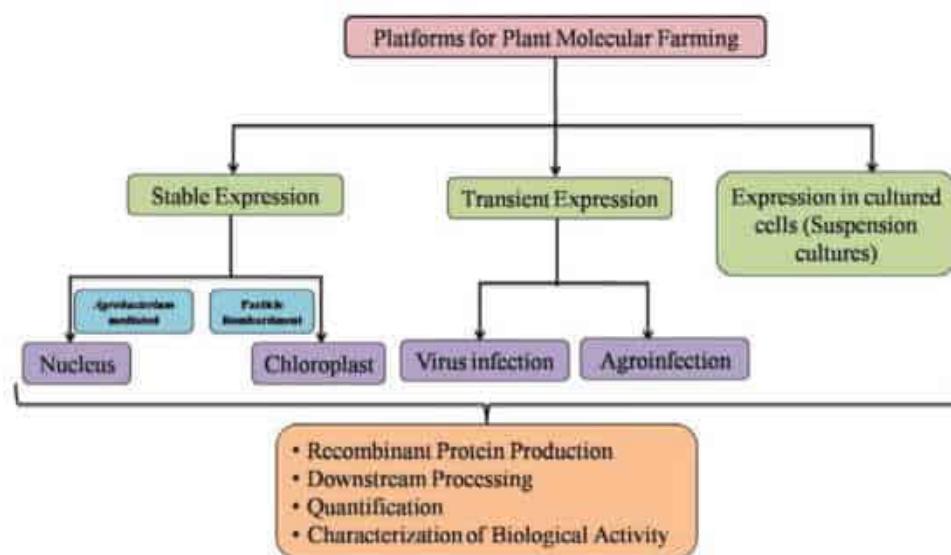
The transformation methods are divided into direct or indirect categories depending on their mode of action. Indirect methods usually require biological agents such as bacteria for the delivery of transgene into the host plant whereas

direct methods are usually physical, involving direct penetration of DNA into the target host cell.

### Plants as a biofactory for heterologous protein expression

Significant advances in the last decade leading to the production of proteins derived from plant biopharmaceuticals, such as vaccines, antibodies, human growth regulators, hormone production and blood have led to the emergence of molecular pharming. (Fischer and Emans, 2000; Giddings, 2001).

## 2.6. Expression Types



(Figure 2.6.1. Schematic of a strategic approach to transgenic crops for protein production in plants<sup>[79]</sup>.)

The key to producing high transcription levels are regulatory genetic elements, particularly the promoter and polyadenylation site. The expression level depends on the high level of translation, the correct folding, the right target, and protein stability.

Table 2.6.1. comparison of the types of genetic manipulation expression<sup>[80]</sup>.

Ekspresi Transient	Ekspresi Nuclear	Ekspresi Kloroplas
Manipulasi yang mudah pada vektor virus dan bakteri	Banyak metode manipulasi genetik yang sesuai untuk sistem ini.	Manipulasi yang mudah, resisten terhadap silencing gen
Hasil ekspresi yang tinggi	Membutuhkan waktu yang relatif lama untuk manipulasi genetik	Hasil ekspresi yang tinggi karena banyaknya kloroplas per sel
Resiko kontaminasi yang tinggi dari transfer gen ke lingkungan		Resiko kontaminasi yang rendah dari transfer gen ke lingkungan
Ekspresi yang cepat		Metode transformasi yang terbatas
		Tidak terdapat jalur glikosilasi yang tetap

### 2.6.1 Chloroplast Expression

Chloroplast expression involves the introduction of a transgene into the chloroplast genome using a gene gun. The chloroplast transformation system has more advantages over the nuclear expression system. The chloroplast genome is easier to manipulate. If the chloroplast genome has been sequenced, transgene tapes can be made to insert foreign genes into the spacer regions between the functional chloroplast genes, using two flanking sequences in the chloroplast genome, via homologous recombination (Daniell, Lin, Yu, & Chang, 2016; Danielly, Streatfield., Streatfield, & Wycoff, 2001).

The transformation of the chloroplast genome is more difficult than the transformation of the nuclear genome because of the double membrane obstruction of the chloroplasts and the lack of knowledge of which viruses can infect chloroplasts. However, effective transformations have been obtained using the gene gun method by shooting young plant tissue with gold and tungsten particles coated with DNA (Verma, Samson, Koya, & Daniell, 2008). There are thousands of copies of the chloroplast genome in each leaf cell. High yields have been obtained (more than 70% of the protein is soluble in plant leaves) (Daniell et al., 2016). Another advantage is the low risk of genes contaminating the environment because chloroplasts are maternally inherited.

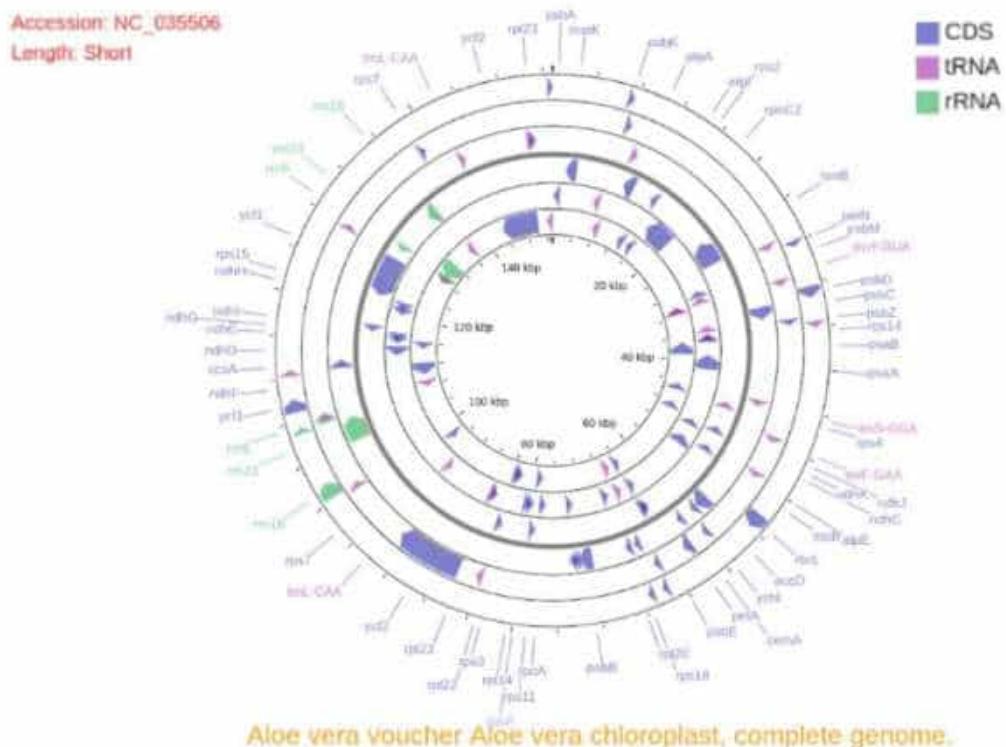
## **Advantage**

Chloroplasts can accumulate transgenic proteins in large amounts up to 46% of the total protein in leaves followed by stable transgene expression and integration (De Cosa et al., 2001). This is mainly due to the polyploid nature of the chloroplast genome where up to 100 copy genomes can exist in one chloroplast. A photosynthetically active plant cell can contain up to 100 chloroplasts, amounting to about 10,000 copies of the genome per functional cell (Verma and Daniell, 2007).

In addition, the transgene inside the chloroplast genome occurs at a specific site via two homologous recombination events. In contrast to the nuclear transformation, chloroplast transformation has a high number of copies because there is no silencing gene mechanism. Another advantage of the chloroplast-based expression system is the capacity of the polycistronic expression that causes the vaccine or multivalent antibody to be expressed from a single mRNA transcript (Quesada-Vargas et al., 2005). Post-translational modifications such as the formation of disulfide bonds (Staub et al., 2000; Ruhlman et al., 2007), lipid modifications (Glenz et al., 2006) also occur in chloroplasts resulting correct folding and of chloroplast protein expression.

## **Chloroplast Genome Organization: Concept of Chloroplast Transformation**

The higher-order plant chloroplast genome consists of double-stranded DNA base units of about 120 to 220kb and can be found in several different configurations such as multimeric, monomeric or linear forms (Palmer, 1985; Lilly et al., 2001).



(Figure 2.6.2. Visualization of the Aloe vera chloroplast genome with CGView)

Foreign DNA can be introduced into the chloroplast by delivery via particle bombardment (Daniell et al., 1990; Sanford et al., 1993) or via PEG permeation (Golds et al., 1993; O'Neill et al., 1993). The foreign DNA is then integrated into the chloroplast genome via two site-directed homologous recombinations between the homologous flanking sequences in the transformation vector and the endogenous sequences in the genome. This process is facilitated by the same RecA-type mechanism as the *E. coli* DNA repair mechanism (Cerutti et al., 1992). Therefore, flanking homologous recombination in the chloroplast transformation vector must be designed specifically according to the genome sequence selected as transgene insertion (image above the design). 1kb sequences are often chosen to be more specific and prevent the incorporation of transgenes at non-specific sites (Verma and Daniell, 2007).

Transformation is achieved when the transgene is successfully inserted into multiple copies of the chloroplast genome, followed by 25 to 30 cycles of cell division under the pressure of antibiotic selection to remove the unformed chloroplasts. As a result, the chloroplast population becomes

homogeneous with the transformed genome achieved, referred to as the homoplastic population.

In the initial chloroplast transformation experiment, the transgene was targeted to the spacer region transcriptionally silent to prevent any possibility of endogenous chloroplast gene disruption (Svab and Maliga, 1993). Although transgene expression succeeded in achieving integration of the transgene into the active spacer region the transcription proved to be more effective and offers several unique advantages such as reducing

reliance on untranslated regulatory elements (UTR) 5' and 3' or promoters for the regulation of transgene expression (Verma and Daniell, 2007). So far, the active transcription spacer region between the trnI and trnA genes within the rrn operon located at either end of the IR region closest to the SSC of the chloroplast genome is the most commonly used site for integration of the transgene. The transgene integrated and expressed from this site has recorded the highest expression level compared to other integration sites in the tobacco chloroplast genome (De Cosa et al., 2001). Thus, based on the results of a literature study it can be concluded that the integration of the transgene into the spacer region of the active transcription region may prove to be very useful for high levels of transgene expression.

### **Regulation of Chloroplast Gene Expression**

Fundamental aspects of chloroplast gene expression including gene transcription, post-transcriptional mRNA processing, mRNA stability modulation, and mRNA translation have been extensively reviewed by Gillham et al., 1994; Mayfield et al., 1995; Rochaix, 1996; Sugita and Sugiura, 1996; Barkan and Goldschmidt-Clermont, 2000; Monde et al., 2000; Zerges, 2000.

## **2.7 Transformation Method**

All transformation methods primarily aim at achieving one common goal, namely, the transfer of deoxyribonucleic acid (DNA) molecules into the target cells.

## 2.7.1 Indirect Transformation Method

### *Agrobacterium tumefaciens*-mediated transformation

Frequently used indirect transformation methods involve the use of *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes* for the transfer of transgene in plasmid form into the plant nuclear genome.

*Agrobacterium tumefaciens* and *Agrobacterium rhizogenes* belong to the Rhizobiaceae family. Rhizobiaceae is a family inhabited by various nitrogen-fixing bacteria that have mutualism symbiosis with peanuts. In contrast to other members of the Rhizobiaceae family (Wikipedia). *A. tumefaciens* is a type of phytopathogenic bacteria that can cause crown gall disease. Crown gall disease is caused by bacterial Ti-plasmid that carries T-DNA to plants via the bacterial type IV secretion system (T4SS).

The size of the plasmid used for *Agrobacterium* transformation is generally between 5 and 12kb including the plasmid backbone which is essential for plasmid replication. Plasmids for *Agrobacterium* transformation are mainly based on tumour-inducing (Ti) plasmids which cause tumour formation and opine synthesis via the transfer of genes involved in auxin and cytokinin synthesis and nopaline synthase into the core genome of infected plants (Gelvin, 1990).

*Agrobacterium*-mediated gene transformation of gene delivery is usually associated with low copy counts and stable integration of the transgene into the host nucleus genome (Newell, 2000). Common disadvantages include low transformation efficiency, random integration of transgene, species dependence and strict optimization requirements for transformation protocols.

## 2.7.2 Direct Transformation Method

### Particle Bombardment-mediated transformation

In genetic engineering, a gene gun or biolistic particle bombardment delivery system is a tool used to deliver exogenous DNA (transgene), RNA, or protein to cells. By coating heavy metal particles with the desired genes and firing these micro-projectiles into the cells using mechanical force, integration of the desired genetic information can be induced into the cells.

The techniques involved with sending micro-projectiles of DNA are often referred to as biolistics<sup>[177]</sup>.

Particle bombardment, also known as gene gun, was developed to allow penetration of the cell wall so that genetic material containing the desired gene can be transferred into the cell. This system can also transform almost all types of cells and is not only limited to the nucleus, but also for the transformation of organelles, including plastids (chloroplasts) and mitochondria.

Particle bombardment or biolistics is the method most widely used in direct plant transformation. DNA material can be DNA, in the form of plasmids or in viral or bacterial vectors that are attached to an inert micron to sub microcarriers of micron size from gold or tungsten and bombarded into plant tissues (Stanford, 1988; Klein et al., 1988).

The particle bombardment method begins by coating the tungsten or gold particles (micro-projections) with plasmid DNA. The coated particles are coated on a macro projectile, which is accelerated by air pressure and fired into the plant tissue on a petri dish (Gan, 1989), as shown in Figure 2.7.1. Perforated plates are used to stop macro projectiles while allowing micro-projections to pass through the cells on the other side. When the micro-projections enter the cell, the transgenes are released from the surface of the particles and can join into the chromosomal DNA of the cell. Selectable markers are used to identify cells that take the transgene. The transformed plant cells are then regenerated into whole plants using tissue culture<sup>[181]</sup>.

### **Aplication**

Teknik particle Bombardment kebanyakan digunakan pada tanaman. Namun, juga dapat diaplikasikan pada manusia dan hewan. Target dari gene gun biasanya berupa kalus dari sel tumbuhan. Setelah partikel emas berlapis DNA dikirim ke sel, DNA digunakan sebagai template untuk transkripsi (ekspressi transien) dan terkadang terintegrasi ke dalam kromosom tumbuhan (transformasi 'stabil')<sup>[178]</sup>.

Partikel logam yang sudah ditempeli gen target dimasukkan ke dalam senjata gen (Gene Gun) dan kemudian ditembakkan pada sel tanaman. Partikel akan mempenetrasi dinding sel tanaman sehingga masuk ke dalam sitoplasma. Partikel akan tercuci oleh cairan sel sehingga gen target yang menempel akan lepas dan kemudian masuk ke dalam nukleus dan menyisip pada kromosom sel tanaman<sup>[178]</sup>.

### **Advantage**

Biolistics has proven to be a versatile and widely used method of genetic modification in plant genetic engineering<sup>[179]</sup>. Plastid transformation mediated particle bombardment has had great success compared to other techniques, such as Agrobacterium-mediated transformation, which has difficulty targeting vectors and expressing stably in chloroplasts<sup>[178,179]</sup>. In addition, chloroplasts did not silence the transgenes introduced with a gene gun<sup>[181]</sup>. And, with just one shot, a skilled technician can produce two transgenic organisms<sup>[180]</sup>.

The advantages of particle bombardment include species independence, the ability to regenerate stable transgenic lines, rapid transformation and possible transformation in the nuclear and plastid genomes. The disadvantages of this method include the tendency to produce a high number of copies, the possibility of damage to cells, and the relatively high cost compared to Agrobacterium-mediated transformation.

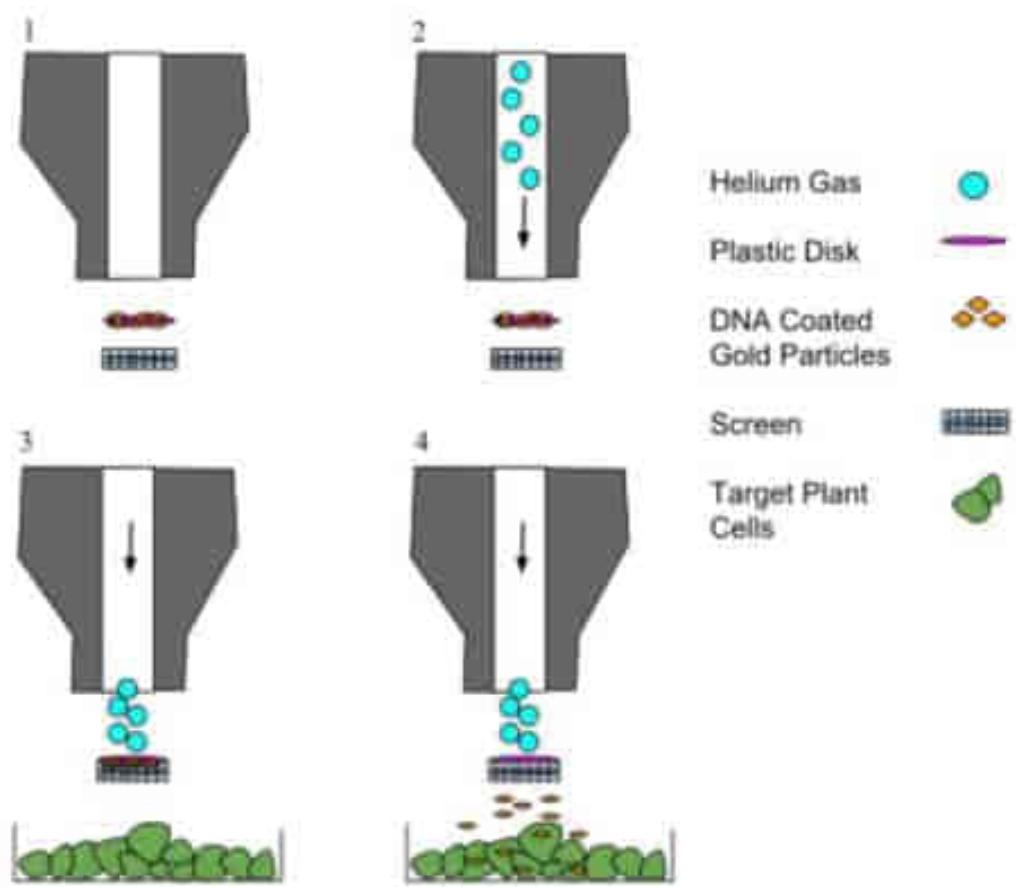


Figure 2.7.1 Gene gun. (Source: AP\_Biology\_Bapst\_2015)

## 2.8 Biolistic Construction Design

Biolistic transformation involves the integration of functional DNA fragments known as DNA constructs into the target cell. A gene construct is a DNA cassette containing all the regulatory elements necessary for proper expression in the target organism<sup>[179]</sup>. While the gene construct can vary in design depending on the desired outcome of the transformation procedure, all constructs usually contain a combination of the promoter sequence, the terminator sequence, the desired gene, and the reporter gene.

### 2.8.1 Promotor

The promoter controls the location and magnitude of gene expression and functions as the "pedal wheel and accelerator" of a gene<sup>[179]</sup>. Promoters are the main factors affecting the level and stability of transgenic expression (Curtis et

al. 1994). The promoter precedes the desired gene in DNA construction and can be altered through laboratory designs to enhance transgene expression.

The promoter often used in dicot plants is CaMV 35S from Cauliflower mosaic virus (Ma et al., 2003), a strong constitutive promoter that can be enhanced by duplicating enhancers (Kay, Chan, Daly, & McPherson, 1987).

In this study we used 2 promoters:

- **psbA Promoter (as a promoter for gene of interest insertion into chloroplasts)**

The chloroplast psbA promoter is the strongest promoter for foreign gene expression in chloroplasts (Daniell and McFadden, 1987; Daniell, 1993). The expression of foreign genes is usually checked in *E. coli* before continuing chloroplast transformation because the transcription and translation machines in *E. coli* are similar to plant chloroplasts (Daniell, 1993).

The psbA promoter is effective in the transcription of foreign genes in *E. coli* and does not require IPTG induction. It has been demonstrated that the psbA bean promoter contains sequences 10 and 35 bp upstream of the transcription initiation site homologous to the prokaryotic promoter sequence and that the transcripts produced *in vivo*, *in vitro*, and *E. coli* are similar (Boyer and Mullet, 1986). This promoter also contains the eukaryotic CAAT and TATA consensus box sequences required for the initiation of transcription in the eukaryotic nucleus (Cornelissen and Vandewiele, 1989).

Most of the plastid promoters recognized by RNA polymerase (PEP) are similar to bacteria similar to the *E. coli*  $\sigma$ 70 type promoter which consists of the elements “-35” and “-10”. Among them, the psbA promoter is unique in bringing additional elements between the conserved -35 and -10 elements. PsbA promoter activity is maintained differently in mature chloroplasts where the activity of most PEP promoters is decreased.

There have been studies of two types of PEP activities on wheat germ [Satoh et al. (1999) J. Plants 18: 407]; The PEP present in mature chloroplasts from the tip of the leaf (tip-type PEP) can initiate transcription

from the defective psbA promoter -35, but the -35 element is essential for transcription by PEP present in immature chloroplasts from the base of the leaf (base-type PEP). The promoter element -35 is important for the initiation of transcription in the psbA promoter in all types of plastids, including chloroplasts in mature leaves, leucoplasts in roots, etioplasts in etiolated cotyledons<sup>[149]</sup>.

- **CamV 35s promoter (As a selectable marker promoter)**

The 35S promoter, which is derived from the cauliflower mosaic virus (CaMV), is a component of the transgenic construction of more than 80% of genetically modified (GM) plants. The 35S RNA promoter is a very powerful constitutive promoter responsible for the transcription of the entire CaMV genome. It is well known for its use in plant transformation. Recent studies have shown that the CaMV 35S promoter also functions in some animal cells, although the promoter elements used are different from those in plants.

Structure :

CaMV contains a circular double-stranded DNA molecule of approximately 8.0 kilobases, interrupted by the notches resulting from the action of RNase H during reverse transcription. These incisions come from Met-tRNA, and the two RNA primers used in reverse transcription. After entering the host cell, these single-strand "nicks" in the viral DNA are repaired, forming supercoiled molecules that bind to the histones.

## 2.8.2 Polyadenylation

Polyadenylation is a post-transcriptional modification of Ribonucleic Acid (RNA) which is found in all cells and in organelles<sup>[193]</sup>. Polyadenylation from the 3' ends occurs before messenger-RNA (mRNA) leaves the nucleus. This polyadenylate tail, about 100-200 nucleotides in length, protects mRNA from the degradation action of phosphatases and nucleases<sup>[194]</sup>. Later the pre-mRNA is cleaved at the polyadenylation site (A) and a polyadenylation tail (A) is added, which is a necessary process for the formation of normal mRNA. Genes with multiple poly (A) sites can undergo alternative polyadenylation (APA),

producing different mRNA isoforms with 3' untranslated regions (3' UTR) and also in some cases different coding regions<sup>[195]</sup>.

### 2.8.3 Terminator

The terminator sequence is required for proper gene expression and is placed after the region of the gene coding of interest in DNA construction. The terminator commonly used in biolistic transformation is the NOS terminator which is derived from *Agrobacterium tumefaciens*.

This method has the advantage of ease and speed in constructing constructs because the desired gene can be inserted into a binary vector that already contains the promoter and the inverted nos domain in one cloning step, and does not require any knowledge of DNA sequences<sup>[97]</sup>.

### 2.8.4 Marker gene

In nuclear biology and molecular biology, marker genes are genes used to determine whether a nucleic acid sequence has been successfully inserted into an organism's DNA. In particular, there are two subtypes of this marker gene: selectable marker and screenable marker.

#### *Selectable Marker*

Selectable markers allow the selection of cells that have transformed. Generally, these markers provide resistance to phototoxic compounds such as antibiotics and herbicides. Is a dominant gene that is stable and is an integral part of the transformation vector. The criterion is a positive selection for a resistant gene without which cells that are not transformed will die when the selection agent is applied. Ideally, the selection agent should not affect the transformed cells and should be efficient even at lower concentrations. Examples of selectable markers are *painter*, *nptII* (resistant to kanamycin), *hpt* (resistant to hygromycin), *bar*, *manA*, and others<sup>[98]</sup>.

#### **Hygromycin phosphotransferase (HPT) (*hpt*)**

*hpt* (or *aph IV*) is a gene from *Escherichia coli* that encodes the HPT enzyme (E.C. 2.7.1.119) and is resistant to the antibiotic hygromycin B (Waldron et al., 1985). When hygromycin occupies the ribosomal binding site of elongation factor 2 (EF-2) in prokaryotic cells, as a result,

polypeptide chain elongation is inhibited and protein synthesis is interrupted, causing the same symptoms described for other aminoglycoside antibiotics.

In plant cells, this antibiotic exerts its effects on mitochondria and chloroplasts, working in the same way by impairing protein synthesis. These organelles have ribosomes similar to those found in bacteria and are also susceptible to aminoglycoside antibiotics. Therefore, in the presence of antibiotics, plant tissue will show chlorosis, which is caused by a lack of chlorophyll synthesis and stunted growth.

## 2.9 Choosing a Model Plant

Table 2.9.1 Comparison of several model plants (adapted from Spok and Karner 2008, European Communities and other sources).

Features/Crop	Organ	Yield	Storage-Protein stability	Transformation	Production Costs	Specialty
Tobacco	Leaf	High	Limited	Well established	Good	Nonfood/feed
Arabidopsis	Leaf	High	Limited	Established	Good	Homogenous N glycosylation; uses atmospheric N <sub>2</sub>
Wheat	Seed	Good	Optimal	Inefficient	Optimal	
Maize	Seed	High	Optimal	Established	Optimal	
Pea	Seed	Good	Optimal	Limited	Good	
Rapeseed	Seed	Good	Optimal	Established	Optimal	Fusion with oleosin for easy purification
Potato	Tuber	Good	Good	Well established	Good	
Banana	Fruit	Good	Good	Inefficient	Good	Can be eaten raw
Aloe vera	Leaves	High	Good	Efficient	Optimal	As a food, medicine, cosmetic plant

### 2.9.1 Aloe vera

*Aloe vera* is a succulent plant native to the Arabian Peninsula. Currently, *Aloe vera* has spread to various parts of the world. According to Dowling (1985), only 3 types of *Aloe vera* are cultivated commercially in the world, namely: Curacao aloe (*Aloe barbadensis Miller*), Cape aloe (*Aloe ferox Miller*), and Socotrine (*Aloe perryi baker*).

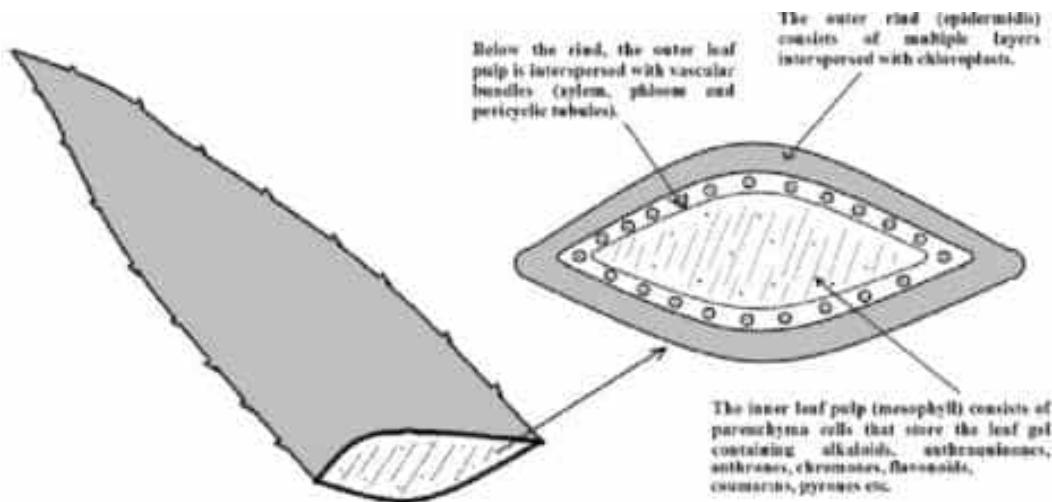
The type most widely cultivated is *Aloe barbadensis Miller*, while in Indonesia the most widely developed species is *Aloe chinensis Baker*. This species was cultivated in Kalimantan and is better known as the Pontianak Aloe Vera. The characteristics of this plant are orange flowers, the midrib is

light green, the upper midrib is slightly concave and has white spots on the midrib when the plant is young<sup>[46]</sup>.

### ***Aloe vera* Physiology**

#### **Metabolism**

From a physiological perspective, *Aloe vera* belongs to the CAM (Crassulacean Acid Metabolism) group. CAM plants only open their stomata at night to absorb CO<sub>2</sub> and water, in contrast to plants which open their stomata during the day and night. This has an advantage for the CAM plant, as it can tightly seal its stomata during the day and reduce water evaporation. Therefore, CAM plants can live in dry areas. Examples of CAM plants other than *Aloe vera* are cactus and pineapple.



(Gambar 2.9.1 Skema daun *Aloe vera* (Cock 2015))

#### **Transformation and Genetic Engineering**

Since the end of the 20th century, *Aloe vera* began to receive attention for genetic engineering. This is due to the various benefits that *Aloe vera* offers. *Aloe vera* components are reported to have antifungal, antiseptic, antiviral, antibacterial, anti-inflammatory, antioxidant, and wound healing properties<sup>[7]</sup>. However, there are some obstacles to the culture of transgenic *Aloe vera*. These obstacles include:

1. browning and oxidative bursts, especially in liquid medium cultures, often cause the death of *Aloe* explants (Natali et al, 1990);

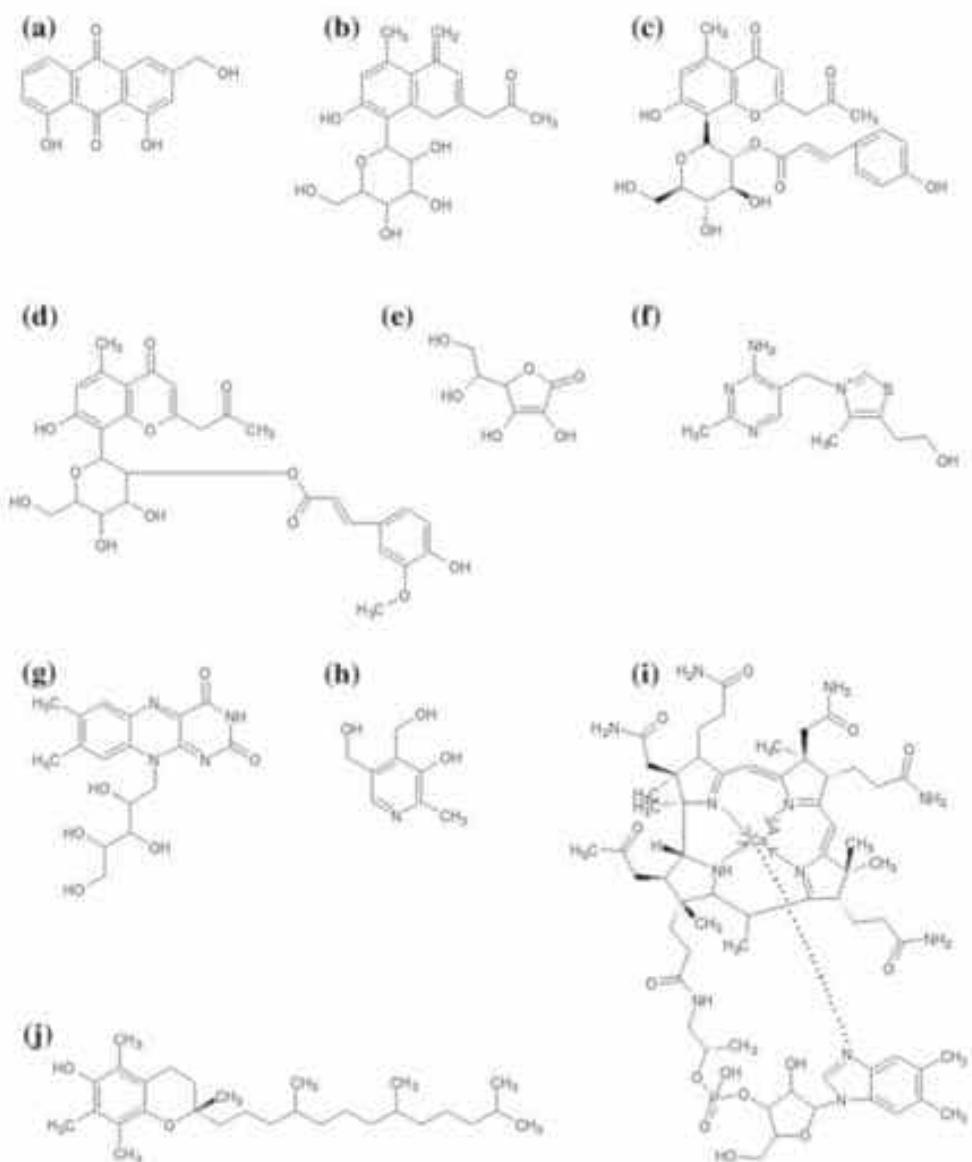
2. cytokinins, which are commonly used to promote regeneration, have been reported to cause increased phenol secretion in *Aloe* cells (Natali et al. 1990; Meyer and van Staden 1991; Roy and Sarkar 1991);
3. *Agrobacterium* which is an additional factor in the occurrence of necrosis on explants after cultivation.
4. Excretion of polyphenols has previously been reported to limit plant transformation in other species (Lagrimini 1992; Perl et al. 1996; Dan 2008)<sup>[63]</sup>.

Pretreatment with antioxidants has shown a reduction in oxidative bursts and necrosis of explants (Enriquez-Obregon et al. 1999; Dan 2008). In addition, dithiothreitol (DTT) was reported to be more effective than other types of antioxidants (Apostol et al. 1989)<sup>[63]</sup>.

### Nutritional Content

Various kinds of nutritional content in Aloe vera leaves include<sup>[46]</sup>:

1. Vitamins, namely A, B1, B2, B3, B12, C, E, Choline, Inositol, Folic Acid
2. Minerals, namely Calcium, Magnesium, Potassium, Sodium, Iron, Zinc, Chromium
3. Enzymes, namely Amylase, Catalase, Cellulose, Carboxy Epilase, Carboxy Cellulose, Bradykinase
4. Amino Acids, namely Arginine, Asparagine, Aspartic Acid, Alanine, Serine, Glutamate, Threonine, Glycine, Phenylalanine, Histidine, Isoleucine



(Figure 2.9.2 The chemical structure of molecules with antioxidants identified in *Aloe* sp. Leaf gel: a) Aloe-emodin, b) Aloesin, c) 2'-Op-Coumaroyl Aloesin, d) 2'-O-Feruloyl Aloesin, e) Acid ascorbate (vitamin C), f) Vitamin B1 (thiamine), g) Vitamin B2 (riboflavin), h) Vitamin B6 (pyridoxal phosphate), i) Vitamin B12, j)  $\alpha$ -Tocopherol (vitamin E) (Cock 2015). )

### Distribution of *Aloe vera*

*Aloe vera* is native to the Arabian Peninsula, however, since the 17th century, *Aloe vera* began to be cultivated in China and Europe. Now *Aloe vera* is widely distributed in tropical, subtropical, and dry regions in the Americas, Asia and Australia (Wikipedia).



(Figure 2.9.3 Distribution map of Aloe vera [www.discoverlife.org](http://www.discoverlife.org))

### **Benefits of *Aloe vera***

#### **Traditional Medicine**

*Aloe vera* has been known as a medicinal plant for decades. The earliest historical records of the use of aloe are in the Ebers Papyrus from 16th century BC Egypt<sup>[48]</sup>. 1st century AD, its use is recorded in the De Materia Medica by the Greek physician Pedanius Dioscorides, and the Naturalis Historia by the Roman writer Pliny the Elder<sup>[48]</sup>. In the 6th century AD Byzantium, the use of this plant is recorded in the Juliana Anicia Codex<sup>[47]</sup>. In Ayurvedic medicine, this plant is called kadhala (same as the agave plant)<sup>[49]</sup>.

#### **Health Products**

There are two substances from *Aloe vera* that are often used, namely gel (clear) and latex (yellow). *Aloe vera* gel is used as a topical medicine for various skin symptoms, such as burns, wounds, inflammation, frostbite, psoriasis, herpes labialis, or too dry skin. *Aloe vera* latex is made into a product (either alone or in combination with other ingredients) for ingested drugs to cure constipation<sup>[50,51]</sup>.

#### **Cosmetic Products**

As stated above, *Aloe vera* has many benefits for the skin. Therefore, many studies have been conducted to find out the properties of *Aloe vera*

and produce it. In addition, there have been several studies regarding transgenic *Aloe vera* for the production of protein proteins which can be used for cosmetic purposes<sup>[52]</sup>.

### **Cultivation of *Aloe vera***

*Aloe vera* has been cultivated in Indonesia since 1980, especially in the Siantan Hulu area, Pontianak. The type of aloe that is cultivated is *Aloe chinensis Baker* or better known as Pontianak *Aloe Vera*. The involvement of the government and medium-sized industry in the cultivation of *Aloe vera* has made it an “icon” of Pontianak<sup>[46]</sup>. The data listed below comes from the website of the Pontianak Food, Agriculture and Fisheries Department and SOP for *Aloe vera* planting, the directorate of vegetables and medicinal plants, the directorate general of horticulture, the ministry of agriculture 2019

### **Planting Land**

The ideal planting area for *Aloe vera* meets the following requirements: Altitude 0-10 m asl; Rainfall 2500-4000 mm / year; Air temperature 24 - 36 ° C; Soil pH 3.5-5.5; Podzolic soil type, organosol with good drainage; Lowlands with good sunshine; The soil is fertile, loose, and contains organic matter.

### **Seed, Planting, and Maintenance**

The seeds used are 25-30 cm high, with 3-4 fronds. The seeds are planted in rows 70-80 cm, with a distance of 100-150 cm between rows. Pre-planting fertilization is carried out 3-4 days before planting, then continued when the plants are 1-2 months old, and further fertilization can be done every 2-3 months.

### **Harvesting**

Harvesting can be done after the plants are 8-12 months old, or after having 15 fronds. Harvesting can be done until the plants are 2-3 years old, after which rejuvenation is carried out to extend the life of the plants.

## Yields

No.	Tahun	Luas	Luas Panen	Produktivitas	Produksi	Keterangan
		Tanam (Ha)	(Ha)	(Kg/Ha)	(Ton)	
1.	2007	70	20	51	2.458	1 Tahun = 24 X Panen/Ha
2.	2008	67	45	51	5.530	Produktivitas 1 X Panen
3.	2009	44	43	51	5.284	
4.	2010	37	37	51	4.546	
5.	2011	46	46	51	5.652	
6.	2012	79	46	58	6.359	
7.	2013	84	84	40	7.879	

(Table 2.5.1 Data on *Aloe vera* yields in Pontianak,  
<https://pertanian.pontianakkota.go.id/produk-unggulan-detil/4-lidah-buaya.html>)

### 2.9.2 *Aloe vera* Excellence

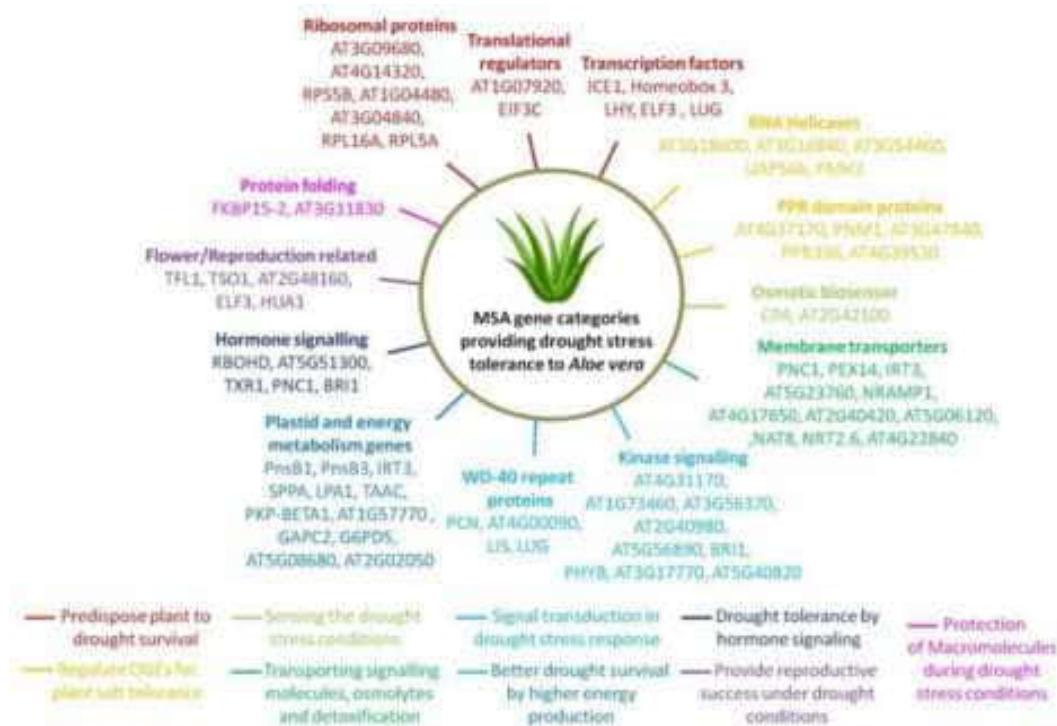
- Offers ease of extraction as the desired protein is localized in the central gel.
- High production
- Easy to cultivate in tropical climates like Indonesia

### 2.10 *Aloe vera* Genetic Substance

*Aloe vera* is a succulent and drought-tolerant plant belonging to the genus family *Aloe Asphodelaceae*<sup>[81]</sup>. More than 400 species are known in the genus

*Aloe*, four of which have medicinal properties with *Aloe vera* being the most potent species<sup>[82]</sup>. *Aloe vera* is a tropical perennial plant with succulent and elongated leaves consisting of a transparent mucilaginous tissue of parenchyma cells in the centre called *Aloe vera* gel<sup>[83]</sup>. This herb is widely used as a herb in traditional practice in several countries, and in cosmetics and skincare products because of its pharmacological properties including anti-inflammatory, anti-tumour, anti-viral, anti-ulcer, fungicidal, etc.<sup>[84,85]</sup>. *Aloe vera* also contains chemicals such as anthraquinones, vitamins, minerals, enzymes, sterols, amino acids, salicylic acids, and carbohydrates<sup>[86,87]</sup>. This property makes it commercially important, with a global market valued at 1.6 billion<sup>[88]</sup>.

*Aloe vera* (L.) (*Aloe barbadensis Miller*) is a perennial succulent xerophyte, which develops a water-retaining network in leaves to survive dry environmental conditions. Several widely recognized genes for stress tolerance were isolated from the *Aloe vera* genome and shown to also express DREB1. HVA22 is a stress-induced gene isolated from the aleurone layer of *Hordeum vulgare*. HVA22 homologs have been found in 354 plant species. HVA22 expression in vegetative tissues can be induced by ABA and environmental stresses, such as cold and drought. *Aloe vera* is known to survive without water for a long time. This stress-causing gene from *Aloe vera* can also be isolated. *Aloe vera* plantlets given ABA were used to isolate the cDNA pool. HVA22 homologs were obtained using RT-PCR which was verified by sequencing and bioinformatics.



(Fig. 2.10.1 The MSA gene in *Aloe vera* which causes a response to drought stress<sup>[89]</sup>.)

Some research on transgenic *Aloe vera* and some achievements in the field of genetic modification.

Table 1. a) Several studies on transgenic *Aloe vera*. b). Some achievements on GMO *Aloe vera*.

a)

Plants	Target	Transformation method	Product	References
Aloe vera	Apoplast	Biolistic/gene gun	Active human interferon alpha 2	<a href="https://doi.org/10.1007/s11248-012-9616-0">https://doi.org/10.1007/s11248-012-9616-0</a>
Aloe vera	Chloroplast	Biolistic/gene gun	Polisakarida (karbohidrat) pada AP-2 dan AP-3	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6515206/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6515206/</a>
Aloe vera	Chloroplast	Biolistic or infection with Agrobacterium t.	Some transgenic protein	<a href="https://patents.google.com/patent/US8008546B2/en">https://patents.google.com/patent/US8008546B2/en</a> Paten no : US8.008.546 B2

Aloe vera	Chloroplast	Agrobacterium t.	Introduced TaDREB gene isolated from wheat into Aloe by Agrobacterium.	Zhao et al. (2009)
Aloe vera	Chloroplast	Agrobacterium	Transformed with otsA gene mediated by Agrobacterium t.	Chen et al. (2007)
Aloe vera	Chloroplast	Biostatic	To improve the resistance to cold stress of Aloe plant by introduced otsA gene into its genome via microprojectile bombardment	Chen et al. (2005)

b)

Plants	Achievements in Genetic Engineering	References
Aloe vera	Assessment of genetic stability and instability of tissue culture propagated plantlets	Rathore et al. (2011)
Aloe vera	Assessments of genetic fidelity of micropagated plants	Gantait et al. (2010)
Aloe vera	Assessments of genetic similarity of micropagated plantlets with their mother plant	Gantait et al. (2011)

### 2.10.1 Sekuensing dari genome *Aloe vera* dan transcriptome

The estimated genome size of *Aloe vera* is 16.04 Gbp with a ploidy diploid level containing 14 (2n) chromosomes [2229], and to cover this large genome size, a total of 506.4 Gbp (~ 32X) short reading data and 123.5 Gbp (~ 7.7X) long readout data were generated using the Illumina and nanopore platforms, respectively [90, 59]. For the transcriptome, the total RNA-seq data of 6.6 Gbp

and 7.3 Gbp were generated from the leaves and roots, respectively. The transcriptome data from this study and the available RNA-seq for general data from the previous study<sup>[91]</sup> were combined, yielding a total of 37.1 Gbp RNA-seq data for *Aloe vera*, which was used for analysis. All genomic and RNA-seq reading data were trimmed and filtered using Trimmomatic, and only high-quality data was used to construct the final genome and transcriptome set.

### 2.10.2 *Aloe vera* Chloroplast Genome

Many studies have shown that the chromosome number for somatic cells for most *Aloe vera* is  $2n = 14$  and that the haploid genome set consists of three short and four long chromosomes (Brandham and Doherty, 1998; Ji et al., 2002; Alam and Khanam, 2005. ).

Many studies show that most plants have the *Arabidopsis*-type telomeres consisting of multiple repeated copies of the 5'-TTTAGGG-3' sequence (Adams et al., 2000a), however, this telomere type is not found in *Aloe*, similar to *Allium*, *Nothoscordum* and *Tulbaghia*. *Aloe* lacks the recurrent telomeric type *Arabidopsis* but has a vertebrate-like telomeric sequence (T2AG3) and according to a report by Weiss and Scherthan (2002). Tetraploid and hexaploid *Aloe* can be induced from diploid *Aloe* (Ren et al., 2007). In a report by Wang et al. (2001), the highest induction rate increased to 50% after exposure to 0.06% colchicine for 12 hours.

The chloroplast genome has a typical quadripartite structure containing a large single-copy region (LSC) of 83.505bp, a small single-copy region (SSC) of 16.178bp and a pair of inverted repeating regions (IR) of 26,596 bp. The nucleotide composition of the chloroplast genome as a whole was: 47.185bp A (30.8%), 48.123bp T (31.5%), 29.326bp C (19.2%), 28.241bp G (18.5%) and G content.  $\beta$  C 37.7% total. Then, 131 genes were discovered which included 85 protein-coding genes (PCG), 38 transfer RNA (tRNA) and 8 ribosomal RNA (rRNA). Phylogenetic analysis showed that *A. vera* is closely related to *A. maculata* in the phylogenetic relationship of the Asphodelaceae family using the Maximum-Likelihood (ML) method<sup>[92]</sup>.

### **2.10.3 *Aloe vera* as a suitable candidate for this research**

*Aloe* offers an attractive protein production system that has the advantage of producing biologically active proteins that are associated with cosmetics. Proteins can be secreted into easily accessible pulp thereby increasing the efficiency of isolation and offering the potential for direct topical application of minimally processed pulp extracts. The original component of the pulp makes *Aloe vera* an excellent means of delivery which is well suited for topical application and absorption through the skin.

The transformation systems we describe can also be used to express plant protein to enhance or develop certain *Aloe vera* properties. *Aloe vera* is a hardy and easy-to-grow plant that can be processed up to four times a year, supplying 3-4 leaves weighing 1 kg each (12-16 leaves total per plant per year). This is equivalent to approximately 6–8 kg of pulp per year, which indicates that the potential for *Aloe vera* production is high.

The accessibility of secreted protein and inherent biomass makes *Aloe vera* an attractive production platform for molecular pharming. Furthermore, *Aloe vera* is not a traditional food crop and unlike some other plant models used for genetic modification, there is no inherent genetic risk from cross-pollination to the associated food crop.

This is one of the advantages of using *Aloe vera*. The leaves of the transgenic aloe plant can facilitate the extraction of carbohydrates, protein, and water contained in the gel, which mainly comes from the pith which is generally located in the middle of *Aloe vera*. Various compounds in the gel have been shown to treat various ailments. In addition, the gel/pith helps stabilize transgenic proteins produced by transgenic *Aloe vera* and localized in gels and/or pith<sup>[150]</sup>.

### **2.11 Restriction and Ligases Enzymes**

Restriction enzymes and ligases are two important components in plant transformation. The reason for this has been explained in chapter 2.7 Transfer DNA Binary System. However, to review, the restriction and ligation enzymes are tasked with cutting and reconnecting DNA fragments. This allows part or all of

the DNA to be replaced with other DNA, provided they both have the same restriction site.

### 2.11.1 Restriction Enzymes

Restriction enzymes, also known as restriction endonuclease or restrictases, are enzymes that cut DNA into fragments at specific recognition sites known as restriction sites<sup>[165-167]</sup>. Restriction enzymes are a class of a larger group of enzymes known as endonuclease enzymes. Restriction enzymes are divided into four types based on the structure and location of the cut. To cut DNA, the restriction enzymes perform two cuts, once on each double-stranded DNA sugar-phosphate backbone.

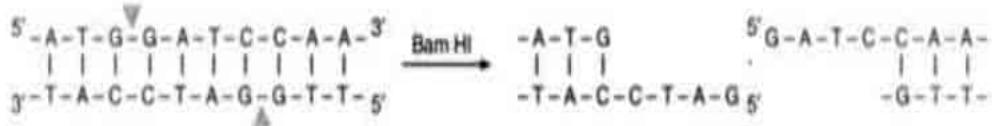
#### Restrictions Site

Restriction enzymes recognize specific nucleotide sequences<sup>[166]</sup> and produce double-stranded pieces in DNA. The recognition sequence can be classified according to the number of bases at the restriction site, usually between 4 and 8 bases. In addition, the number of bases on a restriction site also determines how likely it is that the same restricted site will reappear. For example, a sequence of 4 bp will theoretically appear every 44 bp, which is 256 bp, while a sequence of 6 bp will appear every 46 bp, which is 4,096 bp, and a sequence of 8 bp will appear every 48 bp, namely 65,536 bp.

Many of the sequences for the introduction of restriction enzymes are palindromes, ie, base sequences that can be read the same from the front or back<sup>[175]</sup>. In theory, there are two types of palindromes that can appear in DNA. The first type is a mirror-like palindrome, which is a sequence that reads the same from the front and back and is in a single strand. Meanwhile, the second is inverted repeat palindrome, which is a segment that reads the same from the front and back but is in the complementary double-strand DNA strand. For example, GTATAC is a complementary sequence to CATATG<sup>[176]</sup>. Of the two, inverted repeat palindrome is more common and has a greater biological effect,

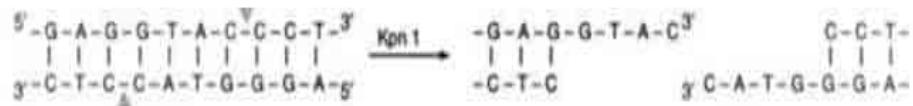
Based on the restriction site, there are 3 types of DNA cutting results, namely:

### 1. 5' sticky-end overhangs



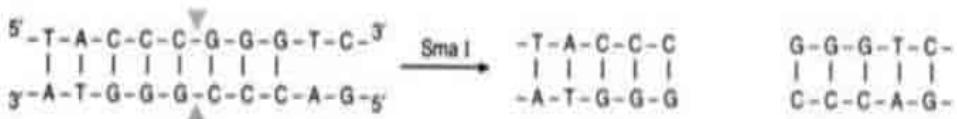
(Figure 2.11.1 5'overhangs as a result of the restriction enzyme Bam HI)

### 2. 3' sticky-end overhangs



(Figure 2.11.2 3'overhangs as a result of the Kpn 1 restriction enzyme)

### 3. blunt-ends



(Figure 2.11.3 Blunt-ends as a result of SMA 1 restriction enzymes)

## Types

There are 4 types of restriction enzymes that can be found naturally. The four types are distinguished by the location of their cutting.

### Type I

Type 1 restriction enzyme was the first restriction enzyme identified and identified in two strains (K-12 and B) of *E. coli*<sup>[177]</sup>. Enzymes of this type cut DNA at random locations with a minimum distance of 1000 bp from the restriction site. These cuts at random locations follow the translocation of DNA, which suggests that this enzyme is also a molecular motor. The restriction site is asymmetrical and consists of two specific parts. The first part contains 3-4 nucleotides, and the second

contains 4-5 nucleotides, both of which are separated by a non-specific spacer about 6-8 nucleotides in size.

This enzyme is multifunctional and can perform DNA cutting and modification, depending on the methylation status of the target DNA. The activity of this enzyme requires the cofactor S-Adenosylmethionine (AdoMet), hydrolyzed adenosine triphosphate (ATP), and magnesium ion ( $Mg^{2+}$ ). Type I restriction enzymes have three subunits, namely HsdR, HsdM, and HsdS; HsdR required for cutting; HsdM for adding a methyl group to the host DNA (methyltransferase), and HsdS for the introduction of the DNA attachment site<sup>[171,178]</sup>.

## Type II

Type II restriction enzymes differ from type I in several ways. Their differences are<sup>[175]</sup>:

1. Type II forms homodimers, with restriction sites that are usually joined and consist of 4–8 nucleotides;
2. They recognize and chop DNA at the same site, and only need  $Mg^{2+}$  as a cofactor.

This enzyme cuts double-stranded DNA phosphodiester bonds, it can cut in the middle of both strands, produce blunt-ends or at different positions and produce overhangs<sup>[179]</sup>. In the 1990s and early 2000s, new enzymes belonging to this family were discovered that did not match the general characteristics of this class of enzymes. Thus, it was decided to divide this class into several categories based on their general characteristics<sup>[175]</sup>.

The division of this class is<sup>[175]</sup>:

1. Type IIB (for example, Bcgl and Bpll) are multimers, containing more than one subunit. They cut DNA on both sides of the restriction site to remove the restriction site. They require the AdoMet and  $Mg^{2+}$  cofactors;
2. Type IIE (for example, Nael) truncates DNA after an interaction between two copies of the recognition sequence occurs. The first

- site is the cutting target, while the second site is an allosteric effector which accelerates cutting;
3. The IIF type (for example, NgoMIV) is the same as the IIE, but the two truncations occur simultaneously;
  4. Type IIG (for example, Eco57I) has 1 subunit but requires AdoMet to be active;
  5. Type IIM (eg, CpnI) can recognize and cut methylated DNA<sup>[175,180,181]</sup>;
  6. Type IIS (eg, fokI) cuts DNA at a predetermined distance from non-palindromic asymmetric restriction sites. This characteristic is widely used in in-vitro cloning techniques such as golden gate cloning, and this enzyme can also act as a dimer.
  7. The IIT type (for example, Bpu10I and BsII) consists of two distinct subunits. Some recognize palindromic sequences, while the rest recognize asymmetric sites.

### Type III

Type III restriction enzymes (for example, EcoP15) recognize two distinct non-palindromic sequences that are reversed oriented. They cut DNA about 20–30 bp after the restriction site<sup>[182]</sup>. This enzyme contains more than one subunit and requires the AdoMet and ATP cofactors<sup>[183]</sup>. They are components of the prokaryotic DNA restriction-modification mechanism that protects organisms from foreign DNA. Type III enzymes are multifunctional and hetero-oligomeric proteins consisting of two subunits, Res (P08764) and Mod (P08763). The Mod subunit recognizes a specific DNA sequence and is a methyltransferase modification. Whereas Res is required for cutting DNA, although it does not have its enzymatic activity.

Type III enzymes recognize short asymmetric DNA sequences measuring 5-6 bp, and chop 25-27 bp towards the 3' end and leave short, single 5' strand protrusions. They require the presence of two non methylated, reverse-oriented recognition sites for truncation to occur. This enzyme only methylates one strand of DNA, at the N-6 position of

the adenyl residue, so that the new DNA only has one methylated strand, this is enough to protect the DNA from cutting restrictions.

The type III enzyme belongs to the beta-subfamily N6 adenine methyltransferase and has 9 motives characterizing this family, including motif I, Adomet binding pocket (FXGXG), and motif IV, catalytic region (S / D / N (PP) Y / F )<sup>[184,185]</sup>.

#### Type IV

Type IV enzymes recognize modified DNA, usually methylated and exemplified by the McrBC and Mrr systems of *E. coli*.

Enzyme	Source	Recognition Sequence	Cut
<i>EcoRI</i>	<i>Escherichia coli</i>	5'GAATTC 3'CTTAAAG	5'---G AATTC---3' 3'---CTTAA G---5'
<i>BamHI</i>	<i>Bacillus amyloliquefaciens</i>	5'GGATCC 3'CCTAGG	5'---G GATCC---3' 3'---CCTAG G---5'
<i>HindIII</i>	<i>Haemophilus influenzae</i>	5'AAGCTT 3'TTCGAA	5'---A AGCTT---3' 3'---TTCGA A---5'
<i>TaqI</i>	<i>Thermus aquaticus</i>	5'TCGA 3'AGCT	5'---T CGA---3' 3'---AGC T---5'
<i>HinfI</i>	<i>Haemophilus influenzae</i>	5'GANTC 3'CTNAG	5'---G ANTC---3' 3'---CTNA G---5'
<i>Sau3A</i>	<i>Staphylococcus aureus</i>	5'GATC 3'CTAG	5'--- GATC---3' 3'---CTAG ---3'
<i>PovII</i>	<i>Proteus vulgaris</i>	5'CAGCTG 3'GTCGAC	5'---CAG CTG---3' 3'---GTC GAC---5'
<i>HaeIII</i>	<i>Haemophilus eggytius</i>	5'GGCC 3'CCGG	5'---GG CC---3' 3'---CC GG---5'
<i>AluI</i>	<i>Arthrobacter luteus</i>	5'AGCT 3'TCGA	5'---AG CT---3' 3'---TC GA---5'
<i>EcoRV</i>	<i>Escherichia coli</i>	5'GATATC 3'CTATAG	5'---GAT ATC---3' 3'---CTA TAG---5'
<i>SalI</i>	<i>Streptomyces albus</i>	5'GTCGAC 3'CAGCTG	5'---G TCGAC---3' 3'---CAGCT G---5'
<i>ScalI</i>	<i>Streptomyces caespitosus</i>	5'AGTACT 3'TCATGA	5'---AGT ACT---3' 3'---TCA TGA---5'

(Figure 2.11.4 Examples of restriction enzymes and their sources<sup>[10]</sup>.)

#### 2.11.2 Ligase Enzymes

The enzyme ligase, or DNA ligase, is an enzyme that facilitates the joining of DNA strands by catalyzing the formation of phosphodiester bonds. DNA

ligase plays a role in repairing single-strand breaks in DNA, but there are several types of ligase enzymes that can repair double-strand breaks (damage to the two complementary strands of DNA). DNA ligase corrects single-strand breaks using a complementary double helix strand as a template<sup>[186]</sup>. DNA ligase also plays a role in DNA replication and the formation of recombinant DNA.

### Enzymatic Mechanism

The mechanism of DNA ligase is to form two covalent phosphodiester bonds between the 3' hydroxyl end of one nucleotide (acceptor) and the other 5' phosphate nucleotide end (donor). Two ATP molecules are used for each formation of a phosphodiester bond, and AMP is needed for the ligase reaction, which can be divided into four stages<sup>[187]</sup>, namely:

1. Reorganization of activity sites such as Okazaki fragments in DNA;
2. Adenylation (addition of AMP) lysine residue to the active centre of the enzyme, and releasing pyrophosphate;
3. AMP transfers to the 5' phosphate and forms a pyrophosphate bond;
4. Formation of a phosphodiester bond between 3' hydroxyl and 5' phosphate.

### Types

#### *E. coli*

*E. coli* DNA ligase is encoded by the lig gene, and uses the energy obtained from cutting nicotinamide adenine dinucleotide (NAD) to form phosphodiester bonds, much like most DNA ligases in other prokaryotes<sup>[188]</sup>. It cannot ligate DNA blunt-ends except under molecular crowding conditions with polyethylene glycol, and cannot efficiently incorporate RNA to DNA.

The activity of *E. coli* DNA ligase can be enhanced by DNA polymerase at the appropriate concentration. The increase only occurs when the concentration of DNA polymerase 1 is lower than the DNA fragment to be ligated. When the DNA polymerase concentration was higher, it had a negative effect on *E. coli* DNA ligase<sup>[189]</sup>.

## T4

DNA ligase from the T4 bacteriophage, and is the DNA ligase most frequently used in the laboratory<sup>[190]</sup>. It can lyse both sticky-ends and blunt-ends of DNA, oligonucleotides, RNA, and RNA-DNA hybrids, but it does not gel single-chain nucleic acids. Unlike *E. coli*, T4 DNA ligase cannot use NAD and has an absolute requirement for ATP as a cofactor.

Several ways have been done to increase the in vitro activity of T4 DNA ligase, one of which is to test the fusion of T4 DNA ligase with other DNA binding proteins. The results of this experiment showed that the fusion construction of T4 DNA ligase with p50 or NF- $\kappa$ B was 160% more effective at blunt-ends ligation compared to wild T4 DNA ligase<sup>[191]</sup>. A typical reaction for inserting fragments into a plasmid vector uses 0.01 unit sticky-ends and 1 unit blunt-ends. The maximum incubation temperature for T4 DNA ligase is 16 ° C.

Mutant T4 DNA ligase has a higher sensitivity to UV irradiation<sup>[192,193]</sup> and methyl methanesulfonate than wild type<sup>[194]</sup>. This shows that DNA ligase plays a role in repairing DNA damage caused by this.

## Mammals

There are four specific types of DNA ligase in mammals, namely:

1. DNA ligase I: New DNA ligase on the lagging strand after Ribonuclease H removes RNA primers from the Okazaki fragments;
2. DNA ligase II: Used in repair, and is formed from alternative splicing of proteolytic fragments of DNA ligase III and does not have its genes. Therefore, it is considered to be virtually identical to DNA ligase III;
3. DNA ligase III: Complexes with DNA repair protein XRCC1 to help seal DNA in the nucleotide excision repair process and recombinant fragments. DNA ligase III is also the only mammalian DNA ligase found in mitochondria;

4. DNA ligase IV: Complex with XRCC4 that catalyzes the final stage of non-homologous end-joining DNA double-strand break repair pathway. It is also required for V (D) J recombination, a process that produces a variety of immunoglobulin loci and T-cell receptors during immune system development.

DNA ligase pada eukariota dan beberapa mikroorganisme menggunakan ATP daripada NAD<sup>[188]</sup>.

### **Thermostable**

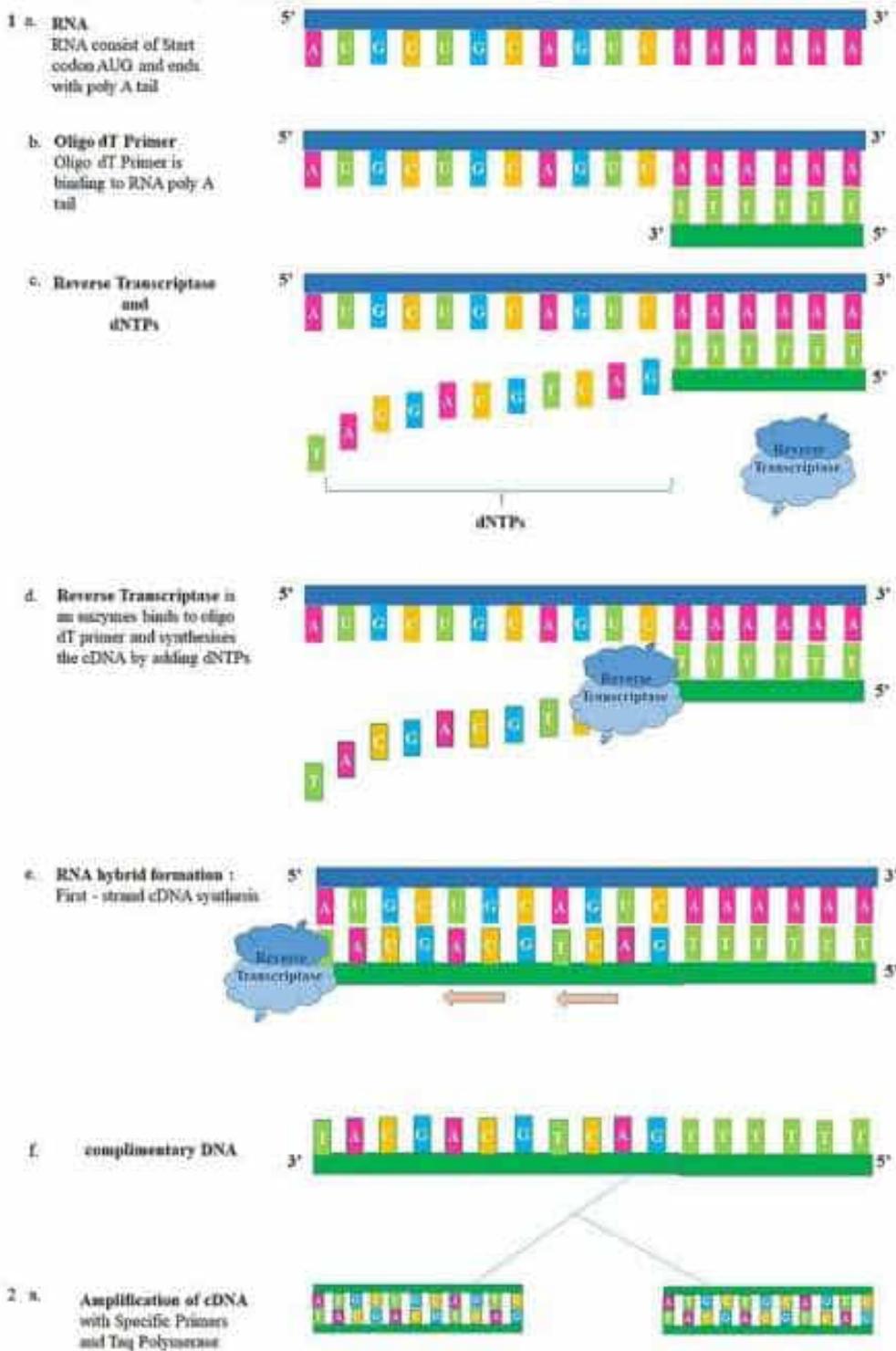
Obtained from thermophilic bacteria, and is an enzyme that is stable at much higher temperatures than conventional DNA ligase. It has a half-life of 48 hours at 65 ° C and more than 1 hour at 95 ° C. DNA ligase amplification has been shown to be active for 500 thermal cycles (94 ° C / 80 ° C) or 16-hour cycles<sup>[195]</sup>. This excellent thermostability allows for very high hybridization tightness and ligation specificity<sup>[195]</sup>.

## **2.12 RT PCR (Reverse Transcription Polymerase Chain Reaction)**

Reverse transcription polymerase chain reaction (RT-PCR) is a laboratory technique that combines reverse transcription of RNA into DNA (in this context it is called complementary DNA or cDNA) and the amplification of specific DNA targets using polymerase chain reaction (PCR)<sup>[145]</sup>. It is mainly used to measure specific amounts of RNA. This is achieved by monitoring the amplification reaction using fluorescence, a technique called real-time PCR or quantitative PCR (qPCR).

#### 4.8 Reverse transcription polymerase chain reaction (RT-PCR)

In RT-PCR, the RNA population is converted to cDNA by reverse transcription (RT), and then the cDNA is amplified by the polymerase chain reaction. The cDNA amplification step provides opportunities to further study the original RNA species, even when they are limited in amount or expressed in low abundance. Common applications of RT-PCR include detection of expressed genes, examination of transcript variants, and generation of cDNA templates for cloning and sequencing.



(Figure 2.12.1 RT-PCR) (source: Lokesh Thimmana, under the guidance of Dr. G. Mallikarjuna, Assistant Professor, Molecular Biology, Agri Biotech Foundation.)

### **Work Principle:**

In RT-PCR, the RNA template is first converted into complement DNA (cDNA) using reverse transcriptase. The cDNA is then used as a template for exponential amplification using PCR. QT-NASBA is currently the most sensitive method for detecting RNA available<sup>[146]</sup>. The use of RT-PCR to detect RNA transcripts has revolutionized gene expression studies in the following important ways:

1. Theoretically, it is possible to detect transcripts of almost all genes<sup>[147]</sup>;
2. Enables sample amplification and eliminates the need for abundant starting materials required when using northern blot analysis<sup>[18,19]</sup>;
3. Provides tolerance for RNA degradation as long as the RNA covering the primer is intact<sup>[148]</sup>.

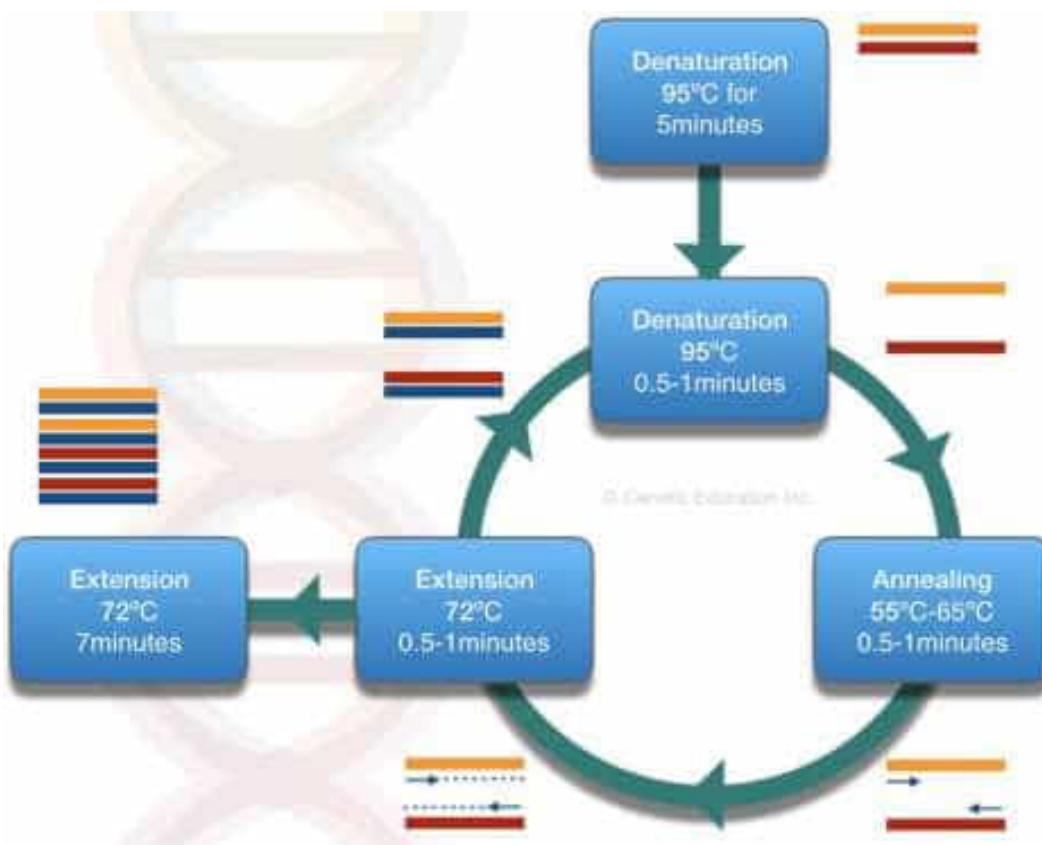
### **Application:**

Exponential amplification via reverse transcription polymerase chain reaction provides a very sensitive technique in which very low copy number of RNA molecules can be detected. RT-PCR is widely used in the diagnosis of genetic diseases and, semi-quantitatively, in determining the abundance of different specific RNA molecules in cells or tissues as a measure of gene expression.

## **2.13 PCR**

Polymerase Chain Reaction (PCR) is an in vitro nucleic acid amplification technique (Guyer and Koshland, 1989). PCR allows researchers to amplify specific DNA sequences exponentially into thousands or even millions of copies. The PCR technique was invented by American researcher Kary Mullis in 1984.

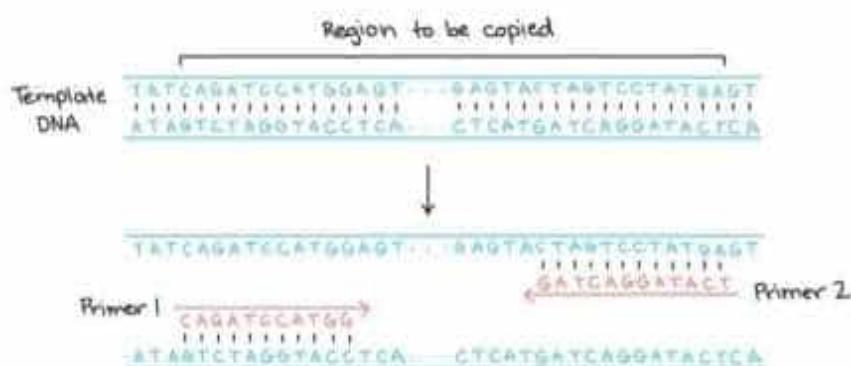
The majority of PCR methods rely on thermal cycling. Thermal cycling exposes the reactants to repeated heating and cooling cycles. This allows for a temperature-dependent reaction to occur. PCR consists of 2 main components, primers (short DNA fragments called oligonucleotides and complementary to the target DNA site), and DNA Polymerase.



(Figure 2.13.1 Polymerase Chain Reaction)

### 2.13.1 Primer

A primer is a short sequence that complements the sequence to be amplified. In PCR, two primary types are used, forward and reverse primary.



(Figure 2.13.2 Primary forward and reverse)

In general, PCR primers are DNA. DNA has several significant advantages over RNA for PCR. Some of these advantages are:

1. DN is more stable than RNA and does not degrade at high temperatures.
2. The DNA polymeration process is unidirectional to PCR so that there is no possibility of a loss of short RNA primers after polymerization is complete.
3. DNA polymerase 1 helps to remove short RNA primers on in vivo replication that are not present in PCR.

Each enzyme requires a cofactor and substrate to carry out the reaction, therefore, Taq DNA polymerase requires a free 3'OH end to initiate polymerization. The primer provides a free 3'OH end for the polymerase and serves as a substrate for the enzyme to work<sup>[38]</sup>.

### **Melting Temperature**

The temperature at which the primary template half split was influenced by the GC and AT DNA content. The reason is that primers attach to parts of the DNA which have the same melting temperature [38].

The formula for calculating the melting temperature is as follows:

$$\mathbf{Tm=4(G+C)+2(A+T)}$$

### **Annealing Temperature**

The temperature that allows the primer to bind to its complement template DNA. It is about 5 ° C lower than the melting temperature, and ranges from 56 ° C to 65 ° C. Temperatures outside this range can inhibit amplification.

If the temperature is too low, the primer can stick anywhere, resulting in an unspecified adhesion. On the other hand, if the temperature is too high, the primer cannot stick to the template DNA<sup>[38]</sup>.

### **Primer Length**

Primers between 18-23 bp in length usually give the best results in PCR. Primers shorter than 18 bp may result in nonspecific adhesion. This happens because the shorter a primer is, the smaller the annealing temperature will be.

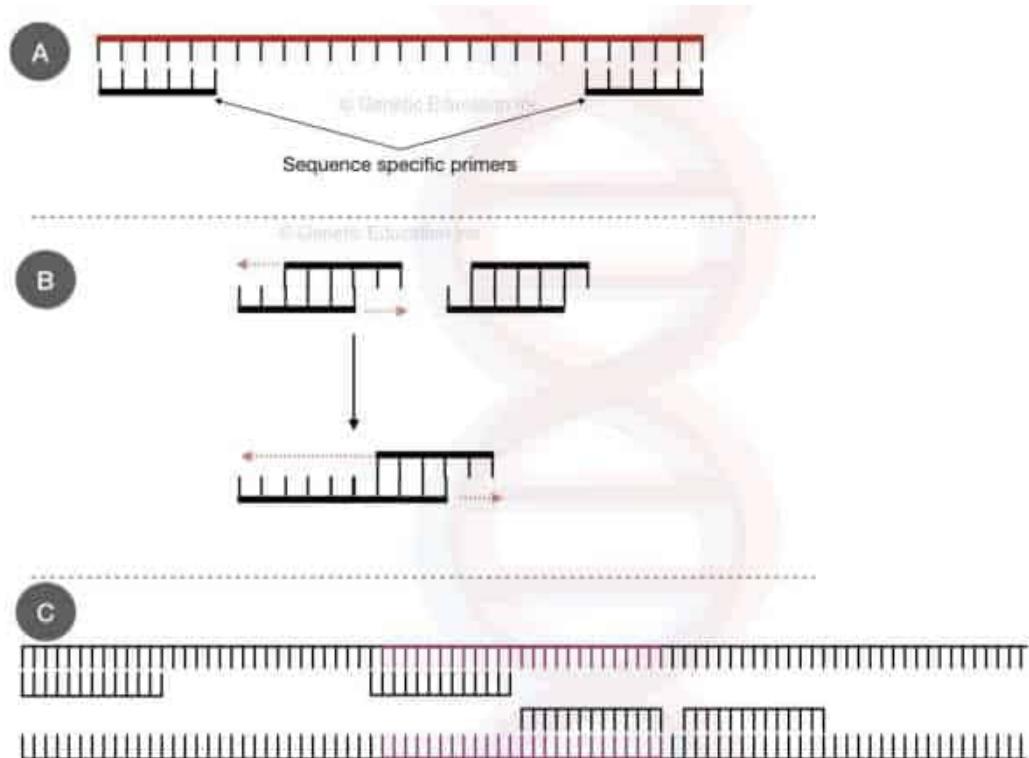
If the primer is too long, the annealing temperature will be higher, and it can reduce the amplification efficiency<sup>[38]</sup>.

## Primer GC content

Recommended GC content ranges from 40% -60%. GC-rich regions are nonspecific, in other words, the probability of nonspecific amplification in GC-rich regions is higher than in AT-rich regions<sup>[38]</sup>.

## Forward and Reverse Primer Complementation

Primary forward and reverse must not match or complement each other. If appropriate, the two primers will bond and form a dimer primer. Dimming primers can be amplified easily in PCR because they are formed by short sequences<sup>[38]</sup>.



(Figure 2.13.3; a) Specific primer attachment; b) Primary dimer formation; c) Adhesion is not specific)

## 2.13.2 PCR Steps

PCR consists of 3 steps, namely:

### 1. Denaturation

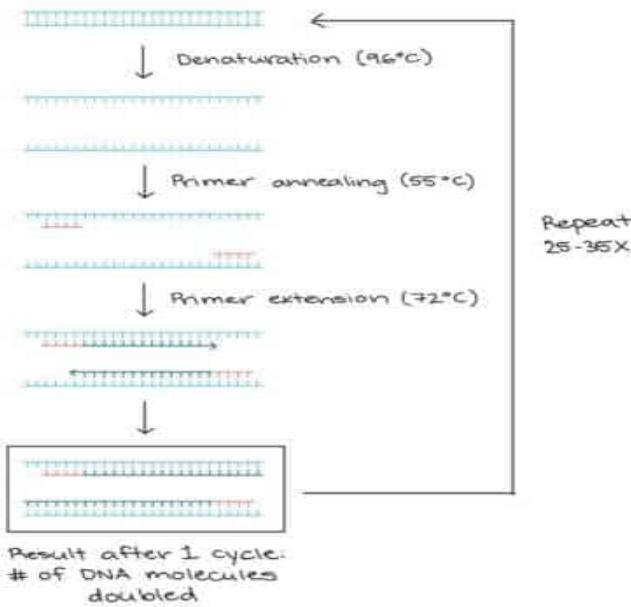
DNA is heated to about 95 ° C, to break the DNA into two single-strand DNA molecules by breaking their hydrogen bonds.

## 2. Annealing

The temperature is reduced to between 50 ° C - 60 ° C, this allows primary annealing with a 3' end of the DNA.

## 3. Elongation

The temperature is raised again until it reaches between 70 ° C-80 °C so that the DNA polymerase enzyme can work. DNA polymerase forms a new DNA strand which is complementary to the template DNA strand from the 5 'to 3' direction.



(Figure 2.13.4 Stages of PCR)

### 2.13.3 Primer3

Primer3 is an online software used to form primers automatically based on DNA templates. Primary 3 can be accessed on the website <http://bioinfo.ut.ee/primer3/>.

#### How to Use Primer3

First of all, open primer3, then enter the template DNA sequence.



(Figure 2.13.5 Primer3)

Then, select the options for primary forward and reverse, as indicated by the arrows in the image. Next, hit the Pick Primers button indicated by the blue box in the image.

## Primer3 Output

PRIMER PICKING RESULTS FOR

Template masking not selected  
No mispriming library specified  
Using 1-based sequence positions  
OLIGO start len tm gc% any\_th 3'\_th hairsns seq  
LEFT PRIMER 364 20 56.81 55.00 0.00 0.00 0.00 GGGAAACGAGGCTTTCTA  
RIGHT PRIMER 617 20 59.09 55.00 5.85 5.85 0.00 CTGAGATTCTGACTGGCCA  
SEQUENCE SIZE: 1431  
INCLUDED REGION SIZE: 1431

PRODUCT SIZE: 250, PAIR ANY\_TH COMPL: 0.00, PAIR 3'\_TH COMPL: 0.00

1 AGCCCTCCAGGACAGGCTGCATCAGAAQAGGCCATCAAGCAGGTCTGTTCCAAGGGCCTT

61 TGCCTCAGGTGGGCTCAGGATTCCAGGGTGGCTGGACCCCAGCCCCAGCTCTGCAGCG

121 GGAGGACGTGGCTGGCTCGTGAAGCATGTGGGGGTGAGCCCCAGGGGCCCCAAGGCAGGG

181 CACCTGGCTTCAGCTGCCTCAGCCCTGCCTGTCTCCAGATCACTGTCTTCTGCCAT

241 GGCCTGTGGATGCCCTCTGGCCCTGCTGGGGCTGCTGGCCCTCTGGGACCTGACCC

301 AGCCGCAGCCTTGTGAACCAACACCTGTOCGGCTCACACCTGGTGAAGCTCTCACCT

361 AGTGTGGGGAAACGAGGCTTCTTACACACCCAAAGACCCGCGGGGAGGAGAGGACCT  
>>>>>>>>>>>

421 GCAGGGTGAGCCAACCTGCCATTGCTGCCCTGGCGCCCCAGCCACCCCTGCTCTG

481 GCGCTCCACCCAGCATGGGCAGAAGGGGGCAGGGGGCTGCCACCCAGCAGGGGGTCAGG

541 TGCACTTTTAAAAAGAAGTTCTTGGTCAGTCCTAAAGTGACCAAGTCCCTGTGG

<<<

601 CCCAGTCAGAATCTAGCCTGAGGACGGTGTGGCTTCGGCAGCCCCAGATAATCAGA  
<<<<<<<<<<<<

(Figure 2.13.6 Primary result3)

Primary result3 will look like the image above. Look at the top, there are the primary forward (left) and reverse (right) specifications. Now consider the red-underlined text, this primer produces a yield of 250 bp, so when doing PCR based on this primary criterion, the result should be 250 bp.

The arrows (>>>>>>>>>) and <<<<<<<<<<<) indicate the primary annealing location during PCR. This software also provides other primer possibilities at the bottom (see image below).

```

1201 ACGTCCTGGCAGTGGGGCAGGTGGAGCTGGGCGGGGCCCCGTGGTGCAGGCAGCCTGCAGC
1261 CCTTGCCCCCTGGAGGGTCCCTGCAGAACGCTGGCATTGTGGAACAATGCTGTACCAAGCA
1321 TCTGCTCCCTTACCAAGCTGGAGAACTACTGCAACTAGAEGCAGCCCGAGGEAGCCCCA
1381 CACCCGCCCTCTGCACCGAGAGATGGAATAAAGCCCTGAAACCAGC

KEYS (in order of precedence):
>>>> left primer
<<<< right primer

ADDITIONAL OLIGOS:
  start len      tm      gc%  any_th  3'_th hairpin seq
1 LEFT PRIMER      984  20  58.86  55.00   0.00   0.00   0.00 TCAGGAGATGGGGAGATGC
RIGHT PRIMER     1070  20  58.80  60.00   0.00   0.00   0.00 CAGGTTAGAGGGAGGGTCAC
PRODUCT SIZE: 167, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00
2 LEFT PRIMER      683  20  58.75  50.00   0.00   0.00   0.00 ACTCGCCCTCAAAACAAATG
RIGHT PRIMER     923  20  58.86  55.00   0.00   0.00   0.00 GCATCTTCCCCATCTCTGA
PRODUCT SIZE: 241, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00
3 LEFT PRIMER     1141  20  58.96  55.00   0.00   0.00   0.00 TGACTGTGTCTCTGTGTC
RIGHT PRIMER    1353  20  59.39  55.00   27.87  27.87   0.00 TGCACTAGTTCTCCAGCTGG
PRODUCT SIZE: 213, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00
4 LEFT PRIMER      22  20  59.01  50.00   5.01   0.00   0.00 TCAGAAGAGGCCATCAAGCA
RIGHT PRIMER    239  20  58.35  50.00   0.00   0.00   0.00 TGGCAGAAGGACAGTGATCT
PRODUCT SIZE: 218, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

Statistics
  con too in in not no tm tm high high high high
  sid many tar excl ok bad GC too too any_th 3'_th hair poly end
  ered Ns get reg reg GC% clamp low high compl compl pin X stab ok
Left 7787 0 0 0 0 2791 0 637 2756 0 0 17 58 0 1528
Right 7790 0 0 0 0 2982 0 719 2574 0 0 19 56 0 1440
Pair Stats:
considered 1585, unacceptable product size 1578, primer in pair overlaps a primer in a better pair 156, ok 7
libprimer3 release 2.4.0

(primer3_results.cgi release 4.1.0)

```

(Figure 2.13.7 Another possible primary)

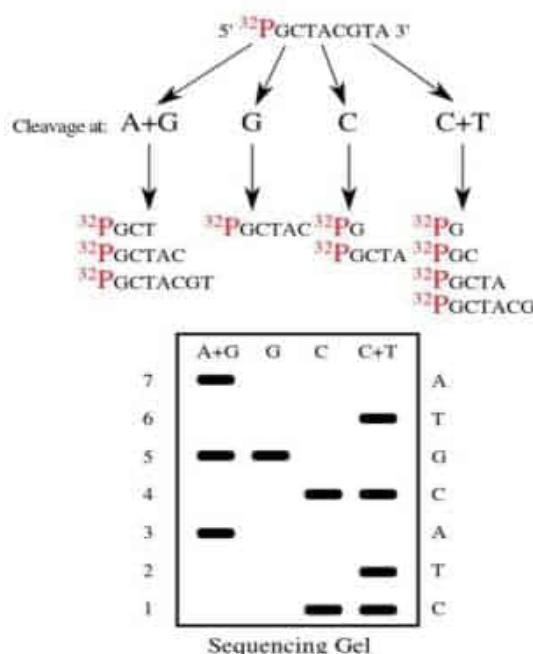
## 2.14 Sequencing

DNA sequencing or DNA sequencing is a process or technique of determining the sequence of nucleotide bases in a DNA molecule<sup>[14]</sup>. DNA sequencing can be used to determine the identity and function of genes or other DNA fragments by comparing the sequence with other known DNA sequences<sup>[15]</sup>. This technique is used in basic biological research as well as in various applied fields such as medicine, biotechnology<sup>[16]</sup>, and anthropology<sup>[17]</sup>.

### 2.14.1 Maxam-Gilbert Sequencing

Maxam-Gilbert sequencing required radioactive labelling of the 5' end of the DNA and purification of the DNA fragments for sequencing. The chemical treatment then produces pauses in a small proportion of one or two of the four nucleotide bases in each of the four reactions (G, A + G, C, C + T). Thus a

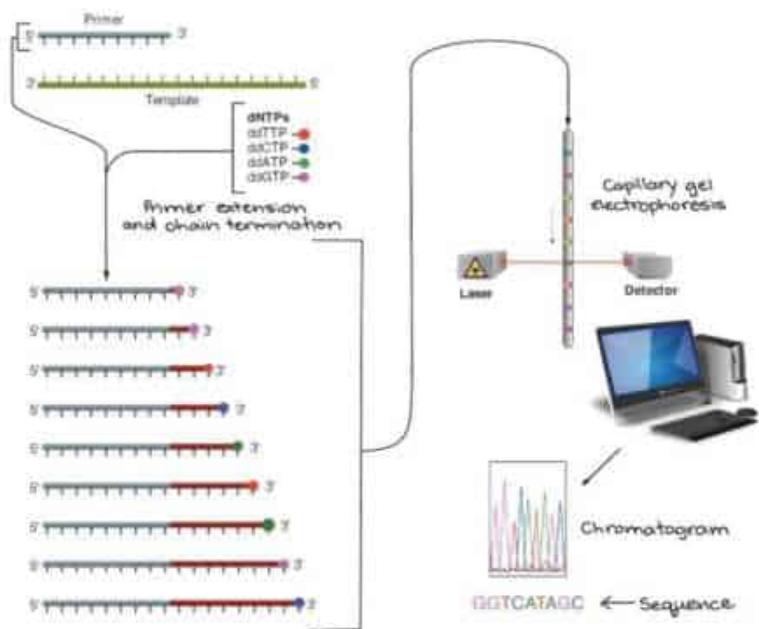
series of labelled fragments are generated, from the radiolabeled end to the first "cut" site in each molecule. The fragments in four reactions coexist electrophoretically in denaturing acrylamide gel for size separation. To visualize the fragments, the gel is exposed to an X-ray film for autoradiography, producing a series of dark bands each of which corresponds to a radiolabel labelled DNA fragment, from which the sequence can be inferred<sup>[18]</sup>.



(figure 2.14.1 Maxam-Gilbert Sequencing)

## 2.14.2 Sanger Sequencing

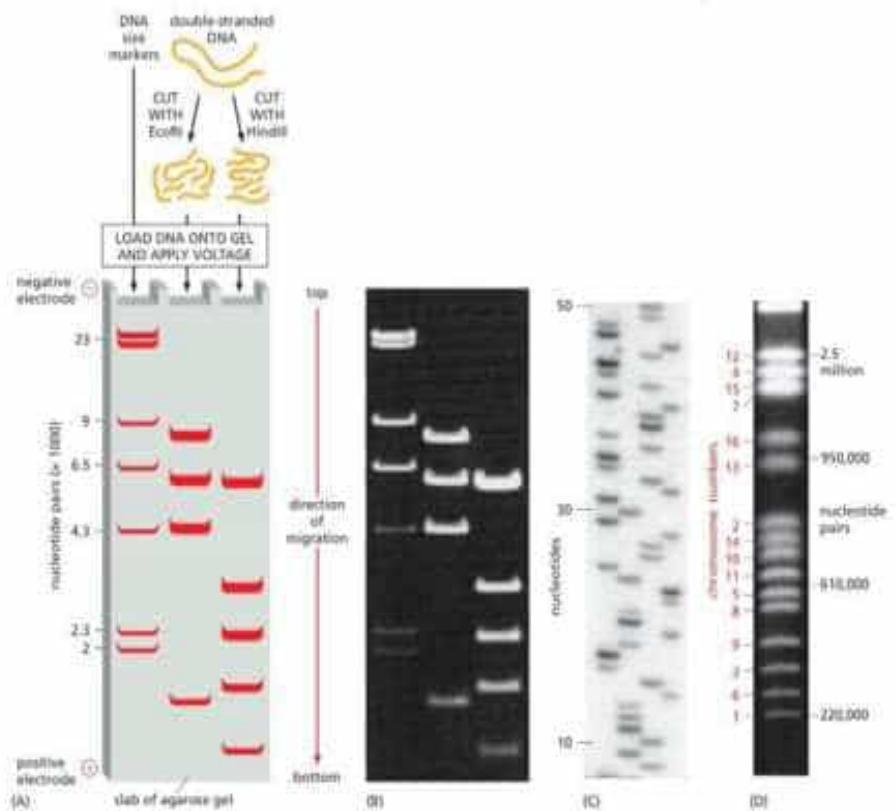
Sanger sequencing is a DNA sequencing method based on the selective fusion of dideoxynucleotide chains that are broken down by DNA polymerase during in vitro DNA replication<sup>[19]</sup>.



(Figure 2.14.2 Sanger Sequencing)

## 2.15 Electrophoresis

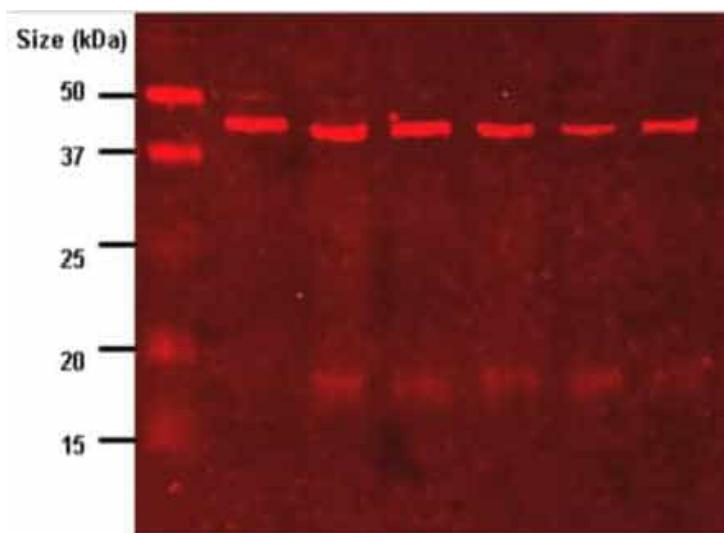
Electrophoresis is a technique of separating charged components or molecules based on the difference in their migration rates in an electric field. This technique makes use of the electric charge on macromolecules, such as DNA which is negatively charged. If a negatively charged molecule is passed through a medium, an electric current is applied from one pole to the opposite pole, the molecule will move from the negative pole to the positive pole. The speed of motion of these molecules depends on the ratio of charge to mass and also depends on the shape of the molecule<sup>[20]</sup>.



(Figure 2.15.1 Elektroforese)

## 2.16 Western Blot

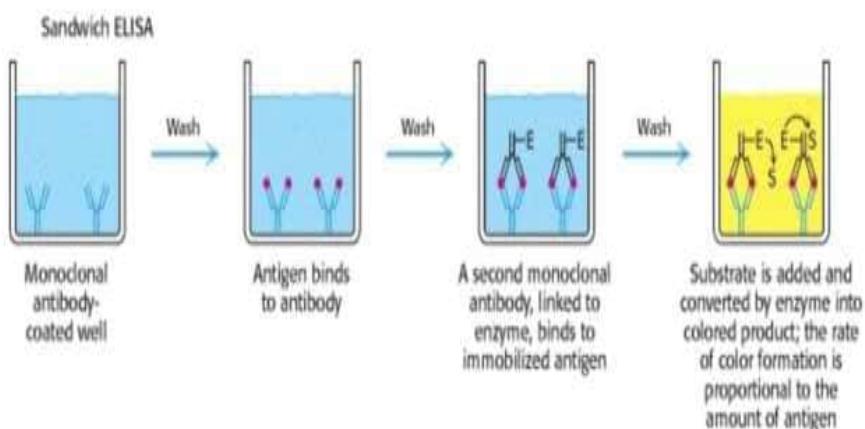
Western blot is a method for detecting protein in tissue samples. Western blot uses gel electrophoresis to separate the original protein or change by polypeptide spacing or by the protein's 3D structure. These proteins are delivered to the membrane, where they are detected using antibodies to target the protein<sup>[22]</sup>.



(Figure 2.16.1. Western blot using anti-lipoic acid primary antigen and IR labelled secondary antigen in the primary extract of Leishmania. Vickers team, 2012)

## 2.17 ELISA

The enzyme-linked immunosorbent assay (ELISA) is a commonly used analytical biochemical test, first described by Engvall and Perlmann in 1971. This test uses a solid-phase Enzyme Immunoassay (EIA) to detect the presence of ligands (mostly proteins) in a liquid sample using antibodies directed against the protein to be measured. ELISA has been used as a diagnostic tool in medicine, plant pathology, and biotechnology, as well as in quality inspection in a variety of industries. Initially, the sample adheres to the surface. Then, a suitable antibody is applied to the surface so that it can bind to the antigen. These antibodies are linked to the enzyme, and in the final step, the substance containing the enzyme substrate is added. The subsequent reaction produces a detectable signal, for example, a colour change<sup>[23]</sup>.



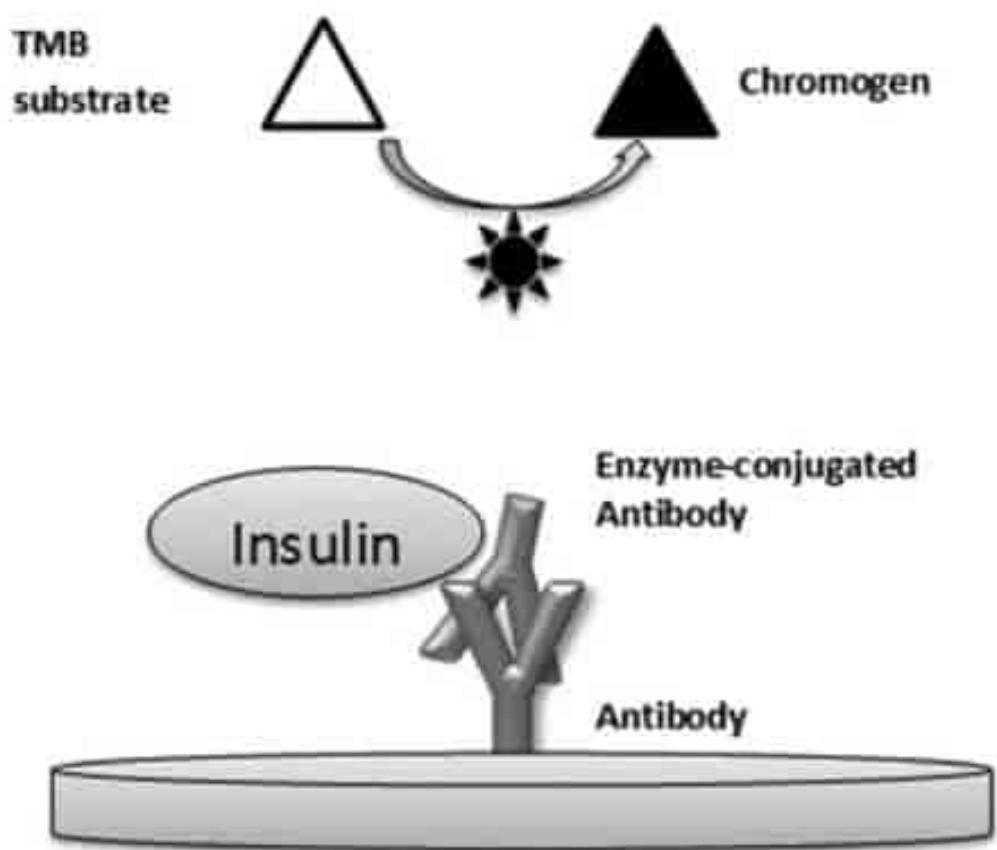
(Figure 2.17.1 Sandwich ELISA. Pratiksha, Molecular Test, 2015)

### 2.17.1 Insulin ELISA kit

The purpose of ELISA is to detect whether a target antigen (proinsulin) is present in the sample. The sandwich-type ELISA is the most popular for detecting the presence of proinsulin and contains two antibodies: a trapping antibody (anti-insulin antibody) and a detecting antibody (peroxide labelled anti-insulin antibodies)<sup>[151]</sup>. During incubation, the insulin in the sample reacts with the immovable anti-insulin antibody on the plate (Figure 2.1.1)<sup>[152]</sup>. The

unbound antibody was rinsed with a rinse solution and 3,3',5,5'-tetramethylbenzidine (TMB) added as a chromogenic substrate to react with insulin/proinsulin to form a conjugate bond<sup>[153]</sup>. After the reaction is stopped by adding the acid, the coloured product is measured at 450 nm using a spectrophotometer and measures the signal corresponding to the analyte concentration<sup>[153]</sup>. Commercially available ELISA kits for proinsulin/insulin detection show a detection limit (LOD) of 3-12pmol / L and an upper limit of 600pmol / L. [154] The LOD is 0.73 µIU mL<sup>-1</sup> and 4.9 µIU mL<sup>-1</sup>, and the upper limit

200 µIU mL<sup>-1</sup> and 324 µIU mL<sup>-1</sup> (Even et al. 2007 and Abellan et al. (2009)<sup>[155,156]</sup>. ELISA has a good selection system for insulin and minimizes interference from other proteins, making it suitable for research purposes where the samples contain physiological solutions in the form of salts and albumin<sup>[157]</sup>. In addition, there are ELISA modifications such as the amplified luminescent proximity homogenous assay (AlphaLISA) and the homogeneous time-resolved fluorescence (HTRF) which have been developed to detect insulin/proinsulin further.



(Figure 2.17.2 Direct sandwich ELISA for detection of Insulin / Proinsulin <sup>[159]</sup>)

## BAB 3

### Methodology

The methodology used in this research is in silico technique.

#### 3.1 In Silico

In silico is a type of research that is mainly based on computer simulation and analysis. In this study, we used the In silico technique because of several considerations such as the time needed for the experiment is faster, more economical, etc. than in vivo and in vitro.

#### 3.2 Data and Analysis Methods

This research is focused on analysis, simulation, and computation of data in order to research and develop transgenic aloe vera which is capable of producing proinsulin. Computational data and simulation results will be collected as a basis for real-world experiments. We used various bioinformatics tools that we got from the internet, and built several others using biopython. The data we use we get from various online databases such as GenBank, UCSC Genome Browser, and others. It should be noted that this research is still in the simulation trial stage, so it is almost certain that this research will undergo many changes in the future, which are related to the results and data from the simulations we have done.

#### 3.3 Procedure

1. Identifying and gathering information
  - Gathering information about diabetes
  - Analyzing from a literature review of constructing the system and basic theory
  - Observe the most suitable approach for applying genetic techniques.
  - Collecting transgenic journals and data from the database.
2. Mengobservasi dan mengumpulkan data
  - Constructing transgene modification to increase insulin production in Aloe vera.
  - Collecting information from journals and databases:
    - Genome sequence (Sequence).
    - Metabolic pathway
    - Comparing specific platforms and targets

- Gene sequence construction
- Selecting restriction and ligation enzymes used in research
- Finding a suitable transformation method
- Searching for supporting theories needed in research
  - Insulin synthesis and transgenic pathways
  - Constructing a suitable clone vector design.

### 3. Software simulation

- Identifying restriction site with Restriction Mapper 3.0;
- Optimizing proinsulin codon with dnachisel and biopython;
  - Splicing;
  - Codon usage table;
  - Optimising with dnachisel;
  - Pairwise alignment with EMBL-EBI<sup>[201]</sup>.
- PCR Optimization.
- Transgene construction with snapgene and benchling.

### 4. Experimental design (GMO modification approach)

- Laboratorium processes
- Isolasi promoter
- Perakitan vektor
- Kloning
- Transformasi
- Identifikasi tumbuhan transgenik
- Laboratory Tests / Experiment Results
  - Extraction and molecular characterization of Aloe vera
  - Transgenic gene identification
  - Test confirmation of the expression
  - Identification of GMO crops
  - Insulin extraction and western blotting
  - Identify the level of insulin produced

## 3.4 Identifying and Gathering Information

Determining the research method between *in vivo*, *in vitro*, and *in silico*, Determining the expression platform and system (transformation target), and the transformation technique used. So that the results obtained using the *in silico* approach, using plants as an expression platform (specifically Aloe vera), using a chloroplast expression system, and a transformation technique using particle bombardment.

## **3.5 Observing and Collecting Construction Data**

### **3.5.1 Collecting and Selecting Sequences**

There are three sequences that we will use in this research. The three sequences are the promoter, gene of interest, and terminator. Since the aim of this study is to design transgenic Aloe vera that can produce proinsulin, the gene of interest we use is the human preproinsulin sequence from Genbank NC\_000011 REGION: complement (2159779..2161209). Then we will optimize the insulin sequence to increase the results of the expression.

#### **Determining the Promoter**

As explained in chapter 2.8.1, the promoter is one of the most important parts of gene expression, because the promoter functions as a kind of pedal wheel for gene expression. Therefore, determining a promoter according to the needs and demands of this experiment is a very crucial part.

In choosing the promoter that we will use, we have several considerations, these considerations are:

- Which promoter is not toxic to humans
- Which promoter matches the chloroplast expression

And after various considerations, we decided to use the psbA promoter, a specific promoter for chloroplasts. We got the sequence for the psbA promoter from UniProt P83755.

#### **Determining the Terminator**

Apart from the promoter, the terminator is also a crucial part of a gene. Terminator functions as a marker of the end of the expression. The terminator has been described in chapter 2.8.3. The terminator we use is the NOS terminator, with the consideration that this terminator is the terminator commonly used in transformations mediated by particle bombardment. We got the NOS terminator sequence from pCAMBIA 1305.1 GenBank AF354045.1

### **3.5.2 Selecting the Expression Vector**

After determining the exact sequence to be used, the next step is to choose the appropriate expression vector. After conducting further research on various

journals and online discussion forums such as researchgate, we decided to use vectors from the pCAMBIA family. The next step is to decide which vector we want from the pCAMBIA cloning vector family. Some of the things we are considering are:

- Vectors must not contain substances that are toxic to humans.
- The vector must not contain genes that can inhibit expression in the chloroplasts of monocot plants such as Aloe vera.
- The vector must have the genes coding for resistance to antibiotics.

Most pCAMBIA vectors have a GUS reporter gene, unfortunately, the  $\beta$ -glucuronidase enzyme, which is encoded by the GUS gene, is toxic to humans<sup>[196]</sup>. Therefore, in the end, we chose to use the pCAMBIA 1300 cloning vector. pCAMBIA 1300 is a cloning vector of the pCAMBIA family that does not have the GUS gene and has genes coding for resistance to kanamycin and hygromycin. pCAMBIA 1300 has a CaMV 35S promoter before the HygR sequence encoding resistance to hygromycin.

### 3.6 In Silico Steps

- **Data and Analysis Methods**

This project focuses on computational analysis and construction methods in the development of transgenic aloe vera producing proinsulin. The computational data and simulation results obtained from several sources and using several algorithms and Python programming language. Detailed experimental plans and methods are determined based on the results of in silico, theoretical basis, the data obtained, and also guided by previous studies. In addition, in the experimental section, the research will be carried out next to compare the data between the computational results and the real experimental results.

- **Softwares**

1. Genbank, UCSC Genome
  - Genome database
2. KEGG: Kyoto Encyclopedia of Genes and Genomes
  - Molecular properties, metabolic pathways of *Aloe vera*

- 3. Restriction Mapper 3.0
  - Restriction site mapping
- 4. Snapgene
  - Vector construction, PCR
- 5. Benchling
  - Virtual electroforesis
- 6. Python dnachisel
  - Codon optimization
- 7. Python biopython
  - Bioinformatics
- 8. Jupyter Notebook
  - Data science module

### 3.6.1 Restriction Sites Identification

We get data about restriction sites with Restriction Mapper 3 software (<http://restrictionmapper.org>). We use the pCAMBIA1300 vector.

- MCS pCAMBIA1300 unique restriction sites:

HindII, SmaI, AccI, ApoI, AvaI, BamHI, BspMI, EcoRI, HindIII, PleI, SalI, XbaI, XhoII, KpnI, NspI, PstI, SacI, SduI, SphI, Sse8387.

Table 3.6.1.1 side of restrictions on each gene (source: Restriction Mapper Version 3.0)

Gen	Non Cutter
psbA Promoter	AarI, AatII, AbsI, AccI, AcII, AfIII, AfIII, AgsI, AjuI, AlfI, AloI, AlwNI, ApaI, ApaLI, ApoI, ArsI, AscI, AsuII, AvaI, AvaiII, AvrII, BaeI, BamHI, BarI, BbvCI, BciVI, BclI, BdI, BfI, BglII, BplI, Bpu10I, BsaAI, BsabI, BsaXI, BseRI, BseSI, BsgI, BsmAI, Bsp1407I, BspHI, BspMI, BsrDI, BstXI, BtrI, BtsI, C1aI, CspCI, DraII, DraIII, DrdI, Eam1105I, Eco31I, Eco47III, Eco57I, Eco57MI, EcoNI, EcoRI, EcoRV, Esp3I, Fall, FseI, FspAI, GsuI, HaeIV, HgaI, Hin4I, HindII, HindIII, HpaI, Hpy99I, KpnI, MauBI, MfeI, MluI, NaeI, NheI, NmeAIII, NotI, OliI, PacI, PasI, PfoI, PleI, PmaCI, PmeI, PpiI, PpuMI, PshAI, PsiI, PI-PspI, PspXI, PsI, PstI, PvuI, PvuII, RsrII, SacI, SalI, SanDI, SapI, ScaI, PI-SceI, SexAI, SfiI, Sgfl, SgrDI, SmaI, SmlI, SnaBI, SpeI, SrfI, Sse8387I, SspI, StuI, SwaI, TaqII, TatI, TspRI, TstI, Tth111I, VspI, XbaI, XcmI, XhoI, Xmni
Proinsulin	AatII, AbsI, AccI, AcII, AfIII, AfIII, AgeI, AgsI, AjuI, AlfI, AloI, ApaLI, ApoI, ArsI, AscI, AsuII, AvrII, BaeI, Ball, BamHI, BarI, BbvCI, BccI,

	BcgI, BciVI, BclI, BdAI, BfI, BglII, BplI, Bpu10I, BsaAI, BsaBI, BseMII, BsePI, BsmI, BsmAI, Bsp1407I, BspHI, BsrI, BsrBI, BsrDI, BstEII, BstXI, BtgZI, CfrI, ClaI, CspCI, DrdI, Eam1105I, EcI, Eco31I, Eco47III, Eco57I, EcoRI, EcoRV, Esp3I, Fall, FseI, FspAI, HaeIV, HgaI, Hin4I, HindII, HindIII, HpaI, Hpy99I, KpnI, MauBI, MfeI, MluI, MmeI, MsII, NcoI, NdeI, NheI, NotI, NruI, NspI, OliI, PacI, PasI, PflMI, PfoI, PleI, PmaCI, PmeI, PpiI, PshAI, PsiI, PI-PspI, PspXI, PsrI, PvuI, RsrII, SacI, SacII, SalI, SanDI, SapI, Scal, PI-SceI, SfaNI, SgfI, SgrAI, SgrDI, SmII, SnaBI, SpeI, SphI, SspI, StuI, StyI, SwaI, TaqII, TatI, TfiI, TsoI, Tsp45I, TspDTI, TspGWI, TstI, Tth111I, VspI, XbaI, XcmI, XhoI, XhoII, Xmni
NOS	AarI, AatII, AbsI, AccI, AcI, AcyI, AgeI, AjuI, AlfI, AloI, AlwNI, ApaI, ApaLI, ApoI, ArsI, AsCI, AsuII, Aval, Avall, BaeI, Ball, BamHI, BarI, BbvI, BbvCI, BcgI, BciVI, BclI, BdAI, BfI, BglII, BplI, Bpu10I, BsaAI, BsaXI, BseMII, BseRI, BseSI, BseYI, BsgI, BsmI, BsmAI, Bsp1407I, BspHI, BspMI, BsrI, BsrBI, BsrDI, BstEII, BstXI, BtrI, BtsI, CfrI, ClaI, CspCI, DraII, DraIII, DrdI, Eam1105I, EcI, Eco31I, Eco47III, Eco57I, Eco57MI, EcoNI, EcoP15I, EcoRI, EcoRII, EcoRV, Esp3I, Fall, FokI, FseI, FspAI, GsuI, HaeII, HgaI, HindII, HindIII, HpaI, HphI, Hpy99I, KpnI, MboII, MfeI, MluI, MmeI, MsII, NaeI, NarI, NcoI, NdeI, NheI, NmeAIII, NotI, NruI, OliI, PacI, PasI, PflMI, PfoI, PmaCI, PmeI, PpiI, PpuMI, PshAI, PsiI, PI-PspI, PspXI, PsrI, PstI, PvuI, PvuII, RsrII, SacI, SacII, SalI, SanDI, SapI, Scal, PI-SceI, SduI, SexAI, SfaNI, SfiI, SgfI, SgrAI, SgrDI, SmaI, SnaBI, SpeI, SphI, SrfI, Sse8387I, SspI, StuI, StyI, SwaI, TaqII, TatI, TauI, TseI, TsoI, Tsp45I, TspDTI, TspGWI, TspRI, TstI, Tth111I, VspI, XbaI, XcmI, XhoI, XhoII, Xmni

### 3.6.2 Optimization of Proinsulin Codons with dnachisel

Our desired proinsulin sequence resides in two distinct exons, the first at base 239..425, and the second at base 1213..1358, separated by an intron of 787 bp. In order to obtain a proinsulin sequence that is not separated by an intron, we decided to take the following steps:

1. Transcription to RNA
2. Splicing
3. Reverse Transcriptase PCR to DNA

After the steps above are complete, you will get the following sequence:

239 at

241 ggcctgtgg atgcgcctcc tgcccctgct ggcgctgctg gccctctggg gacctgaccc  
 301 agccgcagcc ttgtgaacc aacacctgtg cggctcacac ctggtgaaag ctctctacct  
 361 agttgtcggg gaacgaggct tcttctacac acccaagacc cgccgggagg cagaggacct  
 421 gcaggtggggcag gtggagctgg gcgggggccg tggtgccaggc agcctgcagc  
 1261 cttggccct ggaggggtcc ctgcagaagc gtggcattgt ggaacaatgc tgtaccagca

1321 tctgctccct ctaccagctg gagaactact gcaactag

The blue color represents the first exon, and the red color represents the second exon.

Translation results:

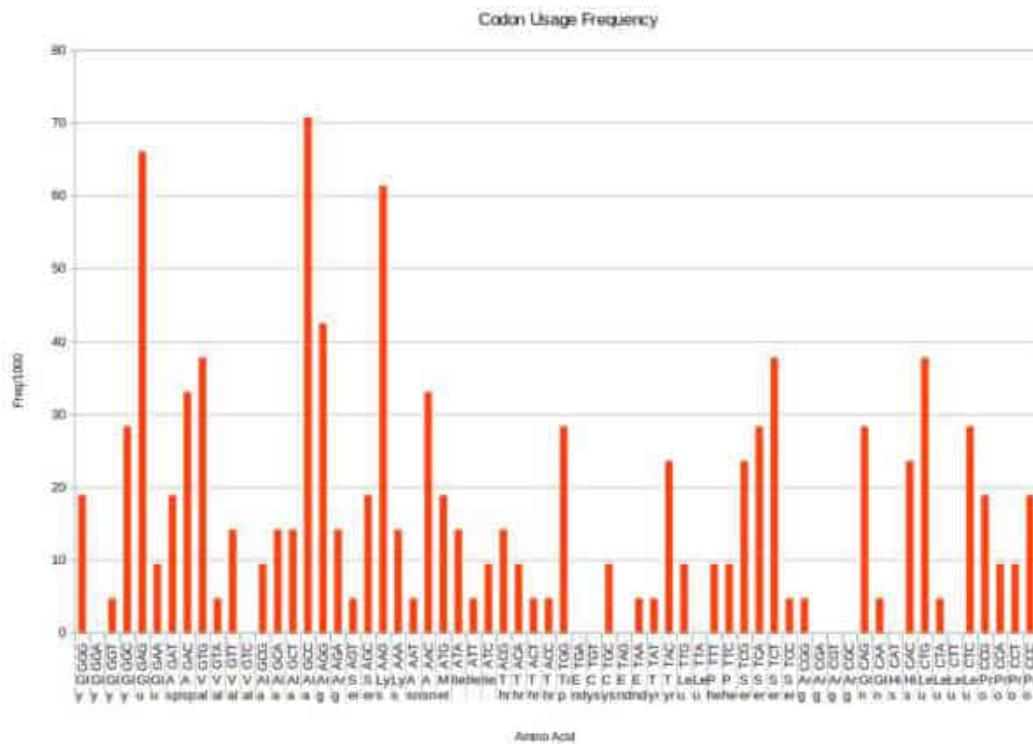
MALWMRLLPLLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGER  
GFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALE GSLQKRGIVEQ  
CCTSICSLYQLENYCN\*

To improve expressions, we codon optimization using python with the dnachisel and biopython modules. Our optimization codon is based on the species index 34199.chloroplast of kazusa.

Codon usage table:

Table 3.6.1 Codon chloroplast Aloe vera (gbpln) 1 CDS (516 codons), kazusa species 34199.chloroplast

chloroplast Aloe vera [gbpln]: 1 CDS's (516 codons)							
fields: [triplet] [amino acid] [fraction] [frequency: per thousand] ([number])							
UUU F 0.69 65.9 {	34)	UCU S 0.35 34.9 {	18)	UAU Y 0.91 58.1 {	30)	UGU C 0.08 13.6 {	7)
UUC F 0.31 29.1 {	15)	UCC S 0.08 7.8 {	4)	UAC Y 0.09 5.8 {	3)	UGC C 0.12 1.9 {	11)
UUA L 0.33 38.8 {	20)	UCA S 0.31 31.8 {	16)	UAA * 0.00 0.0 {	9)	UGA * 1.00 1.9 {	11)
UUG L 0.18 21.3 {	11)	UCG S 0.10 9.7 {	5)	UAG * 0.00 0.6 {	8)	UGG W 1.00 11.6 {	6)
CUU L 0.28 32.9 {	17)	CCU P 0.36 9.7 {	5)	CAU H 0.89 46.5 {	24)	CGU R 0.15 11.6 {	6)
CUC L 0.05 5.8 {	3)	CCC P 0.14 3.9 {	2)	CAC H 0.11 3.8 {	3)	CGC R 0.05 3.9 {	2)
CUA L 0.15 17.4 {	9)	CCA P 0.29 7.8 {	4)	CAA Q 0.74 27.1 {	14)	CGA R 0.38 29.1 {	15)
CUG L 0.02 1.9 {	1)	CCG P 0.21 5.8 {	3)	CAG Q 0.26 9.7 {	5)	CGG R 0.08 5.8 {	3)
AUU I 0.58 58.4 {	26)	ACU T 0.44 15.5 {	8)	AAU N 0.88 38.8 {	28)	AGU S 0.14 13.6 {	7)
AUC I 0.16 13.6 {	7)	ACC T 0.22 7.8 {	4)	AAC N 0.20 9.7 {	5)	AGC S 0.02 1.9 {	11)
AUA I 0.27 23.3 {	12)	ACA T 0.22 7.8 {	4)	AAA K 0.01 40.7 {	21)	AGA R 0.28 21.3 {	11)
AUG M 1.00 11.6 {	6)	ACG T 0.11 3.9 {	2)	AAG K 0.19 9.7 {	5)	AGG R 0.05 3.9 {	2)
GUU V 0.48 19.4 {	10)	GCU A 0.50 13.6 {	7)	GAU D 0.07 25.2 {	13)	GGU G 0.45 9.7 {	5)
GUC V 0.18 3.9 {	2)	GCC A 0.87 1.9 {	1)	GAC D 0.13 3.9 {	2)	GGC G 0.09 1.9 {	11)
GUA V 0.38 15.5 {	8)	GCA A 0.29 7.8 {	4)	GAA E 0.05 44.6 {	23)	GGA G 0.18 3.9 {	2)
GUG V 0.05 1.9 {	1)	GCG A 0.14 3.9 {	2)	GAG E 0.15 7.8 {	4)	GGG G 0.27 5.8 {	3)



Graph 3.6.1 Frequency of Codon Usage kazusa species 34199.chloroplast

The optimization results based on the codon usage table above are as follows:

ATGGCCCTGTGGATGAGGCTGCTGCCCTGCTGGCCCTGCTGGCCCT  
GTGGGGCCCGAACCCGCCGCCGCCTTGTGAACCAGCACCTGTGC  
GGCTCTCACCTGGTGGAGGCCCTGTACCTGGTGTGCGGCGAGAGGG  
GCTTCTTCTACACGCCAAGACGAGGAGGGAGGCCGAGGACCTGCA  
GGTGGGCCAGGTGGAGCTGGCGGCCGGCGCCGGCTCTCTG  
CAGCCCCTGCCCTGGAGGGCTCTCTGCAGAACAGAGGGCATAGTGG  
AGCAGTGCTGCACGTCTATGCTCTGTACCAAGCTGGAGAACTA  
CTGCAACTAA

Translation results after optimization:

MALWMRLPLLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGER  
GFFYTPKTRREAEDLQVGQVELGGPGAGSLQPLALEGSLQKRGIVEQ  
CCTSICSLYQLENYCN\*

Sequence alignment before and after optimization:

NP_001278826.	1 ATGGCCCTGTGGATGCGCCTCTGCCCTGCTGGCGCTGCTGGCCCTCTG        . .   .       . .	50
NP_001278826.	1 ATGGCCCTGTGGATGAGGCTGCTGCCCTGCTGGCCCTGCTGGCCCTGTG        . .   .       . .	50
NP_001278826.	51 GGGACCTGACCCAGCCGCAGCCTTGTGAACCAACACCTGTGCGGCTCAC      .   .     .       .     .	100
NP_001278826.	51 GGGCCCGAACCCGCCGCCGCCCTTGTGAACCAGCACCTGTGCGGCTCTC      .   .     .       .	100
NP_001278826.	101 ACCTGGTGGAAAGCTCTACCTAGTGTGCGGGGAACGAGGCTTCTTCTAC        .   .   .     .       .   .	150
NP_001278826.	101 ACCTGGTGGAGGCCCTGTACCTGGTGTGCGGGAGAGGGGCTTCTTCTAC        .   .   .     .       .   .	150
NP_001278826.	151 ACACCCAAGACCGCCGGGAGGGCAGAGGACCTGCAGGTGGGGCAGGTGGA    .       .   .   .     .       .	200
NP_001278826.	151 ACGCCAAGACGAGGGAGGGAGGCCGAGGACCTGCAGGTGGGCCAGGTGGA      .   .     .   .     .       .	200
NP_001278826.	201 GCTGGCGGGGGCCCTGGTCAGGCAGCCTGCAGCCCTGGCCCTGGAGG        .     .   .   .    ...       .	250
NP_001278826.	201 GCTGGCGGCCGCCGGCGCCGGCTCTGCAGCCCTGGCCCTGGAGG        .     .   .   .     .       .	250
NP_001278826.	251 GGTCCCTGCAGAAGCGTGGCATTGTGAAACATGCTGTACAGCATCTGC    .       .   .     .     .   .   .	300
NP_001278826.	251 GCTCTCTGCAGAAGAGGGCATAGTGGAGCAGTGTGCACGTCTATATGC      .       .       .       .   .	300
NP_001278826.	301 TCCCTCTACCAGCTGGAGAACTACTGCAACTAG    .   .       .       .   .	333
NP_001278826.	301 TCTCTGTACCAAGCTGGAGAACTACTGCAACTAA      .       .       .   .	333

Figure 3.6.12 Pairwise alignment of proinsulin sequences before and after codon optimization

Alignment hasil translasi sebelum dan sesudah optimasi:

NP_001278826.	1 MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFY        . .   .       . .	50
NP_001278826.	1 MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFY        . .   .       . .	50
NP_001278826.	51 TPKTRREAEDLQVGQVELGGPGAGSLQPLALEGSLQKRGIVEQCCTSIC        . .   .     .     .       .	100
NP_001278826.	51 TPKTRREAEDLQVGQVELGGPGAGSLQPLALEGSLQKRGIVEQCCTSIC        . .   .     .     .       .	100
NP_001278826.	101 SLYQLENYCN* 111        . .   .       .   .	
NP_001278826.	101 SLYQLENYCN* 111        . .   .       .   .	

Figure 3.6.3 Pairwise alignment of translation results before and after codon optimization

### 3.6.3 PCR Optimization

This system is intended to help evaluate the quantity of primers and nucleotides for the optimal PCR reaction and determine the number of cycles required. This estimate is basic. The default parameters relate to the amplification of the 2Kb fragment of 0.5 µg of human DNA. So it is assumed :

1. The conditions of the reaction are close to normal, so there is no need to worry about the theory (Most polymerases can cause nonspecific amplification, most primers can cause dimer-primers);

2. A, T, and G, C are equivalent to the PCR results;
3. There are no primary-dimers;
4. Taq polymerase does not lose its activity during the reaction.

If:

- The length of the PCR product is "L" [kbp];
- The dNTP's concentration is "c" [mM];
- The primary quantity is "q" [pmol];
- The quantity of Taq polymerase is "a" [u];
- The reaction volume is "V" [ $\mu$ l];
- The elongation time is "t" [min];
- Template quantity is "mo".

Then:

1. The maximum result is the minimum result of the two evaluations below:

- a. If all nucleotides are used:

$$mn = 4[\text{nucleotides}] \times c[\text{mmol/l}] 324.5 [\text{g/mol}] \times V [\mu\text{l}] = 1300 cV [\text{ng}]$$

- b. If all primers are used :

$$mp = q[\text{pmol}] \times 2[\text{strands}] 324.5[\text{g/mol}] \times L[\text{kbp}] = 650qL [\text{ng}]$$

2. The maximum quantity of PCR products per one cycle depends on two factors:

- a. Taq polymerase speed: 2-4 [kbp / min];

- b. Taq polymerase activity (1 u is the sum of the enzymes, which incorporate 10 nmol of the four dNTPs in 30 min at 72°C).

$$mcycle = 10[\text{nmol}] \times 324.5[\text{g/mol}] \times a [\text{u}] t[\text{min}] / 30[\text{min}] = 108at [\text{ng}]$$

3. The number of cycles that are useful for the synthesis of "mmax" PCR products are:

$$mmax = 2^n \times mo \Rightarrow n = \ln(mmax/mo)/\ln$$

4. The relationship between mass and quantity of moles is:

$$m[\mu\text{g}] = 649[\text{g/mol}] \times q[\mu\text{mol}] \times L[\text{kbp}] \times 1000$$

### 3.6.4 Transgene Construction Using Snapgene

The first step we do is look for primers and restriction sites based on the results of the restriction mapper. Taking into account the data from the

restriction mapper, we decided to use four restriction enzymes to connect the three genes. The four restriction enzymes are KpnI, BamHI, SphI and HindIII, ordered from left to right in the final vector. The four restriction sites were chosen because the four of them are non-cutters for the three genes. To connect the three genes with MCS pCAMBIA1300, we decided to use the insertion cloning technique with the enzyme T4 ligase.

The stages we do are as follows:

1. Adding restriction sites at both ends of the gene of interest with PCR;
2. Benchling PCR virtual electrophoresis;
3. Perform insertion cloning.

### **Phase 1:**

We performed virtual PCR in snapgene to generate genes sandwiched by restriction enzymes. The primers we use are designed automatically by snapgene, while the restriction sites we add manually using data from the restriction mapper as a consideration.

Here is a list of the primers we use:

1. psbA Promoter
  - a. Forward: 5'-(KpnI)ATGACCGCGATTCTGGAACG-3'
  - b. Reverse: 5'-(BamHI)GCCGTTGGTGCTCGGC-3'
2. Proinsulin
  - a. Forward: 5'-(BamHI)ATGGCCCTGTGGATGAGGC-3'
  - b. Reverse: 5'-(SphI)TTAGTTGCAGTAGTTCTCCAGCTGGT-3'
3. NOS Terminator
  - a. Forward:  
5'-(SphI)GATCGTTCAAACATTGGCAATAAAGTTCTTAAGA-3'
  - b. Reverse: 5'-(HindIII)GATCTAGAACATAGATGACACCGCGC-3'

### **Phase 2:**

After the PCR was completed, we performed electrophoresis and virtual digest with the help of benchling. The results of the virtual digest are as follows:

Enzymes		Cuts	Temp.	1.1	2.1	3.1	4/CS
BamHI		1	37°C	75*	100*	100	100*
Start	End	Length	Left Cutter	Left Overhang	Right Cutter	Right Overhang	
1	5	5	None	blunt	KpnI	3'	
6	1066	1061	KpnI	3'	BamHI	5'	
1067	1071	5	BamHI	5'	None	blunt	

Figure 3.6.3 Virtual digest psbA Promoter with BamHI and Kpn I enzymes (Benchling)

Enzymes		Cuts	Temp.	1.1	2.1	3.1	4/CS
BamHI		1	37°C	75*	100*	100	100*
Start	End	Length	Left Cutter	Left Overhang	Right Cutter	Right Overhang	
1	1	1	None	blunt	BamHI	5'	
2	344	343	BamHI	5'	SphI	3'	
345	345	1	SphI	3'	None	blunt	

Figure 3.6.4 Virtual digest of Proinsulin with BamHI and SphI enzymes (Benchling)

Enzymes		Cuts	Temp.	1.1	2.1	3.1	4/CS
HindIII		1	37°C	25	100	50	50
Start	End	Length	Left Cutter	Left Overhang	Right Cutter	Right Overhang	
1	5	5	None	blunt	SphI	3'	
6	260	255	SphI	3'	HindIII	5'	
261	265	5	HindIII	5'	None	blunt	

Figure 3.6.5 Virtual digest NOS Terminator with HindIII and SphI enzymes (Benchling)

Results of virtual electrophoresis:

Ladder	Life 1 kb Plus
1	psbA Promoter - BamHI KpnI
2	Optimized Human Insulin - BamHI SphI
3	NOS Terminator - HindIII SphI

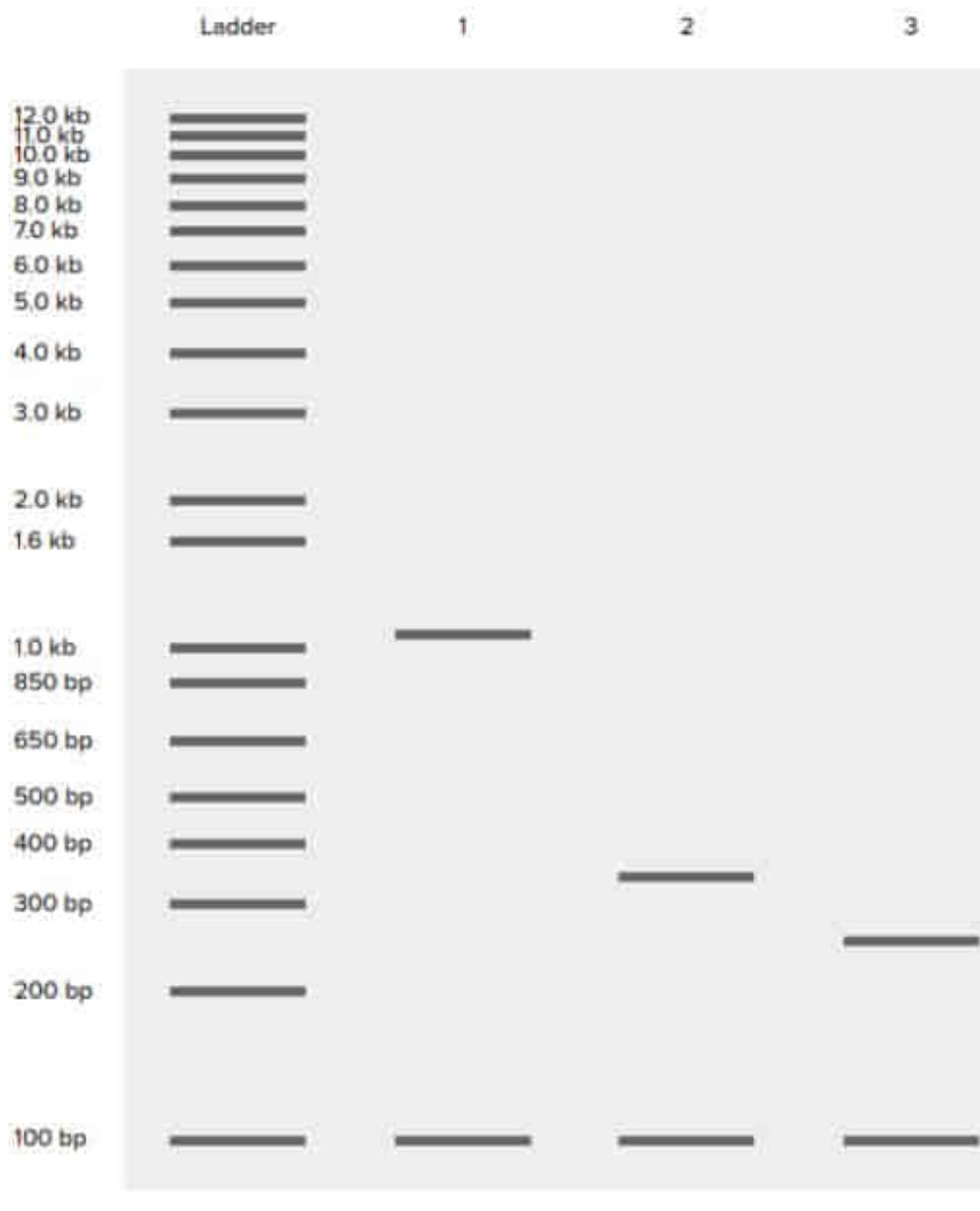


Figure 3.6.6 Virtual electrophoresis of the psbA promoter, insulin, NOS (benchling)

### Phase 3:

After virtual digest and electrophoresis were completed, we entered the three sequences into pCAMBIA using insertion cloning with the help of snapgene software.

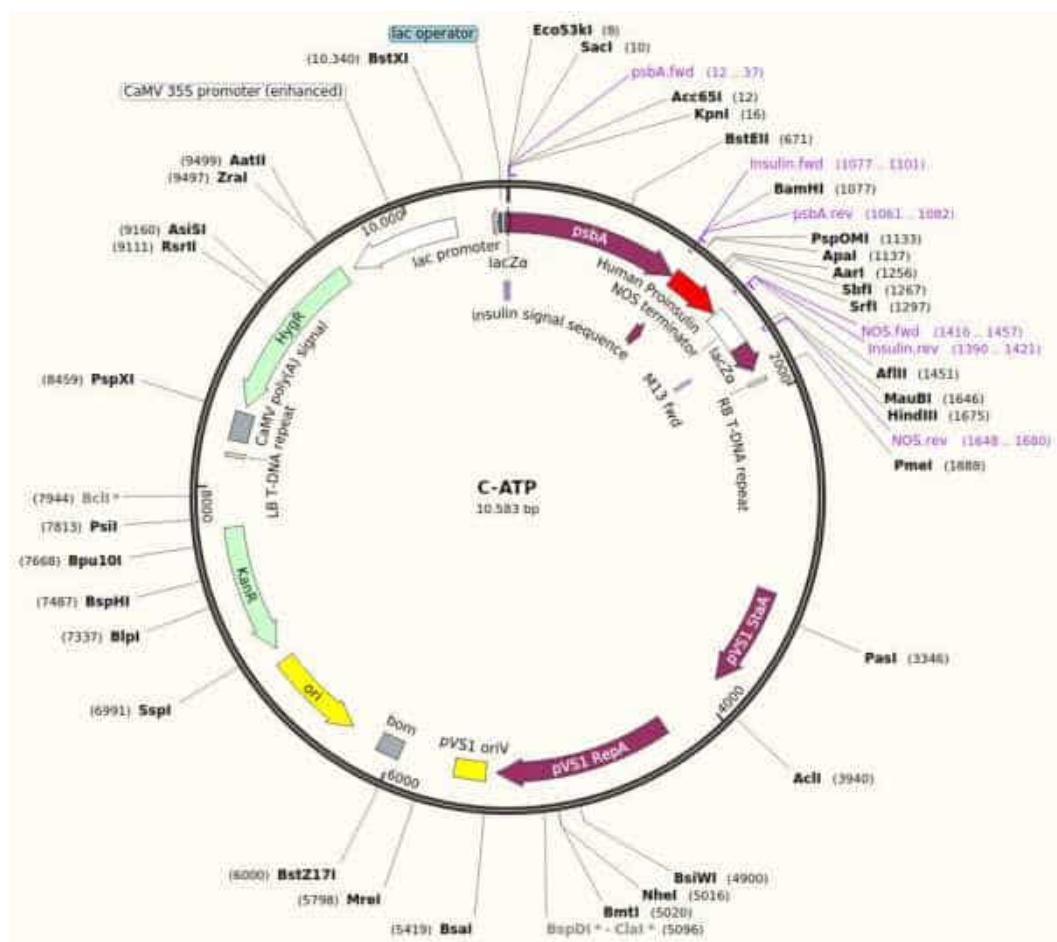


Figure 3.6.7 The final plasmid from insertion cloning (Benchling)

### 3.6.5 Protein Visualization

After the final plasmid is formed, the next step is to visualize the protein into a 3D model for further analysis using a computer. For visualization, we will use the rosetta software which is available online at the rosetta.bakerlab.org website. The 3D model will then be analyzed using PyMol<sup>[205-207]</sup>.

## Bab 4

### Results and Discussion

#### 4.1 Transgene Construction

##### 4.1.1 Codon Optimization

Codon optimization is done to increase the yield of expressions. The genetic code consists of 64 different codons coding for 21 types of amino acids and several stop codons, this means that there are several different codons coding for the same amino acid. Therefore, even though the amino acid yield is the same, each species tends to use a certain codon, compared to other codons. This is what causes Codon Usage Bias, which is the difference in the frequency of synonym codons (codons encoding the same amino acid) in DNA. For example, leucine acid is coded by 6 different codons, some of which are very rare. By adjusting the use of the codons based on the species preference, the leucine codons that are rarely used are replaced with the leucine codons that are used more frequently. This is thought to increase the yield of protein expression<sup>[197]</sup>.

Even so, until now, there is still no method that can predict the results of expressions accurately. Although codon optimization can improve the expression yield, it does not guarantee a successful expression, and the increase in the expression yield by codon optimization will vary depending on the protein type and the organism. Also, many other factors influence the expression results such as tRNA copy<sup>[198]</sup>, mRNA stability<sup>[199]</sup>, protein folding kinetics<sup>[200]</sup>, protein stability, protein transport, protein toxicity in the environment of cell expression, and various other factors. Therefore, the codon optimization results need to be verified by experiment.

We used DnaChisel to perform codon optimization, and after that, we performed pairwise alignment with the Smith-Waterman algorithm provided by EMBL-EBI<sup>[201]</sup>.

Sequence alignment before and after optimization:

NP_001278826.	1 ATGGCCCTGTGGATGCGCCTCTGCCCTGCTGGCGCTGCTGGCCCTCTG        . .   .       .	50
NP_001278826.	1 ATGGCCCTGTGGATGAGGCTGCTGCCCTGCTGGCCCTGCTGGCCCTGTG        . .   .       .	50
NP_001278826.	51 GGGACCTGACCCAGCCGCAGCCTTGTGAACCAACACCTGTGCGGCTCAC      .   .     .       .       .	100
NP_001278826.	51 GGGCCCGAACCCGCCGCCGCTTGTGAACCAGCACCTGTGCGGCTCTC        .       .       .	100
NP_001278826.	101 ACCTGGTGGAAAGCTCTACCTAGTGTGCGGGGAACGAGGCTTCTTCTAC        .   .   .     .       .   .	150
NP_001278826.	101 ACCTGGTGGAGGCCCTGTACCTGGTGTGCGGGAGAGGGGCTTCTTCTAC        .       .       .   .	150
NP_001278826.	151 ACACCCAAGACCGCCGGGAGGGCAGAGGACCTGCAGGTGGGGCAGGTGGA    .       .   .   .     .       .       .	200
NP_001278826.	151 ACGCCAAGACGAGGGAGGGAGGCCGAGGACCTGCAGGTGGGCCAGGTGGA        .       .       .       .	200
NP_001278826.	201 GCTGGCGGGGGCCCTGGTCAGGCAGCCTGCAGCCCTGGCCCTGGAGG        .     .   .   .    ...       .       .	250
NP_001278826.	201 GCTGGCGGCCGCCGGCGCCGGCTCTGCAGCCCTGGCCCTGGAGG        .       .       .       .	250
NP_001278826.	251 GGTCCCTGCAGAACAGCTGGCATTGTGAAACATGCTGTACAGCATCTGC    .       .   .     .       .   .     .   .	300
NP_001278826.	251 GCTCTCTGCAGAACAGGGGATAGTGAGCAGTGTGCACTGTATATGC        .       .       .       .	300
NP_001278826.	301 TCCCTCTACCAGCTGGAGAACTACTGCAACTAG    .   .       .       .	333
NP_001278826.	301 TCTCTGTACCAAGCTGGAGAACTACTGCAACTAA        .       .	333

Figure 4.1.1 Pairwise alignment of proinsulin sequences before and after codon optimization

#### Alignment of translation results before and after optimization:

NP_001278826.	1 MALWMRLPLALLALWGPDPAAFVNQHLCGSHLVEALYLVCGERGFFY        .	50
NP_001278826.	1 MALWMRLPLALLALWGPDPAAFVNQHLCGSHLVEALYLVCGERGFFY        .	50
NP_001278826.	51 TPKTRREAEDLQVGQVELGGPGAGSLQPLALEGSLQKRGIVEQCCTSIC        .	100
NP_001278826.	51 TPKTRREAEDLQVGQVELGGPGAGSLQPLALEGSLQKRGIVEQCCTSIC        .	100
NP_001278826.	101 SLYQLENYCN*        .	111
NP_001278826.	101 SLYQLENYCN*        .	111

Figure 4.1.2 Pairwise alignment of translation results before and after codon optimization

Based on the picture above, we can see that there is a difference between the proinsulin sequence before and after the codon optimization, but the translation results remain the same. Therefore, we conclude that the proinsulin sequence can be expressed in Aloe vera cells. However, since there is currently no method that can predict the optimization result accurately, further experiments are needed to ensure the success of the expression.

#### **4.1.2 Virtual Digest and Virtual Electrophoresis**

Virtual digest was carried out in order to simulate the cutting at the restriction site by restriction enzymes during transgene construction. The virtual digest will predict the virtual cutting location, as well as the length of the cut. After that, we performed virtual electrophoresis to obtain data on the size of each sequence that could be used as a comparison in in vivo laboratory tests.

#### **4.1.3 Transgenic Plasmids**

Next, we combine the three desired sequences into the plasmid. The plasmid we chose was pCAMBIA 1300 Cloning Vector. We chose pCAMBIA because it is one of the most commonly used plasmids, so it is more laboratory tested than other plasmids. Then, we chose pCAMBIA 1300 because it does not have the toxic GUS reporter gene, and it also has the hygromycin resistance and kanamycin resistance genes, which code for resistance to hygromycin and kanamycin.

Then we entered the three genes of interest into the multiple cloning site (MCS) of the pCAMBIA 1300 clockwise, sequentially from the psbA promoter, proinsulin, and NOS terminator. Overall, the insertion process added 1659 base pairs, consisting of 1061 psbA promoter bases, 343 proinsulin bases, and 255 NOS terminator bases. In addition, insertion also removed as many as 35 base pairs from the restriction sites in the MCS that were truncated at insertion. In total, there was an increase of 1624 base pairs, from 8959 base pairs of pCAMBIA 1300 to 10,583 C-ATP base pairs.

Here's a comparison of pCAMBIA 1300 and plasmid C-ATP:

## pCAMBIA1300 (8959 bp)

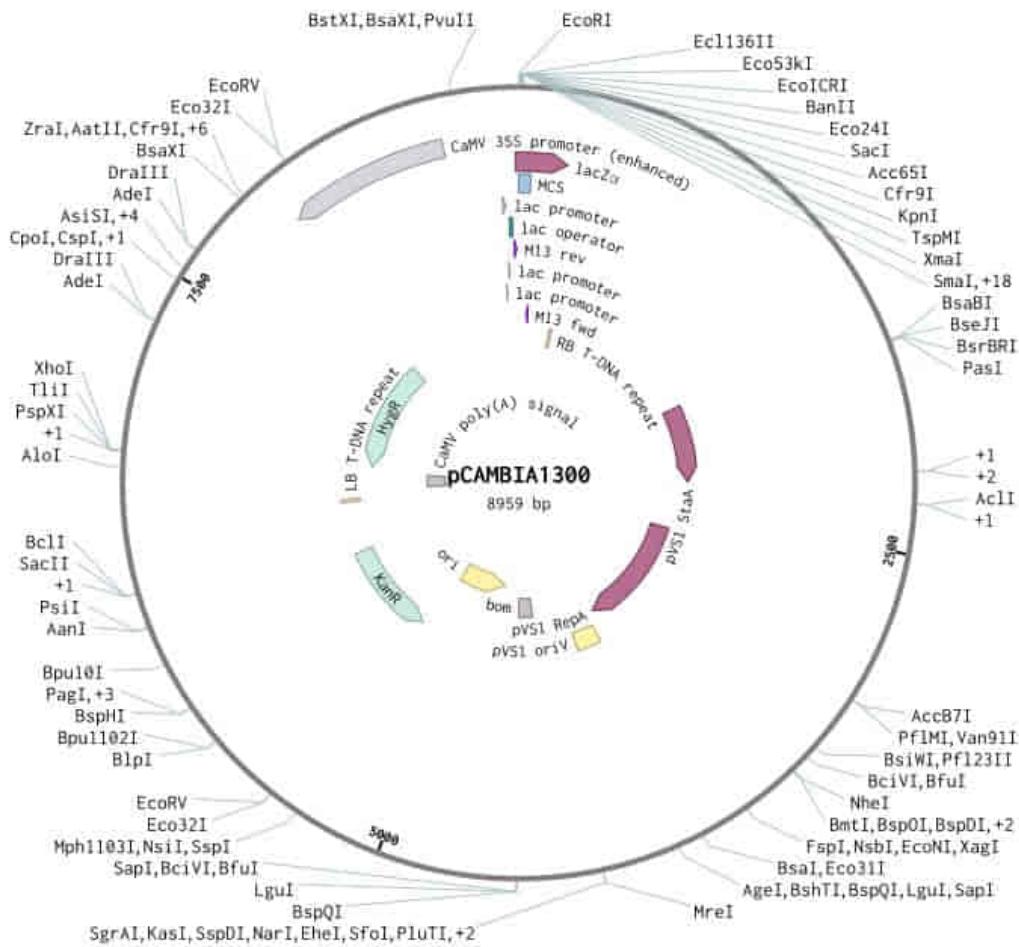


Figure 4.1.3 pCAMBIA1300 Cloning Vector

C-ATP (10583 bp)

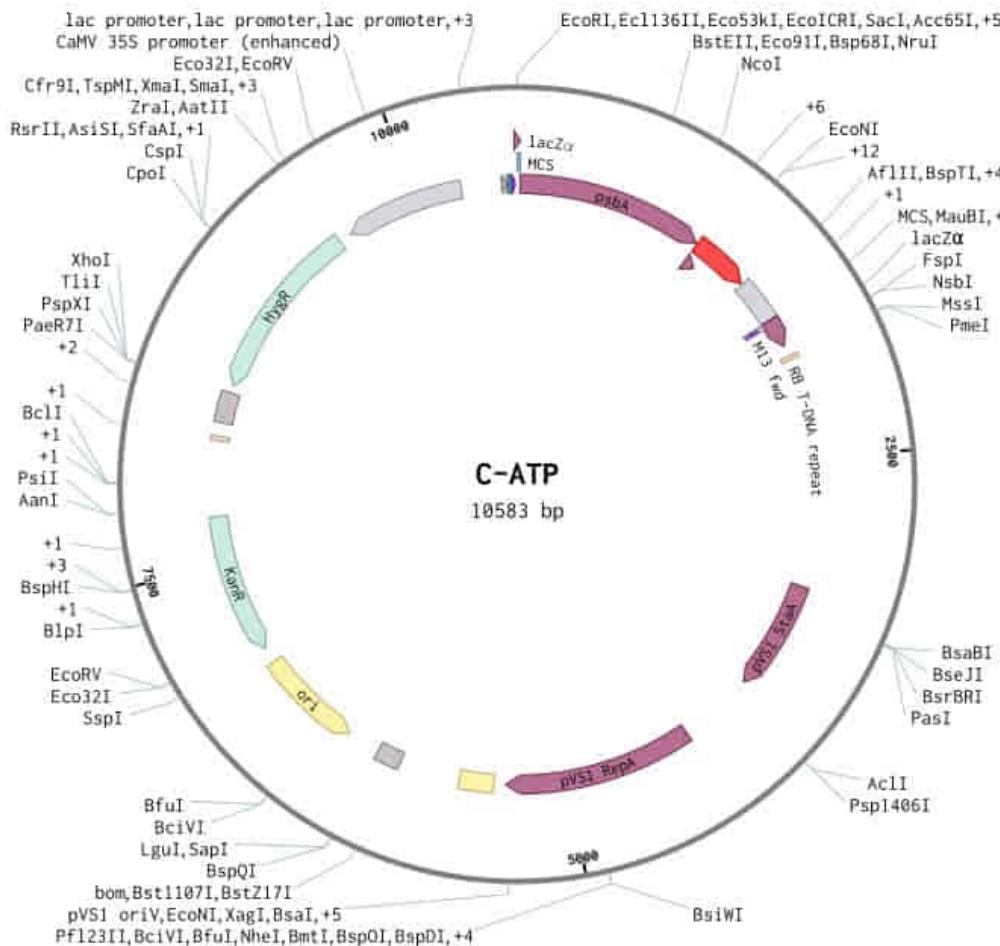


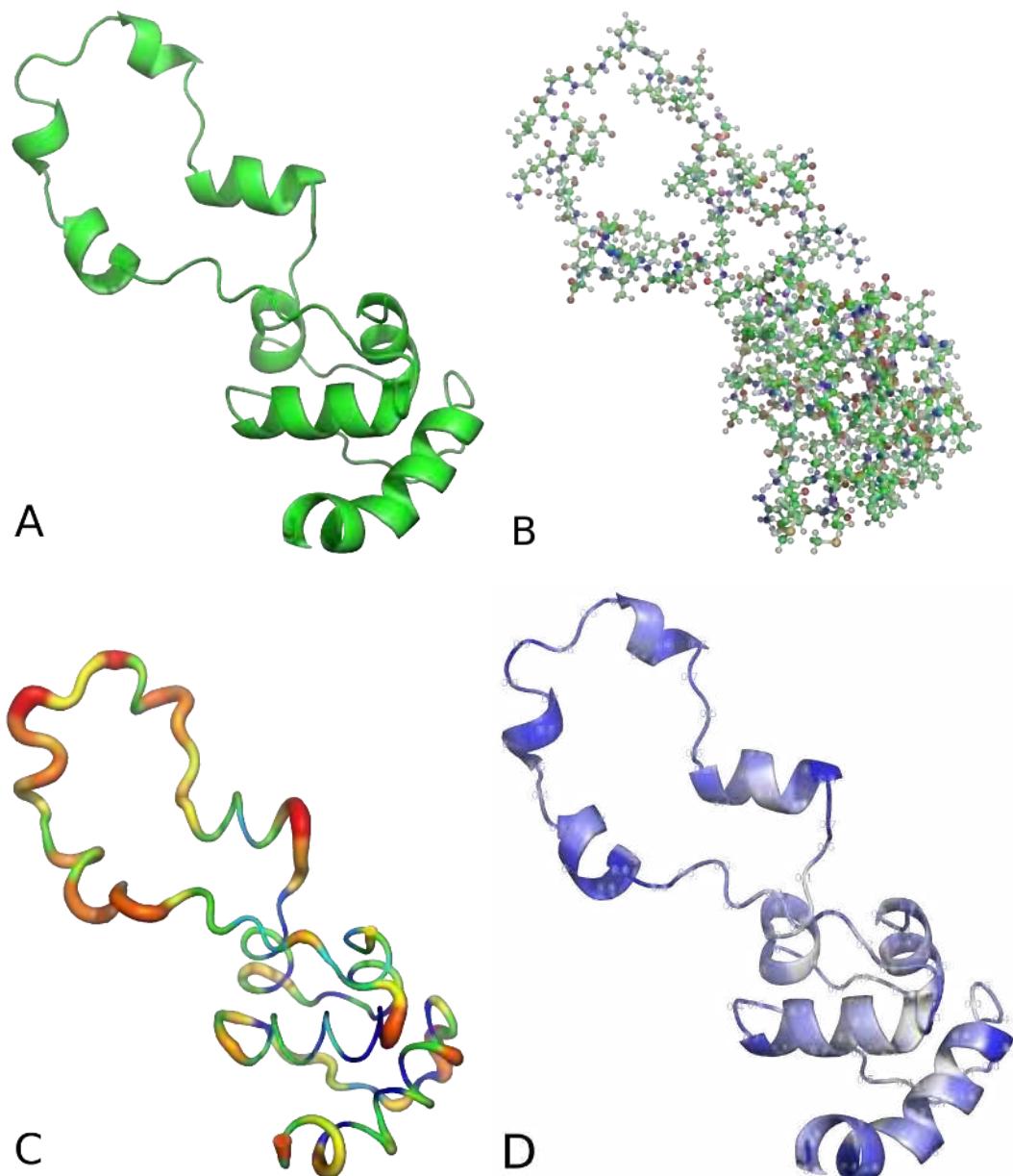
Figure 4.1.4 Plasmid C-ATP

## 4.2 Visualization

The next stage is the visualization of proteins. We use robeta software to generate a pdb model, which is then visualized using PyMol. It should be noted that currently, there is no method that can accurately visualize the protein in silico experiments, the robeta software itself uses the existing protein database and predicts the model based on similarities to other proteins. In addition, robeta is evaluated continuously by CAMEO<sup>[202-204]</sup>.

In total, there were 4 different predictive models for the structure of our transgenic insulin. We then analyzed the four models using PyMol<sup>[205-207]</sup>.

#### 4.2.1 Model 1



(Figure 4.2.1. A). Protein interface B). Ball and Stick model C). β-Factor D). Surface per residue)

Table 4.2.1 Data on the number of atoms and molecular weight model 1

Model	Atom Count	Molecular Weight
Full Model	1671 atom	11971.7169 u
Polimer	1671 atom	11971.7169 u
Organik	0 atom	0.0000 u

Solvent	0 atom	0.0000 u
Polar Hydrogen	173 atom	174.3736 u
Non-Polar Hydrogen	659 atom	664.2325 u
Donor	148 atom	2096.9040 u
Acceptor	154 atom	2461.9149 u
Surface Atom	648 atom	3610.7895 u
C-Alpha	110 atom	1321.1770 u

Table 4.2.2 Charge data model 1

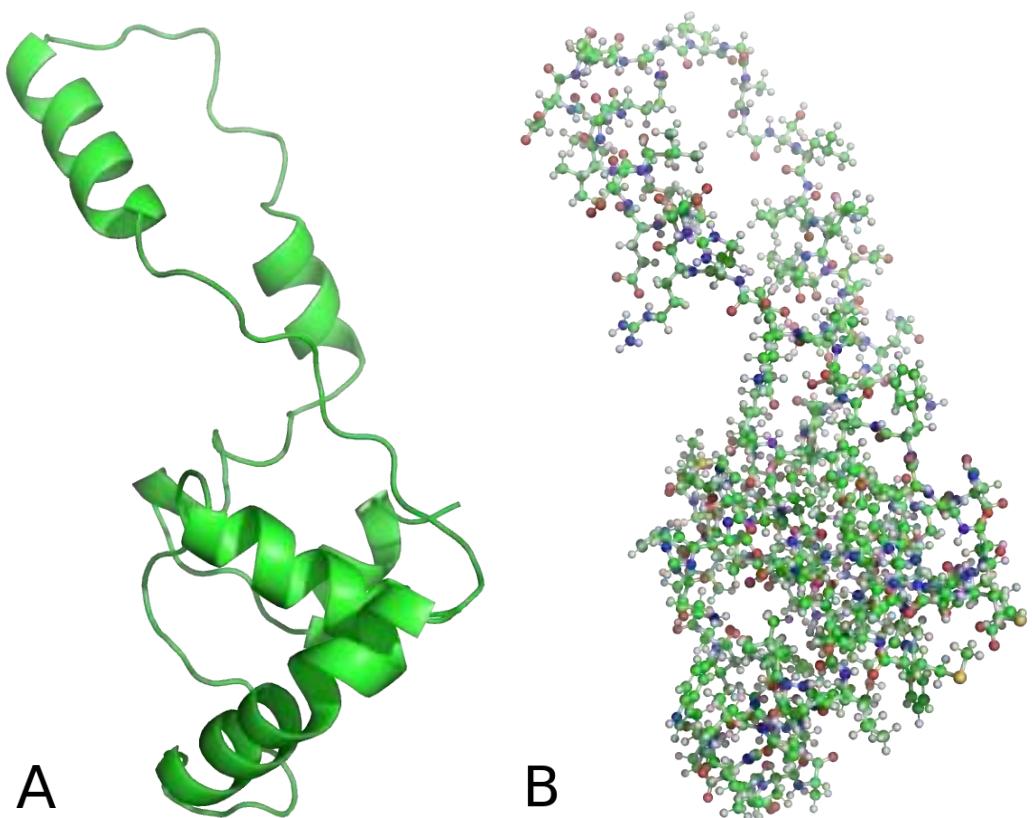
Model	Formal Charge	Partial Charge
Full Model	-4	0.0000
Polimer	-4	0.0000
Organik	0	0.0000
Solvent	0	0.0000
Polar Hydrogen	0	0.0000
Non-Polar Hydrogen	0	0.0000
Donor	7	0.0000
Acceptor	-11	0.0000
Surface Atom	-6	0.0000
C-Alpha	0	0.0000

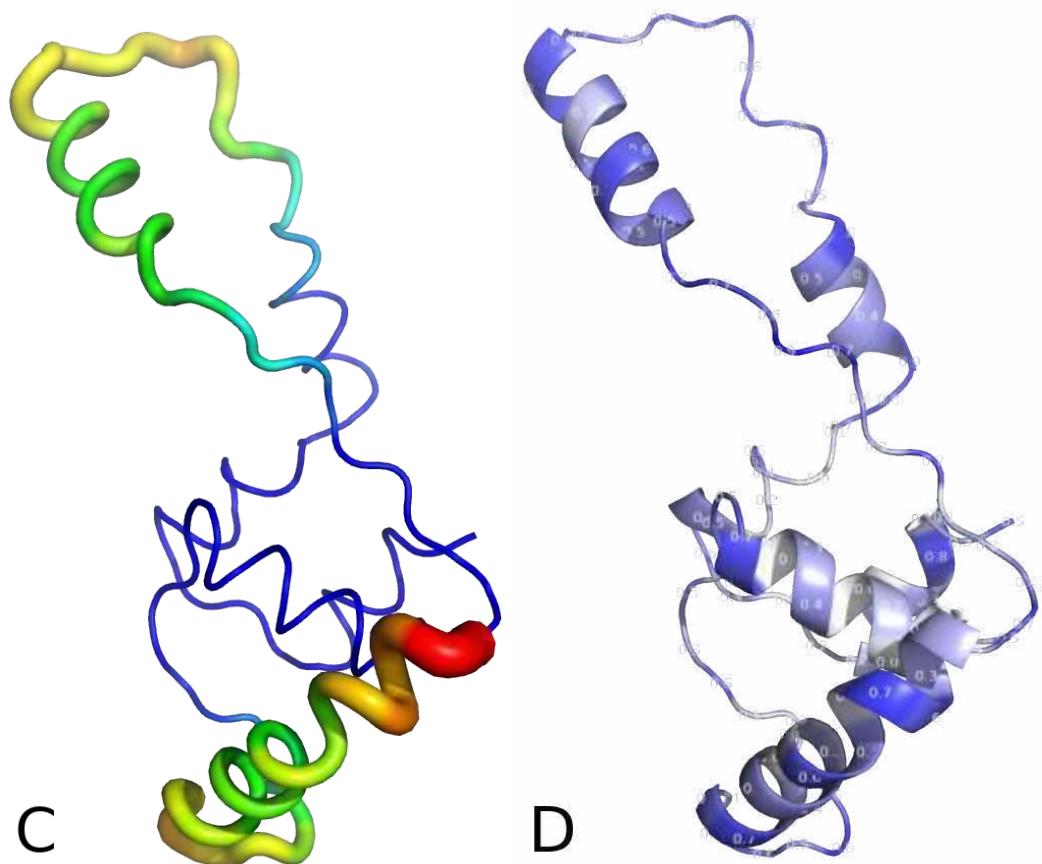
Table 4.2.3 Model surface area data 1

Model	Molecular Surface Area	Solvent Accessible Surface Area
Full Model	12442.948 Angstroms <sup>2</sup>	8148.517 Angstroms <sup>2</sup>
Polimer	12442.948 Angstroms <sup>2</sup>	8148.517 Angstroms <sup>2</sup>
Organik	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Solvent	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>

Polar Hydrogen	1147.248 Angstroms <sup>2</sup>	1182.374 Angstroms <sup>2</sup>
Non-Polar Hydrogen	5052.217 Angstroms <sup>2</sup>	4292.863 Angstroms <sup>2</sup>
Donor	841.471 Angstroms <sup>2</sup>	385.156 Angstroms <sup>2</sup>
Acceptor	2254.160 Angstroms <sup>2</sup>	1674.963 Angstroms <sup>2</sup>
Surface Atom	6046.378 Angstroms <sup>2</sup>	7797.255 Angstroms <sup>2</sup>
C-Alpha	332.974 Angstroms <sup>2</sup>	19.984 Angstroms <sup>2</sup>

#### 4.2.2 Model 2





(Gambar 4.2.2. A). Protein interface B). Ball and Stick model C).  $\beta$ -Factor D). Surface per residue)

Table 4.2.4 Data on the number of atoms and molecular weight model 2

Model	Atom Count	Molecular Weight
Full Model	1671 atom	11971.7169 u
Polimer	1671 atom	11971.7169 u
Organik	0 atom	0 u
Solvent	0 atom	0 u
Polar Hydrogen	173 atom	174.3736 u
Non-Polar Hydrogen	659 atom	664.2325 u
Donor	148 atom	2096.9040 u
Acceptor	154 atom	2461.9149 u
Surface Atom	668 atom	3602.9581 u
C-Alpha	110 atom	1321.1770 u

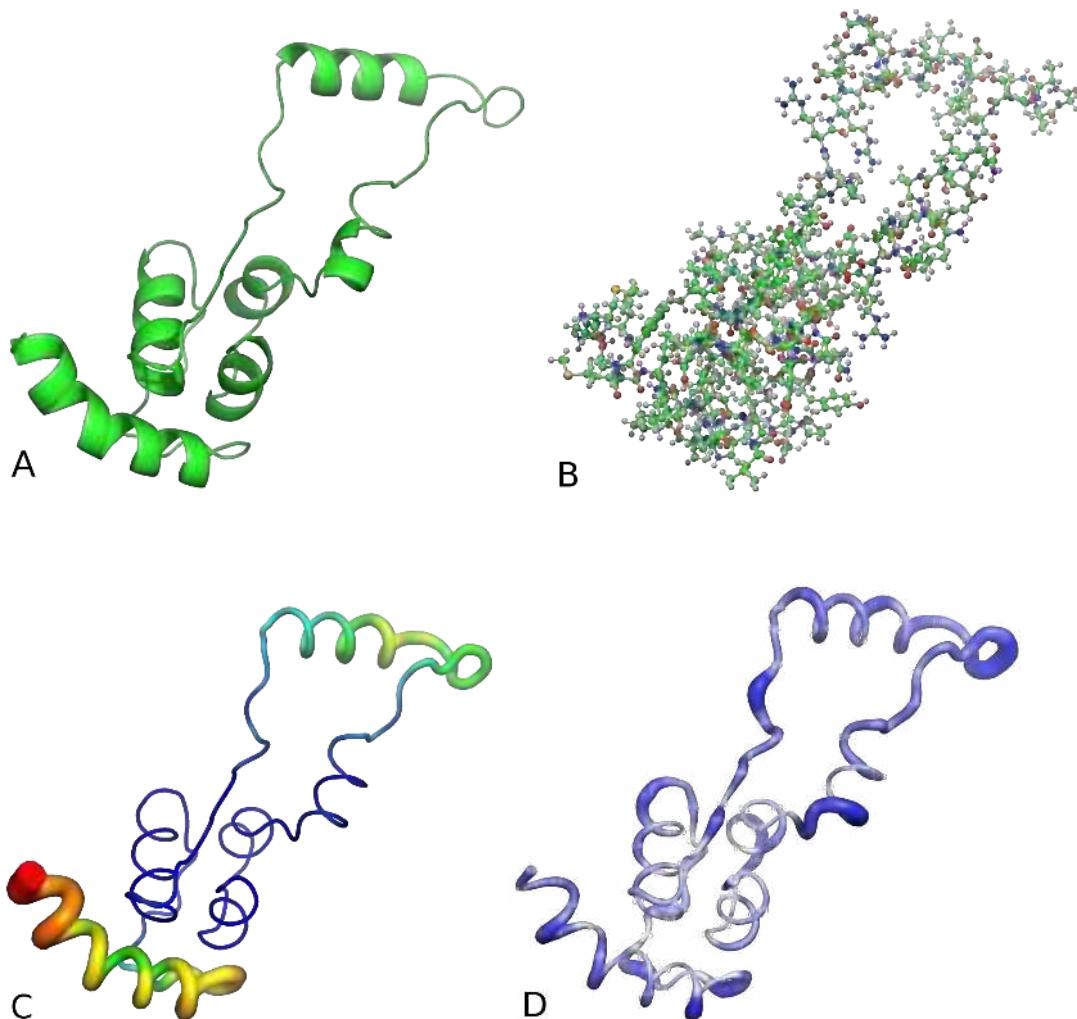
Table 4.2.5 Charge data model 2

Model	Formal Charge	Partial Charge
Full Model	-4	0.0000
Polimer	-4	0.0000
Organik	0	0.0000
Solvent	0	0.0000
Polar Hydrogen	0	0.0000
Non-Polar Hydrogen	0	0.0000
Donor	7	0.0000
Acceptor	-11	0.0000
Surface Atom	-6	0.0000
C-Alpha	0	0.0000

Table 4.2.6 Model surface area data 2

Model	Molecular Surface Area	Solvent Accessible Surface Area
Full Model	12401.562 Angstroms <sup>2</sup>	8540.121 Angstroms <sup>2</sup>
Polimer	12401.562 Angstroms <sup>2</sup>	8540.121 Angstroms <sup>2</sup>
Organik	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Solvent	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Polar Hydrogen	1159.096 Angstroms <sup>2</sup>	1150.314 Angstroms <sup>2</sup>
Non-Polar Hydrogen	5026.537 Angstroms <sup>2</sup>	4561.397 Angstroms <sup>2</sup>
Donor	825.722 Angstroms <sup>2</sup>	363.157 Angstroms <sup>2</sup>
Acceptor	2249.460 Angstroms <sup>2</sup>	1813.615 Angstroms <sup>2</sup>
Surface Atom	6216.088 Angstroms <sup>2</sup>	8228.303 Angstroms <sup>2</sup>
C-Alpha	331.766 Angstroms <sup>2</sup>	18.612 Angstroms <sup>2</sup>

### 4.2.3 Model 3



(Gambar 4.2.3. A). Protein interface B). Ball and Stick model C).  $\beta$ -Factor D). Surface per residue)

Table 4.2.7 Data on the number of atoms and molecular weight model 3

Model	Atom Count	Molecular Weight
Full Model	1671 atom	11971.7169 u
Polimer	1671 atom	11971.7169 u
Organik	0 atom	0 u
Solvent	0 atom	0 u
Polar Hydrogen	173 atom	174.3736 u
Non-Polar Hydrogen	659 atom	664.2325 u

Donor	147 atom	2082.8973 u
Acceptor	155 atom	2475.9216 u
Surface Atom	642 atom	3534.7366 u
C-Alpha	110 atom	1321.1770 u

Table 4.2.8 Charge data model 3

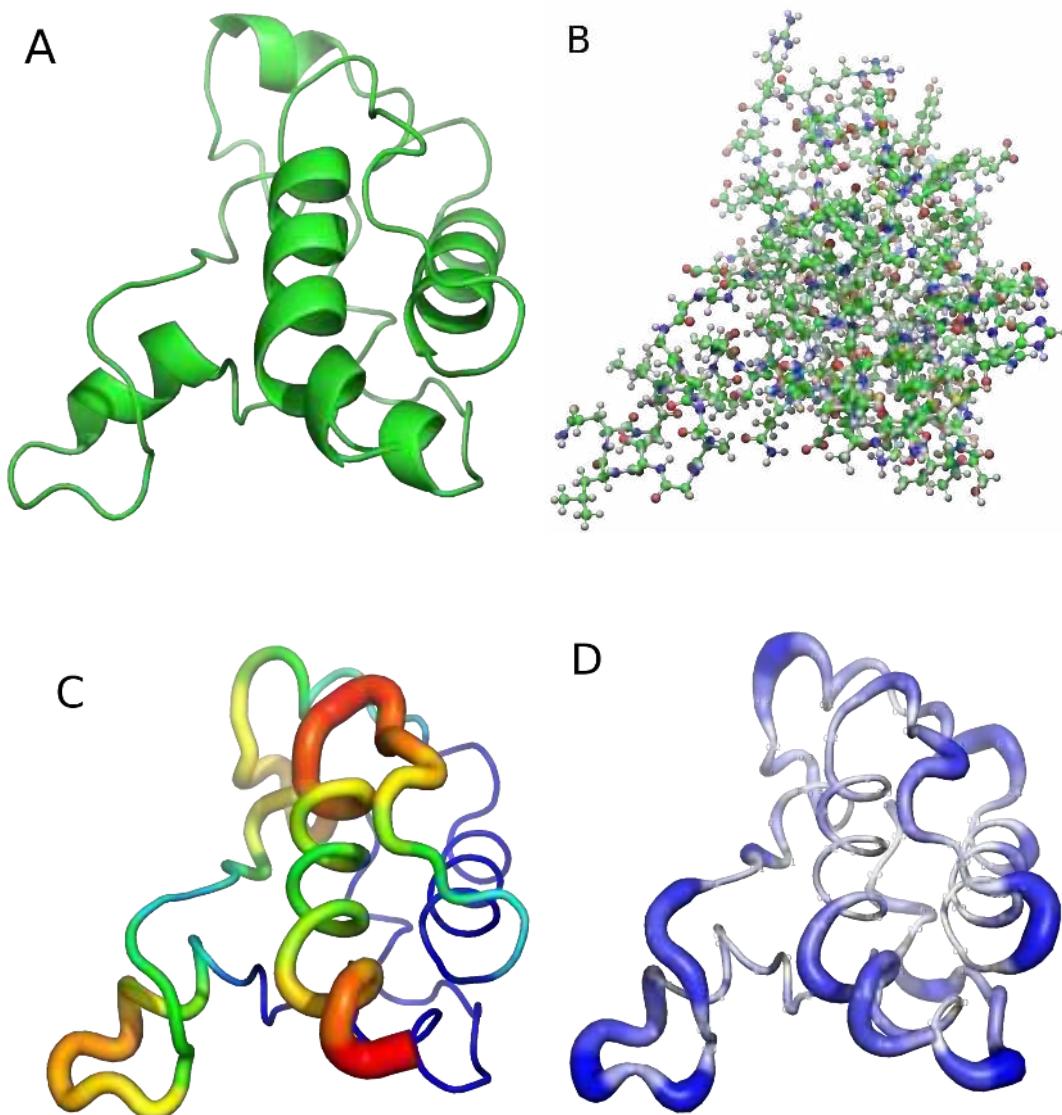
Model	Formal Charge	Partial Charge
Full Model	-4	0.0000
Polimer	-4	0.0000
Organik	0	0.0000
Solvent	0	0.0000
Polar Hydrogen	0	0.0000
Non-Polar Hydrogen	0	0.0000
Donor	7	0.0000
Acceptor	-11	0.0000
Surface Atom	-6	0.0000
C-Alpha	0	0.0000

Table 4.2.9 Model surface area data 3

Model	Molecular Surface Area	Solvent Accessible Surface Area
Full Model	12379.827 Angstroms <sup>2</sup>	7946.748 Angstroms <sup>2</sup>
Polimer	12379.827 Angstroms <sup>2</sup>	7946.748 Angstroms <sup>2</sup>
Organik	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Solvent	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Polar Hydrogen	1154.202 Angstroms <sup>2</sup>	1148.252 Angstroms <sup>2</sup>
Non-Polar Hydrogen	5025.376 Angstroms <sup>2</sup>	4295.495 Angstroms <sup>2</sup>
Donor	823.042 Angstroms <sup>2</sup>	321.154 Angstroms <sup>2</sup>

Acceptor	2236.014 Angstroms <sup>2</sup>	1606.604 Angstroms <sup>2</sup>
Surface Atom	5983.686 Angstroms <sup>2</sup>	7635.942 Angstroms <sup>2</sup>
C-Alpha	330.181 Angstroms <sup>2</sup>	16.935 Angstroms <sup>2</sup>

#### 4.2.4 Model 4



(Gambar 4.2.4. A). Protein interface B). Ball and Stick model C).  $\beta$ -Factor D). Surface per residue)

Table 4.2.10 Data on the number of atoms and molecular weight model 4

Model	Atom Count	Molecular Weight

Full Model	1675 atom	11975.7486 u
Polimer	1675 atom	11975.7486 u
Organik	0 atom	0 u
Solvent	0 atom	0 u
Polar Hydrogen	177 atom	178.4054 u
Non-Polar Hydrogen	659 atom	664.2325 u
Donor	149 atom	2110.9107 u
Acceptor	153 atom	2447.9082 u
Surface Atom	534 atom	3045.8878 u
C-Alpha	110 atom	1321.1770 u

Table 4.2.11 Charge data model 4

Model	Formal Charge	Partial Charge
Full Model	-4	0.0000
Polimer	-4	0.0000
Organik	0	0.0000
Solvent	0	0.0000
Polar Hydrogen	0	0.0000
Non-Polar Hydrogen	0	0.0000
Donor	7	0.0000
Acceptor	-11	0.0000
Surface Atom	-7	0.0000
C-Alpha	0	0.0000

Table 4.2.12 Model surface area data 4

Model	Molecular Surface Area	Solvent Accessible Surface Area
Full Model	12305.951 Angstroms <sup>2</sup>	6627.765 Angstroms <sup>2</sup>

Polimer	12305.951 Angstroms <sup>2</sup>	6627.765 Angstroms <sup>2</sup>
Organik	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Solvent	0.000 Angstroms <sup>2</sup>	0.000 Angstroms <sup>2</sup>
Polar Hydrogen	1135.465 Angstroms <sup>2</sup>	1096.120 Angstroms <sup>2</sup>
Non-Polar Hydrogen	4962.301 Angstroms <sup>2</sup>	3276.135 Angstroms <sup>2</sup>
Donor	831.292 Angstroms <sup>2</sup>	340.515 Angstroms <sup>2</sup>
Acceptor	2215.719 Angstroms <sup>2</sup>	1408.931 Angstroms <sup>2</sup>
Surface Atom	5043.382 Angstroms <sup>2</sup>	6346.040 Angstroms <sup>2</sup>
C-Alpha	335.935 Angstroms <sup>2</sup>	10.105 Angstroms <sup>2</sup>

### 4.3 In-Silico Research Continuation Plan

Future in-silico research will focus more on the analysis and simulation of proteins while in the human body. One way that might be done is docking, a method used to predict the orientation of a molecule to other molecules when they bind to form stable complexes<sup>[211]</sup>. We haven't done docking yet because it is beyond our problem limits. After all, our problem limitation is reaching the insulin production design in Aloe vera only.

### 4.4 In-Vivo Experimental Design

The next step after completing the in-silico experiment, is to conduct in-vivo experiments, because the results of the in-silico experiments are just predictions that can be used as a guide and comparison for in-vivo and in-vitro experiments, so that the in-vivo or in-vitro is needed to ensure the results of in-silico. Due to time constraints and current conditions that do not allow in-vivo experiments to be carried out, we only make designs, as a guide for the future.

There are a variety of different factors that can influence the final outcome of an experiment. This factor is divided into two, namely biological factors and environmental factors. Biological factors that need to be considered are the construction of genes in cloning vectors, such as promoters, genes of interest, terminators, gene reporters, and marker genes, each of which is important for the

transformation process as stated in chapter 2.8. Other biological factors include tissue type, cell size, culture age, mitotic phase, and so on.

While environmental factors include temperature, humidity, and light intensity, each of which has a direct effect on the physiological tissue culture. Moisture is also important in the preparation and firing of the microcarrier. Too high humidity can cause the microcarriers to clump, thereby reducing the rate of transformation.

In this experiment, we will insert the plasmid we have assembled, as described in chapter 4.1, into the genes for the Aloe vera chloroplast. As for the transformation method, we will use the biolistic method or gene gun. We need to emphasize again that the procedure below is still a rough design based on research studies concerning the transformation of Aloe vera with particle bombardment<sup>[150]</sup>.

Our experimental design is as follows:

- 1) Plant cell preparation and culture
- 2) Preparation of DNA coated microcarriers
- 3) Particle bombardment
- 4) Embryo selection, germination, and regeneration
- 5) Extraction, purification and analysis of insulin

#### **4.4.1 Culture and Preparation**

Aloe callus was initiated from seeds (Thompson and Morgan, UK). Seeds were surface sterilized in 70 % (v/v) ethanol for 30 s followed by 10 min in dilution of bleach for an effective concentration of 1 % (v/v) sodium hypochlorite and then soaked overnight in a solution of 0.5 % (v/v) plant preservative medium. Isolated embryos were plated on MS media supplemented with 0.25 % (w/v) gelrite, 3 % (w/v) sucrose, 0.01 % (v/v) PPM, 300 mg/L myo-inositol, 2 mg/L NAA, 0.2 mg/L BA, 0.01 mM AgNO<sub>3</sub> and 2 mM CaCl<sub>2</sub>. Adjust to pH 5.4 post autoclave. Isolated embryos were grown in the dark at 25 C. Developing shoots and roots were removed. Friable callus was selected after 4 weeks. Somatic embryogenesis was initiated on MS media containing 0.2 mg/L NAA, 0.2 mg/L BA without AgNO<sub>3</sub> and exposed to 14 h light, 10 h dark cycle provided by cool white fluorescent tubes. Developing shoots were transferred to MS media with 0.2 mg/L NAA to develop roots in magenta vessels, hardened and transferred to soil.

#### **4.4.2 DNA *Coated Microcarrier Preparation***

##### **Transgene Preparation**

The transgene was prepared in the same way as described in chapter 3.6. Genetic material can be ordered from various supply companies for biotechnology, or it can be isolated independently or synthesized.

##### **Gold Particle and *Microcarrier Preparation***

Steps to prepare DNA coated microcarrier is as follow<sup>[208]</sup>:

- 1) Twenty-five-milligram gold particles were placed in a 1.5-ml microfuge tube to which was added 50 ml of 0.05 M spermidine and 50 ml of DNA (1 mg/ml) for transfection.
- 2) The tube was vortexed for 10 s, and then, whilst still vortexing, 50 ml of 1 M CaCl was added in a dropwise manner.
- 3) The gold/DNA suspension was incubated at room temperature for 5 min with brief pulse vortexing every 30 s, and then centrifuged at 3000 rpm for 10 s.
- 4) The supernatant was removed and the gold pellet resuspended in 200 ml polyvinyl-pyrrolidine (PVP; 0.075 mg/ml in ethanol) from a 3.5-ml aliquot, and transferred to a 5-ml sterile polypropylene tube. As the transfer is inefficient this procedure was repeated with 200-ml aliquots of PVP solution until all the gold particles had been transferred (4–6 aliquots), and then the remaining PVP solution added to the tube.
- 5) The gold/PVP solution was vortexed briefly to ensure an even suspension of gold. It was then drawn up with a 10-ml syringe into a 75-cm piece of Tefzel tubing; this had previously been dried in the Tubing preparation station with nitrogen (0.3–0.4 LPM) for 10 min.
- 6) The loaded tube was inserted into the Tubing Prep Station with the syringe still attached, and the slurry immediately removed slowly from the Tefzel tubing using the syringe.
- 7) The syringe was detached from the Tefzel tubing, and the tubing rotated for 30–40 s.

- 8) The gold/DNA in the tubing was dried using a flow of nitrogen (0.3–0.4 LPM) for 5–10 min, cut to desired length (1 cm for Bio-Rad Gene Gun), and stored desiccated at 4 °C until required.

#### **4.4.3 Particle Bombardment**

After the microcarrier has been prepared, the next step is to carry out particle bombardment. The general steps for performing particle bombardment are as follows<sup>[180]</sup>:

Time	Activity (in sequential steps)
<b><i>Week prior to bombardment</i></b>	
(-) 6 d	Sterilize supplies (Whatman and Sharkskin filter papers, funnels, flasks, water, etc.).
(-) 5 d	Prepare media needed for transformation procedure, GM+NOA suspension culture medium. 1/2 MS-HF bombardment medium with osmotica. 1/2 MS-HF medium without osmotica. 1/2 MS-HF selective medium.
(-) 4 d	Subculture or refresh medium of embryogenic cell suspensions.
<b><i>Week of bombardment</i></b>	
(-) 1 d	Set gene gun parameters (distances as described in Fig. 2). Weigh gold particles (microcarriers) and place in an oven overnight. Sterilize macrocarriers, holders and stopping screens. Assemble macrocarriers into holders.
Key d	Bombardment day (suggested day, Tuesday). Examine embryogenic cell suspension for contamination using a microscope. Prepare cells on filter paper for bombardment. Sterilize microcarriers. Coat microcarriers with DNA. Bombard cells. Incubate cells in the dark at $23 \pm 1^\circ\text{C}$ .
(+) 1 d	Transfer cells to medium without osmotica. First transfer approx 16 h after bombardment. Second transfer approx 24 h after bombardment.
(+) 2 d	Transfer cells to selective medium. Analysis of reporter gene (i.e., GUS assay) for transient expression.
(+) 3 d	Examine GUS-positive blue spots per filter paper.
<b><i>Postbombardment weeks</i></b>	
(+) 30 d	Transfer cells to fresh selective medium. Reporter gene assay for transient expression.
(+) 60 d	Check plates for development of embryos. Transfer embryos to germination medium. Transfer remaining cells to fresh selective medium. Reporter gene assay for long-term expression.
(+) 90 d	Items and procedure as in (+) 60 d. Transfer germinated embryos to plant growth medium.

(Figure 4.4.1 Transformation flowchart using particle bombardment in general)

### ***Aloe vera Transformation***

Aloe callus were treated for 4 h in osmotic medium (MS media supplemented with 36.4 g/L sorbitol and 36.4 g/L mannitol as recommended by et al.<sup>[209]</sup>, before being transformed with the Biolistic PD-1000/He

particle delivery system (BioRad, Hercules, CA) using 1 nm Au particles, 1,100 psi Helium pressure and 9 cm target distance. Gold particles were coated with plasmid DNA with the protocol previously described. The expression vector was constructed by cloning the PCR amplified maize ubiquitin promoter as was previously described by Christensen and Quail<sup>[210]</sup> to pCAMBIA1300<sup>[150,160]</sup>.

Cells were returned to osmotic medium (further supplemented with 200 mg/ L glutathione and 1 mM DTT) for 1 h before being transferred to MS supplemented with antioxidants. Within 2 weeks cells were transferred to MS media supplemented with 50 mg/L Kan for selection once secondary globules began to develop. Cells were transferred to regeneration media and developing shoots were isolated and assayed for transgene expression<sup>[150]</sup>.

#### **4.4.4 Embryo Selection, Germination, and Regeneration**

After the transformation is complete, the next step is to identify the transgenic plants. One way to identify transgenic plants is to use a medium containing antibiotics. The antibiotic is thought to kill the cell culture that grows in the medium. The transgene that is wanted to be developed usually has the genes coding for resistance to this antibiotic, in our case hygromycin (positions 8498-9187 forward) and kanamycin (positions 6936-6980 forward).

The transformed plants are then acclimatized in preparation for transplanting to the ground. After the acclimatization is complete, the plants are then transferred to the ground and grown until they are ready for harvest. The harvested insulin will then be analyzed and tested further.

#### **4.4.5 Extraction, Purification, and Analysis**

##### **Extraction and Purification**

Protein extraction can be done from the total biomass or certain tissues. Extraction can use physical methods and/or chemical methods. Extraction methods can involve ultracentrifugation, chromatography, etc. Further explanation regarding protein extraction and purification can be seen in chapter 2.3.6

## **Insulin Analysis**

Several methods that can be used to analyze protein are western blot and ELISA, which we have described in chapters 2.16 and 2.17.

## Bab 5

### Conclusions and Recommendations

#### 5.1 Conclusions

*Aloe vera* is a succulent plant that has the potential to be used as material for transgenic plant research. The physiology of *Aloe vera* which allows it to grow in a variety of climates, does not require a lot of water, and has a fast production time, coupled with *Aloe vera*'s popularity in the world of cosmetics, makes it an ideal plant for mass production. *Aloe vera* also offers various advantages, including ease in the cultivation and extraction process so that it becomes more economical, low risk of contamination by toxic substances, and requires relatively short time with high production yields, thus increasing productivity, especially in tropical climates like Indonesia.

On the other hand, diabetes is one of the top 10 causes of death in the world and is the third-largest cause of death in Indonesia<sup>[1]</sup>. Currently, the number of diabetics has increased significantly, and with insulin as the only effective treatment technique, it is predicted that the demand for insulin will increase, and this can lead to insulin scarcity.

Therefore we researched in silico regarding transgenic *Aloe vera* and made a design for *Aloe vera* which can produce insulin. The procedure we use has been described in chapter 3, and we have discussed it in chapter 4. Based on our results, we conclude that *Aloe vera* has the possibility to be used as an insulin producer. We conclude this from the data showing that the translation results of *Aloe vera* insulin are the same as human insulin, as we have described in chapter 4.1. Even so, the results of our in silico experiment are still in the form of predictions that can be used as a guide and comparison of in vivo experiments, so further research and experiments in the real world are still needed to determine the ability of *Aloe vera* as an insulin producer.

#### 5.2 Recommendations

The next in silico research can begin to focus on examining the performance of transgenic insulin produced in *Aloe vera* in the human body, while the next in

vivo research may begin to develop transgenic aloe vera which can produce insulin.

## Reference

- [1] Diabetes atlas, international Diabetes Federation, 2019
- [2] Dong YM, Zhu SQ. Shanghai: Aloe Refinery Tech and Its Rules. Shanghai: Shanghai Science-Tech Press, 2002 (in Chinese).
- [3] Gu YX, Zhu SZ. Plant and Application of Aloe. Shanghai: Shanghai Science-Tech Press, 1999 (in Chinese).
- [4] Shi XJ. Discussion on aloe and its exploitation. Southwest Horticulture, 2001, 29: 14–17.
- [5] Sustiprijatno. Jagung Transgenik dan Perkembangan Penelitian di Indonesia. Balai Besar Penelitian dan Pengembangan Bioteknologi dan Sumberdaya Genetik Pertanian,  
Bogor.[http://balitsereal.litbang.pertanian.go.id/wp-content/uploads/2016/11/duade\\_lapan.pdf](http://balitsereal.litbang.pertanian.go.id/wp-content/uploads/2016/11/duade_lapan.pdf). Diunduh pada 22 April 2020.
- [6] Cells and beyond Austin, 2016.  
<https://cellsandbeyond.wordpress.com/2016/04/28/using-e-coli-for-insulin-production/>.
- [7] Francisco Pouci, 2017. Aloe vera for Tissue Engineering Application. Journal of Functional Biomaterials. DOI: [10.3390/jfb8010006](https://doi.org/10.3390/jfb8010006)
- [9] Bowen RA, Austgen L, Rouge M. 2002. Biology and Activity of Restriction Endonucleases. [http://www.vivo.colostate.edu/hbooks/genetics/biotech/enzymes/re\\_enzymes.html](http://www.vivo.colostate.edu/hbooks/genetics/biotech/enzymes/re_enzymes.html) [10 Januari 2020].
- [10] Pingoud A, Jeltsch A. 2001. Structure and function of type II restriction endonucleases. Nucleic Acids Res 29(18): 3705-27
- [11] Molecular Biology of the Cells, Bruce Albert, 6<sup>th</sup> edition, 2015
- [12] add gene. <https://www.addgene.org/mol-bio-reference/promoters/>
- [13] Molecular Biology of the Cells, Bruce Albert, 6<sup>th</sup> edition, 2015
- [14] Rogers, K., ed. (2011), New Thinking about Genetics, New York: Britannica Educational Publishing, hlm. 132
- [15] Glick, B.R., Pasternak, J.J., Patten, C.L. (2010). Molecular Biotechnology: Principles and Applications of Recombinant DNA (edisi ke-4). Washington, DC: ASM Press. hlm. 117–118.
- [16] Sweet, K.M., Michaelis, R.C. (2011). The Busy Physician's Guide to Genetics, Genomics and Personalized Medicine. Dordrecht: Springer. hlm. 76.

- [17] Bieber, F.R. (2004), "Science and Technology of Forensic DNA Profiling: Current Use and Future Directions", dalam Lazer, D. (Penyunting), DNA and the Criminal Justice System: the Technology of Justice, Cambridge, MA: MIT Press, hlm. 30.
- [18] Maxam AM, Gilbert W (February 1977). "[A new method for sequencing DNA](#)". Proc. Natl. Acad. Sci. USA. 74 (2): 560–64. Bibcode:1977PNAS...74..560M. DOI:[10.1073/pnas.74.2.560](https://doi.org/10.1073/pnas.74.2.560). PMC [392330](#). PMID [265521](#).
- [19] Sanger F; Coulson AR (May 1975). "A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase". J. Mol. Biol. 94 (3): 441–8. DOI:[10.1016/0022-2836\(75\)90213-2](https://doi.org/10.1016/0022-2836(75)90213-2). PMID [1100841](#).)
- [20] Westermeier. 2004. Electrophoresis in Practice: A Guide to Theory and Practice. New Jersey: John Wiley & Sons inc.  
Yuwono T. 2005. Biologi Molekuler. Jakarta: Erlangga.
- [21] U. Albrecht et al., J. Gen. Virol. 75:3353-3363, 1994; C, courtesy of Leander Lauffer and Peter Walter; D, from D. Vollrath and R.W. Davis, Nucleic Acids Res. 15:7865–7876, 1987. With permission from Oxford University Press. Molecular Biology of the cells, bruce Albert,2015)
- [22] Renart J, Reiser J, Stark GR. "Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure", Proc Natl Acad Sci U S A, 1979 Jul;76(7):3116-20. PMID: [91164 abstract](#)\
- [23] Engvall, E (1972-11-22). "[Enzyme-linked immunosorbent assay, Elisa](#)". The Journal of Immunology. 109 (1): 129–135. ISSN [0022-1767](#). PMID [4113792](#).
- [24] researchgate.net/figure/Comparison-of-in-vivo-in-vitro-and-in-silico-methods-for-UV-A-assessment\_tbl3\_280623647
- [25] Cong Fen He, Jiaxing Zhang, Jie Chen. 2007. Genetic Transformation of Aloe barbadensis Miller by Agrobacterium tumefaciens. Journal of Genetics and Genomics. College of Chemistry, Beijing. Institute of crops science, Beijing.
- [26] Sahdev S, Khattar SK. Production of active eukaryotic proteins through bacterial expression system: a review of the existing biotechnology strategies. Mol Cell Biochem. 2008;307: 249-264.

- [27] Ferrer-Miralles N, Villaverde A. Bacterial cell factories for recombinant protein production; expanding the catalogue. *Microb Cell Fact*.
- [28] Gellissen G, Hollenberg CP: A comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* - a review. *Gene*. 1997, 190: 87-97. doi:10.1016/S0378-1119(97)00020-6.
- [29] Porro D, Sauer M, Branduardi P, Mattanovich D: Recombinant protein production in yeasts. *Mol Biotechnol*. 2005, 31: 245-259. doi:10.1385/MB:31:3:245.
- [30] Staub JM, Garcia B, Graves J, Hajdukiewicz PT, Hunter P, Nehra N, Paradkar V, Schlittler M, Carroll JA, Spatola L, Ward D, Ye G, Russell DA: High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nat Biotechnol*. 2000, 18: 333-338. doi:10.1038/73796.
- [31] Lower et al. "transgenic aloe plants for production of protein and related methods"
- [32] Jefferson, R. A. (1987). *Assaying chimeric genes in plants: The GUS gene fusion system*. *Plant Molecular Biology Reporter*, 5(4), 387–405. doi:10.1007/bf02667740
- [33] William Lowther • Kevin Lorick • Susan D. Lawrence • Wen-Shuz Yeow Expression of biologically active human interferon-alpha 2 in Aloe vera
- [34] Plotkin, JB, Kudla G. (2011). Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet*, 12:32–42.
- [35] Wildt, S., & Gerngross, T. U. (2005). The humanization of N-glycosylation pathways in yeast. *Nature Reviews Microbiology*, 3(2), 119–128. doi:10.1038/nrmicro1087
- [36] Baeshen, N. A., Baeshen, M. N., Sheikh, A., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A. I., ... Redwan, E. M. (2014). Cell factories for insulin production. *Microbial Cell Factories*, 13(1). doi:10.1186/s12934-014-0141-0
- [37] Boyhan D, Daniell H: Low-cost production of proinsulin in tobacco and lettuce chloroplasts for injectable or oral delivery of functional insulin and C-peptide. *Plant Biotech J* 2011, 9:585–598
- [38] PCR primer design guidelines
- [39] Campbell RK, White JR., Jr Insulin therapy in type 2 diabetes. *J Am Pharm Assoc (Wash)* 2002;42:602–11

- [40] Rolla AR, Rakel RE. Practical approaches to insulin therapy for type 2 diabetes mellitus with premixed insulin analogues. *Clin Ther*. 2005;27:1113–25.
- [41] Yaturu S. Insulin therapies: Current and future trends at dawn. *World J Diabetes* 2013;4:1-7
- [42] Benedict C, Frey WH 2nd, Schiöth HB, Schultes B, Born J, Hallschmid M. Intranasal insulin as a therapeutic option in the treatment of cognitive impairments. *Exp Gerontol* 2011;46:112-5.
- [43] Craft S, Baker LD, Montine TJ, Minoshima S, Watson GS, Claxton A, et al. Intranasal insulin therapy for Alzheimer disease and amnestic mild cognitive impairment: A pilot clinical trial. *Arch Neurol* 2012;69:29-38.
- [44] Kim H.K., Chung H.J., Park T.G. Biodegradable polymeric microspheres with “open/closed” pores for sustained release of human growth hormone. *J Control Release*. 2006;112:167–174.
- [45] Eilleia, S. Y., Soliman, M. E., Mansour, S., & S. Geneidi, A. (2018). Novel technique of insulin loading into porous carriers for oral delivery. *Asian Journal of Pharmaceutical Sciences*, 13(4), 297–309. doi:10.1016/j.ajps.2018.03.003
- [46] Dinas Pangan, Pertanian, dan Perikanan Pontianak - Lidah Buaya.
- [47] Reynolds, Tom (Ed.) (2004) *Aloes: The genus Aloe (Medicinal and Aromatic Plants - Industrial Profiles)*. CRC Press. ISBN 978-0415306720
- [48] Quattrocchi, Umberto (2012) *CRC World Dictionary of Medicinal and Poisonous Plants: Common Names, Scientific Names, Eponyms, Synonyms, and Etymology* (5 Volume Set) CRC Press.ISBN 978-1420080445
- [49] Food Drug Administration, HHS (2002). "Status of certain additional over-the-counter drug category II and III active ingredients. Final rule". *Fed Regist*. 67 (90): 31125–7.PMID 12001972
- [50] National Center for Complementary and Integrative Health, US National Institutes of Health. 1 September 2019. Diakses tanggal 24 August 2019.
- [51] Mayo Clinic. 17 September 2017. Diakses tanggal 21 Januari 2020.
- [52] Transgenic aloe plants for production of proteins and related methods
- [53] Borgoño, C. A., & Zinman, B. (2012). *Insulins: Past, Present, and Future. Endocrinology and Metabolism Clinics of North America*, 41(1), 1–24. doi:10.1016/j.ecl.2012.03.002

- [54] Hagedorn HC, Jensen BN, Krarup NB, et al. Protamine insulinate. *JAMA* 1936; 106:177–80.
- [55] Scott DA, Fisher AM. Studies on insulin with protamine. *J Pharmacol Exp Therapeut* 1936;58:78–92.
- [56] Himsworth HP. Protamine insulin and zinc protamine insulin. *Br Med J* 1937;1: 541–6.
- [57] Krayenbuhl C, Rosenberg T. Crystalline protamine insulin. *Rep Steno Mem Hosp Nord Insulin Lab* 1946;1:60–73.
- [58] Hallas-Moller K. The lente insulins. *Diabetes* 1956;5:7–14.
- [59] Oakley W, Hill D, Oakley N. Combined use of regular and crystalline protamine (NPH) insulins in the treatment of severe diabetes. *Diabetes* 1966;15:219–22.
- [60] Sanger F. Chemistry of insulin; determination of the structure of insulin opens the way to greater understanding of life processes. *Science* 1959;129:1340–4.
- [61] Markussen J, Damgaard U, Jorgensen KH, et al. Human monoclonal insulin. Chemistry and characteristics. *Acta Med Scand Suppl* 1983;671:99–105.
- [62] Keen H, Glynne A, Pickup JC, et al. Human insulin produced by recombinant DNA technology: safety and hypoglycaemic potency in healthy men. *Lancet* 1980;2:398–401.
- [63] Velcheva, M., Faltin, Z., Vardi, A., Hanania, U., Eshdat, Y., Dgani, O., ... Perl, A. (2010). Aloe vera transformation: the role of Amberlite XAD-4 resin and antioxidants during selection and regeneration. *In Vitro Cellular & Developmental Biology - Plant*, 46(6), 477–484. doi:10.1007/s11627-010-z
- [66] ^ Norris, Sonya (4 July 2005). "Molecular farming". Library of Parliament. Parliament of Canada. PRB 05-09E. Diakses tanggal 2008-09-11.
- [67] ^ Kaiser, Jocelyn (25 April 2008). "Is the Drought Over for Pharming?" (PDF). *Science*. 320 (5875): 473–5. doi:10.1126/science.320.5875.473. PMID 18436771.
- [68]. Vieira Gomes, A.M.; Souza Carmo, T.; Silva Carvalho, L.; Mendonca Bahia, F.; Parachin, N.S. Comparison of yeasts as hosts for recombinant protein production. *Microorganisms* 2018, 6, 38. [CrossRef]

- [71]. Shanmugaraj, B.; Ramalingam, S. Plant expression platform [74]. Sahdev, S.; Khattar, S.K.; Saini, K.S. Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies. *Mol. Cell. Biochem.* 2008, 307, 249–264. [CrossRef]
- [75]. Strasser, R.; Stadlmann, J.; Schahs, M.; Stiegler, G.; Quendler, H.; Mach, L.; Glossl, J.; Weterings, K.; Pabst, M.; Steinkellner, H. Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol.*
- [77]. Montero-Morales, L.; Steinkellner, H. Advanced plant-based glycan engineering. *Front. Bioeng. Biotechnol.* 2018, 6, 81. [CrossRef] [PubMed]
- [79]. Menary, J.; Hobbs, M.; Mesquita de Albuquerque, S.; Pacho, A.; Drake, P.M.W.; Prendiville, A.; Ma, J.K.; Fuller, S.S. Shotguns vs Lasers: Identifying barriers and facilitators to scaling-up plant molecular farming for high-value health products. *PLoS ONE* 2020, 15, e0229952. [CrossRef]
- [80]. Floss, D.M.; Sack, M.; Stadlmann, J.; Rademacher, T.; Scheller, J.; Stöger, E.; Fischer, R.; Conrad, U. Biochemical and functional characterization of anti-HIV antibody–ELP fusion proteins from transgenic plants. *Plant Biotechnol. J.* 2008, 6, 379–391. [CrossRef]
- [81]. Silva H, Sagardia S, Seguel O, Torres C, Tapia C, Franck N, Cardemil L: Effect of water availability on growth and water use efficiency for biomass and gel production in Aloe Vera (*Aloe barbadensis* M.). *Industrial Crops and Products* 2010, 31(1):20-27.
- [82]. Grace OM, Buerki S, Symonds MR, Forest F, van Wyk AE, Smith GF, Klopper RR, Bjørå CS, Neale S, Demissew S: Evolutionary history and leaf succulence as explanations for medicinal use in aloes and the global popularity of Aloe vera. *BMC evolutionary biology* 2015, 15(1):29.
- [83]. Reynolds T, Dweck A: Aloe vera leaf gel: a review update. *Journal of ethnopharmacology* 1999, 68(1-3):3-37.
- [84]. Gupta VK, Malhotra S: Pharmacological attribute of Aloe vera: Revalidation through experimental and clinical studies. *Ayu* 2012, 33(2):193.
- [85]. Raksha B, Pooja S, Babu S: Bioactive compounds and medicinal properties of Aloe vera L.: An update. *Journal of Plant Sciences* 2014, 2(3):102-107.
- [86]. Hamman JH: Composition and applications of Aloe vera leaf gel. *Molecules* 2008, 13(8):1599-1616.

- [87]. Joseph B, Raj SJ: Pharmacognostic and phytochemical properties of Aloe vera linn an overview. International journal of pharmaceutical sciences review and research 2010, 4(2):106-110.
- [88]. Choudhri P, Rani M, Sangwan RS, Kumar R, Kumar A, Chhokar V: De novo sequencing, assembly and characterisation of Aloe vera transcriptome and analysis of expression profiles of genes related to saponin and anthraquinone metabolism. BMC genomics 2018, 19(1):427.
- [89]. Zonneveld BJ: Genome size analysis of selected species of Aloe (Aloaceae) reveals the most primitive species and results in some new combinations. Bradleya 2002, 2002(20):5-12.
- [90]. Kadirvelu Sharon1, Sharma Suvarna. 2017. Cloning of HVA22 Homolog from Aloe vera and Preliminary Study of Transgenic Plant Development. ISSN: 2320 – 7051. DOI: <http://dx.doi.org/10.18782/2320-7051.6033>.
- [91]. Jia-Ju Rena, Jian Wang et al. 2020. The complete chloroplast genome of Aloe vera from China as a Chinese herb. aSchool of Nursing, Beijing University of Chinese Medicine, Beijing, China. Taylor & Francis. VOL. 5, NO. 1, 1092–1093 <https://doi.org/10.1080/23802359.2020.1726229>
- [91]. Shubham, Abhisek et al. 2020. The genome sequence of Aloe vera reveals adaptive evolution of drought tolerance  
2 mechanisms. bioRxiv preprint doi: <https://doi.org/10.1101/2020.05.29.122895>.
- [92]. Jia-Ju Ren, Jian wang. 2020. The complete chloroplast genome of Aloe vera from China as a Chinese herb. © 2020 The Author(s). Published by InformaUK Limited, trading as Taylor & FrancisGroup. DOI : VOL. 5, NO. 1, <https://doi.org/10.1080/23802359.2020.1726229>
- [93]. Gellissen G, Hollenberg CP, Application of yeasts in gene expression studies A comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* - a review. Gene. 1997;190:87–97. doi: 10.1016/S0378-1119(97)00020-6. [PubMed] [CrossRef] [Google Scholar] [Ref list]
- [94]. 24. Porro D, Sauer M, Branduardi P, Mattanovich D: Recombinant protein production in yeasts. Mol Biotechnol 2005, 31:245–259.

- [95]. Kjeldsen T, Pettersson AF, Hach M: Secretory expression and characterization of insulin in *Pichia pastoris*. *Biotechnol Appl Biochem* 1999, 29:79–86.
- [96]. Mansur M, Cabello C, Hernandez L, Pais J, Varas L, Valdes J, Terrero Y, Hidalgo A, Plana L, Besada V, Garcia L, Lamazares E, Castellanos L, Martinez E: Multiple gene copy number enhances insulin precursor secretion in the yeast *Pichia pastoris*. *Biotechnol Lett* 2005, 27:339–345.
- [97]. "nopaline synthase terminator: Topics by Science.gov". [www.science.gov](http://www.science.gov). Retrieved 2019-11-20.
- [98] Carrer, Helaine; Hockenberry, Tish Noel; Svab, Zora; Maliga, Pal (October 1993). "Kanamycin resistance as a selectable marker for plastid transformation in tobacco". *MGG Molecular & General Genetics*. 241-241 (1-2): 49–56.
- [99] Vande Broek, A; Lambrecht, M; Vanderleyden, J (1998). "Bacterial chemotactic motility is important for the initiation of wheat root colonization by *Azospirillum brasilense*". *Microbiology*. 144 (9): 2599–606. doi:10.1099/00221287-144-9-2599. PMID 9782509.
- [100] Sundberg, C., Meek, L., Carroll, K., Das, A., & Ream, W. (1996). VirE1 protein mediates export of the single-stranded DNA-binding protein VirE2 from *Agrobacterium tumefaciens* into plant cells. *Journal of Bacteriology*, 178(4), 1207–1212. doi:10.1128/jb.178.4.1207-1212.1996
- [101] Sundberg CD, Ream W (November 1999). "[The Agrobacterium tumefaciens chaperone-like protein, VirE1, interacts with VirE2 at domains required for single-stranded DNA binding and cooperative interaction](#)". *Journal of Bacteriology*. 181 (21): 6850–5. [doi:10.1128/JB.181.21.6850-6855.1999](https://doi.org/10.1128/JB.181.21.6850-6855.1999)
- [106] Damerval C, Vienne DD, Zivy M, Thiellement H. Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. *Electrophoresis* 1986; 7: 52–4
- [107] Shaw, M. M., & Riederer, B. M. (2003). *Sample preparation for two-dimensional gel electrophoresis*. *PROTEOMICS*, 3(8), 1408–1417. doi:10.1002/pmic.200300471

- [108] Novák, P., & Havlíček, V. (2016). *Protein Extraction and Precipitation. Proteomic Profiling and Analytical Chemistry*, 51–62. doi:10.1016/b978-0-444-63688-1.00004-5
- [109] Koontz, L. (2014). *TCA Precipitation. Laboratory Methods in Enzymology: Protein Part C*, 3–10. doi:10.1016/b978-0-12-420119-4.00001-x
- [110] He, C.-F., & Wang, Y.-M. (2008). *Protein Extraction From Leaves of Aloe vera L., A Succulent and Recalcitrant Plant, for Proteomic Analysis. Plant Molecular Biology Reporter*, 26(4), 292–300. doi:10.1007/s11105-008-0040-9
- [111] Hurkman, W. J., & Tanaka, C. K. (1986). *Solubilization of Plant Membrane Proteins for Analysis by Two-Dimensional Gel Electrophoresis. PLANT PHYSIOLOGY*, 81(3), 802–806. doi:10.1104/pp.81.3.802
- [112] Faurobert, M., Pelpoir, E., & Chaïb, J. (n.d.). Phenol Extraction of Proteins for Proteomic Studies of Recalcitrant Plant Tissues. *Plant Proteomics*, 9–14. doi:10.1385/1-59745-227-0:9
- [113] Carpentier, S. C., Witters, E., Laukens, K., Deckers, P., Swennen, R., & Panis, B. (2005). *Preparation of protein extracts from recalcitrant plant tissues: An evaluation of different methods for two-dimensional gel electrophoresis analysis. PROTEOMICS*, 5(10), 2497–2507. doi:10.1002/pmic.200401222
- [114] Saravanan, R. S., & Rose, J. K. C. (2004). *A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. PROTEOMICS*, 4(9), 2522–2532. doi:10.1002/pmic.200300789
- [115] [https://www.researchgate.net/post/Protein\\_denaturation\\_vs\\_degradation](https://www.researchgate.net/post/Protein_denaturation_vs_degradation)
- [116] Anderson, N. L., & Anderson, N. G. (1998). *Proteome and proteomics: New technologies, new concepts, and new words. Electrophoresis*, 19(11), 1853–1861. doi:10.1002/elps.1150191103
- [117] Blackstock, W. P., & Weir, M. P. (1999). *Proteomics: quantitative and physical mapping of cellular proteins. Trends in Biotechnology*, 17(3), 121–127. doi:10.1016/s0167-7799(98)01245-1
- [118] Komari, T., Takakura, Y., Ueki, J., Kato, N., Ishida, Y., & Hiei, Y. (n.d.). Binary Vectors and Super-binary Vectors. *Agrobacterium Protocols*, 15–42. doi:10.1385/1-59745-130-4:15

- [119] Lee, L.-Y., & Gelvin, S. B. (2007). *T-DNA Binary Vectors and Systems*. *PLANT PHYSIOLOGY*, 146(2), 325–332. doi:10.1104/pp.107.113001
- [120] Hoekema, A., Hirsch, P. R., Hooykaas, P. J. J., & Schilperoort, R. A. (1983). *A binary plant vector strategy based on separation of vir- and T-region of the Agrobacterium tumefaciens Ti-plasmid*. *Nature*, 303(5913), 179–180. doi:10.1038/303179a0
- [121] Fraley, R. T., Rogers, S. G., Horsch, R. B., & Gelvin, S. B. (1986). *Genetic transformation in higher plants*. *Critical Reviews in Plant Sciences*, 4(1), 1–46. doi:10.1080/07352688609382217
- [122] Christensen, A. H., Sharrock, R. A., & Quail, P. H. (1992). *Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation*. *Plant Molecular Biology*, 18(4), 675–689. doi:10.1007/bf00020010
- [123] Zhang, W. (1991). *Analysis of Rice Act1 5[prime] Region Activity in Transgenic Rice Plants*. *THE PLANT CELL ONLINE*, 3(11), 1155–1165. doi:10.1105/tpc.3.11.1155
- [124] GARFINKEL, D. (1981). *Genetic analysis of crown gall: Fine structure map of the T-DNA by site-directed mutagenesis*. *Cell*, 27(1), 143–153. doi:10.1016/0092-8674(81)90368-8
- [125] Zambryski P, Joos H, Genetello C, Leemans J, Montagu MV, Schell J (1983). Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. *EMBO J*. 1983; 2(12):2143-50.
- [126] Fraley, R. T., Rogers, S. G., Horsch, R. B., Eichholtz, D. A., Flick, J. S., Fink, C. L., ... Sanders, P. R. (1985). *The SEV System: A New Disarmed Ti Plasmid Vector System for Plant Transformation*. *Nature Biotechnology*, 3(7), 629–635. doi:10.1038/nbt0785-629
- [127] De Framond, A. J., Barton, K. A., & Chilton, M.-D. (1983). *Mini-Ti: A New Vector Strategy for Plant Genetic Engineering*. *Nature Biotechnology*, 1(3), 262–269. doi:10.1038/nbt0583-262
- [128] Pansegrau, W., Lanka, E., Barth, P. T., Figurski, D. H., Guiney, D. G., Haas, D., ... Thomas, C. M. (1994). *Complete Nucleotide Sequence of*

*Birmingham IncPa Plasmids. Journal of Molecular Biology*, 239(5), 623–663.  
doi:10.1006/jmbi.1994.1404

[129] Okumura, M. S., & Kado, C. I. (1992). *The region essential for efficient autonomous replication of pSa in Escherichia coli*. *MGG Molecular & General Genetics*, 235(1), 55–63. doi:10.1007/bf00286181

[130] Hellens, R. P., Edwards, E. A., Leyland, N. R., Bean, S., & Mullineaux, P. M. (2000). *Plant Molecular Biology*, 42(6), 819–832.  
doi:10.1023/a:1006496308160

[131] Jouanin, L., Vilaine, F., d' Enfert, C., & Casse-Delbart, F. (1985). *Localization and restriction maps of the replication origin regions of the plasmids of Agrobacterium rhizogenes strain A4*. *MGG Molecular & General Genetics*, 201(3), 370–374. doi:10.1007/bf00331325

[132] Deblaere, R., Reynaerts, A., Höfte, H., Hernalsteens, J.-P., Leemans, J., & Van Montagu, M. (1987). [16] *Vectors for cloning in plant cells. Recombinant DNA Part D*, 277–292. doi:10.1016/0076-6879(87)53059-2

[133] Sambrook, J. and Russell, D. W. (2001) Molecular Cloning, A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.  
ISBN: 978-087969577-4

[134] Lemos, M. L., & Crosa, J. H. (1992). *Highly preferred site of insertion of Tn7 into the chromosome of Vibrio anguillarum*. *Plasmid*, 27(2), 161–163.  
doi:10.1016/0147-619x(92)90016-4

[135] Pang, S. Z., DeBoer, D. L., Wan, Y., Ye, G., Layton, J. G., Neher, M. K., ... Fromm, M. E. (1996). *An Improved Green Fluorescent Protein Gene as a Vital Marker in Plants*. *Plant Physiology*, 112(3), 893–900. doi:10.1104/pp.112.3.893

[136] OW, D. W., DE WET, J. R., HELINSKI, D. R., HOWELL, S. H., WOOD, K. V., & DELUCA, M. (1986). *Transient and Stable Expression of the Firefly Luciferase Gene in Plant Cells and Transgenic Plants*. *Science*, 234(4778), 856–859. doi:10.1126/science.234.4778.856

[137] Bevan, M. (1984). *Binary Agrobacterium vectors for plant transformation*. *Nucleic Acids Research*, 12(22), 8711–8721. doi:10.1093/nar/12.22.8711

- [138] Chung, S.-M., Frankman, E. L., & Tzfira, T. (2005). *A versatile vector system for multiple gene expression in plants*. *Trends in Plant Science*, 10(8), 357–361. doi:10.1016/j.tplants.2005.06.001
- [139] Tzfira, T., Tian, G.-W., Lacroix, B., Vyas, S., Li, J., Leitner-Dagan, Y., ... Citovsky, V. (2005). *pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants*. *Plant Molecular Biology*, 57(4), 503–516. doi:10.1007/s11103-005-0340-5
- [140] Goderis IJWM, De Bolle MFC, Francois IEJA, Wouters PFJ, Broekaert WF, Cammue BPA (2002) A set of modular plant transformation vectors allowing flexible insertion of up to six expression units. *Plant Mol Biol* 50 17–27 doi:10.1023/a:1016052416053
- [141] Citovsky, V., Lee, L.-Y., Vyas, S., Glick, E., Chen, M.-H., Vainstein, A., ... Tzfira, T. (2006). *Subcellular Localization of Interacting Proteins by Bimolecular Fluorescence Complementation in Planta*. *Journal of Molecular Biology*, 362(5), 1120–1131. doi:10.1016/j.jmb.2006.08.017
- [142] Hamilton, C. M. (1997). *A binary-BAC system for plant transformation with high-molecular-weight DNA*. *Gene*, 200(1-2), 107–116. doi:10.1016/s0378-1119(97)00388-0
- [143] Liu, Y.-G., Shirano, Y., Fukaki, H., Yanai, Y., Tasaka, M., Tabata, S., & Shibata, D. (1999). *Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning*. *Proceedings of the National Academy of Sciences*, 96(11), 6535–6540. doi:10.1073/pnas.96.11.6535
- [144] Liu, Y.-G., Nagaki, K., Fujita, M., Kawaura, K., Uozumi, M., & Ogihara, Y. (2000). *Development of an efficient maintenance and screening system for large-insert genomic DNA libraries of hexaploid wheat in a transformation-competent artificial chromosome (TAC) vector*. *The Plant Journal*, 23(5), 687–695. doi:10.1046/j.1365-313x.2000.00827.x
- [145] Freeman WM, Walker SJ, Vrana KE (January 1999). "Quantitative RT-PCR: pitfalls and potential". *BioTechniques*. 26 (1): 112–22, 124–5. doi:10.2144/99261rv01. PMID 9894600.

- [146] Schmittgen TD, Zakrajsek BA, Mills AG, Gorn V, Singer MJ, Reed MW (October 2000). "Quantitative reverse transcription-polymerase chain reaction to study mRNA decay: comparison of endpoint and real-time methods". *Anal. Biochem.* 285 (2): 194–204. doi:10.1006/abio.2000.4753. PMID 11017702. S2CID 258810.
- [147] Deepak S, Kottapalli K, Rakwal R, et al. (June 2007). "Real-Time PCR: Revolutionizing Detection and Expression Analysis of Genes". *Curr. Genomics*. 8 (4): 234–51. doi:10.2174/138920207781386960. PMC 2430684. PMID 18645596.
- [148] Bustin SA (August 2002). "Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems". *J. Mol. Endocrinol.* 29 (1): 23–39. doi:10.1677/jme.0.0290023. PMID 12200227.
- [149] Keiko, Takashi. 2003. A Role of the –35 Element in the Initiation of Transcription at psbA Promoter in Tobacco Plastids
- [150] Lowther, Lorick. 2012. Expression of biologically active human interferon alpha 2 in Aloe vera
- [151] T. Cruz, T. Kanashiro, A. de Castro, C. Baldin, L. Richtzenhain and J. Araujo Jr., Pesqui. *Vet. Bras.*, 2016, 36, 1171–1177.
- [152] T. Csont, Hungary-Romania cross Border co-operation programme, 2013, vol. 2007–2013, pp. 3–4.
- [153] I. A. Darwish, *Int. J. Biomed. Sci.*, 2006, 2, 217–235.
- [154] S. E. Manley, I. M. Stratton, P. M. Clark and S. D. Luzio, *Clin. Chem.*, 2007, 53, 922–932.
- [155] M. S. Even, C. B. Sandusky, N. D. Barnard, J. Mistry and M. K. Sinha, *Clin. Biochem.*, 2007, 40, 98–103.
- [156] R. Abellán, R. Ventura, I. Palmi, S. di Carlo, R. di Giovannandrea, M. Bellver, R. Olive, J. A. Pascual, R. Pacifici, J. Segura, P. Zuccaro and S. Pichini, *J. Pharm. Biomed. Anal.*, 2009, 49, 793–799.
- [157] Z. J. Farino, T. J. Morgenstern, J. Vallaghe, N. Gregor.
- [158] Ushimaru, T., Terada, H., Tsuboi, K., Kogou, Y., Sakaguchi, A., Tsuji, G., & Kubo, Y. (2010). *Development of an efficient gene targeting system in*

*Colletotrichum higginsianum* using a non-homologous end-joining mutant and *Agrobacterium tumefaciens*-mediated gene transfer. *Molecular Genetics and Genomics*, 284(5), 357–371. doi:10.1007/s00438-010-0572-1

[159] T. Csont, Hungary-Romania cross Border co-operation programme, 2013, vol. 2007–2013, pp. 3–4.

[160] Zahra, Ravi. 2016. A novel aseptic technique for micropropagation of Aloe vera mill. herbmed.skums.ac.ir

[161] Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, 77(1), 51–59. doi:10.1016/0378-1119(89)90358-2

[162] MICHELSE, C., ARENTSHORST, M., RAM, A., & VANDENHONDEL, C. (2005). -mediated transformation leads to improved gene replacement efficiency in. *Fungal Genetics and Biology*, 42(1), 9–19. doi:10.1016/j.fgb.2004.06.009

[163] Idnurm, A., Giles, S. S., Perfect, J. R., & Heitman, J. (2006). Peroxisome Function Regulates Growth on Glucose in the Basidiomycete Fungus *Cryptococcus neoformans*. *Eukaryotic Cell*, 6(1), 60–72. doi:10.1128/ec.00214-06

[164] Li, L., Shao, Y., Li, Q., Yang, S., & Chen, F. (2010). Identification of *Mga1*, a G-protein α-subunit gene involved in regulating citrinin and pigment production in *Monascus ruber* M7. *FEMS Microbiology Letters*, no–no. doi:10.1111/j.1574-6968.2010.01992.x

[165] Roberts, R. J., & Murray, K. (1976). *Restriction Endonuclease. CRC Critical Reviews in Biochemistry*, 4(2), 123–164. doi:10.3109/10409237609105456

[166] Kessler, C., & Manta, V. (1990). Specificity of restriction endonucleases and DNA modification methyltransferases — a review (edition 3). *Gene*, 92(1-2), 1–240. doi:10.1016/0378-1119(90)90486-b

[167] Pingoud A, Alves J, Geiger R (1993). "Chapter 8: Restriction Enzymes". In Burrell M (ed.). *Enzymes of Molecular Biology*. Methods of Molecular Biology. 16. Totowa, NJ: Humana Press. pp. 107–200. ISBN 0-89603-234-5.

- [168] Luria SE, Human ML (October 1952). "A nonhereditary, host-induced variation of bacterial viruses". *Journal of Bacteriology*. **64** (4): 557–69. doi:10.1128/JB.64.4.557-569.1952. PMC 169391. PMID 12999684
- [169] Bertani G, Weigle JJ (February 1953). "Host controlled variation in bacterial viruses". *Journal of Bacteriology*. **65** (2): 113–21. doi:10.1128/JB.65.2.113-121.1953. PMC 169650. PMID 13034700.
- [176] Clark DP (2005). *Molecular biology*. Amsterdam: Elsevier Academic Press. ISBN 0-12-175551-7.
- [177] O'Brien, John A.; Lummis, Sarah CR (2011). "Nano-biolistics: A method of biolistic transfection of cells and tissues using a gene gun with novel nanometer-sized projectiles". *BMC Biotechnology*. **11**: 66. doi:10.1186/1472-6750-11-66. PMC 3144454. PMID 21663596.
- [178] Sanford, John C. (1990). "Biolistic plant transformation". *Physiologia Plantarum*. **79** (1): 206–209. doi:10.1111/j.1399-3054.1990.tb05888.x. ISSN 1399-3054.
- [179] Slater, Adrian; Scott, Nigel; Fowler, Mark (2008). *Plant Biotechnology: the genetic manipulation of plants* (2 ed.). Oxford, New York, USA: Oxford University Press Inc. ISBN 978-0-19-928261-6.
- [180] Kikkert, Julie; Vidal, Jose; Reisch, Bruce (2005). Stable transformation of plant cells by particle bombardment/bioistics. *Methods in Molecular Biology*. 286. pp. 61–78. doi:10.1385/1-59259-827-7:061. ISBN 978-1-59259-827-4. PMID 15310913. S2CID 44395352.
- [181] Gan C. (1989) Gene Gun Accelerates DNA-Coated Particles To Transform Intact Cells. *The Scientist* 3(18):25.
- [182] Baldy MW, Strom B, Bernstein H (March 1971). "Repair of alkylated bacteriophage T4 deoxyribonucleic acid by a mechanism involving polynucleotide ligase". *Journal of Virology*. **7** (3): 407–8. doi:10.1128/JVI.7.3.407-408.1971. PMC 356131. PMID 4927528.
- [183] Baldy, M. W. (1970). *The UV sensitivity of some early-function temperature-sensitive mutants of phage T4*. *Virology*, **40**(2), 272–287. doi:10.1016/0042-6822(70)90403-4

- [184] Baldy MW (1968). "Repair and recombination in phage T4. II. Genes affecting UV sensitivity". *Cold Spring Harbor Symposia on Quantitative Biology*. **33**: 333–8.doi:10.1101/sqb.1968.033.01.038. PMID 4891973.
- [185] Wilson, R. H., Morton, S. K., Deiderick, H., Gerth, M. L., Paul, H. A., Gerber, I., ... Patrick, W. M. (2013). *Engineered DNA ligases with improved activities in vitro*. *Protein Engineering Design and Selection*, **26**(7), 471–478. doi:10.1093/protein/gzt024
- [186] Pascal, J. M., O'Brien, P. J., Tomkinson, A. E., & Ellenberger, T. (2004). *Human DNA ligase I completely encircles and partially unwinds nicked DNA*. *Nature*, **432**(7016), 473–478. doi:10.1038/nature03082
- [187] Lehman, I. R. (1974). *DNA Ligase: Structure, Mechanism, and Function*. *Science*, **186**(4166), 790–797. doi:10.1126/science.186.4166.790
- [188] Foster JB, Slonczewski J (2010). *Microbiology: An Evolving Science* (Second ed.). New York: W. W. Norton & Company. ISBN 978-0-393-93447-2.
- [189] Yang, Y., & LiCata, V. J. (2018). *Pol I DNA polymerases stimulate DNA end-joining by Escherichia coli DNA ligase*. *Biochemical and Biophysical Research Communications*, **497**(1), 13–18. doi:10.1016/j.bbrc.2018.01.165
- [190] "Ligases" (PDF). *Enzyme Resources Guide*. Promega Corporation. pp. 8–14.
- [191] m-edukasi. Kemdikbud. Pembentukan hormon-hormon Endokrin Pankreas
- [192] Richard. Insulin Synthesis and Secretion. VIVO Pathophysiology.
- [193] Régnier, Philippe; Marujo, Paulo E. (2013). Polyadenylation and Degradation of RNA in Prokaryotes
- [194] "Polyadenylation - an overview | ScienceDirect Topics". www.sciencedirect.com.
- [195] Rehfeld, Anders; Plass, Mireya; Krogh, Anders; Friis-Hansen, Lennart (2013). "[Alterations in Polyadenylation and Its Implications for Endocrine Disease](#)". *Frontiers in Endocrinology*
- [196] Beaud, D. (2005). Genetic characterization of the -glucuronidase enzyme from a human intestinal bacterium, *Ruminococcus gnavus*. *Microbiology*, **151**(7), 2323–2330. doi:10.1099/mic.0.27712-0

- [197] Plotkin JB, Kudla G. (2011) [Synonymous but not the same: the causes and consequences of codon bias.](#)
- [198] Cannarozzi G, Schraudolph NN, et al. (2010) [A role for codon order in translation dynamics.](#)
- [199] Wang Y, Liu CL, et al. (2002) [Precision and functional specificity in mRNA decay.](#)
- [200] Zhang G, Ignatova Z. (2011) [Folding at the birth of the nascent chain: coordinating translation with co-translational folding.](#)
- [201] The EMBL-EBI search and sequence analysis tools APIs in 2019 [Fábio Madeira](#), [Young mi Park](#), [Joon Lee](#), [Nicola Buso](#), [Tamer Gur](#), [Nandana Madhusoodanan](#), [Prasad Basutkar](#), [Adrian R N Tivey](#), [Simon C Potter](#), [Robert D Finn](#), and [Rodrigo Lopez](#)
- [202] Haas J., Gumienny R., Barbato A., Ackermann F., Tauriello G., Bertoni M., Studer G., Smolinski A., Schwede T. Introducing "best single template" models as reference baseline for the Continuous Automated Model Evaluation (CAMEO). *Proteins.* (2019), 87, 1378-1387. [[DOI: 10.1002/prot.25815](#)]
- [203] Haas J., Barbato A., Behringer D., Studer G., Roth S., Bertoni M., Mostaguir K., Gumienny R., Schwede T. Continuous Automated Model EvaluatiOn (CAMEO) complementing the critical assessment of structure prediction in CASP12. *Proteins.* (2018), 86, 387-398. [[DOI: 10.1002/prot.25431](#)]
- [204] Haas J., Roth S., Arnold K., Kiefer F., Schmidt T., Bordoli L. and Schwede T.(2013). The Protein Model Portal - a comprehensive resource for protein structure and model information. Database [[PMID: 23624946](#)].
- [205] DeLano, W. L. (2002). PyMOL. DeLano Scientific, San Carlos, CA, 700.
- [206] DeLano, W. L. (2009). The PyMOL Molecular Graphics System; DeLano Scientific: San Carlos, CA, 2002.
- [207] DeLano, W. L. (2002). Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography*, 40, 82-92.
- [208] O'Brien, J., & Lummis, S. C. R. (2002). An improved method of preparing microcarriers for biolistic transfection. *Brain Research Protocols*, 10(1), 12–15. doi:10.1016/s1385-299x(02)00175-7

- [209] Frame, B. R., Zhang, H., Coccilone, S. M., Sidorenko, L. V., Dietrich, C. R., Pegg, S. E., ... Wang, K. (2000). Production of transgenic maize from bombarded type II callus: Effect of gold particle size and callus morphology on transformation efficiency. In Vitro Cellular & Developmental Biology - Plant, 36(1), 21–29. doi:10.1007/s11627-000-0007-5
- [210] Christensen, A. H., & Quail, P. H. (1996). Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Research, 5(3), 213–218. doi:10.1007/bf01969712
- [211] Lengauer T, Rarey M (Jun 1996). "Computational methods for biomolecular docking". Current Opinion in Structural Biology. 6(3): 402–6. [doi:10.1016/S0959-440X\(96\)80061-3](https://doi.org/10.1016/S0959-440X(96)80061-3). PMID [8804827](#).