

2

BIOTECNOLOGIA APLICADA À SAÚDE

FUNDAMENTOS E APLICAÇÕES



organizador

RODRIGO RIBEIRO RESENDE

colaborador

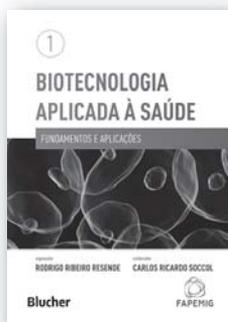
CARLOS RICARDO SOCCOL

Blucher


FAPEMIG

BIOTECNOLOGIA APLICADA À SAÚDE

Coleção **Biotecnologia Aplicada à Saúde**



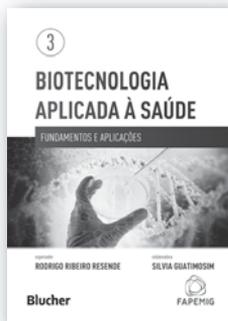
Volume 1
Biotechnologia Aplicada à Saúde
Fundamentos e Aplicações

ISBN: 978-85-212-0896-9
623 páginas



Volume 2
Biotechnologia Aplicada à Saúde
Fundamentos e Aplicações

ISBN: 978-85-212-0921-8
1192 páginas



Volume 3
Biotechnologia Aplicada à Saúde
Fundamentos e Aplicações

Pré-lançamento



Volume 4
Biotechnologia Aplicada à Agro&Indústria
Fundamentos e Aplicações

Pré-lançamento

Blucher



MATERIAL DE APOIO
www.blucher.com.br

BIOTECNOLOGIA APLICADA À SAÚDE
FUNDAMENTOS E APLICAÇÕES
VOLUME 2

RODRIGO RIBEIRO RESENDE

ORGANIZADOR

CARLOS RICARDO SOCCOL

COLABORADOR

Biotecnologia aplicada à saúde: fundamentos e aplicações – vol. 2
(coleção Biotecnologia Aplicada à Saúde, vol. 2)

© 2015 Rodrigo Ribeiro Resende (organizador)
Editora Edgard Blücher Ltda.

Blucher

Rua Pedroso Alvarenga, 1245, 4º andar
04531-934, São Paulo – SP – Brasil
Tel.: 55 11 3078-5366
contato@blucher.com.br
www.blucher.com.br

Segundo o Novo Acordo Ortográfico, conforme 5ª ed.
do *Vocabulário Ortográfico da Língua Portuguesa*,
Academia Brasileira de Letras, março de 2009.

É proibida a reprodução total ou parcial por quaisquer
meios, sem autorização escrita da Editora.

Todos os direitos reservados pela Editora Edgard Blücher Ltda.

FICHA CATALOGRÁFICA

Biotecnologia aplicada à saúde: fundamentos e
aplicações, volume 2 / organizado por Rodrigo
Ribeiro Resende; colaboração de Carlos Ricardo
Soccol. – São Paulo: Blucher, 2015.

ISBN 978-85-212-0921-8

1. Biotecnologia 2. Clonagem 3. DNA recombinante 4. Terapia
gênica 5. Organismos transgênicos - Animais I. Resende,
Rodrigo Ribeiro II. Soccol, Carlos Ricardo

15-0508

CDD 620.8

Índices para catálogo sistemático:

1. Biotecnologia

AGRADECIMENTOS

Esta obra não poderia ter iniciado sem a dedicação de cada um que participou em sua elaboração, desde os professores, alunos, editores, corretores, diagramadores, financiadores, amigos, esposa, irmão, pais, leitores até o desejo de tornar o conhecimento acessível para todos. Obrigado!

Prof. Rodrigo R. Resende

APOIO



Agradecemos pelo apoio financeiro aos projetos científicos da Fapemig, CNPq, Capes, Instituto Nacional de Ciência e Tecnologia em Nanomateriais de Carbono, Rede Mineira de Toxinas com Ação Terapêutica, Instituto Nanocell.

CARTA AO LEITOR

Parece brincadeira quando vejo as estatísticas do governo federal ou de agências internacionais sobre a qualidade do ensino. Elas informam que 70% dos brasileiros são analfabetos funcionais. A Organização para a Cooperação e Desenvolvimento Econômico (OCDE) divulgou em maio de 2015 um *ranking* mundial de qualidade de educação. Dentre os 76 países avaliados, o Brasil ocupa a 60ª posição.

O Analfabetismo Funcional constitui um problema silencioso e perverso que afeta as empresas e o ensino superior. Não se trata de pessoas que nunca foram à escola. Elas sabem ler, escrever e contar; chegam a ocupar cargos administrativos, mas não conseguem compreender a palavra escrita.[...] Elas preferem ouvir explicações da boca de colegas. "Calcula-se que, no Brasil, os analfabetos funcionais representam 68% da população economicamente ativa, e, somados aos 7% totalmente analfabetos, temos que 75% da população não possui o domínio pleno da leitura, da escrita e das operações matemáticas" *.

E, no afã de melhorar a educação e o Brasil para que se torne uma Nação, nada melhor do que reunir os líderes nacionais de renome internacional para escreverem, junto comigo, uma obra alinhada com a perspectiva de construção de uma nação forte e indelével, produtora de conhecimento e tecnologia. Para isso, uma obra técnica com linguagem para leigos e a disponibilização de tecnologias de ponta, apresentadas ao final de cada capítulo desta coleção, ao alcance de todos, desde o aluno, passando pelo funcionário, até o presidente de uma pequena ou grande companhia de biotecnologia. Uma maneira de todos nós, professores e cientistas, darmos ao povo brasileiro a oportunidade de crescer pelo do conhecimento!

Prof. Rodrigo R. Resende
Presidente do Instituto Nanocell
Presidente da Sociedade Brasileira de Sinalização Celular

* Fonte: Flavia Leite, Analfabetismo funcional, uma realidade atual e intrigante. Webartigos. 24 jul. 2011. Disponível em: <<http://www.webartigos.com/artigos/analfabetismo-funcional-uma-realidade-atual-e-intrigante/72269/>>.

CONTEÚDO

| | |
|---|----|
| <i>Prefácio – Helena B. Nader – SBPC</i> | 9 |
| <i>Apresentação – Willibaldo Schmidell – UFSC</i> | 11 |

Vetores e Clonagem

| | |
|---|-----|
| 1. Novas Técnicas de Recombinação em Microrganismos: Engenharia Genética pela Tecnologia do DNA Recombinante | 15 |
| 2. Vetores Plasmidiais e Vetores de Fago: Estrutura Básica e Funções | 77 |
| 3. Vetores de Expressão Bacteriana: Tipos e Usos | 115 |
| 4. Clonagem Gênica: Fundamentos e Aplicações | 137 |

Proteínas Recombinantes

| | |
|---|-----|
| 5. Expressão de Proteínas Recombinantes em Bactérias: Fundamentos Básicos e Aplicações | 177 |
| 6. Produção de Proteínas Recombinantes em Leveduras e Fungos Filamentosos: Fundamentos Básicos e Rotina | 217 |
| 7. Baculovírus para Expressão de Proteínas Recombinantes em Células de Insetos | 255 |
| 8. Expressão de Proteínas Recombinantes em Células de Mamíferos | 307 |
| 9. Mutagênese Sítio-Dirigida em Bactérias: Fundamentos Básicos e Aplicações | 339 |
| 10. Anticorpos Poli e Monoclonais como Ferramenta Biotecnológica: Produção e Usos | 361 |
| 11. Clonagem e Expressão de Genes de Anticorpos: Métodos e Aplicações | 435 |
| 12. Anticorpos Humanizados | 481 |
| 13. Plantas como Biorreatores: Fundamentos, Métodos e Aplicações | 551 |
| 14. Animais Domésticos como Biorreatores: Produção e Usos | 609 |
| 15. Planta <i>Biofarming</i> : Perspectivas Inovadoras para a Expressão de Peptídeos em Sistemas Heterólogos | 637 |

Terapia Gênica

| | |
|--|-----|
| 16. Derivação de Novas Linhagens de Células-Tronco Embrionárias: Evolução da Metodologia | 673 |
| 17. Proteína <i>Prion</i> : Biologia Molecular, Aspectos Fisiológicos e Patológicos, Diagnóstico e Interesse Sanitário | 713 |
| 18. Células-Tronco Mesenquimais Adultas de Diversas Origens: Uma Visão Geral Multiparamétrica para Aplicações Clínicas | 745 |
| 19. Produção e Diferenciação de Células-Tronco Induzíveis | 815 |
| 20. Reprogramação Nuclear de Células para um Estado Pluripotente Usando Três Estratégias | 879 |
| 21. Mecanismos de Entrada dos Vírus na Célula: Métodos de Estudo e Aplicações | 903 |
| 22. Vetores Virais: Tipos, Diversidades, Usos e Aplicações | 921 |
| 23. Modelos de Terapia Gênica Baseados na Expressão de Hormônio de Crescimento em Células Humanas e na Administração de DNA Plasmidial em Camundongos Anões | 981 |

Animais Transgênicos

| | |
|---|------|
| 24. Produção de Animais <i>Knockouts</i> e <i>Knockins</i> | 1015 |
| 25. Sistema <i>Cre-Lox</i> : Transgênicos Tecido e Tempo Programados | 1037 |
| 26. Sistema Duplo-Híbrido em Levedura: Conceitos e Aplicações | 1073 |
| 27. Estratégias Transgênicas para a Expressão Combinatória de Proteínas Fluorescentes | 1107 |
| 28. Transplante de Espermatogônias-Tronco como Abordagem Biotecnológica para a Produção de Animais Transgênicos | 1131 |
| Autores | 1173 |

PREFÁCIO

Esta é uma obra audaciosa. De um lado, ela contém a saudável ousadia de congregar o conhecimento científico e tecnológico disponível no campo da biotecnologia em suas aplicações nas áreas da saúde e da agroindústria. Do outro, é o resultado do trabalho de 376 autores, todos com destacada atuação na vida acadêmica brasileira e internacional.

Os quatro volumes de *Biotecnologia Aplicada à Saúde: Fundamentos e Aplicações* e *Biotecnologia Aplicada à Agro&Indústria: Fundamentos e Aplicações*, ao reunir uma centena de artigos, expressam diversidade temática de tal grandeza que desaconselha sua sumarização neste espaço. Este conjunto ainda conseguiu agregar, por meio de textos sucintos e objetivos, os passos mais recentes da biotecnologia com os conhecimentos clássicos que estavam registrados em diferentes publicações. Em cada capítulo são apresentadas as raízes que sustentam as técnicas e os modelos apresentados pelos autores. Pensando no aprendizado de estudantes – e também nos passos de profissionais que se interessem em introduzir-se na biotecnologia utilizando técnicas recentes e sem restrições de modelos celulares –, os textos apresentam as metodologias e os procedimentos para a realização de experimentos.

Assim, temos em mãos uma obra didática que vem a contribuir com nossos cursos de graduação e de pós-graduação, de alguma maneira dedicados à biotecnologia e/ou às áreas da saúde e da agroindústria.

Há ainda mais uma grande virtude neste trabalho organizado pelos professores Rodrigo Resende e Carlos Ricardo Soccol: sua sintonia com as necessidades atuais da ciência brasileira em uma área estratégica para o país.

A biotecnologia é uma das disciplinas nas quais mais facilmente podemos enxergar o impacto social da ciência. Por exemplo, o diagnóstico e tratamento de muitas doenças são fruto do uso da biotecnologia na saúde humana e animal, assim como a redução no uso de agrotóxicos na agricultura.

Nas últimas duas décadas, a ciência brasileira deu um grande salto. Resume esse avanço o fato de termos passado de posição intermediária para posição de relativo destaque na produção científica mundial. Contudo, três grandes desafios se apresentam correntemente: aumentar o impacto da ciência brasileira na ciência mundial; dar à ciência e tecnologia posição de protagonista do desenvolvimento econômico, social e cultural do país; expandir o sistema de produção científica, hoje baseado no mundo acadêmico, para o mundo empresarial.

Considerando ainda que a biotecnologia é uma área das mais promissoras para o nosso desenvolvimento sustentado, está mais do que entendida a importância desta obra em todos os contextos em que ela se encaixa.

Ademais, *Biotecnologia Aplicada à Saúde: Fundamentos e Aplicações* e *Biotecnologia Aplicada à Agro&Indústria: Fundamentos e Aplicações* revelam a abrangência da biotecnologia brasileira e sua inserção dinâmica em nossa vida científica. Os 376 autores estão distribuídos por 39 universidades e 27 institutos de pesquisa, e atuam em mais de 150 programas de pós-graduação.

Como presidente de uma instituição cuja principal missão é trabalhar para o progresso da ciência no Brasil, entendo que a chegada desta obra deve ser comemorada, e seu conteúdo – pertinente e atual – certamente vai ajudar a biotecnologia a acelerar a ciência brasileira e seu protagonismo na vida nacional.

Helena B. Nader

Professora Titular de Biologia Molecular,
Escola Paulista de Medicina, UNIFESP
Presidente da Sociedade Brasileira
para o Progresso da Ciência (SBPC)

APRESENTAÇÃO

Antes de qualquer consideração, devo destacar a enorme satisfação em ter sido convidado para efetuar a presente “Apresentação”, motivo pelo qual agradeço aos organizadores Rodrigo R. Resende e Carlos Ricardo Soccol pela grata oportunidade.

Trata-se de uma sequência de quatro livros, sendo três deles (volumes I ao III) intitulados *Biotechnology Aplicada à Saúde: Fundamentos e Aplicações*, e o quarto (Volume IV), *Biotechnology Aplicada à Agro&Indústria: Fundamentos e Aplicações*. O Volume I está dividido em 4 partes que compreendem 19 capítulos; o Volume II, 4 partes e 28 capítulos; o Volume III, 5 Partes e 26 capítulos; e o Volume IV, 6 partes e 27 capítulos, ou seja, um global de 19 grandes tópicos e 100 capítulos. Certamente o esforço dos organizadores foi muito significativo, pois acionaram nada menos do que 376 autores de capítulos, que militam em 69 instituições de ensino e pesquisa no Brasil.

Parte significativa desta sequência de livros se dedica ao desenvolvimento de bioprocessos para a produção de produtos de alto valor agregado, utilizados, por exemplo, na área de saúde humana e animal, sendo parte deles através do emprego de células alteradas geneticamente, de forma a se contar com agentes de conversão mais adequados, bem de acordo com os atuais desenvolvimentos relacionados à moderna biotecnologia.

Deve-se ressaltar, como algo de grande importância, a preocupação nos vários capítulos com a descrição de procedimentos experimentais detalhados, objetivando um mais rápido aprendizado das técnicas disponíveis na moderna biotecnologia.

Assim sendo, tenho a certeza de que este conjunto de quatro livros irá em muito acelerar o desenvolvimento da biotecnologia em nosso país. Trata-se de uma obra necessária para o ensino e a pesquisa da biotecnologia, tanto na graduação quanto na pós-graduação. Cabe, portanto, cumprimentar os organizadores pela iniciativa e esforço para concretizar a presente obra.

Prof. Dr. Willibaldo Schmidell

Professor Titular (Aposentado) na Escola Politécnica da USP

Professor Visitante

Depto. Engenharia Química e Eng. de Alimentos

Universidade Federal de Santa Catarina (UFSC)

VETORES E CLONAGEM

NOVAS TÉCNICAS DE RECOMBINAÇÃO EM MICRORGANISMOS: ENGENHARIA GENÉTICA PELA TECNOLOGIA DO DNA RECOMBINANTE

Edson Júnior do Carmo
Márcia Neiva
Spartaco Astolfi-Filho

1.1 INTRODUÇÃO

A elucidação da estrutura do DNA por Watson e Crick¹ em 1953 foi um marco para a biologia molecular, e os anos que se seguiram foram ricos em descobertas sobre sua replicação, transcrição e expressão. Nesse período estabeleceram-se as bases para a síntese química de DNA e desvendaram-se também o código genético e o processo de regulação da expressão gênica em seres procariontes (bactérias). Nessas duas décadas de 1950 e 1960, extremamente profícuas nas descobertas que desvendaram o dogma central da

biologia molecular, essa área de pesquisa era sem dúvida a mais acadêmica das áreas das ciências biológicas. Entretanto, foi nesse período que as bases foram estabelecidas para o surgimento da Tecnologia do DNA Recombinante (TDR), a qual também pode ser denominada de engenharia genética moderna.

No início dos anos 1970, dois grupos do estado da Califórnia (Estados Unidos) inventaram dois processos diferentes de se realizar recombinação de DNA de espécies diferentes *in vitro*. Em 1971, o grupo de Paul Berg (da Stanford University) construiu *in vitro* a primeira molécula de DNA recombinante, um híbrido de DNA de bacteriófago lambda e do vírus SV-40², sendo que por razões de ordem ética e de biossegurança essa molécula híbrida não foi introduzida na bactéria *Escherichia coli*. Em 1973, Herbert Boyer (University of California at San Francisco) e Stanley Cohen (Stanford University) construíram moléculas híbridas de plasmídeos bacterianos³ e também introduziram DNA ribossomal de sapo em um plasmídeo, e essas moléculas híbridas foram então introduzidas na bactéria *E. coli*, em que mostram a capacidade de replicação conferida pelos plasmídeos.

O conjunto de ferramentas e técnicas utilizadas no âmbito da TDR veio sanar limitações metodológicas existentes até então e tornou possível o isolamento de genes específicos a partir do DNA das mais diversas fontes. Assim, possibilitou que se estudasse melhor a regulação desses genes, em especial os provenientes de eucariotos. Além disso, essas técnicas viabilizaram a recombinação gênica entre espécies distintas, abrindo a possibilidade de criação de organismos não encontrados na natureza, os denominados transgênicos. Esses organismos, por sua vez, podem ser utilizados para diversos estudos ou para a área de biotecnologia industrial, especialmente nos setores agropecuário e farmacêutico.

Até o surgimento dessa tecnologia, no setor farmacêutico industrial as principais fontes disponíveis para a obtenção de proteínas e outras moléculas para uso terapêutico eram animais e plantas. Nesse contexto, a insulina humana que era obtida por processo de extração de pâncreas de boi e porco passou a se produzida por *E. coli* geneticamente manipulada, resultando no primeiro produto recombinante de uso clínico aprovado pela Food and Drug Administration (FDA, a agência regulatória de medicamentos e alimentos dos Estados Unidos) no início da década de 1980, e abriu as portas para uma nova era na obtenção de proteínas em larga escala. Desde então, o termo biofármaco começou a ser utilizado para designar os produtos terapêuticos obtidos com o uso da TDR e iniciou-se um intenso processo de desenvolvimento técnico, objetivando a utilização de outros microrganismos como leveduras e fungos filamentosos.

Dados recentes mostram que no mercado existem 151 produtos recombinantes aprovados para uso terapêutico pela FDA e pela European Medicines Agency (EMA). Aproximadamente 50% deles são obtidos com o uso de microrganismos, sendo que 30% dessas proteínas são produzidas por *E. coli*⁴, que ainda é o sistema mais amplamente utilizado tanto para clonagem quanto para a expressão de proteínas heterólogas. Isso se dá pela facilidade de manuseio, crescimento rápido, fácil mudança de escala de produção e custo relativo baixo. No entanto, embora a *E. coli* ainda ocupe esse posto de destaque, a demanda pela produção eficiente de proteínas mais complexas, com modificações pós-traducionais, que não são feitas em bactérias, principalmente glicoproteínas e em especial anticorpos monoclonais (do inglês *monoclonal antibodies*, mAbs), levou ao desenvolvimento de sistemas eficientes de produção desses tipos de moléculas baseados em células de mamíferos em cultura. Para produção de mAbs em larga escala utilizam-se, atualmente, células de mamífero da linhagem CHO (do inglês *chinese hamster ovary*), geneticamente programadas contendo genes de anticorpos monoclonais.

Hoje, cerca de quarenta anos após o início da engenharia genética (TDR), analisando a evolução da área, podemos dizer que a TDR foi uma evolução natural da biologia molecular, pois um de seus maiores trunfos atuais é a possibilidade de sintetizar quimicamente genes naturais ou desenhados por nós. Essa técnica foi desenvolvida em grande parte pelo grupo liderado por H. G. Khorana⁵, que trabalhou em diversas Instituições nos Estados Unidos, inclusive no Massachusetts Institute of Technology (MIT). Essa facilidade, turbinada pelo advento da automatização da reação em cadeia da polimerase (do inglês *polymerase chain reaction* – PCR)⁶, popularizou outra forma de denominar o conjunto das técnicas atuais avançadas de engenharia genética, engenharia de proteínas e engenharia metabólica: “biologia sintética”. Além disso, podemos dizer que graças a essa tecnologia, a biologia molecular deixou de ser essencialmente acadêmica e passou a ser a área das ciências biológicas com maiores possibilidades de aplicações no setor industrial.

Neste capítulo, veremos uma síntese das principais ferramentas utilizadas na Tecnologia do DNA Recombinante, seu uso e evolução, assim como exemplos de aplicação na área farmacêutica.

1.2 AS FERRAMENTAS DA TDR

De modo sucinto, a TDR pode ser definida como um conjunto de técnicas que possibilita a manipulação e combinação de sequências gênicas

provenientes de diferentes organismos com os mais diversos fins. No laboratório, quando estamos isolando genes e construindo células/seres transgênicos, também dizemos que estamos realizando “Clonagem Molecular de Genes”. O processo de clonagem molecular baseia-se principalmente em quatro passos principais, a saber:

- 1) Seleção e isolamento de uma sequência de DNA ou um gene a ser clonado.
- 2) Seleção de um elemento transportador dessa sequência clonada para uma célula hospedeira, denominado vetor genético.
- 3) Ligação da sequência a ser clonada ao vetor genético e inserção do conjunto gene/vetor na célula hospedeira.
- 4) Seleção de células hospedeiras possuidoras da marca do vetor que indica que tais células incorporaram o DNA de interesse.

A molécula recombinante é introduzida em uma célula hospedeira (Figura 1.1), na qual se multiplica autonomamente ou em paralelo ao genoma hospedeiro. À medida que a célula hospedeira receptora do DNA recombinante vai se dividindo, o material recombinante também se divide e é transmitido às células filhas, formando, assim, um clone de células iguais. Como consequência, o gene ou o fragmento de DNA de interesse é amplificado muitas vezes e pode ser isolado e caracterizado.

No processo de clonagem molecular são utilizados diferentes instrumentos (ferramentas) moleculares, que serão apresentados a seguir.

1.2.1 Endonucleases de restrição

Enzimas que clivam moléculas de ácidos nucleicos são chamadas de nucleases. As que clivam as moléculas no seu interior são denominadas de endonucleases, enquanto as que degradam os ácidos nucleicos de suas extremidades para o interior são chamadas de exonucleases.

As endonucleases de restrição são consideradas enzimas-chave no processo de clonagem molecular. São enzimas que naturalmente fazem parte do sistema de defesa de um microrganismo procarionte, protegendo-o contra a invasão de DNA exógenos, por exemplo, de um vírus cujo genoma seja de DNA que queira infectar a célula. Essas enzimas reconhecem sequências específicas do DNA dupla fita e têm a capacidade de cortar seletivamente o DNA estranho em regiões específicas, denominadas sítios de restrição. É essa

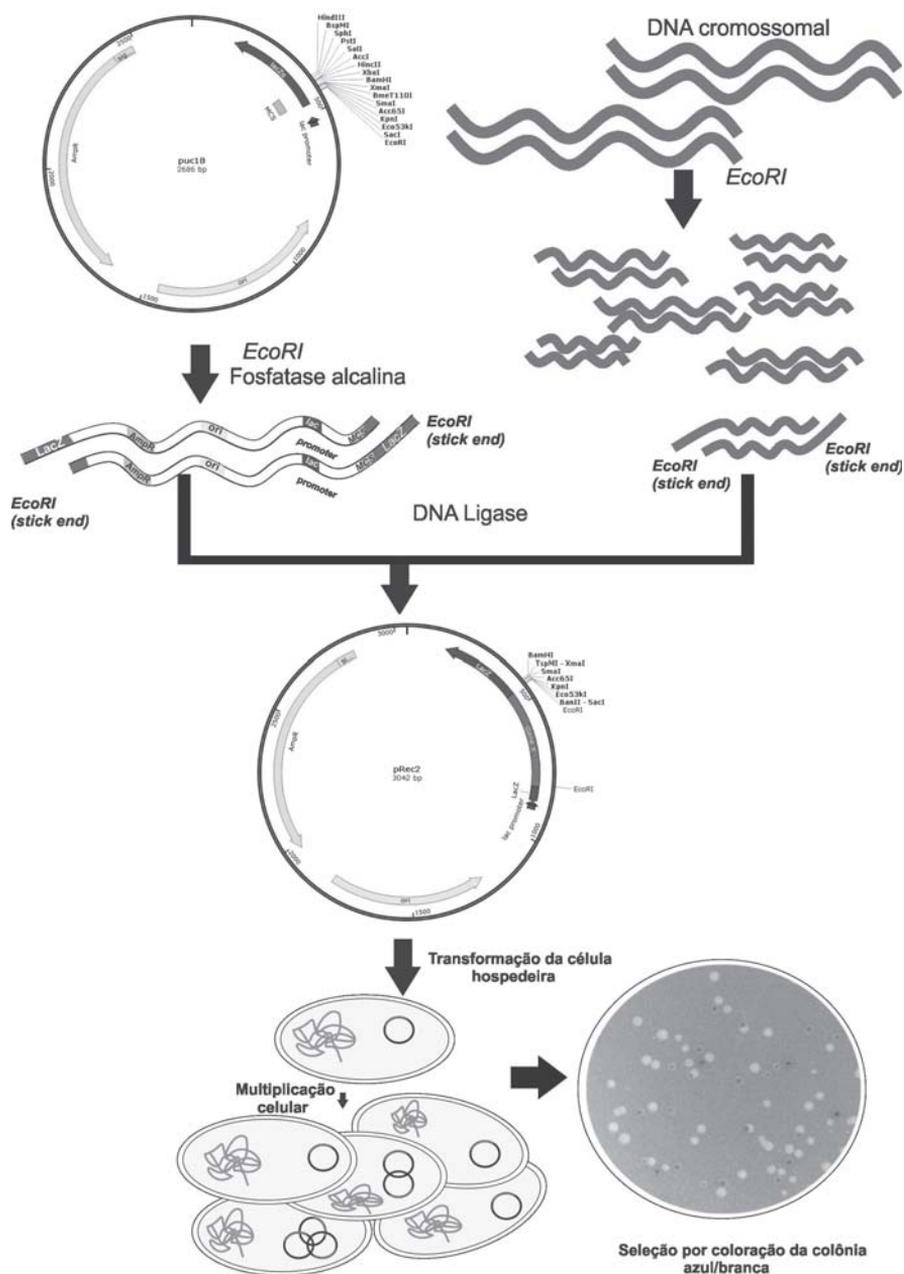


Figura 1.1 Exemplo de clonagem molecular – construção de uma biblioteca genômica. O vetor de clonagem, no caso o plasmídeo pUC18, é digerido com a enzima de restrição *EcoRI* e então tratado com fosfatase alcalina. O DNA cromossomal é digerido com a mesma enzima de restrição, e a seguir os fragmentos são ligados ao plasmídeo pela DNA ligase, resultando em um conjunto de plasmídeos recombinantes pRECs. Esses plasmídeos são, então, inseridos nas células hospedeiras, num processo denominado de transformação genética. Tais células, por sua vez, são plaqueadas em meio seletivo.

habilidade que torna as endonucleases de restrição ferramentas importante na manipulação de ácidos nucleicos.

As endonucleases de restrição utilizadas em engenharia genética são originárias de microrganismos e podem ser classificadas em três tipos (tipos I, II e III), de acordo com a especificidade de ação, cofatores e substratos requeridos. Como os mecanismos de ação das enzimas do tipo I e tipo III são mais complexos e não são muito utilizados em engenharia genética, esses tipos não serão abordados neste capítulo.

As enzimas do tipo II são empregadas para esse fim por reconhecerem sequências nucleotídicas específicas, geralmente palindrômicas. Uma sequência palindrômica, ou simplesmente palíndromo, é um segmento de DNA em que as duas fitas têm a mesma sequência, tanto quando uma fita é lida da esquerda para a direita como quando a outra é lida da direita para esquerda. Ao reconhecer um palíndromo específico, as enzimas de restrição cortam DNA dupla fita, na sequência de reconhecimento, deixando os fragmentos com dois tipos de extremidades possíveis: extremidades coesivas (*stick ends*) ou extremidades abruptas (*blunt ends*).

A enzima de restrição *Bam*HI, como pode ser observado na Figura 1.2, reconhece a sequência 5'-GGATCC-3' e, após o corte, origina extremidades de DNA coesivas (*stick ends*), com extremidades 5' protuberantes. Do mesmo modo, a enzima *Pst*I reconhece a sequência 5'-CTGCAG-3' e o corte origina extremidades também coesivas, mas dessa vez a extremidade protuberante é

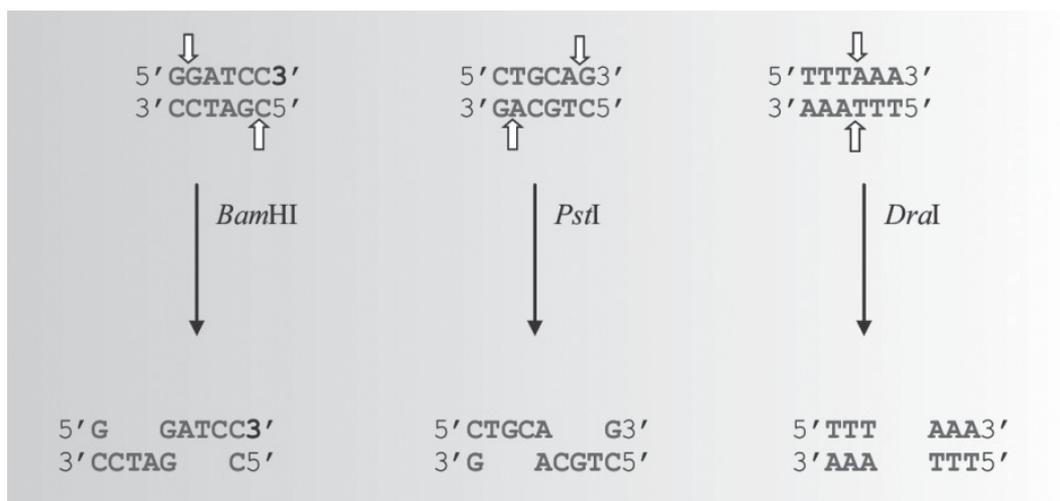


Figura 1.2 Esquema ilustrativo dos tipos de corte de enzimas de restrição. (a) Enzima *Bam*HI: extremidade 5' coesiva, (b) enzima *Pst*I: extremidade 3' coesiva, (c) enzima *Dra*I: extremidade abrupta. Note que as sequências de reconhecimento são palíndromos.

a 3'. Diferentemente, outros tipos de enzimas, como a enzima *DraI*, clivam a sequência de reconhecimento na mesma posição nas duas fitas de DNA, originando extremidades de DNA abruptas em fita dupla (*blunt ends*).

É importante frisar que um dos passos da construção de uma molécula de DNA recombinante é a ligação ou união entre fragmentos, e as extremidades coesivas geradas pelo corte podem ser pareadas por complementaridade e ligadas eficientemente pelas DNA ligases. Nesse sentido, é também possível ligar fragmentos de DNA obtidos por cortes de enzimas diferentes desde que tenham extremidades compatíveis, como no caso dos cortes com *Bam*HI e *Bg*III (Figura 1.3). Fragmentos de DNA de extremidades abruptas podem ser ligados, porém o são com menos eficiência. Para contornar essa limitação, o pesquisador pode aumentar a quantidade de DNA ligase no sistema de ligação.

As sequências de reconhecimento são, na maioria das vezes, de quatro ou seis nucleotídeos. No entanto, há enzimas de cortes raros que reconhecem e clivam um ponto específico de uma sequência de oito nucleotídeos ou de até mais nucleotídeos. Enzimas de cortes raros, como *NotI* e *PacI*, que reconhecem sequências de oito nucleotídeos, são interessantes em estratégias de clonagem molecular, pois têm menos chances de ser encontradas ao longo das sequências do DNA-alvo.

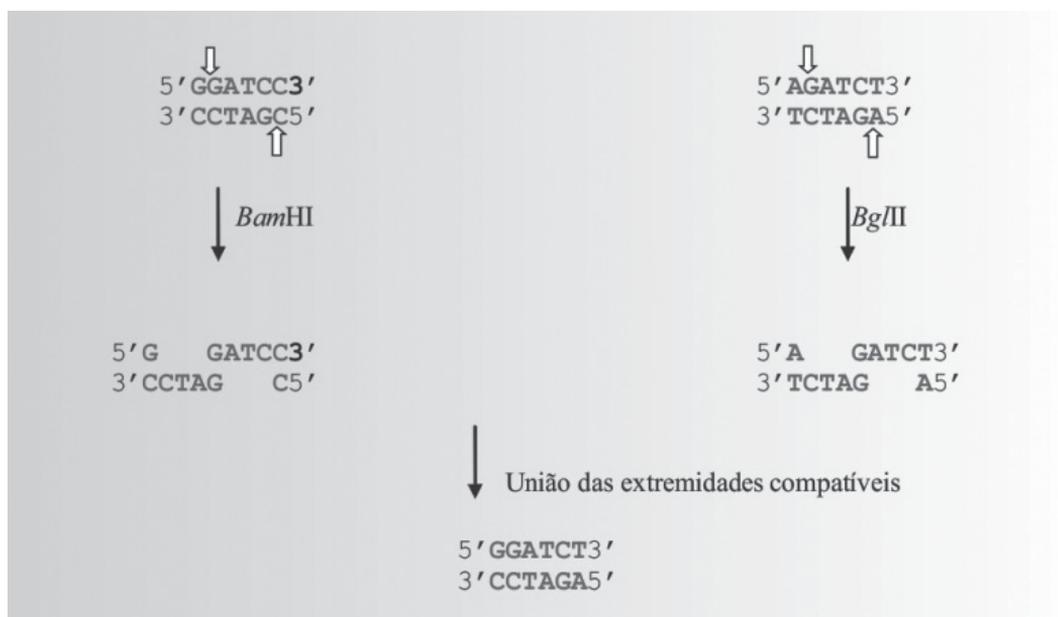


Figura 1.3 Compatibilidade dos sítios de restrição. União de extremidades coesivas geradas por cortes produzidos por enzimas de restrição diferentes.

Uma maneira prática de se determinar a ocorrência de certo sítio na molécula é fazer um experimento de análise de restrição com a enzima e visualizar os fragmentos obtidos por eletroforese em gel de agarose. Na Figura 1.4 pode ser visualizada uma análise de restrição de plasmídeos recombinantes de clones de uma biblioteca de cDNA da glândula do veneno da serpente *Bothrops atrox*. Os cDNAs com extremidades colantes de *Sfi*I foram ligados a moléculas do vetor pDNR-Lib, pré-digerido com a mesma enzima e tratados com fosfatase alcalina, e o conjunto inserido por eletroporação em *E. coli*. Plasmídeos extraídos de doze diferentes clones recombinantes foram digeridos com a enzima de restrição *Sfi*I e analisados por eletroforese em gel de agarose para se determinar o tamanho dos fragmentos clonados. Um experimento dessa natureza pode tanto permitir a determinação dos tamanhos de fragmentos clonados quanto a quantidade de sítios presentes ao longo da molécula.

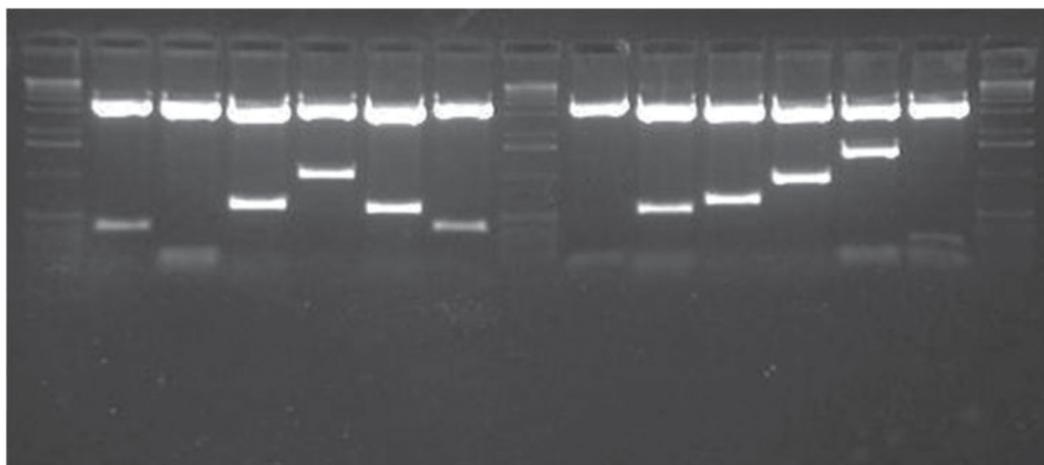


Figura 1.4 Perfil eletroforético em gel de agarose dos fragmentos de restrição. Neste experimento foi possível mostrar que cDNAs de diferentes tamanhos foram clonados no plasmídeo pDNR-Lib. A visualização do DNA foi possível pela coloração com o composto fluorescente brometo de etídio.

Em relação à nomenclatura, a denominação dada a uma enzima de restrição é derivada do nome do microrganismo de onde é isolada, contendo também a identificação da cepa e um número correspondente à ordem cronológica de isolamento de enzimas daquela espécie. Por exemplo, a primeira enzima de restrição isolada da bactéria *Bacillus globigii* recebeu o nome de *Bgl*I; a descoberta de outra enzima na mesma bactéria que reconhece

e clivava outro sítio de restrição foi nomeada de *BglII*. Da mesma forma, enquanto a primeira enzima de restrição de *Escherichia coli* cepa R recebeu o nome *EcoRI*, na ordem cronológica, a quinta enzima isolada dessa cepa foi denominada *EcoRV*. Na Tabela 1.1, a seguir, são apresentadas algumas enzimas de restrição utilizadas em engenharia genética, suas sequências de reconhecimento e o microrganismo de origem.

Tabela 1.1 Algumas das principais enzimas de restrição utilizadas em clonagem molecular

| ENZIMA DE RESTRIÇÃO | SÍTIO DE CLIVAGEM 5'-----3' | TIPO DE CORTE | ORGANISMOS DE ORIGEM |
|---------------------|--------------------------------|---------------|-----------------------------------|
| AluI | AGCT | <i>Blunt</i> | <i>Arthrobacter luteus</i> |
| BamHI | GGATCC | <i>Stick</i> | <i>Bacillus amyloliquefaciens</i> |
| BglII | AGATCT | <i>Stick</i> | <i>Bacillus globigii</i> |
| BspMAI | CTGCAG | <i>Stick</i> | <i>Bacillus species M</i> |
| DraI | TTTAAA | <i>Blunt</i> | <i>Deinococcus radiophilus</i> |
| EcoRI | GAATTC | <i>Stick</i> | <i>Escherichia coli</i> |
| HaeIII | GGCC | <i>Blunt</i> | <i>Haemophilus aegyptius</i> |
| HpaII | CCGG | <i>Stick</i> | <i>Haemophilus parainfluenzae</i> |
| KpnI | GGTACC | <i>Stick</i> | <i>Klebsiella pneumoniae</i> |
| NotI | GCGGCCGC | <i>Stick</i> | <i>Nocardia otitidis-caviarum</i> |
| PacI | TTAATTAA | <i>Stick</i> | <i>Pseudomonas alcaligenes</i> |
| PstI | CTGCAG | <i>Stick</i> | <i>Providencia stuartii</i> |
| PvuI | CGATCG | <i>Stick</i> | <i>Proteus vulgaris</i> |
| PvuII | CAGCTG | <i>Blunt</i> | <i>Proteus vulgaris</i> |
| SnaBI | TACGTA | <i>Blunt</i> | <i>Sphaerotilus natans</i> |
| XhoI | CTCGAG | <i>Stick</i> | <i>Xanthomonas holcicola</i> |

É possível observar na Tabela 1.1 que as enzimas *BspMAI* e *PstI* reconhecem o mesmo sítio de clivagem. Quando duas ou mais enzimas de restrição reconhecem a mesma sequência nucleotídica, são chamadas de isosquizômeros. Em engenharia genética, o emprego de isosquizômeros é vantajoso

principalmente para isolar ou produzir fragmentos de DNA com enzimas alternativas, aumentando as possibilidades de estratégias de clonagem.

1.2.2 DNA ligase

A DNA ligase é uma ferramenta fundamental para a formação de moléculas recombinantes de DNA, pois essa enzima é, na maioria dos casos, responsável por selar eficientemente a união de duas moléculas diferentes. A DNA ligase promove a ligação fosfodiéster entre os grupos fosfato da extremidade 5' e a hidroxila da extremidade 3' de moléculas de DNA numa reação dependente de ATP e magnésio (Mg^{2+}) como cofator. Nas células, a ligase participa do processo de duplicação do DNA da ligação dos fragmentos de Okazaki e do processo de reparo das moléculas de DNA. As reações de ligação são mais eficientes em extremidades coesivas, que podem se emparelhar por complementaridade e deixar as extremidades 3'OH e 5'P mais próximas para ocorrer a ligação. No entanto, extremidades abruptas também podem ser ligadas pela ação catalítica dessa enzima. A Figura 1.5 ilustra os detalhes da ação da T4 DNA ligase extraída de *E. coli* infectada com bacteriófagos T4.

1.2.3 DNA polimerase I

A primeira DNA polimerase isolada foi a DNA polimerase I de *Escherichia coli*. Como todas as DNA polimerases conhecidas, ela catalisa a síntese de DNA utilizando como substrato nucleotídeos trifosfatados (nucleosídeos trifosfatos) e uma molécula de DNA fita simples preexistente a partir de um iniciador (*primer*). A nova fita é polimerizada no sentido 5' para 3', como mostra a Figura 1.6.

Além da atividade de síntese na direção 5' para 3', ela possui atividade de exonuclease tanto na direção 5' para 3' quanto 3' para 5'. Uma das principais utilizações era a obtenção de sondas de DNA marcadas radioativamente, num processo denominado *nick-translation*.

A DNA polimerase I tem sido utilizada *in vitro* para diferentes propósitos, como, por exemplo: marcação de DNA por *nick translation*, síntese da segunda fita de cDNA, amplificação de fragmentos de DNA ou mesmo genes inteiros por meio de uma técnica conhecida como reação em cadeia da polimerase (PCR), que será explicada mais adiante neste capítulo. Para reações de PCR, utiliza-se a DNA polimerase I da bactéria *Thermus*

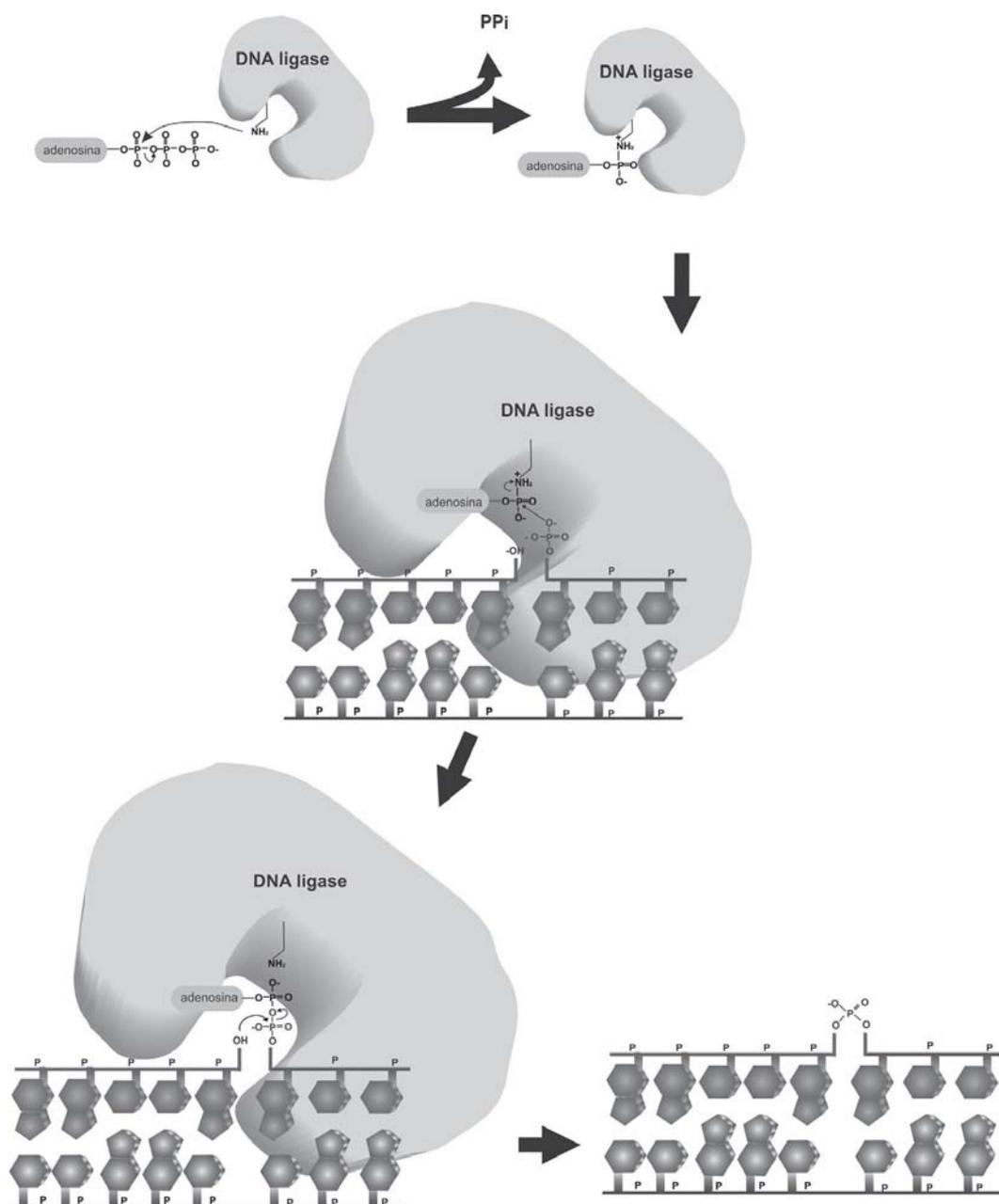


Figura 1.5 Mecanismo de ação da T4 DNA ligase. A enzima é ativada ao se ligar ao nucleosídeo AMP a partir da reação com o ATP. O grupo AMP-enzima liga-se covalentemente ao grupo fosfato 5' exposto, facilitando que o átomo de fósforo sofra ataque nucleofílico a partir do oxigênio da hidroxila da extremidade 3', catalisando a ligação das duas extremidades. Na sequência, o complexo grupo AMP-enzima se solta da molécula de DNA.

aquaticus, que tem alto grau de resistência ao aquecimento e é denominada de *Taq* polimerase.

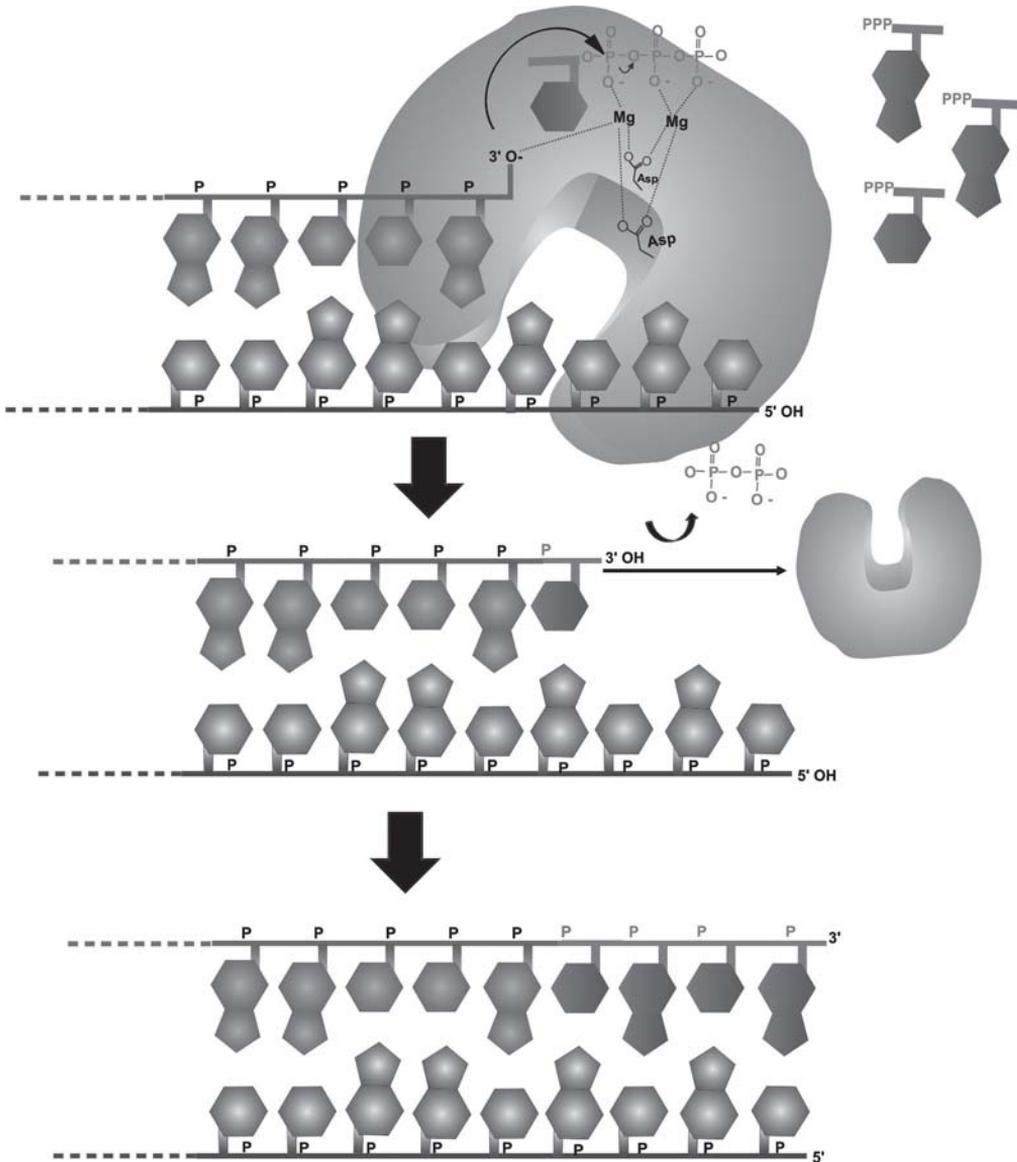


Figura 1.6 DNA polimerase I. Síntese de uma nova fita de DNA pela DNA polimerase I a partir de um *primer* e uma fita molde. No sítio ativo da enzima, o fosfato alfa do nucleosídeo trifosfato (substrato) sofre ataque nucleofílico do par de elétrons da hidroxila 3'OH do *primer*, resultando na adição de um nucleosídeo monofosfato à cadeia nascente e liberação de pirofosfato.

Algumas DNA polimerases termofílicas, como a *Taq* polimerase, adicionam após a síntese um nucleotídeo de adenina em cada extremidade 3'-OH. Essa característica é utilizada para clonagem em outras moléculas com extremidades complementares, terminadas em nucleotídeos de timina. Outras polimerases recombinantes, como a enzima denominada Pfu, não adicionam nucleotídeos nas extremidades, deixando-as propícias à ligação em moléculas de extremidades *blunt ends*.

1.2.4 Enzima Klenow

É um fragmento proteico resultante da clivagem da DNA polimerase I de *E. coli* com a protease subtilisina, que elimina a atividade de exonuclease da direção 5' para 3'. Mantém apenas a atividade polimerásica dependente de *primer* e a atividade exonucleásica do sentido 3' para o 5', que pode ser inibida na presença de excesso de nucleotídeos.

É utilizada nas ocasiões em que se deseja obter a síntese de novas fitas de DNA, mas sem correr o risco de que suas extremidades 5' sejam degradadas, como por exemplo: para realizar reações de sequenciamento pelo método clássico de Sanger, para preencher extremidades coesivas (*stick ends*) de sítios de enzimas de restrição transformando-as em extremidades abruptas (*blunt ends*) e para obter sondas de DNA marcadas radioativamente pelo método de *random primers*.

Na marcação de DNA pelo método de *random primers*, o DNA dupla fita a ser marcado é misturado com um conjunto de *primers* de cinco ou seis nucleotídeos sintetizados com sequências aleatórias (randômicas) e, então, desnaturado por aquecimento em banho-maria. Em seguida, é colocado no gelo para que os *primers* se anelem ao DNA fita simples em sequências complementares. A seguir, adicionam-se os quatro nucleotídeos trifosfatados, sendo pelo menos um deles marcado na posição alfa com P³² e a enzima Klenow. A temperatura é então elevada a 16 °C, para que catalise a polimerização e ocorra a conseqüente marcação radioativa.

1.2.5 T4 DNA polimerase

Esta enzima também catalisa a síntese de DNA na direção 5' para 3', requerendo uma fita molde de nucleotídeos trifosfatados e um *primer*. Ao contrário da DNA polimerase I, porém, a T4 DNA polimerase não possui

atividade exonucleásica do sentido 5' para 3'; no entanto, sua atividade exonucleásica no sentido 3' para 5' é consideravelmente superior. Por esse motivo, é usada para a produção de extremidades *blunt ends* em moléculas de DNA, retirando nucleotídeos das extremidades 3' das moléculas e recolocando-os com sua atividade polimerásica. Além disso, também é aplicada na síntese da segunda fita quando se usa a técnica de mutagênese sítio-dirigida.

1.2.6 Transcriptase reversa

A transcriptase reversa é uma DNA polimerase RNA dependente, produzida naturalmente por retrovírus, capaz de realizar o processo de transcrição inversa, ou seja, catalisa a formação de uma fita de DNA a partir de um molde de RNA fita simples, de dNTPs e um iniciador (*primer*) com extremidade 3'OH livre. Esta enzima também catalisa a formação de uma fita de DNA a partir de um molde de DNA fita simples e um iniciador. É utilizada, principalmente, para sintetizar DNA complementar (cDNA) a partir de sequências de mRNA, que pode ser utilizado na preparação de bibliotecas de cDNA ou para realizar PCR a partir de um molde inicial de RNA (*RT-PCR*, do inglês *reverse transcription-polymerase chain reaction*).

1.2.7 Topoisomerases

As DNA topoisomerases são nucleases reversíveis, pois são capazes de clivar e refazer as ligações fosfodiéster nos eventos relacionados às mudanças na topologia do DNA durante os processos replicação, transcrição, recombinação e remodelagem da cromatina, pela introdução de uma quebra temporária em ambas as fitas (topoisomerase II) ou numa única fita (topoisomerase I) da hélice dupla do DNA, tanto em eucariotos quanto em procariotos. Durante o processo de duplicação do DNA, essas enzimas promovem o relaxamento das fitas de DNA, retirando a tensão provocada pelo enovelamento da molécula após catalisar sua quebra e posterior rearranjo⁷.

As topoisomerases I catalisam o rompimento da ligação fosfodiéster de uma das fitas da molécula que se encontra espiralada. Parte da energia liberada com o rompimento é, então, conservada no ponto de quebra na forma de ligação covalente entre os grupamentos fosfatos. A fita clivada depois gira sobre a estrutura da fita intacta da molécula e ocasiona o relaxamento da cadeia de DNA, seguido pela restauração da ligação da fita quebrada

utilizando a energia resultante da clivagem da ligação fosfodiéster⁸. Diferentemente, as topoisomerases II catalisam a clivagem das duas cadeias de DNA simultaneamente, podendo tanto introduzir ou retirar a tensão espiral da fita, possibilitando o giro das fitas através do corte e sua posterior religação, similarmente ao que ocorre com a topoisomerase I. Em engenharia genética, as propriedades das topoisomerases são aplicadas em clonagem de fragmentos em vetores sem a utilização da enzima DNA ligase, o que será mais bem explanado quando tratarmos de vetores de clonagem, mais adiante.

1.2.8 Terminal transferase

Terminal desoxinucleotidiltransferase (TdT) é uma enzima polimerase que apresenta a capacidade de adicionar nucleotídeos na extremidade 3'OH de DNA fita simples ou dupla, a partir de desoxirribonucleosídeos trifosfatos, sem necessidade de molde. Esta proteína foi uma das primeiras enzimas com atividade de DNA polimerásica identificada em mamíferos⁹. Durante muito tempo não se sabia exatamente qual era o seu papel biológico, mas hoje já sabe-se que a TdT é responsável pela adição randômica de nucleotídeos durante os processos de recombinação V(D)J de genes de imunoglobulinas, que levam à diversidade imunológica¹⁰. Além da atividade de transferase, a TdT possui atividade 3'-oligodeoxinucleotidilquinase, catalisando a reação de fosforilação de um oligodeoxinucleosídeo no grupo 3' hidroxil¹¹.

A polimerização com TdT não depende de molde de DNA e depende de íon bivalente como cofator, principalmente Mg^{++} , Zn^{++} , Co^{++} e Mn^{++} . Em engenharia genética, as reações de adição de nucleotídeos a partir de extremidades 3'OH pela TdT ocorrem utilizando Co^{++} ou Mn^{++} , pois na presença desses íons a enzima é bem ativa para a extensão de terminais dupla fita, e, além disso, eventuais nucleases contaminantes não são ativadas por esses íons bivalentes.

Aproveitando tais propriedades, foram desenvolvidas metodologias para a síntese de extremidades homopoliméricas em fragmentos de DNA, fornecendo à reação apenas um tipo de desoxirribonucleosídeo trifosfato como substrato da enzima terminal transferase, como ilustrado na Figura 1.7.

Desse modo, a enzima terminal transferase é utilizada em engenharia genética, principalmente, para adicionar extremidades homopoliméricas complementares em dois fragmentos de DNA, favorecendo a formação de uma molécula quimérica pela hibridação dessas extremidades. Essa ferramenta é usada também para a marcação de extremidades 3' de DNA com

ribo ou desorribonucleotídeos radioativos ou modificados, como biotina-II-dUTP, fluorosceína-dUTP, entre outros.

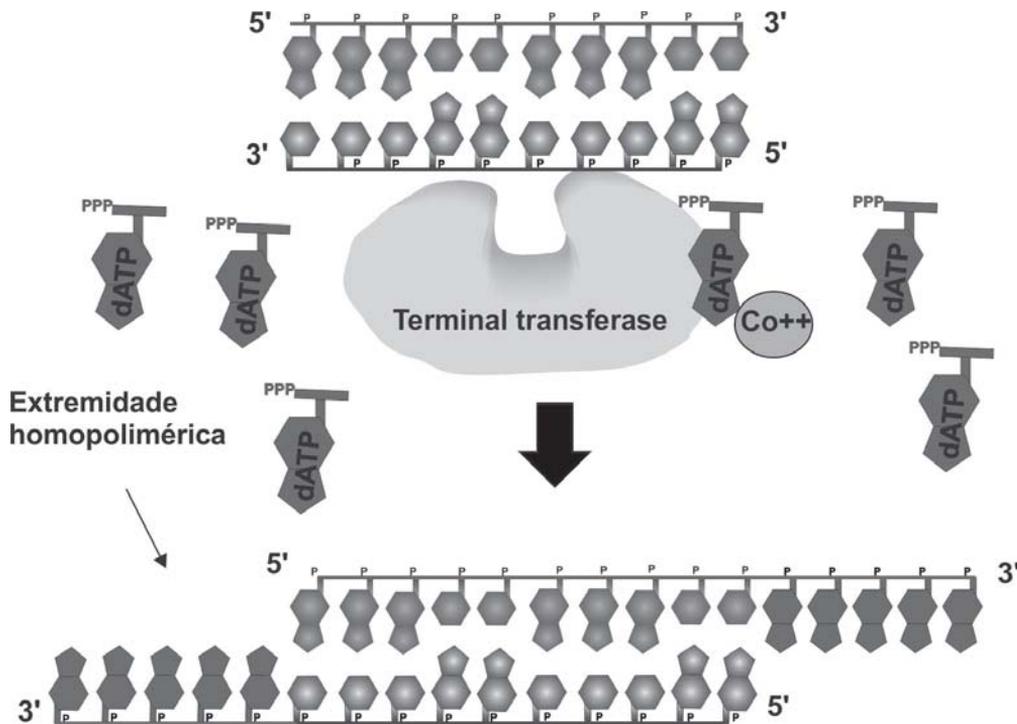


Figura 1.7 Atividade da enzima terminal transferase. Esquema ilustrativo da adição homopolimérica de nucleotídeos nas extremidades 3'-OH de uma cadeia de DNA utilizada em clonagem molecular.

1.2.9 Cre-recombinases

A Cre-recombinase do bacteriófago P1 é uma proteína que catalisa a recombinação entre dois sítios de reconhecimento, chamados de loxP^{12} . O loxP é uma sequência de 34 pb formado por uma região central de 8 bp (*core*) flanqueada por duas regiões palindrômicas de 13 pb (repetidas e invertidas). A Cre-recombinase é um tipo de topoisomerase I sítio-específica e por sua habilidade é capaz de fazer integração ou excisão de sequência de DNA no genoma hospedeiro. A partir dessa possibilidade, em TDR essa enzima é empregada em técnicas de interrupção gênica, deleção via integração, translocação ou excisão de fragmentos-alvo por recombinação da sequência loxP , que pode ser inserida em um vetor ou em um cassete

de expressão linear direcionando a recombinação homóloga, como se observa na Figura 1.8.



Esquema 1.1 Sequência loxP de 34 nucleotídeos. O *core* da sequência (vermelho) é flanqueada por duas sequências invertidas repetidas.

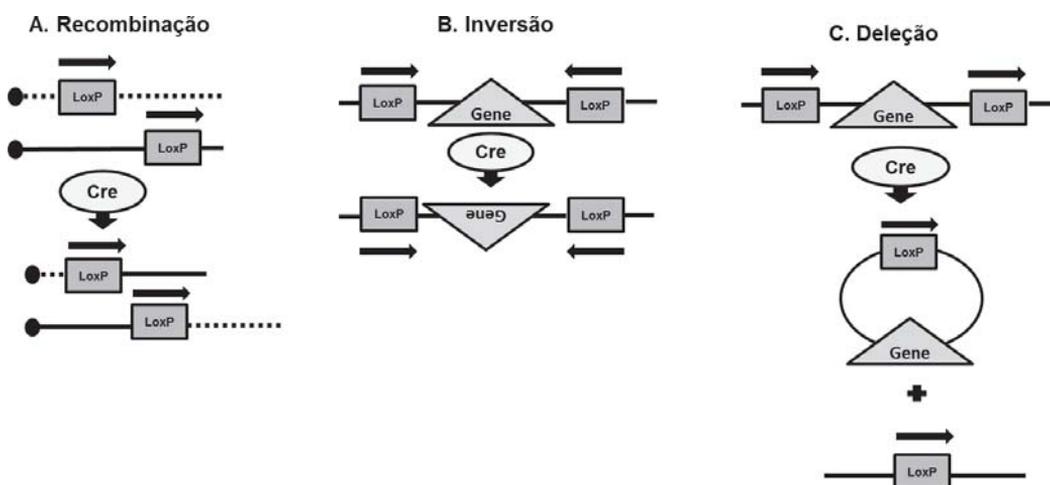


Figura 1.8 Tipos de recombinação promovidos pelo sistema Cre-lox. O resultado de uma recombinação Cre-lox é determinado pela orientação e localização dos sítios loxP. (a) Se os sítios loxP são orientados em direções opostas, a enzima Cre recombinase catalisa a inversão do segmento que estão flanqueados por loxP; (b) se os sítios loxP estão localizados em diferentes cromossomos (arranjo trans), a Cre recombinase catalisa a translocação cromossômica; e (c) se os sítios loxP são orientados na mesma direção, em um segmento de cromossomo (cis arranjo) a Cre recombinase promove deleção do segmento flanqueado por loxP.

1.2.10 Fosfatase alcalina

A fosfatase alcalina catalisa em pHs altos a hidrólise de monoésteres de fosfato de vários tipos de álcoois. Nos ácidos nucleicos, esta enzima remove grupos fosfatos das extremidades 5'-P de ácidos nucleicos, gerando grupos hidroxila 5'. É frequentemente utilizada em processos de clonagem para evitar as ligações fosfodiésteres entre os fragmentos de uma mesma molécula, a qual se deseja unir a outra molécula para gerar a quimera. É também

utilizada em conjunto com a polinucleotídeo quinase para marcação radioativa das extremidades da molécula de DNA.

As fosfatases alcalinas comercializadas são as de *E. coli* (*bacterial alkaline phosphatase* – BAP), de intestino de boi (*calf intestinal phosphatase* – CIP) e de camarão do Ártico (*shrimp alkaline phosphatase* – SAP). Esta última é a mais termossensível, o que é uma propriedade interessante, pois, após seu uso, pode ser inativada a 65 °C por quinze minutos, possibilitando economia de tempo.

1.2.11 Polinucleotídeo quinase

Este tipo de enzima foi descoberta em células de *E. coli* infectadas com fagos T4. Na presença de ATP, a polinucleotídeo quinase obtida de *E. coli* fosforila os grupos hidroxila 5' (5'-OH) de ácidos nucleicos. A reação ocorre com a transferência de grupo terminal fosfato γ do ATP para a extremidade 5' do fragmento. Caso se utilize ATP marcado na posição γ com fosfato ^{32}P , o resultado é a marcação radioativa da extremidade 5'. Além dessa utilização, a enzima também pode ser empregada para a fosforilação das extremidades de oligonucleotídeos de DNA, obtidos por síntese química.

1.3 VETORES

Um vetor, como descrito anteriormente, é uma molécula de DNA capaz de transportar a sequência gênica a ser clonada para o interior da célula.

De maneira geral, um vetor deve ter os seguintes elementos genéticos básicos:

- Uma origem de replicação, a fim de poder se replicar autonomamente dentro da célula hospedeira e, dessa forma, amplificar o DNA de interesse.
- Uma marca de seleção que permita selecionar as células recombinantes, ou seja, as células hospedeiras que incorporarem o vetor.
- Uma região de múltiplos sítios de reconhecimento de enzimas de restrição, onde será inserida a sequência a ser clonada, denominada sítio múltiplo de clonagem (*multiple cloning site*, MCS).

De acordo com a característica, os vetores são divididos segundo sua função em vetores de clonagem, vetores de expressão e vetores integrativos.

1.3.1 Vetores de clonagem

São vetores utilizados em clonagem molecular de genes, seja na construção de bibliotecas de genes ou de cDNAs seja na subclonagem de fragmentos específicos como produtos de amplificação por PCR, ou mesmo um gene, ou fragmento de DNA produzido por síntese química.

Os primeiros vetores empregados na clonagem molecular foram derivados dos plasmídeos naturais de bactérias, principalmente *E. coli*, e são moléculas de DNA circular de dupla fita, geralmente de baixa massa molecular. O vetor pBR322¹³ (Figura 1.9a) foi dos primeiros vetores construídos, o mais utilizado para clonagem molecular até a década de 1980 e deu origem a grande parte dos vetores plasmidiais utilizados atualmente. Esse plasmídeo possui a origem de replicação do plasmídeo ColE1 de *E. coli* e genes de resistência à ampicilina (*ampR*) e à tetraciclina (*tetR*) como marcas de seleção. Apresenta uma interessante forma de seleção por inativação de um dos genes de resistência a antibióticos, pois contém sítios únicos de restrição dentro de cada gene. Assim, de acordo com a enzima de restrição utilizada para a clonagem, o inserto irá inativar um dos dois genes e, ao transformar uma célula hospedeira sensível aos dois antibióticos, a que receber o plasmídeo recombinante passará a ser resistente apenas àquele cujo gene de resistência permaneceu intacto. Além de esse vetor oferecer uma forma fácil de seleção, apresenta o processo de replicação denominado “relaxado”, que permite que ele se duplique inúmeras vezes na célula hospedeira. O número de cópias do PBR322 pode ainda ser amplificado com a adição de cloranfenicol no meio de cultura.

A série de vetores desenvolvida posteriormente é a série de vetores pUC¹⁴ (Figura 1.9b), que são os chamados vetores de seleção direta, ou seja, além de permitirem a seleção de transformantes por resistência à ampicilina, o sítio múltiplo de clonagem está localizado dentro do gene da enzima β -galactosidase (*lacZ*). A seleção é feita com base na clivagem do substrato cromogênico X-gal por essa enzima, que produz colônias de coloração azul índigo.

Uma vez inserido o fragmento de DNA no vetor, o gene (*lacZ*) é inativado, tornando a célula incapaz de produzir a enzima ativa. Dessa forma, ao adicionar o indutor IPTG e o substrato cromogênico da enzima (X-gal) ao meio de cultura sólido, a célula que recebeu o plasmídeo recombinante, por não

produzir a enzima, não converterá o X-gal em azul índigo, ficando a colônia da cor bege-clara, típica da *E. coli*.

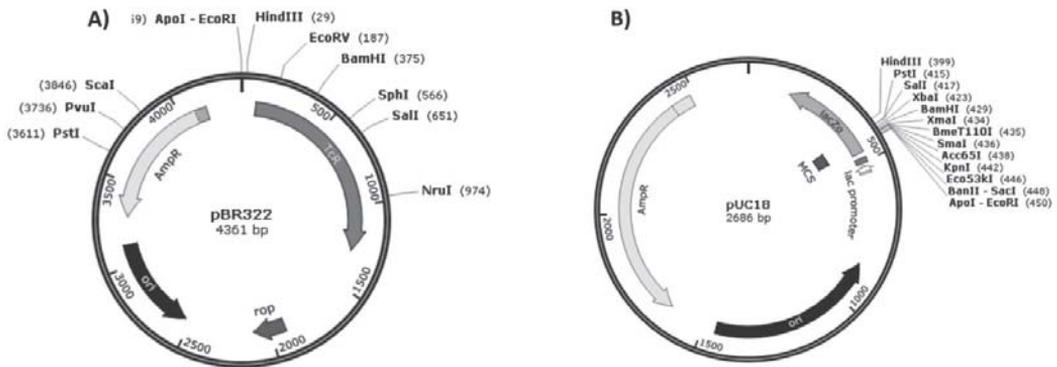


Figura 1.9 Vetores de clonagem. a) pBR322 – Ori: origem de replicação; *Amp^r*: gene de resistência à ampicilina; *Tc^r*: gene de resistência à tetraciclina; *Rop*: região importante para a replicação; b) pUC18 – Ori: origem de replicação; *Amp^r*: gene de resistência à ampicilina; *LacZ*: gene da β -galactosidase; *MCS*: sítio múltiplo de clonagem.

Os plasmídeos vêm sendo ao longo dos anos os vetores mais utilizados em engenharia genética, e a busca de inovações que tornassem o processo de clonagem molecular mais simplificado e eficiente levou ao desenvolvimento dos vetores para clonagem do tipo TA¹⁵, para clonagem direta de insertos de DNA obtidos a partir da amplificação por PCR. A partir do conhecimento de que a *Taq* DNA polimerase também possui a atividade de terminal transferase não dependente de molde e, dessa forma, adiciona uma desoxiadenosina monofosfato à extremidade 3' do produto de PCR, foi desenvolvido um tipo de vetor já na forma linearizada, contendo em sua extremidade 3' uma desoxitimidina monofosfato, os assim denominados *T-vectors*. Dessa forma, ao misturar o produto de PCR com o vetor na presença da DNA ligase, obtém-se o plasmídeo circular covalentemente ligado, pronto para ser utilizado na transformação genética.

Mais recentemente, foi desenvolvido um tipo de *T-vector* já acoplado com a topoisomerase I do vírus *Vaccinia*, e foi desenvolvida toda uma família de vetores denominados TOPO[®]. Esse vetor também é comercializado na forma linearizada e representou uma grande inovação, pois possibilitou a ligação direta de produtos de PCR dispensando o uso da DNA ligase, devido à adição em suas extremidades de topoisomerasas ativadas que catalisam a ligação do inserto ao vetor (Figura 1.10). Desde então, essa tecnologia vem sendo

amplamente utilizada, e já existem também vetores com topoisomerasas capazes de fazer a ligação de insertos com extremidades abruptas (*blunt end*).

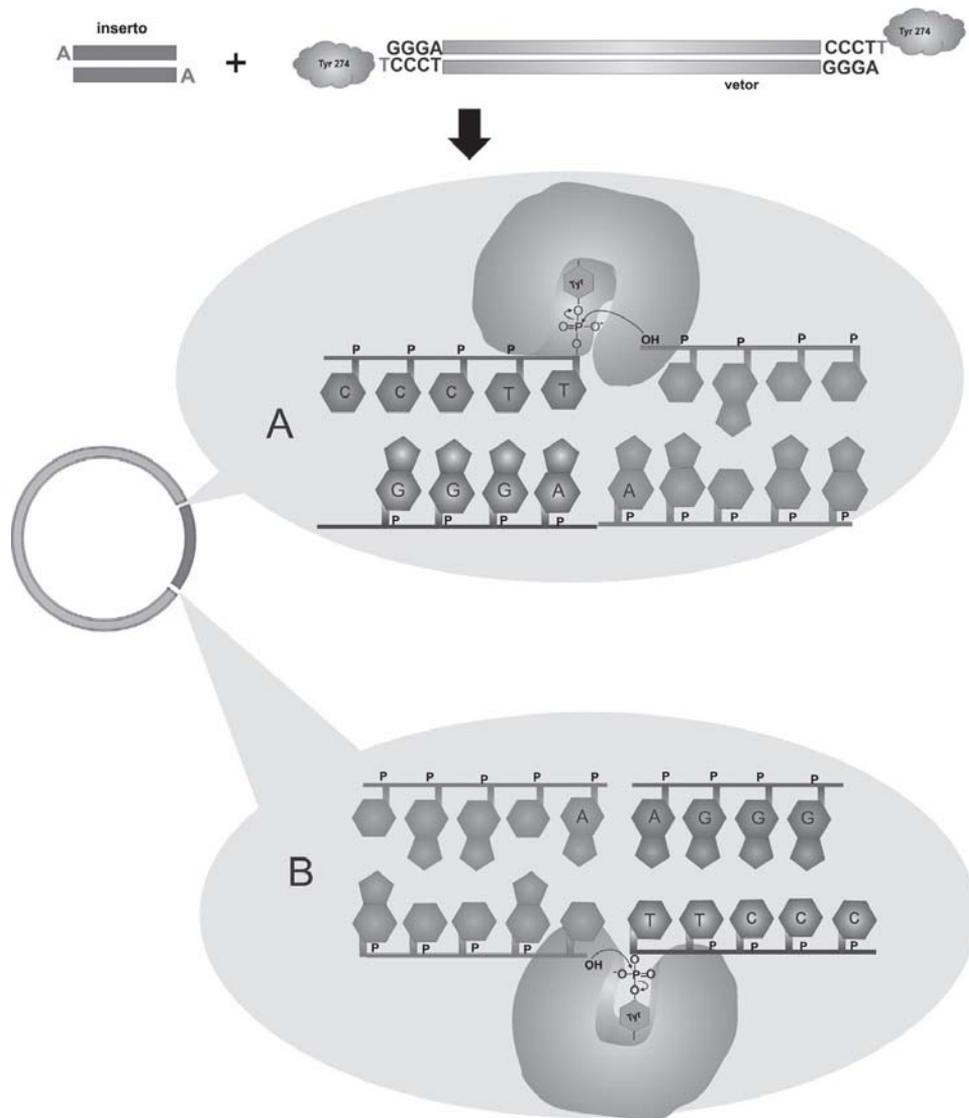


Figura 1.10 Vetor com topoisomerasas I utilizado em clonagem molecular. Esquema mostrando o vetor de clonagem com extremidades em T (*T-vector*) com topoisomerasas ativadas em suas extremidades e a ligação com o inserto. No detalhe ampliado, o mecanismo enzimático de ligação do 5'-OH via ataque nucleofílico do par de elétrons da hidroxila ao grupo fosfato da extremidade T do vetor que está ligado ao resíduo tirosina da topoisomerase. É importante lembrar que amplicons cujos *primers* utilizados na PCR não foram fosforilados na posição 5' permanecem mesmo após a PCR na forma 5'-OH.

As marcas de seleção mediadas pela resistência a antibióticos e/ou inativação de função do gene da β -galactosidase, embora eficientes no processo de clonagem molecular, mesmo após tratamento com fosfatase alcalina, certa porcentagem do vetor plasmidial recirculariza sem o inserto. Assim, visando a minimizar essas ocorrências, foi desenvolvido um novo tipo de seleção que denominamos positiva (ou Kamikaze). Citamos como exemplo a inserção do gene *ccdB*, que codifica uma proteína letal para *E. coli*, em forma de fusão com o gene *lacZ*¹⁶. A seleção é feita baseada na interrupção da função dos genes fusionados após a ligação do inserto. Caso haja a recircularização do plasmídeo sem o inserto, a proteína letal é expressa e causa a morte da *E. coli* hospedeira, permitindo apenas o crescimento de clones recombinantes.

Ao longo dos anos, as limitações, como número de cópias e tamanho da sequência de DNA a ser clonada, nortearam a evolução dos vetores de clonagem no sentido de superá-las. Para facilitar o processo de clonagem molecular e expressão de genes heterólogos em hospedeiras de difícil manipulação foram desenvolvidos os vetores bifuncionais (*shuttle vectors*), capazes de se duplicar e serem selecionados em dois hospedeiros diferentes, sendo um deles a *E. coli* hospedeira, de manipulação muito fácil. Além disso, vetores foram desenvolvidos com o objetivo de garantir a expressão ou expressão/secreção das proteínas recombinantes em altos níveis em diferentes tipos de células hospedeiras para aplicações biotecnológicas. A seguir, serão apresentados importantes tipos de vetores.

1.3.1.1 Vetores derivados de bacteriófagos λ

Devido à necessidade de clonar fragmentos de DNA grandes e, ainda assim, introduzi-los em *E. coli* eficientemente para a construção de bibliotecas de genes de seres eucariontes (com genomas grandes), desenvolveram-se derivados de bacteriófagos λ de cujo interior do DNA era possível retirar cerca de 20 Kb (não essencial para o ciclo lítico do fago) com digestão com enzimas de restrição e substituir essa sequência por DNA exógeno, reconstituindo-se um DNA recombinante de cerca de 50 Kb. Para a introdução desse DNA recombinante na *E. coli* hospedeira, o DNA era empacotado com um extrato de proteínas do capsídeo viral e, em seguida, colocado em contato com a hospedeira para que a infecção ocorresse. Para ocorrer o empacotamento adequadamente, além de ter tamanho da ordem de 50 Kb +/- 5 Kb, o DNA deve conter uma sequência de cerca de dezessete nucleotídeos que

fica na região das extremidades coesivas (*stick ends*), denominada de COS (*cohesive ends*)¹⁷.

Dessa forma, já no final da década de 1970 era possível introduzir DNA na célula hospedeira com eficiência de 10^8 clones/ μg de DNA, o que viabilizou a construção das primeiras bibliotecas genômicas humanas completas, das quais foram isolados os primeiros genes humanos completos.

1.3.1.2 Vetores derivados de bacteriófagos M13

Os bacteriófagos tipo M13 são vírus de DNA fita simples com cerca de 6 Kb. Quando empacotado, seu genoma se apresenta em forma de fita simples, porém, durante sua fase de replicação no interior da *E. coli* seu genoma se mantém em forma duplex (dupla fita). Essa característica despertou o interesse dos pesquisadores, pois facilmente seria possível isolar o DNA do fago em forma de fita simples a partir dos fagos obtidos no exterior da hospedeira (no meio de cultura) e em forma duplex no interior da *E. coli*. A forma duplex pode servir para a digestão com enzimas de restrição e clonagem de DNA heterólogo em seu interior como se fosse um plasmídeo, e a forma de fita simples foi usada por um bom tempo para o sequenciamento pelo método dideoxi de Sanger e, até hoje, para realizar mutagênese sítio-dirigida¹⁸. Fagos M13 têm sido usados também para a construção de bibliotecas de peptídeos randômicos ligados às proteínas do fago, em um processo denominado de *phage display*. Há limitação de tamanho do inserto a ser clonado no interior do genoma viral, pois são fagos tipo bastonetes e, caso se clone um inserto muito grande (maior que 5-6 Kb), a estrutura do fago pode não suportar e quebrar.

1.3.1.3 Cosmídeos

Antes do advento da eletroporação, para se realizar transformação genética de *E. coli* (introdução de DNA nas células) utilizava-se o método do CaCl_2 . Por esse procedimento, quanto maior o tamanho do fragmento a ser clonado, menos eficiente era o processo. Por isso, idealizou-se uma estratégia que permitia introduzir DNA plasmidial com de cerca de 50 Kb no interior da *E. coli* com alta eficiência. Para tanto, introduziu-se o sítio COS do fago λ no interior do vetor plasmidial de clonagem¹⁹. Caso o vetor tivesse cerca de 5 Kb, como era o caso do pBR322, era possível clonar no interior de um

de seus genes de resistência a antibióticos fragmentos de até 45 Kb, empacotar com proteínas do capsídeo do fago λ e, então, introduzi-los nas células da hospedeira bacteriana com altíssima eficiência. No interior da célula, a molécula se comportava como um plasmídeo. Esse tipo de vetor tem sido de grande relevância na produção de bibliotecas genômicas para fins de seleção de genes específicos e para a determinação da sequência genômica completa.

1.3.1.4 Cromossomos artificiais de bactéria (bacterial artificial chromosomes – BAC)

São vetores derivados do plasmídeo conjugativo F de bactérias, que são muito grandes, (de 100 a 200 Kb. Separa-se do plasmídeo F sua origem de replicação, que é de baixo número de cópias, pois atua de forma sincronizada com a hospedeira, e os genes necessários para sua estabilização e partição, juntando-se a genes marcadores de seleção, como de resistência a antibióticos e/ou da β -galactosidase. Dessa forma o vetor plasmidial fica pequeno, da ordem de 10 Kb, e dessa forma pode receber insertos enormes (de 100 a 200 Kb e permanecer estável no interior da *E. coli* após sua introdução, que ocorre geralmente por eletroporação. Dessa forma, pode-se fazer com relativa facilidade bibliotecas genômicas de grandes fragmentos de DNA.

1.3.1.5 Cromossomos artificiais de leveduras (yeast artificial chromosomes – YACs)

A levedura de cerveja, a *Saccharomyces cerevisiae*, é o microrganismo eucarionte mais estudado até o momento, e já nos anos 1980 suas estruturas cromossômicas importantes já haviam sido clonadas e caracterizadas como origens de replicação cromossômicas (*autonomously replicating sequences* – ARS), telômeros (TEL) e centrômeros, além de diversos genes de vias metabólicas importantes que foram utilizados como marcadores auxotróficos (URA3, TRP1, LEU2). Nesse contexto foi possível construir vetores equivalentes a minicromossomos da levedura, contendo esses importantes elementos ligados à parte de um vetor plasmidial bacteriano para ser manipulado facilmente na *E. coli* durante sua construção. Nesse tipo de vetor era possível clonar fragmentos grandes de DNA, de até 1 milhão de pares de bases.

1.3.2 Vetores de expressão

Os vetores utilizados para expressar o peptídeo ou proteína codificados pelo segmento de DNA inserido, denominados vetores de expressão, são veículos que, como podemos ver na Figura 1.11, a seguir, além da origem de replicação e marca genética de seleção, contêm uma região promotora, uma região codificadora do sítio de ligação ao ribossomo (*ribosomal binding site* – RBS), códon de início da tradução seguido de pelo menos um sítio de restrição para inserção em fase da sequência de inserto, a fim de que a proteína possa ser expressa, além de uma sequência terminadora de transcrição.

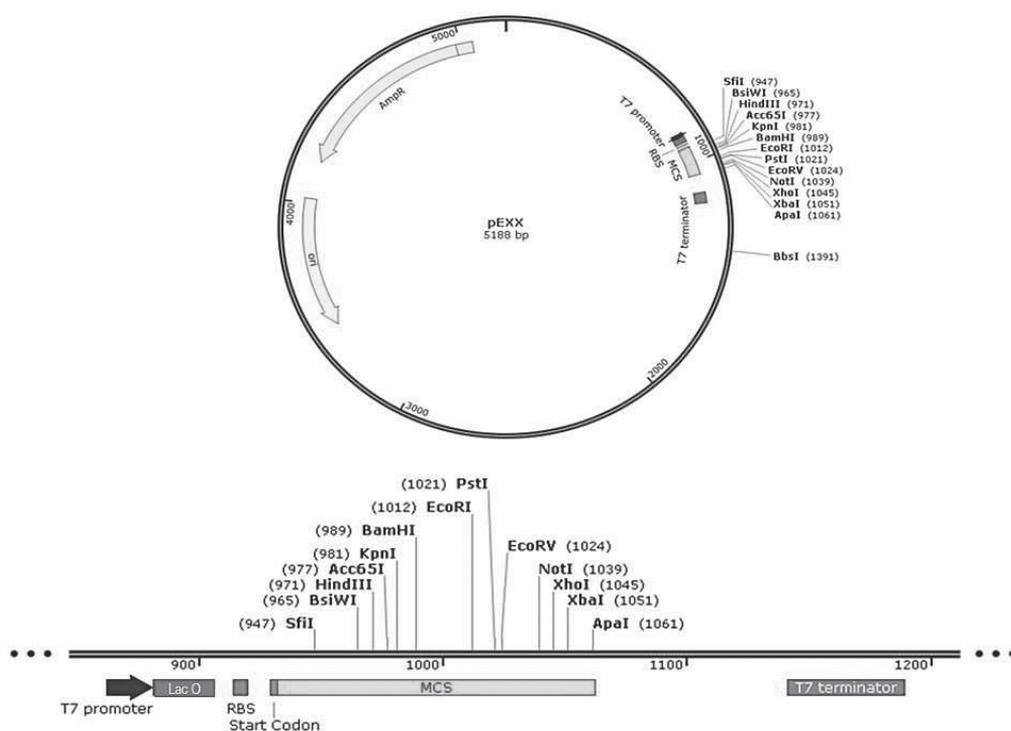


Figura 1.11 Vetor de expressão. Mapa físico de um vetor de expressão genérico (pEXX) contendo origem de replicação (ori), gene de resistência à ampicilina (AmpR) promotor do fago T7, operador Lac (LacO), sítio de ligação ao ribossomo (RBS), *start* códon, sítio múltiplo de clonagem (MCS) e terminador de transcrição. No detalhe abaixo, ampliação do cassete de expressão do vetor.

De maneira geral, o principal objetivo do uso de vetores de expressão é obter altos níveis de RNAs mensageiros estáveis, que serão traduzidos resultando em grandes quantidades da proteína ou peptídeo de interesse. Para

atingir esse objetivo, um elemento importante é que o vetor deve possuir um promotor forte. Os promotores mais utilizados para a expressão em *E. coli* são, geralmente, baseados no promotor do operon *lac*, no promotor PL do fago λ , ou no promotor do fago T7²⁰, que está presente na família de vetores pET, um dos mais utilizados atualmente. Estes promotores podem também ser híbridos de diferentes promotores, por exemplo, o promotor *tac*, que é um híbrido dos promotores *trp* e *lac*²¹. Os promotores utilizados nos vetores de expressão são normalmente indutíveis, ou seja, a síntese de proteínas é induzida apenas quando necessária. O sistema de expressão contém elementos de controle de um operon bacteriano (promotor/operador e gene regulador que codifica a proteína repressora). A expressão do gene heterólogo ocorre então ao adicionar ao meio de cultivo um indutor, tais como a arabinose e o isopropil- β -D-tiogalactopiranosídeo (IPTG). A expressão da proteína, contudo, também pode ser constitutiva em alguns vetores de expressão. Um promotor fraco em geral só é utilizado em casos nos quais a proteína recombinante necessita ficar em menor concentração na hospedeira ou quando é tóxica para a hospedeira, embora um baixo nível de síntese de proteína constitutiva possa ocorrer mesmo em vetores de expressão com os promotores controlados. Há ainda os vetores termoinduzíveis, que são baseados no promotor P_L do fago λ , para os quais foi desenvolvida uma molécula repressora termossensível. No caso de vetores de expressão para leveduras, tem sido utilizado muito para *Pichia pastoris* o promotor da enzima álcool oxidase (AOX1), que é induzível por metanol.

No que se refere ao sítio de ligação do ribossomo (região Shine Dalgarno), o que tem sido mais usado, por ser muito eficiente, é do gene da proteína 10 do fago T7. Em relação a terminadores de transcrição, um bom vetor de expressão combina um promotor forte com um terminador de transcrição também bastante eficiente. Para *E. Coli*, terminadores Rho independentes, como o do operon TRP, têm sido utilizados. Um tipo especial de vetor de expressão é o que, além de programar a expressão da proteína heteróloga, promove também sua secreção. Para tanto, o vetor ou sequência codificadora da proteína deve conter a região que codifica o peptídeo-sinal, que é uma região situada logo na região aminoterminal de proteína, que a direciona em células eucarióticas para o retículo endoplasmático, de onde as proteínas a serem secretadas seguem então o caminho do complexo de Golgi e, por meio de vesículas de secreção, são lançadas ao meio extracelular. Em células procarióticas, o peptídeo-sinal dirige as proteínas para o espaço periplasmático. Os peptídeos-sinais têm tamanhos, com algumas exceções, de 18 a 30 aminoácidos, o que equivale no DNA a 54 a 90 pares de nucleotídeos.

A experiência tem demonstrado que os peptídeos-sinais de um determinado organismo são capazes de funcionar em outros, mesmo que sejam taxonomicamente muito distantes, como eucariontes e procariontes. Isso se deve ao fato de a estrutura do peptídeo-sinal ser mais importante que sua sequência de aminoácidos para o processo de secreção. Os peptídeos-sinais têm normalmente na região inicial (aminoterminal) pelo menos um aminoácido carregado positivamente, um núcleo central hidrofóbico e aminoácido final neutro com uma cadeia lateral pequena. O peptídeo-sinal é reconhecido pela maquinaria de translocação da célula e, logo após ser translocado para o retículo endoplasmático, é clivado do restante da proteína pela peptidase-sinal. Diferentes peptídeos-sinais têm sido utilizados para dirigir para a secreção peptídeos heterólogos. Os mais comuns são: o da proteína A de *Staphylococcus aureus* e de α -amilase de *Bacillus subtilis*, o do fator α de *S. cerevisiae* para leveduras e o de imunoglobulinas G (IgGs) para secreção em células de mamíferos em cultura.

1.3.2.1 Vetores de expressão de proteínas com cauda de fusão

Após a expressão do produto do gene, geralmente é necessário purificar a proteína expressa. O processo de separação da proteína de interesse, a partir da grande maioria das proteínas da célula hospedeira, pode ser trabalhoso e demorado. A fim de tornar o processo de purificação mais fácil, uma sequência peptídica (TAG) eleita pode ser adicionada à região aminoterminal ou carboxiterminal da proteína recombinante, alterando-se a sequência de seu gene. Essa sequência ligada à proteína de interesse será utilizada para a purificação da proteína recombinante por cromatografia de afinidade e depois será removida pela ação de uma protease específica. Por exemplo, pode-se adicionar à proteína recombinante seis histidinas (HIS6) e, com isso, purificá-la por cromatografia de afinidade com resinas que tenham átomos de níquel imobilizados. A Tabela 1.2 a seguir mostra diferentes *tags* utilizadas para facilitar a purificação e as resinas utilizadas nos processos cromatográficos de afinidade.

1.3.3 Vetores integrativos

São vetores bifuncionais que possuem marcadores genéticos para dois tipos de hospedeiras diferentes, porém com uma única origem de replicação

Tabela 1.2 Diferentes tipos de processos cromatográficos utilizados na purificação de proteínas de fusão

| TAG DE AFINIDADE | MATRIX CROMATOGRÁFICA | CONDIÇÕES DE ELUIÇÃO |
|--------------------------------------|-----------------------------|--|
| Poli-His (HIS)6 | Ni ⁺⁺ - NTA | 100 – 500 mM imidazol |
| Poli-Arg (ARG)5 | Resina trocadora de cátions | Gradiente de NaCl de 0 a 0,5 M em pH>8,0 |
| Domínio de ligação à celulose | Celulose | Cloridrato de guanidina ou ureia (>4 M) |
| c-myc | Anticorpo monoclonal (Mab) | pH baixo |
| FLAG (DYKDDDDK) | Mab anti FLAG | pH baixo ou EDTA 2-5 mM |
| SBP | Estreptoavidina | Biotina |
| Glutathiona-S-transferase | Glutathiona | Glutathiona reduzida 10 mM |
| Domínio de ligação à maltose | Amilose <i>cross-linked</i> | Maltose 10 mM |

funcional em apenas uma das hospedeiras. Isso faz com que transforme e replique em uma das hospedeiras, mas na outra para que ocorra transformação genética há necessidade do vetor se integrar ao seu genoma. Caso contrário, não há replicação, o marcador de seleção não se expressa adequadamente e nem outros genes que se deseje expressar. Os vetores integrativos são muito comumente utilizados para expressão de proteínas heterólogas em leveduras e contam com a grande capacidade das leveduras de fazer recombinação homóloga. Por essa razão é que nas leveduras pode-se fazer *gene disruption* com facilidade.

1.4 HOSPEDEIROS

Em engenharia genética, hospedeiro é uma célula ou um ser receptor de DNA exógeno. As características e peculiaridades discutidas para os vetores de clonagem e, sobretudo, para os vetores de expressão, podem estar intimamente ligadas ao tipo de hospedeiro do DNA exógeno, podendo este ser hoje em dia praticamente qualquer tipo de célula ou organismo, como bactérias, leveduras, fungos, vírus, células vegetais ou animais.

Um bom hospedeiro deve preencher uma série de requisitos básicos, entre os quais:

- Aceitar o DNA exógeno sem modificá-lo ou modificá-lo ao modo que se deseja.
- Permitir a fácil seleção dos hospedeiros que contêm o DNA exógeno.
- Se organismo for hospedeiro para expressão em altos níveis de proteínas heterólogas, ser deficiente em proteases.
- Apresentar baixo potencial de proliferação no ambiente caso a mensagem exógena a ser clonada apresente “risco” ao meio ambiente.

Para se garantir esses requisitos e assegurar outros incrementos, muitos tipos de organismos foram manipulados geneticamente para se garantir eficiente transformação, manutenção e expressão das moléculas recombinantes veiculadas nesses hospedeiros, desenvolvendo o que chamamos de sistemas de expressão, divididos basicamente em sistemas de expressão procariótico e eucariótico.

1.4.1 Sistema de expressão procariótico

Sistemas bacterianos de expressão de proteínas heterólogas são atrativos devido às suas várias características, como habilidade de rápido crescimento, fisiologia e genética bacteriana bem caracterizadas, baixo custo e disponibilidade de grande número de vetores de expressão e linhagens mutantes de hospedeiros²². Várias proteínas são produzidas com sucesso em procariotos e com consideráveis rendimentos. Geralmente, os produtos expressos não são secretados para o meio extracelular e são acumulados no citoplasma ou no espaço periplasmático.

Embora sejam bons hospedeiros, os procariotos empregam a desvantagem de não possuir uma maquinaria para o processamento transcricional de genes originados de eucariotos, não realizando processamento de mRNA (*splicing*), o que leva muitas vezes à necessidade de clonagem de cDNA para evitar a presença de introns, e também naturalmente não realizam modificações pós-traducionais nas proteínas expressas, como por exemplo fosforilação e glicosilação²³. Essas modificações na molécula proteica são muitas vezes decisivas para a garantia da função e atividade biológica e tornam-se bastante críticas no caso de proteínas para uso terapêutico. Como a capacidade e potencial dos procariotos para modificação pós-traducional é limitada, o uso de hospedeiras procarióticas é, geralmente, restrito à produção de proteínas que são naturalmente não glicosiladas, como insulina, IGF, somatotrofina ou proteínas que são naturalmente glicosiladas, porém que

são farmacologicamente também ativas sem a glicosilação, como várias citocinas (fator de necrose tumoral, interleucinas e interferons).

Outra limitação dos sistemas procarióticos para expressão de proteínas eucarióticas é com relação a códons raros. Muitas bactérias possuem *codon usage* diferente do de células eucarióticas. Em cada hospedeira, a evolução ajustou o conjunto de tRNAs (RNAs transportadores) de acordo com a necessidade de tradução de suas proteínas. Principalmente quando se necessita de altos níveis de expressão, a existência no mRNA da proteína heteróloga de códons pouco utilizados na hospedeira (códons raros) pode acarretar baixos níveis de expressão, exatamente pela insuficiência dos tRNAs correspondentes a esses códons raros. Para potencializar a expressão de proteínas eucarióticas ou proteínas que contêm códons raramente usados em hospedeiras bacterianas, muitas linhagens de *E. coli* foram engenheiradas para mitigar esse problema. Essas linhagens foram suplementadas com genes adicionais de tRNA sob controle de promotores de *E. coli*. Hoje, com a facilidade com que se pode fazer a síntese química de genes, pode-se, ao desenhá-los, evitar colocar códons raros, respeitando-se o *codon usage* da hospedeira²⁴.

Dentre os hospedeiros procarióticos, a bactéria *Escherichia coli* é o organismo de escolha por ser, sem dúvida, o mais bem estudado em termos de genética e bioquímica. Além disso, é de fácil manutenção e manipulação genética em laboratório e possui variados sistemas de promotores altamente regulados, o que a levou a ser um sistema de expressão bacteriano bem estabelecido até os dias atuais.

Além dos sistemas baseados em *E. coli*, houve o desenvolvimento e uso de outros hospedeiros procariotos que seriam promissores para sanar os problemas de expressão. Espécies de bactéria como *Bacillus subtilis*, *Streptomyces lividans*, *Anaena sp*, *Staphylococcus carnosus* são descritas como hospedeiras para expressão de proteínas heterólogas e, entre elas, espécies representantes do gênero *Bacillus* são os hospedeiros mais estudados e bem estabelecidos em TDR depois da *E. coli*. As linhagens de *Bacillus* são hospedeiros atraivos, pois, diferentemente da *E. coli*, não apresentam lipopolissacarídeos na membrana externa (endotoxinas) e possuem a capacidade natural de secreção e endereçamento da proteína para o meio extracelular. Os *Bacillus* têm sido usados para a produção de enzimas industriais como, por exemplo, as alfa-amilases e proteases, ao passo que a *E. coli* tem sido utilizada para a produção de proteínas terapêuticas²⁵.

1.4.2 Sistema de expressão eucariótico

Embora os procariotos tenham se estabelecido como poderosos sistemas de expressão de mensagens genéticas heterólogas, os hospedeiros eucariotos apresentam consideráveis vantagens em relação aos procariotos, especialmente para a expressão de proteínas eucarióticas, contornando alguns problemas na expressão em bactérias que poderiam levar a produtos com erros na conformação, função e atividade sem os processamentos finais necessários em proteínas eucarióticas.

Os primeiros grupos de eucariotos empregados como hospedeiros foram as leveduras e os fungos filamentosos, mas com os avanços biotecnológicos outros sistemas alternativos de expressão utilizando células de insetos, de mamíferos e vegetais foram desenvolvidos e bem estabelecidos. No entanto, os hospedeiros alternativos, como células de mamíferos, por exemplo, necessitam de técnicas e manipulações sofisticadas que encarecem os custos de produção. Diferentemente, fungos filamentosos e leveduras que crescem em meios de cultivos simples e baratos levam a altos níveis de expressão, crescem em altas densidades celulares e possuem rígida parede celular, em comparação com células de insetos ou mamíferos, que os tornam mais resistentes a estresse mecânico durante o processo fermentativo²⁶.

Dentre as leveduras, as mais empregadas na produção de proteínas recombinantes são *Saccharomyces cerevisiae*, *Hansenula polymorpha*, *Pichia pastoris* e *Yarrowia lipolytica*.

O eucarioto mais amplamente utilizado para a expressão heteróloga por muito tempo *Saccharomyces cerevisiae*. Isto se deve ao grande conhecimento acerca deste microrganismo, que tem sido utilizado pelo homem há milhares de anos para a obtenção de produtos alimentares e bebidas alcoólicas sem restrição para consumo humano, o que conferiu a essa levedura o *status* GRAS (*generally recognized as safe*, geralmente reconhecida como segura). Sua importância levou a uma gama de estudos, incluindo o campo da genética, desvendando atributos que tornariam essa hospedeira umas das mais importantes em biologia molecular. Assim, foi demonstrado que essa levedura possui um plasmídeo próprio, circular, denominado de “2 μ m” (dois micron), presente em múltiplas cópias no seu citoplasma. A partir do plasmídeo natural 2 μ m desenvolveram-se vetores de clonagem e de expressão para *S. cerevisiae*, como o vetor pMA91, ilustrado na Figura 1.12. Métodos de transformação genética foram desenvolvidos para a inserção de plasmídeos de replicação autônoma derivados do 2 μ m, até o estabelecimento de técnicas de transformação para a integração do material genético por

recombinação homóloga. Os transformantes podem ser selecionados por marcas auxotróficas como *LEU2*, *TRIP1*, *HIS3* e *URA3* ou por marcas de seleção dominante, como as que conferem resistência a G418 ou higromicina B. O genoma dessa levedura foi totalmente sequenciado e está disponível para acesso²⁷.

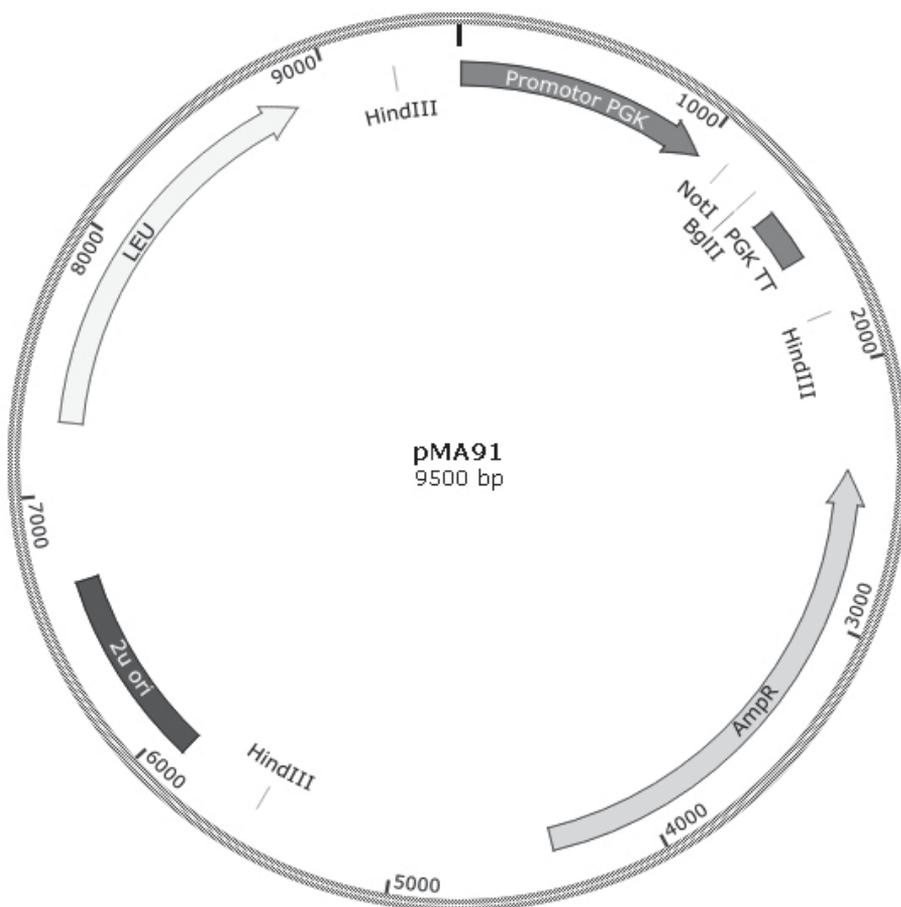


Figura 1.12 Plasmídeo pMA91. Representação esquemática do vetor para expressão em *S. cerevisiae* que contém o promotor do gene da enzima PGK da via glicolítica de *S. cerevisiae*.

O sucesso dessa levedura como hospedeira dá-se pela sua capacidade de realizar muitas das modificações pós-traducionais e dobramentos em proteínas de estruturas complexas, tais como acetilação do terminal amino, acilação, fosforilação e glicosilação⁴⁴. Vários genes de interesse em saúde humana foram clonados e expressos em *Saccharomyces cerevisiae*, como antígenos

virais, insulina humana e anticorpos. No entanto, os níveis de expressão atingidos por essa levedura são muitas vezes inferiores aos de outras leveduras, e por vezes mais baixos que os níveis de expressão observados em *E. coli*, por exemplo. Além disso, o perfil de glicosilação realizado por essa levedura não é idêntico ao perfil de glicosilação realizada em humanos, o que inviabiliza sua utilização na produção de algumas glicoproteínas para aplicação terapêutica devido à possibilidade de induzir imunogenicidade.

Os avanços nos conhecimentos biológicos e a busca incessante por produtos biotecnológicos mais rentáveis e seguros para usos humanos levaram ao desenvolvimento de novos sistemas de expressão heteróloga de genes baseados em outras leveduras. Emergiu o uso de leveduras metilotróficas (que metabolizam metanol), que se tornaram atraentes sistemas para a produção industrial de proteínas recombinantes, pois permitem rígido controle da expressão pela simples manipulação do meio de cultura. Dentre essas leveduras, *Pichia pastoris* mostrou-se uma hospedeira de vantagens consideráveis, tornando-se, nos dias atuais, a levedura de escolha para a expressão de genes heterólogos.

Uma das características marcantes é o forte promotor usado no vetor pPIC9²⁹ (Figura 1.13) para transcrever genes heterólogos, que é o do gene codificador da enzima álcool oxidase (AOX1) de *P. pastoris*. Esse promotor é regulado transcricionalmente por metanol, um indutor relativamente barato. Em células expostas a metanol como única fonte de carbono, o início da transcrição no promotor AOX1 é altamente eficiente e comparável aos promotores derivados dos genes altamente expressos da via glicolítica. Para ser ativado, o promotor AOX1 requer a presença de metanol. Na ausência desse indutor, ele se torna inativo. Além de metanol, o sistema AOX1 necessita da ausência de glicose para ser plenamente ativado. Uma vez que o promotor AOX1 é controlado pela manipulação da fonte de carbono adicionado ao meio de cultura, o crescimento e a indução de cepas de *P. pastoris*, que expressam proteínas heterólogas, são facilmente obtidos em todas as escalas, desde frascos até grandes fermentadores³⁰⁻³¹.

Outra notável característica é o fato de que culturas de *Pichia pastoris* atingem níveis de alta densidade celular, podendo ser facilmente cultivadas a densidades celulares de aproximadamente 100 g/L de peso seco, ou até maiores, as quais não são facilmente obtidas com *S. cerevisiae*³¹.

A maioria das proteínas secretadas por *P. pastoris* são glicosiladas, o que pode não afetar a atividade biológica da proteína recombinante. O tamanho da cadeia de carboidratos adicionados por *P. pastoris* é bem menor que aquele adicionado por *S. cerevisiae*. A estrutura destes oligossacarídeos é

muito similar à adicionada em mamíferos e, por não ser capaz de adicionar manoses terminais com ligações α -1,3, como *S. cerevisiae*, as proteínas produzidas em *P. pastoris* são menos imunogênicas.

Ao longo das últimas duas décadas, a levedura *Pichia pastoris* tornou-se o organismo de escolha para a produção de proteínas heterólogas, justamente pelas suas características citadas acima. Diversos produtos biotecnológicos foram produzidos nessa levedura, incluindo enzimas, proteínas diversas, biofármacos etc. Atualmente, vários produtos terapêuticos produzidos em *P. pastoris* encontram-se em fases avançadas de testes clínicos, como um inibidor de elastase para tratamento de fibrose cística, uma endostatina e uma angiostatina para aplicação como antiangiogênico, um fator de crescimento epidermal para tratamento de feridas causadas pela diabetes, entre outros biofármacos.

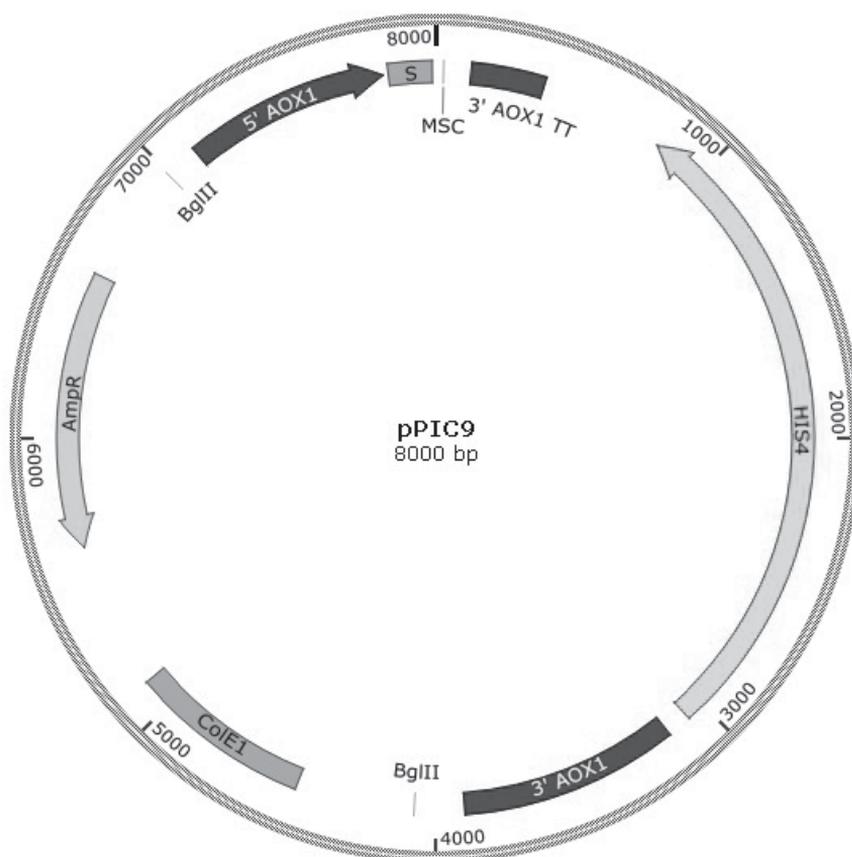


Figura 1.13 Plasmídeo pPIC9. Representação esquemática do vetor para expressão e secreção em *P. pastoris*.

O primeiro produto terapêutico para uso em humanos expresso em *Pichia pastoris* foi liberado pela FDA em 2009. Trata-se do ecallantide, um peptídeo de sessenta aminoácidos com atividade inibidora de caliceína plasmática humana. Foi descoberto a partir de uma biblioteca de *phage display* de fatores de inibição da coagulação humana, recebeu o nome comercial de Kalbitor® e é utilizado no tratamento de angiodema hereditário.

Alternativamente às leveduras, os fungos filamentosos são bons hospedeiros para expressão de genes heterólogos. O desenvolvimento de estratégias de biologia molecular rapidamente favoreceu o aprimoramento genético de fungos tanto para melhorar a capacidade de expressão de genes homólogos dos organismos selvagens quanto para o uso como hospedeiros de expressão heteróloga. Esses organismos vêm demonstrando vantagens sobre as bactérias e leveduras destacando-se pela elevada capacidade de produção e secreção de proteínas heterólogas, além da facilidade de se obter recombinantes mitoticamente estáveis e eficientes modificações pós-traducionais nos produtos expressos³². No entanto, vale ressaltar que há ainda poucos promotores adequados para a expressão de proteínas homólogas e heterólogas, havendo um número limitado de fungos filamentosos explorados como células hospedeiras. A pouca utilização desses hospedeiros deveu-se ao fato de secretarem proteases que hidrolisam as proteínas heterólogas, assim como pela dificuldade em se obter clones recombinantes estáveis com bom nível de expressão de proteínas heterólogas. Os fungos mais empregados como hospedeiros são as espécies *Aspergillus awamori*, *Aspergillus niger*, *Aspergillus oryzae*, *Humicola insolens*, *Trichoderma reesei*, *Penicillium chrysogenum*, *Cephalosporium acremonium* e *Mucor miehei*.

1.5 OUTRAS METODOLOGIAS IMPORTANTES DE MANIPULAÇÃO GENÉTICA

1.5.1 Reação em cadeia da polimerase (Polymerase Chain Reaction – PCR)

A invenção da técnica de amplificação de segmentos DNA *in vitro*, chamada de reação em cadeia da polimerase (PCR), possibilitou a amplificação de segmentos de DNA a partir de uma quantidade minúscula de DNA, o que ocasionou avanços enormes na área de genética forense e revolucionou os procedimentos de clonagem molecular de genes, permitindo o surgimento do

novo paradigma de sequenciamento de DNA de altíssima eficiência, como o pirosequenciamento.

Essa técnica foi desenvolvida por R. K. Saiki e colaboradores no ano de 1985⁶ e, a partir de então, foi automatizada com o uso de DNA polimerases termorresistentes como a *Taq* polimerase extraída da bactéria termofílica *Thermus aquaticus*. O procedimento é de grande simplicidade e consiste basicamente na amplificação exponencial de DNA, cujo sistema de amplificação contém: amostra de DNA que se quer amplificar (molde), um par de *primers* (iniciadores), quatro desoxinucleosídeos trifosfatos, tampão e força iônica adequada e uma DNA polimerase termorresistente.

O procedimento em si ocorre em três etapas que compõem um ciclo que se repete de 25 a 40 vezes, sendo elas:

- 1) Desnaturação e separação da dupla fita de DNA.
- 2) Pareamento ou anelamento dos iniciadores (*primers*) em cada fita.
- 3) Síntese de novas fitas de DNA.

Como descrito na Figura 1.14, a desnaturação, etapa inicial, consiste na incubação do DNA a temperaturas entre 92 °C e 95 °C, cujo objetivo é separar a dupla fita em cadeias simples. A etapa seguinte é feita com a diminuição gradual da temperatura e pareamento ou anelamento dos *primers* em cada fita simples, de forma a flanquear a região a ser amplificada. Durante a diminuição da temperatura, as duas cadeias de DNA que foram desnaturadas não são novamente pareadas porque estão em concentração baixa na reação, já os *primers* vão se parear rapidamente nas suas regiões complementares no DNA-alvo, pois estão em concentrações muito altas. O anelamento ocorre entre 5 °C e 10 °C abaixo da menor T_m (*T melting*, ou temperatura de desnaturação média) dos *primers* e é uma etapa bastante importante para o sucesso da reação. Caso o pareamento não ocorra, ou ocorra em regiões indesejadas, a reação é fadada ao insucesso. A última etapa fecha o ciclo com a síntese de fitas de DNA pela enzima polimerase, normalmente à temperatura de 68 °C a 72 °C. Com os *primers* pareados nas fitas moldes, a DNA polimerase vai sintetizando a cadeia pela adição de nucleosídeos monofosfatos nas extremidades 3' dos *primers*, formando, então, as novas fitas na direção 5' para 3'.

A cada ciclo a quantidade de DNA se duplica, levando à produção de 2ⁿ moléculas de DNA. Como cada ciclo demora apenas poucos minutos, o processo todo é muito rápido, e após os 25 a 40 ciclos obtêm-se de milhões a centenas de milhões de cópias do segmento desejado, geralmente com comprimento de umas poucas centenas a uns poucos milhares de pares de base.

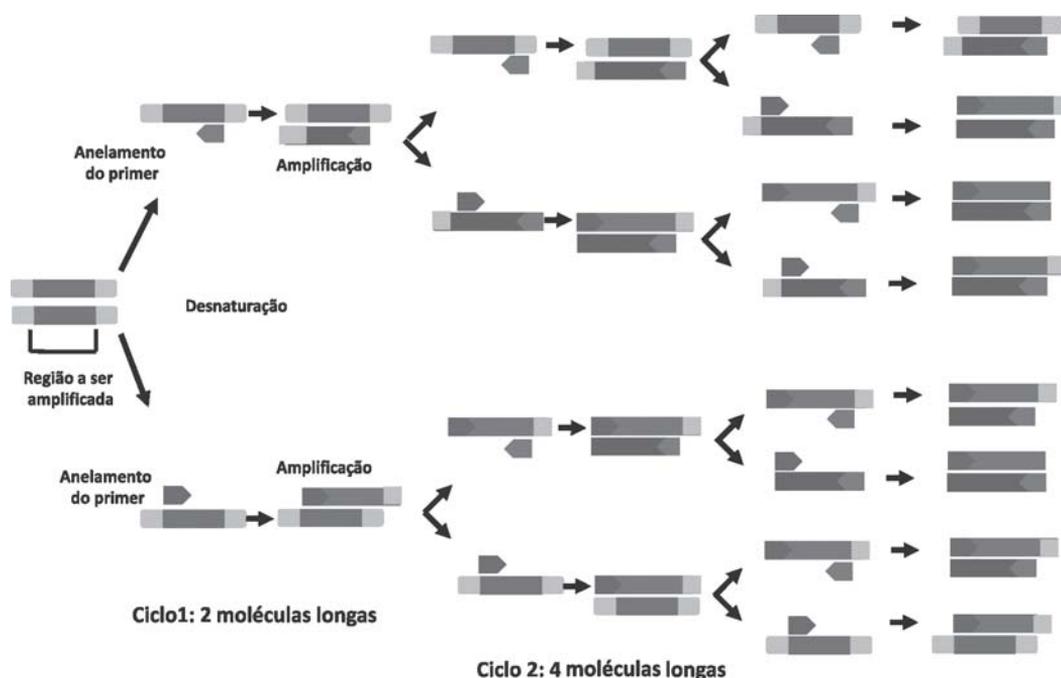


Figura 1.14 Reação em cadeia da polimerase (PCR). Esquema mostrando a reação de amplificação de um segmento de DNA por PCR após quatro ciclos.

1.5.2 Obtenção de cDNA

O dogma central da biologia molecular, na sua essência, é determinado pelo processo de transferência da informação biológica do DNA para RNA e deste para proteína. Entretanto, em alguns casos a informação flui do RNA para o DNA. Alguns vírus (retrovírus), como o HIV, têm o genoma em forma de RNA que, em seu ciclo de vida, é copiado, originando um DNA complementar (cDNA) por uma enzima chamada transcriptase reversa, que é uma DNA polimerase RNA-dependente. O processo é chamado de transcrição reversa e é utilizado pelos biólogos moleculares para produção de cDNA *in vitro* para diversas aplicações:

- Quando se deseja produzir bibliotecas de cDNAs (nesse caso, produzem-se cDNAs a partir de uma população de mRNAs).
- Quando, a partir de biblioteca de cDNA deseja-se isolar um determinado gene específico.

- Quando se deseja expressar um gene eucarioto em hospedeiros procaríotos que não processam os introns dos genes eucarióticos.
- Quando se deseja amplificar determinadas sequências a partir de molde de RNA em um processo denominado de RT-PCR.
- Quando se deseja conhecer toda a população de mRNA expressa em determinada fase de vida de uma célula ou de um tecido (transcriptoma).

No procedimento de obtenção de cDNA atualmente utilizado, o RNA total dos organismos é extraído e a fração contendo o mRNA é purificada através de uma cromatografia de afinidade com colunas contendo oligo-dT celulose ou poli-U sefarose. O mRNA é utilizado como molde para a síntese da primeira fita de cDNA pela ação da enzima transcriptase reversa. Cumpre lembrar que em células eucarióticas o mRNA nascente é processado para remoção de introns (*splicing*), adição de cap na extremidade 5' e de uma cauda poliadenilada – cauda poli(A) – na extremidade 3', produzindo o mRNA maduro. Com base nisso, para a construção da primeira fita de cDNA utiliza-se uma pequena sequência de desoxitimidinas (poli dT) que hibridiza na cauda poli(A) do mRNA e serve como um *primer* para a enzima transcriptase reversa catalisar a síntese da primeira fita do DNA complementar, resultando em uma molécula híbrida formada por DNA/RNA (Figura 1.15). Para a formação da segunda fita do cDNA, a molécula híbrida é tratada com a RNase H, que cliva a fita de mRNA em diversos pontos, introduzindo quebras entre nucleotídeos vizinhos na sequência ribonucleica. Os fragmentos de mRNA servem de *primers* para a DNA polimerase I sintetizar a segunda fita de cDNA utilizando sua atividade polimerásica, o que faz concomitantemente à degradação do mRNA na direção 5'-3' com sua atividade exonucleásica 5'-3'. A fita dupla do cDNA está agora formada e pode ser tratada com a enzima T4 DNA polimerase para tornar as extremidades *blunt end* para ligar em um vetor, também *blunt end*, ou adicionar adaptadores com extremidades coesivas, normalmente geradas por endonucleases de restrição e, então ligar a um vetor tratado com as mesmas endonucleases.

1.5.3 Engenharia genética aplicada à saúde: o caso da insulina humana expressa em bactérias

Será descrita a seguir a primeira experiência brasileira de produção de um biofármaco recombinante, a insulina humana. A tecnologia foi desenvolvida pela Biobrás (Biotecnologia Brasileira S.A.), de Montes Claros (MG),

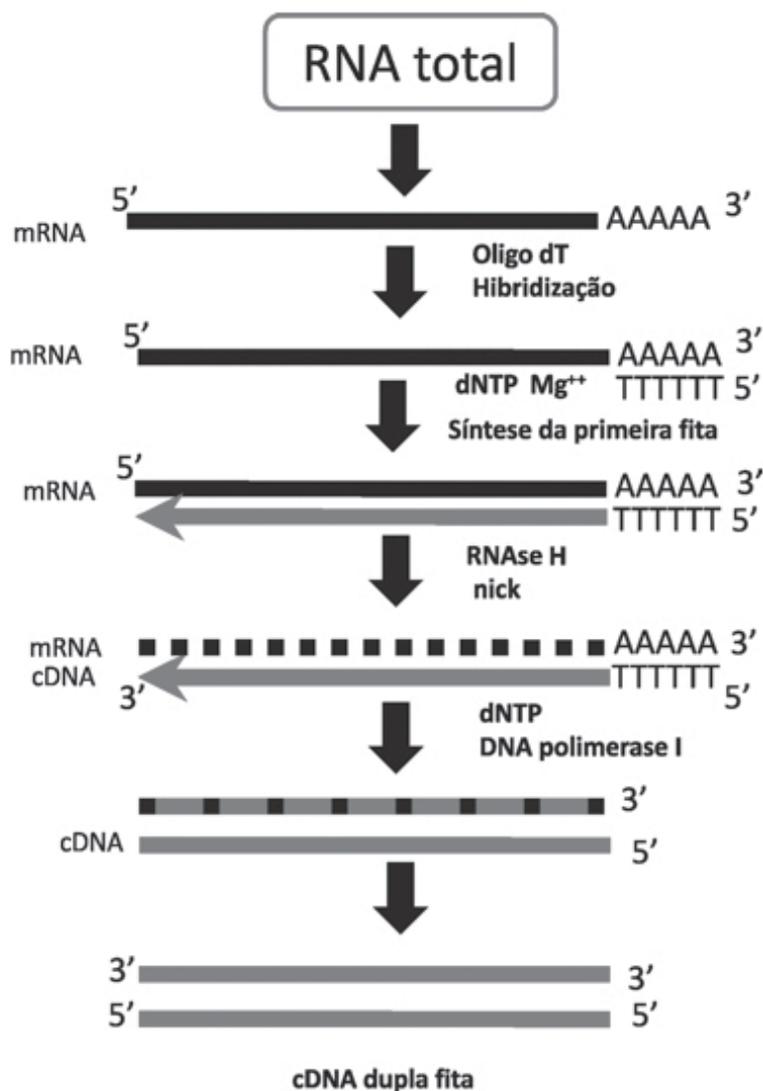


Figura 1.15 Síntese de cDNA. Esquema demonstrativo da obtenção de um cDNA de fita dupla a partir do mRNA com uso das enzimas transcriptase reversa, DNA polimerase I e RNase H.

em colaboração com a Universidade de Brasília (UnB), no período de 1988 a 1999³³.

A insulina é um hormônio proteico produzido por células das ilhotas de Langerhans no pâncreas. É produzida como uma proteína precursora com 110 aminoácidos, chamada de pré-proinsulina, constituída de três cadeias (A, B e C). Para adquirir sua forma ativa, a pré-proinsulina passa

por modificações pós-traducionais (proteólise), resultando na insulina com 51 aminoácidos, com apenas as cadeias A e B.

O peptídeo-sinal da insulina contém 24 aminoácidos, tendo a função de direcionamento da proteína para o retículo endoplasmático, onde é processado e retirado da sequência da insulina gerando a forma proinsulina ainda constituída das cadeias A, B e C. As cadeias A e B, compostas por 21 e 30 aminoácidos, respectivamente, são conectadas pela cadeia C (35 aminoácidos). Durante a maturação, há a formação de duas ligações dissulfeto entre os resíduos de cisteínas das cadeias A e B e entre resíduos internos na cadeia A. A proteína dobra-se e é transportada para o aparelho de Golgi empacotada em vesículas de secreção, onde é clivada pelas enzimas: PC1 e PC2 (“proprotein convertase 1 e 2”), que são enzimas “*trypsin like*” e digerem a proinsulina nas posições R32 e R65, deixando dois resíduos de aminoácidos básicos na extremidade carboxi da cadeia B (R-R ou Arg-Arg) e na extremidade carboxi da cadeia C (K-R ou Lys-Arg). A seguir, a carboxipeptidase F, uma enzima “*carboxipeptidase B like*” retira os dois resíduos básicos de ambas as extremidades, resultando na liberação da cadeia C, de modo que as cadeias A e B permanecem ligadas entre si pelas ligações dissulfeto formando o hormônio insulina, que é então secretado na corrente sanguínea (Figura 1.16).

A anormalidade na secreção ou função da insulina em seres humanos gera sérias alterações no metabolismo da glicose, causando uma desordem chamada de *Diabetes mellitus*, que pode ser controlada com a administração de insulina e controle de ingestão de glicose. Devido a sua importância terapêutica, esse hormônio foi produzido para uso terapêutico em humanos por extração de pâncreas bovinos ou suínos. Esse procedimento, além de ser dispendioso, tem algumas outras desvantagens em relação ao de TDR, como:

- requer grande quantidade de pâncreas animal;
- no processo de purificação a partir de pâncreas animal, o hormônio pode vir contaminado com patógenos de mamíferos, como vírus, viróides e príons;
- como as moléculas têm sequências de aminoácidos que variam um pouco em relação à sequência humana, podem induzir à formação de anticorpos durante o tratamento.

A técnica desenvolvida no Brasil consiste em introduzir e expressar em altos níveis no interior da bactéria *Escherichia coli* a sequência codificadora

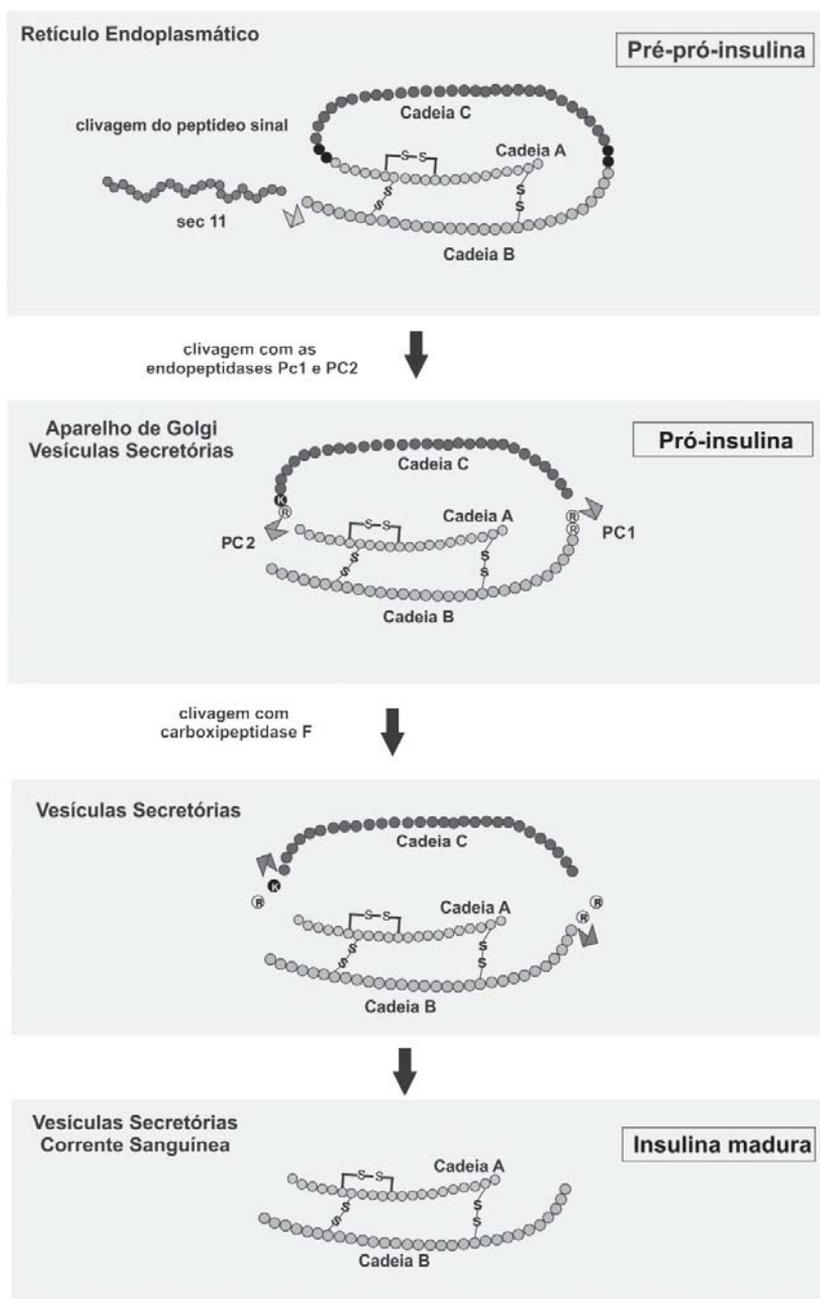


Figura 1.16 Processamento da insulina. A pré-proinsulina é convertida em proinsulina pela clivagem do peptídeo-sinal pela peptidase sec11. A seguir, é translocada para o Aparelho de Golgi, onde a cadeia C é clivada pela ação das peptidases PC1 e PC2 e da carboxipeptidase F, liberando o hormônio para a secreção.

da proinsulina humana produzida por síntese química, conforme descrito a seguir:

- O gene da insulina humana foi predito a partir da sequência de aminoácidos disponível na literatura e construído por síntese química. A vantagem de se usar a síntese química é a possibilidade de definir a estrutura ideal do gene para sua expressão ideal.
- O gene da proinsulina foi construído através do anelamento de sequências oligonucleotídicas sintetizadas separadamente, de forma a conter as cadeias A, B e C do peptídeo original, como observado no esquema da Figura 1.17. Para a construção do gene, quatro oligonucleotídeos foram sintetizados para a cadeia B, quatro para as cadeias C e dois para a A. Os oligonucleotídeos foram anelados entre si. Montaram-se inicialmente as cadeias B, C e A separadamente, e então elas foram ligadas nessa ordem. Após a construção, o gene foi inserido em vetor de clonagem para a realização de sequenciamento de DNA e a confirmação da sequência nucleotídica correta do gene³⁴.

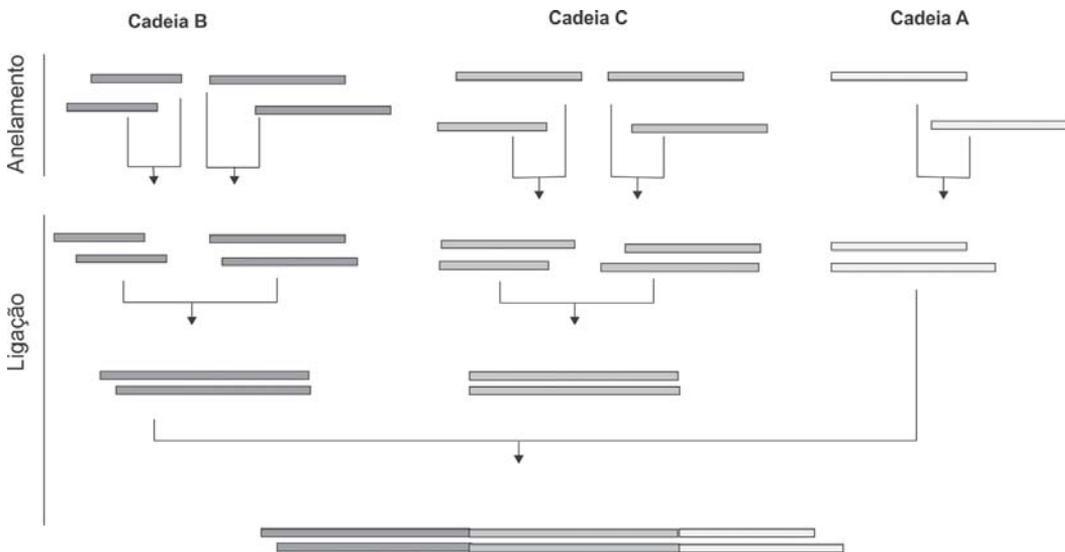


Figura 1.17 Estratégia utilizada para a construção da mensagem genética da proinsulina humana por síntese química para a expressão na bactéria *Escherichia coli*.

Após a sequência do gene da proinsulina ter sido confirmada com o sequenciamento, o próximo passo foi a subclonagem em um vetor de expressão chamado de pLMT8.5, construído especialmente para expressar em altos níveis a pro-insulina em *E. coli*. O vetor foi projetado para conter partes ou sequências específicas (promotor, Shine Dalgarno, terminador de transcrição e gene de resistência à antibiótico) que garantissem alto nível de expressão da proteína recombinante na hospedeira bacteriana.

Para se garantir uma replicação estável foi utilizada a origem de replicação do pUC 18 derivada do plasmídeo pMB1 projetado para aumentar o número de cópias do vetor na célula.

A marca de seleção dominante utilizada foi a resistência à tetraciclina. O gene da tetraciclina foi isolado do plasmídeo RP4 e foi escolhido para garantir pressão seletiva durante a fermentação na fase de produção do inóculo.

O promotor escolhido foi o pL do bacteriófago lambda que pode ser ativado simplesmente pela mudança de temperatura. É um promotor forte e regulado negativamente pelo repressor cI, sensível à temperatura, que é funcional a 28 °C e não funcional a 42 °C.

Para se garantir eficiente expressão, o sítio de ligação de ribossomos (região Shine Dalgarno) foi obtido por síntese química e baseada na região Shine Dalgarno do fago T7, bastante eficiente na tradução de genes expressos em *E. coli*. De maneira similar, uma sequência de término de transcrição Rho-independente, também foi obtida por síntese química, baseando-se na sequência do operon triptofano de *E. coli*.

Por último, para facilitar a clonagem direcional foi inserido entre a região promotora e a região terminadora, um múltiplo sítio de clonagem com sítios únicos para as enzimas de restrição *Nco*I, *Eco*RI, *Stu*I, *Pst*I e *Bam*HI. Estrategicamente, o códon de início de tradução ATG já estava incluído no sítio de *Nco*I. Com essas estratégias construiu-se o vetor pLMT8.5 para superexpressão em *E. coli*, visualizado na Figura 1.18.

O gene da proinsulina foi subclonado no pLMT8.5, e o vetor recombinante foi denominado de pPTA1. Esse plasmídeo foi inserido por transformação em uma linhagem de *E. coli* que possui o repressor termossensível, e a proinsulina foi expressa por meio de indução térmica.

Verificou-se a expressão da proinsulina com aproximadamente 10 mil daltons, que correspondia a cerca de 20% das proteínas totais da bactéria (Figura 1.19). A proteína foi expressa em forma de corpúsculos de inclusão no citoplasma da bactéria, o que auxilia na proteção contra proteólise e facilita na recuperação para a purificação após a lise da célula, centrifugação e solubilização.

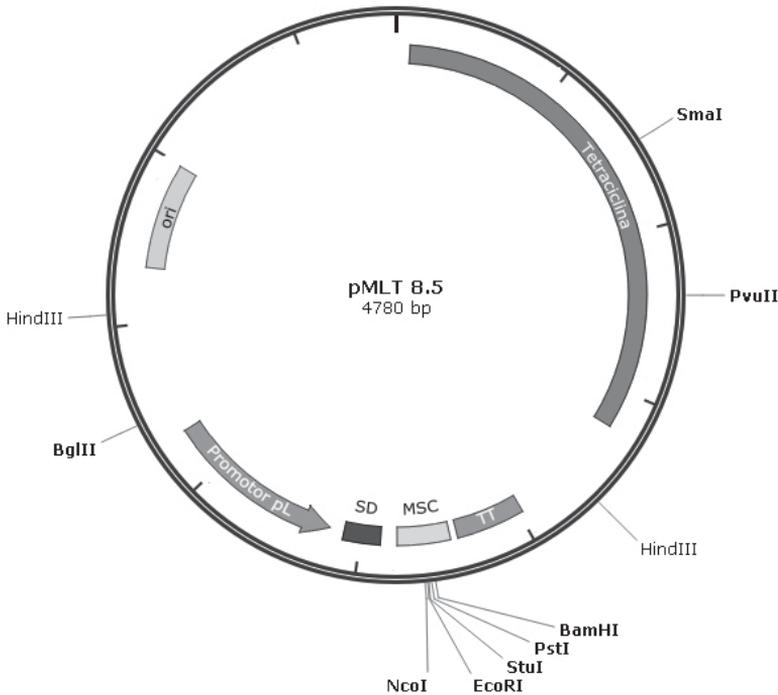


Figura 1.18 Mapa do vetor de expressão pLMT8.5 construído para a produção da proinsulina humana em *E. coli*. *TetR*: gene de resistência a tetraciclina; *ori*: origem de replicação; MCS: região de sítios únicos para enzimas de restrição; pL: promotor pL; SD: sequência Shine-Dalgarno e TT: região de terminação de transcrição.

Após a obtenção das células induzidas com os corpúsculos de inclusão no interior, estas eram coletadas por centrifugação, os corpúsculos de inclusão separados por centrifugação, dissolvidos com alta concentração de ureia em pH alto, a proinsulina era purificada por cromatografia de afinidade com níquel, em seguida era renaturada e a cadeia C eliminada com digestão por duas proteases, tripsina e carboxipeptidase B, e era, então, purificada por procedimentos cromatográficos adicionais.

1.6 POSSIBILIDADES TERAPÊUTICAS: A REVOLUÇÃO NA ÁREA DE SAÚDE VIA ENGENHARIA GENÉTICA

A biotecnologia molecular, por meio da engenharia genética, tem provocado profundas alterações em nosso cotidiano, em diversos segmentos da sociedade moderna, da produção de sementes à colheita, das biorrefinarias

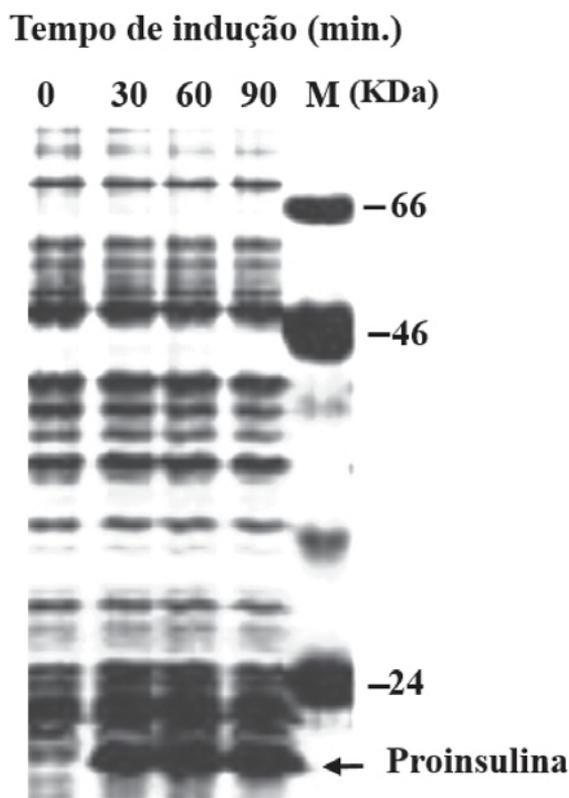


Figura 1.19 Análise da expressão em diferentes tempos de indução da proinsulina em *E. coli*. Gel desnaturante (SDS-PAGE) a 15% dos lisados proteicos totais das células bacterianas, contendo o plasmídeo pPTA1 (pLMT8.5 + gene da proinsulina)³⁴.

aos biocombustíveis e do diagnóstico de doenças ao desenvolvimento de novos medicamentos, assim como no aperfeiçoamento de novas técnicas de tratamento.

As técnicas de manipulação genética que permitiram a construção de moléculas recombinantes possibilitaram grandes avanços no campo da saúde, a partir do momento em que se tornou possível obter moléculas de aplicação terapêutica e investigativa. Nesse sentido, a engenharia genética tem sido um caminho promissor para a cura e prevenção de muitas enfermidades ou desordens fisiológicas em humanos. Doenças diversas, como diabetes, hepatites, câncer e outras anormalidades, têm agora possibilidade de tratamento ou terapia, ou até mesmo cura, pelo uso de moléculas produzidas via TDR, tais como enzimas, hormônios, vacinas, anticorpos monoclonais, entre outras.

1.6.1 Terapia genética

A terapia gênica ou geneterapia baseia-se na introdução de genes nas células e tecidos de indivíduos que possuam uma doença causada pela deficiência desses genes.

Considerando as características dos vírus de infectar células e nelas replicar seu material genético, inferiu-se que os vírus seriam eficientes vetores para carrear genes ao interior de células humanas em procedimentos de terapia genética. Para se desenvolver um vetor viral, normalmente altera-se seu DNA com o objetivo de manter sua capacidade de replicação, mas não seu lado patogênico ou a capacidade de transmitir a doença (em outras palavras, desarma-se o vírus). Pode-se inclusive adicionar múltiplos sítios de clonagem para facilitar a inserção do gene correto a ser introduzido no paciente.

O primeiro caso de sucesso de terapia genética em humanos foi realizado nos Estados Unidos em uma menina chamada Ashanti Silva, que tinha imunodeficiência severa causada pela ausência de uma enzima encontrada em leucócitos, denominada de adenina deaminase (*ADA*). A doença aparece em crianças portadoras homozigotas do gene não funcional, que para sobreviverem têm que viver no interior de uma bolha estéril. O gene *ADA* correto foi introduzido em células da medula óssea utilizando um vetor retroviral. A imunodeficiência de Ashanti foi parcialmente recuperada, o que permite que ela viva uma vida normal.

Em 2003, a China aprovou o primeiro produto comercial para terapia genética, denominado de Gendicine. Trata-se de um produto para tratamento de câncer no qual o gene da proteína anticancerígena P53 é carregado por um vetor adenoviral. O produto mostrou-se ativo como coadjuvante dos tratamentos cirúrgicos e quimioterápicos de tumores sólidos do pescoço e da cabeça. Atualmente, existem no mundo cerca de mil testes clínicos em andamento visando ao desenvolvimento de diversos produtos dessa natureza.

Somente em 2012 foi aprovado um medicamento baseado em terapia genética no mundo ocidental. A aprovação se deu pela European Medicines Agency (EMA) para uso no âmbito da Comunidade Europeia, e o medicamento é o Alipogene Tiparvec (nome fantasia Glybera). Esse medicamento destina-se ao tratamento de deficiência da enzima lipoproteína lipase, que causa grande aumento de gordura no sangue e, conseqüentemente, pancreatite severa. O medicamento consiste do gene correto da enzima carregado por vetor derivado de um vírus adenoassociado (AAV) sorotipo I, que tem propensão em infectar células musculares. O medicamento Glybera é introduzido no corpo por meio de injeção intramuscular, e os testes clínicos

mostraram claramente a diminuição do nível de gordura no sangue e do aparecimento de pancreatite severa.

Alternativamente ao uso de vetores virais, a terapia pode utilizar outros vetores, como lipossomas ou conjugados moleculares, diminuindo possibilidades de infecção viral ou toxicidade e aumentando a possibilidade de efeito local mais eficiente, quando essas partículas adentram as células-alvo.

1.6.2 Vacinas recombinantes

Vacinas, no sentido estrito, são constituídas por macromoléculas específicas que induzem a resposta imunológica protetora contra um agente patogênico. O uso de vacinas é, sem dúvida, um dos avanços da ciência de maior impacto em saúde pública, visto que muitas epidemias foram completamente controladas em várias partes do mundo com o uso de vacinas. Os objetivos de se desenvolver vacinas são, basicamente: promover a proteção dos indivíduos contra infecções, bloquear a transmissão, prevenir sintomas e, de forma mais ambiciosa, levar à erradicação de tais infecções. Assim, foram desenvolvidos vários tipos de vacinas como as que conhecemos hoje, contra doenças como difteria, tétano, febre amarela, coqueluche, poliomielite, varíola, entre tantas outras.

As primeiras vacinas foram desenvolvidas a partir da purificação de antígenos ou da inativação química do agente patogênico extraído de plasma de animais infectados. A modernização da biologia trouxe a possibilidade de melhorias nos métodos de produção e purificação das vacinas já existentes, e os avanços da biologia molecular possibilitaram o desenvolvimento de vacinas recombinantes, a partir da identificação de vários antígenos, clonagem de seus genes e expressão em sistemas heterólogos.

A primeira vacina recombinante produzida foi a vacina contra a hepatite B, no ano de 1986. Utilizando a TDR, o antígeno de superfície do capsídeo do vírus HBV foi clonado e expresso na levedura *Saccharomyces cerevisiae*. O antígeno (AgHB) purificado das células recombinantes de *S. cerevisiae* é formulado com adjuvante adequado (hidróxido de alumínio), gerando então a vacina contra hepatite B. A vantagem de produzir vacinas via engenharia genética é a possibilidade de obter o antígeno em grandes quantidades, além de praticamente zerar o risco de infecção, pois se trata de proteínas recombinantes livres do DNA do patógeno. Esta tecnologia tem sido aplicada na obtenção de outros antígenos para a produção de outras vacinas, como a vacina contra o vírus do papiloma humano (HPV).

Estão em desenvolvimento o que chamamos de vacinas de DNA, que, em vez de conter o antígeno recombinante, contêm o gene do antígeno clonado no vetor de expressão, que irá garantir a expressão do antígeno quando o vetor for introduzido nas células do indivíduo no qual se deseja desencadear a resposta imune protetora. Dentro das células, o vetor garante a expressão ou expressão/secreção do antígeno, que, por ser estranho, desencadeará resposta imune. Essa estratégia está em desenvolvimento, embora mais atrasada, pois tem havido relutância em se aprovar procedimentos que envolvam injeção de DNA exógeno em um animal, pela possibilidade de, em certos casos, acarretar o aparecimento de teratomas ou neoplasias malignas.

1.6.3 Biofármacos recombinantes

A indústria biotecnológica tem desenvolvido uma infinidade de biofármacos recombinantes com excelentes aplicações na área médica. A exemplo da insulina humana, discutida acima, muitos outros produtos farmacêuticos para uso humano têm sido produzidos por fermentação de microrganismos, principalmente *Escherichia coli*.

Moléculas terapêuticas humanas complexas que têm modificações pós-traducionais como glicosilações e fosforilações não podem ser feitas em bactérias, pois estas não fazem tais modificações, e nem em leveduras ou fungos filamentosos, que fazem modificações pós-traducionais não idênticas às de mamíferos e, portanto, podem induzir, com maior chance, resposta imune durante o tratamento. Nesse caso, a hospedeira predileta para a produção de proteínas humanas complexas são as células recombinantes da linhagem CHO (*chinese hamster cells*), que secretam biofármacos como: fator VIII de coagulação sanguínea, anticorpos monoclonais, interferon beta etc. A Tabela 1.3, a seguir, mostra alguns dos principais biofármacos produzidos por engenharia genética, suas aplicações e em que hospedeiro são produzidos.

Existem aproximadamente duzentos biofármacos aprovados para uso e cerca de 460 em triagem clínica, dentre os quais predominam os anticorpos monoclonais recombinantes (MABs)³⁵. No cenário mundial, o Brasil vem tendo consideráveis destaques na produção de moléculas recombinantes, fruto de décadas de pesquisa e do elevado nível de modernização dos centros de pesquisa brasileiros, do conhecimento de cientistas nacionais e da disposição de empresas de capital nacional em investirem nessa área. Dentre os produtos recombinantes de aplicação clínico-terapêutica há o exemplo pioneiro da insulina já comercializada pela Biobrás e o hormônio de

Tabela 1.3 Alguns produtos terapêuticos feitos por engenharia genética

| PRODUTO | APLICAÇÃO | HOSPEDEIRO |
|----------------------------------|---|--|
| Insulina humana | Tratamento de diabetes | Bactérias Leveduras |
| Hormônio de crescimento | Desordem de crescimento/baixa estatura | Bactéria Leveduras |
| Fator de crescimento epidérmico | Cicatrização de feridas | Leveduras |
| Interleucina 2 | Tratamento de câncer | Bactérias Leveduras |
| α -interferon | Antiviral, agente antitumoral | Bactérias |
| β -interferon | Tratamento de esclerose múltipla | Bactérias Células de mamíferos |
| Eritropoetina | Tratamento de certos tipos de anemia | Células de mamíferos |
| Fator VII, VIII e IX | Agentes coagulantes | Células de mamíferos |
| Ativador de plasmogênio tecidual | Dissolver coágulos | Bactérias Leveduras Células de mamíferos |
| Anticorpos monoclonais (MABs) | Diversas doenças. Ex.: rituximabe (antileucêmico), humira (tratamento de doenças autoimunes) | Células de mamíferos |

crescimento humano para o tratamento de nanismo, no momento em fase de testes clínicos pela Cristália Produtos Químicos Farmacêuticos Ltda., projeto desenvolvido em colaboração com a Universidade Federal do Amazonas. Alguns outros produtos, como enzimas, hormônios e anticorpos monoclonais, já tiveram seus genes clonados no país e se encontram em diferentes fases de desenvolvimento.

Recentemente, pela importância que têm assumido os anticorpos monoclonais, principalmente contra o câncer e doenças autoimunes, que têm que ser importados e consomem cerca de 40% de todos recursos disponíveis pelo SUS para a aquisição de medicamentos, o governo brasileiro, por intermédio do Ministério da Saúde tem criado condições por meio de Parcerias Público-Privadas (PPPs) para estimular a produção desses biofármacos em nosso país, a fim de que a população brasileira tenha maior acesso ao medicamento e possamos economizar importantes divisas.

1.7 COMO CONSTRUIR UMA BIBLIOTECA GENÔMICA

Os procedimentos para a construção de uma biblioteca genômica serão apresentados, a seguir, pela sua importância para a tecnologia do DNA recombinante e também porque as técnicas relacionadas requerem o uso de importantes ferramentas que foram apresentadas neste capítulo. A Figura 1.20 a seguir mostra as etapas que são normalmente cumpridas na construção de uma biblioteca genômica, que foram também apresentadas de outra maneira na Figura 1.1, no início deste capítulo.

No conjunto de protocolos apresentados a seguir, utilizaremos como

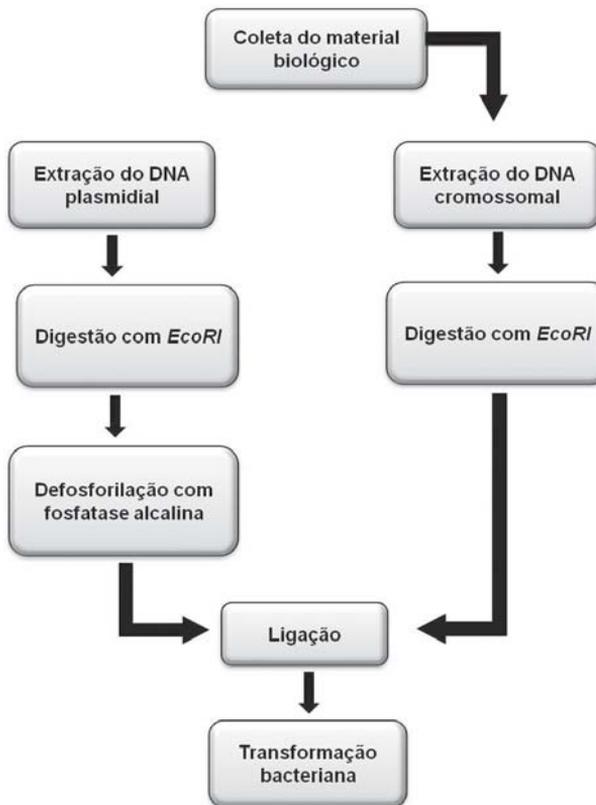


Figura 1.20 Principais etapas do procedimento de construção de uma biblioteca genômica.

exemplo a construção de uma biblioteca genômica a partir de material extraído de um fígado animal. Com as devidas modificações, esses protocolos podem ser utilizados com outros tipos de tecidos e/ou células.

1.7.1 Extração de DNA cromossomal pelo método do fenol/clorofórmio^{36*}

Esse procedimento foi escolhido, pois além de resultar em DNA com bom nível de pureza, é similar ao utilizado por Avery e colaboradores³⁷ quando extraíram DNA de pneumococos e demonstraram ser essa substância o material genético (“princípio transformante”). Na época, em vez de solução com fenol, eles utilizaram para a extração uma mistura de clorofórmio com álcool isoamílico. Se for utilizada uma quantidade de tecido três vezes maior (150 mg), aumentando-se proporcionalmente os volumes das soluções, é possível na etapa 14 adicionar vagarosamente o álcool a -20 °C e coletar o DNA com um bastão de vidro em vez de centrifugação e lavá-lo mergulhando o bastão por alguns minutos em solução de etanol 70% a -20 °C.

Soluções

NE:

- NaCl 150 mM
- EDTA 10 mM pH 8,0

TEN:

- NaCl 100 mM
- EDTA 50 mM
- Tris-HCl 50 mM pH 7,5

Triton X-100:

- 10% (V/V) em água destilada ou milli-Q.

SDS 10% (P/V):

- Em água destilada ou milli-Q.

RNase A (10 mg/mL):

- Dissolver em tampão TR (ferver em banho-maria por 15 minutos)

Proteinase K:

- 10 mg/mL em água milli-Q

TR:

- EDTA 0,2 mM

- Tris-HCl 10 mM pH 8,0

Fenol: fenol hidratado equilibrado com:

- NaCl 100mM; EDTA 1 mM
- Tris-HCl 20 mM pH = 7,5, adicionado de 8-hidroxiquinolina e B-mercaptoetanol para as concentrações finais de 0,05% (P/V) e 1% (V/V), respectivamente

Clorofane:

- Fenol 1V + 1V de clorofórmio hidratado, misturados na hora do uso.

Clorofórmio hidratado:

Clorofórmio saturado em água milli-Q

Procedimento

- 1) Pesar 50 mg de tecido em um microtubo e homogeneizar diretamente em 500 mL de TEN. Prosseguir imediatamente ou manter a amostra a 4 °C.
- 2) Adicionar 60 mL de Triton X-100 10% e homogeneizar suavemente.
- 3) Adicionar 4 mL de RNase A (10 mg/mL).
- 4) Homogeneizar bem (sem movimentos bruscos), durante 5 minutos.
- 5) Incubar a 37 °C por 30 minutos.
- 6) Adicionar 60 mL de SDS 10%.
- 7) Adicionar 5 mL de proteinase K (10 mg/mL).
- 8) Incubar a 50 °C por 60 minutos.
- 9) Adicionar 1 volume de fenol hidratado, agitar suavemente o tubo por inversão por 10 minutos.
- 10) Centrifugar a 2.000 g por 10 minutos.
- 11) Coletar o sobrenadante (fase aquosa) e transferir para um novo microtubo.
- 12) Adicionar à fase aquosa um volume de clorofane, agitando suavemente por inversão por 10 minutos e repetir os procedimentos dos itens 10 e 11.
- 13) Adicionar à fase aquosa recuperada um volume de clorofórmio hidratado agitando suavemente por inversão por 5 minutos e repetir os procedimentos dos itens 10 e 11.
- 14) Adicionar à fase aquosa 0,1 volume de NaCl 3M e 2,5 volumes de etanol 100% a -20 °C e misturar invertendo o tubo várias vezes.
- 15) Deixar precipitando a 0 °C (banho água-gelo) por 30 minutos.
- 16) Centrifugar a 12.000 g por 10 minutos.
- 17) Descartar o sobrenadante.

- 18) Lavar o sedimento, sem ressuspender, com etanol 500 μ L 70% a -20°C , mantendo por 5 minutos no gelo, centrifugar a 12.000 g por 2 minutos e desprezar o sobrenadante.
- 19) Secar em “*speed vac*” ou em fluxo laminar.
- 20) Dissolver em 100 mL a 300 mL de tampão TR, mantendo a 4°C . Para dissolver bem, deixar a 4°C durante a noite.
- 21) Quando estiver totalmente dissolvido, aliquotar e estocar a -20°C .
- 22) Analisar uma alíquota de 5 μ L por eletroforese em gel de agarose a 0,8%. A concentração do DNA pode ser estimada por espectrofotometria na região do UV, considerando-se que uma solução com 50 μ g de DNA por mL apresenta a absorvância igual a 1,0 no comprimento de onda de 260 nm.

1.7.2 Extração de plasmídeos por lise alcalina^{38, 39 *}

Esta metodologia permite separar de forma rápida e barata o DNA plasmidial do DNA cromossomal. As bactérias são lisadas com solução de NaOH e SDS, que também desnatura o DNA, após o que a solução é neutralizada por meio da adição de solução de acetato de potássio em alta concentração com pH 5,8, condições nas quais o DNA cromossomal desnaturado precipita e o plasmídeo, que se renatura mais facilmente recuperando sua grande hidrofiliçidade, permanece em solução.

Se houver necessidade de uma preparação de DNA plasmidial mais pura, é possível a partir da etapa 12 realizar extração pelo método do fenol/clorofórmio de forma similar ao descrito para a purificação do DNA cromossomal. Pode-se também utilizar kits comerciais que incluem, após a etapa 6, uma etapa adicional que consiste na ligação do DNA especificamente em uma resina seguida de lavagem dos contaminantes e então sua eluição da resina.

Soluções

Solução I:

- Tris-HCl 100 mM pH 7,5
- EDTA 10 mM pH 8,0
- Lisozima 2 mg/mL

* Modificado.

Solução II:

- NaOH 0,2 N
- SDS 1%
- Observação: diluir primeiro o NaOH e depois adicionar o SDS. Essa solução não deve ser estocada por mais que um mês).

Solução III:

- Acetato de potássio 3 M (P/V) pH 5,8

TR:

- EDTA 0,2 mM
- Tris-HCl 10 mM pH 8,0

Procedimento

- 1) Semear em placas as bactérias contendo o plasmídeo pUC18 no dia anterior, de forma a obter-se colônias isoladas.
- 2) Transferir uma colônia isolada para um tubo estéril contendo 5 mL de meio LB contendo ampicilina 100 ug/mL e incubar a 37 °C por 16 horas, agitação vigorosa (180 rpm).
- 3) Transferir 1,5 mL da cultura para um microtubo e centrifugar a 12.000 g por 1 minuto. Repetir o passo 2 para a obtenção de maior quantidade de *pellet* celular.
- 4) Ressuspender o *pellet* em 200 µL da solução I gelada. Agitar rapidamente no agitador de tubos (vórtex). Acrescentar 1µL de RNase A (10 mg/mL) e incubar a temperatura ambiente por 10 minutos.
- 5) Adicionar 360 µL da solução II e misturar rapidamente, invertendo o tubo, e incubar no gelo por 5 minutos.
- 6) Adicionar 300 µL da solução III gelada, misturar por inversão do tubo até precipitar bem (deve aparecer um precipitado claro bem definido).
- 7) Centrifugar a 12.000 g por 5 minutos (para sedimentar os restos celulares e o DNA cromossomal precipitado). Transferir o sobrenadante para um tubo novo.
- 8) Adicionar um volume de isopropanol misturar bem por inversão e deixar 5 minutos a temperatura ambiente. Centrifugar a 12.000 g por 5 minutos.
- 9) Descartar o sobrenadante e remover excesso de isopropanol com auxílio de micropipeta.
- 10) Adicionar 1 mL de etanol 70% gelado a -20 °C.
- 11) Centrifugar por 5 minutos a 12.000 g. Remover o sobrenadante.

- 12) Secar o *pellet* em “*speed vac*” por 5 minutos ou em corrente de fluxo laminar por 15 minutos.
- 13) Dissolver o DNA em 50 μL de Tampão TR.
- 14) Analisar 5 μL por eletroforese em gel de agarose 0,8%.

1.7.3 Digestão dos DNAs com a enzima de restrição *EcoRI*

Nesta etapa, o DNA cromossomal e o plasmídeo pUC18 serão digeridos com a enzima *EcoRI* a fim de gerar extremidades compatíveis para posterior ligação. Por definição, uma unidade de enzima (U) é a menor quantidade capaz de digerir 1 μg de DNA do fago λ em uma hora a 37 °C em volume total de reação de 50 μL . Geralmente é recomendada, para os procedimentos de clonagem, uma reação de digestão exaustiva com cinco a dez vezes essa quantidade de enzima por μg de DNA. Assim, usamos como exemplo uma reação que utiliza 1 μg de DNA, cujo volume dependerá da concentração obtida na extração. Sempre tomar cuidado para que o volume de enzima não ultrapasse 10% do volume final da reação, a fim de evitar a atividade inespecífica da enzima devido ao excesso de glicerol (>5%), pois normalmente as enzimas são estocadas em glicerol 50%.

Procedimento

Durante a preparação, manter os microtubos em gelo.

De acordo com a Tabela 1.4, misturar em microtubos os volumes dos reagentes e incubar a 37 °C por 2 horas. Analisar 5 μL de cada reação de digestão por eletroforese em gel de agarose a 1%. Se as reações de digestão estiverem adequadas, incubar a 65 °C por 20 minutos para inativar a enzima de restrição *EcoRI*.

1.7.4 Defosforilação do vetor linearizado com fosfatase alcalina

Nesta etapa, o vetor digerido com a enzima *EcoRI* será tratado com a enzima fosfatase alcalina de camarão do Ártico (*shrimp alkaline phosphatase* – SAP) para remoção dos grupos fosfato da extremidade 5' do DNA, pois essa enzima pode ser inativada termicamente evitando a necessidade de, após a defosforilação, eliminar a enzima por tratamento com protease e extração com fenol/clorofórmio, que é um processo muito mais trabalhoso.

Tabela 1.4 Reações de digestão do DNA cromossomal e do vetor.

| REAGENTE | PUC18 (VETOR) + ECORI (μL) | CONTROLE PUC18 – ECORI(μL) | DNA CROMOS. + ECORI (μL) | CONTROLE DNA CROMOS. – ECORI (μL) |
|------------------------------------|---|---|--|--|
| Água milli-Q | q.s.p. | q.s.p. | q.s.p. | q.s.p |
| Tampão 10X | 5 | 5 | 5 | 5 |
| DNA | ___(1 μg) | ___(1 μg) | ___(1 μg) | ___(1 μg) |
| <i>EcoRI</i> (10U/ μL) | 1 | - | 1 | - |
| Volume final | 50 | 50 | 50 | 50 |

Procedimento

- 1) O DNA do vetor de clonagem pUC18 digerido com *EcoRI* deverá ser precipitado com sal e álcool, lavado com etanol 70% e seco.
- 2) Em um microtubo, dissolver 0,5 μg a 1,0 μg do DNA do pUC18 digerido em 10 μL de tampão de reação TSAP.
- 3) Adicionar 0,5 U de fosfatase alcalina (SAP) para cada 1 μg de pUC18 (ou cerca de 1 pmol de extremidades 5' fosfato).
- 4) Incubar a 37 °C por 30 minutos.
- 5) Incubar a 65 °C por 15 minutos para inativar a enzima (SAP).
- 6) Utilizar para ligação ou estocar a -20 °C para uso posterior.

Solução

TSAP:

- 10 mM MgCl_2
- Tris-HCl 20 mM pH 8,0

1.7.5 Ligação dos fragmentos de DNA ao vetor

Antes da ligação, algumas considerações são necessárias. A enzima *EcoRI* é uma enzima de corte raro e produz fragmentos grandes. Dessa forma, consideraremos como tamanho médio de inserto 1 kb. O vetor pUC18 tem tamanho de aproximadamente 3 kb. Assim, para a ligação será utilizada

uma razão de quantidade inserto/vetor de 3:1 em uma reação de volume final de 20 μL .

Procedimento

- 1) Em um microtubo, misture gentilmente os reagentes relacionados na Tabela 1.5.
- 2) Incubar a 16 °C durante a noite.

Tabela 1.5 Reação de ligação do inserto ao vetor.

| REAGENTE | VOLUME (μL) |
|---------------|--------------------------|
| Água milli-Q | q.s.p |
| Tampão 10X | 2 |
| DNA (Vetor) | ___ (50 ng) |
| DNA (Inserto) | ___ (150 ng) |
| T4 DNA Ligase | 1 |
| Volume final | 20 |

1.7.6 Transformação bacteriana

Após a ligação, as moléculas recombinantes serão, então, inseridas em células da bactéria *Escherichia coli* por transformação por choque térmico. Para isso, as células bacterianas devem ser tornadas competentes para a transformação genética por tratamento com cloreto de cálcio. Alternativamente, as células podem ser transformadas geneticamente pelo método da eletroporação.

Procedimento

Preparo de células competentes para a transformação genética

- 1) No final da tarde, inocular 5 mL de meio (LB) com uma única colônia isolada da linhagem escolhida de *E. coli* e incubar durante a noite a 37 °C.

- 2) Inocular em 50 mL de meio de cultura com 1 mL do pré-inóculo obtido durante a noite. Deixar crescer a 37 °C com agitação (180 rpm) até a densidade óptica (OD) de 0,5 a 0,6 a 600 nm.
- 3) Transferir 45 ml da cultura para um tubo de centrifugação tipo falcon estéril e centrifugar por 10 minutos a 5.000 g e 4 °C e, em seguida, cuidadosamente descartar todo o sobrenadante.
- 4) Ressuspender cuidadosamente as células em 10 mL de solução de TCa gelada (banho gelo-água) e manter em gelo durante 30 minutos.
- 5) Coletar as células por centrifugação por 10 minutos a 2.000 g e 4 °C e cuidadosamente descartar o sobrenadante.
- 6) Ressuspender cuidadosamente as células em 1 mL de solução de TCa gelada.
- 7) As células estão prontas para a transformação por choque térmico (ou para congelamento e armazenamento). Observação: A eficiência da transformação aumenta a partir do preparo das células, se adequadamente estocadas entre 0 °C e 4 °C, até cerca de 15 horas, quando começa a diminuir devido à morte das células. Se desejar transformar com baixa eficiência, é possível até 48 horas após o preparo. Para o armazenamento a -80 °C, adicionar glicerol estéril para uma concentração final de 12,5% V/V e congelar em lotes de 100 µL em microtubos de 1,5 mL.

1.7.7 Transformação genética

Soluções e meios de cultura

TCa:

- CaCl₂ 100mM
- Tris-HCl 10mM pH= 7,5 (esterilizar por autoclavagem)

Meio LB:

- peptona 10g/L
- extrato de levedura 5 g/L
- NaCl 4 g/L pH = 7,0
- Esterilizado por autoclavagem

Meio LB sólido:

- Meio LB ao qual se adicionou 15 g/L de ágar bacteriológico
- Esterilizado por autoclavagem

Ampicilina:

- 100 mg/mL (solução 1000X) em água milli-Q estéril

IPTG (isopropil-β-D-tiogalactopiranosídeo):

- 100 mM em água milli-Q estéril

X-gal (5-bromo-4-cloro-3-indolil-β-galactopiranosídeo):

- 20 mg/mL dissolvidos em dimetilformamida

Procedimento

- 1) Adicionar de 1 μL a 10 μL da reação de ligação contendo de 50 ng a 200 ng de DNA em um microtubo contendo 100 μL da suspensão de células competentes.
- 2) Deixar no gelo por 30 minutos.
- 3) Dar um choque térmico transferindo o tubo do gelo para um banho-maria a 42 °C por 45 segundos (ou alternativamente a 37 °C por 5 minutos).
- 4) Deixar no gelo por 3 minutos e adicionar 900 μL de meio LB.
- 5) Incubar à 37 °C por 1 hora sob agitação.
- 6) Plaquear diversos volumes (por exemplo, 20 μL, 50 μL, 100 μL e 200 μL) com alça de vidro em meio LB sólido contendo ampicilina 100 μg/mL, IPTG 0,1 mM e Xgal 40 μg/mL e incubar as placas invertidas em estufa a 37 °C por 14 a 16 horas.

Análise dos resultados

Como o pUC18 é um vetor de clonagem de seleção direta, se o inserto for inserido em sítio localizado no “*polilinker*”, como é o caso do sítio da *EcoRI*, as colônias que ficarem com a coloração natural de *E. coli* (bege) são as recombinantes e as de coloração azul em princípio receberam o plasmídeo cujas extremidades se religaram sem o inserto, e portanto produzem β-galactosidase que converte Xgal (substrato cromogênico) em azul índigo. Pode-se contar as colônias e estimar a eficiência de transformação, o número de clones recombinantes, bem como a porcentagem dos clones recombinantes da biblioteca.

REFERÊNCIAS

1. Watson JD, Crick FHC. A Structure for Deoxyribose Nucleic Acid. *Nature*. 1953;171:737-8.
2. Jackson DA, Symons RH, Berg P. Biochemical method for inserting new genetic information into DNA of simian virus 40. *Proc Natl Acad Sci USA*. 1972;69:2904-9.
3. Cohen SN, Chang ACY, Boyer HW, Helling RB. Construction of biologically functional bacterial plasmids in vitro. *Proc Natl Acad Sci USA*. 1973;70:3240-4.
4. Huang CJ, Lin H, Yang X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol*. 2012;39:383-99.
5. Khorana HG. Total synthesis of a gene. *Science*. 1979;4381:614-25.
6. Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich H. A. Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA polymerase. *Science*. 1988;239:487-91.
7. Liu LF, Liu CC, Alberts BM. Type II DNA topoisomerases: enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break. *Cell*. 1980;19:697-707.
8. Shuman S. Recombination Mediated by Vaccinia Virus DNA Topoisomerase I in *Escherichia coli* is Sequence Specific. *Proc Natl Acad Sci USA*. 1991;88:10104-8.
9. Bollum FJ. Oligodeoxyribonucleotide Primers for Calf Thymus Polymerase. *The Journal of Biological Chemistry*. 1960;235:18-20.
10. Motea EA, Berdis, AJ. Terminal deoxynucleotidyl transferase: the story of a misguided DNA polymerase. *Biochimica et Biophysica Acta*. 2010;1804:1151-66.
11. Krayevskya AA, Victorova LS, Arzumanova AA, Jaskoa MV. Terminal deoxynucleotidyl transferase: catalysis of DNA (oligodeoxynucleotide) phosphorylation. *Pharmacology & Therapeutics*. 2000;85:165-73.
12. Nagy A. Cre recombinase: the universal reagent for genome tailoring. *Genes J. Genet Dev*. 2000;26:99-109.
13. Bolivar F, Rodriguez RL, Greene PJ, Betlach MC, Heyneker HL, Boyer HW, Crosa JH, Falkow, S. Construction and characterization of new cloning vehicles II. A multipurpose cloning system. *Gene*. 1977;2:95-113.
14. Vieira J, Messing J. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene*. 1982;19:259-68.
15. Holton TA, Graham MWA. Simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucleic Acids Research*. 1991;19:1156.
16. Bernard P, Gabant P, Bahassi EM, Couturier M. Positive Selection Vectors Using the F Plasmid ccdB Killer Gene. *Gene*. 1994;148:71-4.
17. Chauthaiwale VM, Therwath A, Deshpande VV. Bacteriophage Lambda as a Cloning Vector. *Microbiological Reviews*. 1992;56:577-91.

18. Messing J. M13 cloning vehicles. Their contribution to DNA sequencing. *Methods Mol Biol* 1993;23:9-22.
19. Collins J, Hohn B. Cosmids: A type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage λ heads. *Proc Natl Acad Sci USA*.1978;9: 4242-46.
20. De Boer HA, Comstock LJ, Vasser M. The tac promoter: a functional hybrid derived from trp and lac promoters. *Proc Natl Acad Sci USA*. 1983;80:21-5.
21. Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol Biol* 1986;189:113-30.
22. Terpe K. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology Biotechnology*. 2006;72:211-22.
23. Baneyx F, Mujacic M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology*. 2004;22:1399-1408.
24. Pan SH, Malcolm BA. Reduced background expression and improved plasmid stability with pET vectors in BL21 (DE3). *Biotechniques*. 2000;29:1234-38.
25. Westers L, Wester H, Quax WJ. *Bacillus subtilis* as a cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochimica et Biophysica Acta*. 2004;1694:299-310.
26. De Pourq K, De Schutter K, Callewaert N. Engineering of glycosylation in yeast and other fungi: current state and perspectives. *Applied Microbiology Biotechnology*. 2010;87:1617-31.
27. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, et al. Life 600 genes. *Science*. 1996;5287:546-67.
28. Scheberg ACG. Elementos de engenharia genética. In: Borzani W, Schmidell W, Lima UA, Aquarone E, editores. *Biotecnologia Industrial*. São Paulo: Blucher; 2001.
29. Cregg JM, Vedvick TS, Raschke WC. Recent Advances in the Expression of Foreign Genes in *Pichia pastoris*. *Nature Biotechnology*. 1993;11:905-10.
30. Cereghino GP L, Cregg, JM. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiology Reviews*. 2000;24:45-66.
31. Torres FAG, Moraes LMP. Proteínas recombinantes produzidas em leveduras. *Biotecnologia Ciência & Desenvolvimento*. 2001;23:20-2.
32. Nevalainen HKM, Téo VSJ, Bergquist PL. Heterologous proteins expression in filamentous fungi. *Trends in Biotechnology*. 2005;23:468-74.
33. Astolfi-Filho, S, Lima BD, Thiemann JE, Souza HRT, Vilela L, inventores. Universidade de Brasília, Biom S.A. Vector for expression of heterologous protein and methods for extracting recombinant protein and for purifying isolated recombinant insulin. United States patent US 6.068.993, 1997.
34. Lima BD. A produção de insulina humana por engenharia genética. *Biotecnologia Ciência & Desenvolvimento*. 2001;23:28-31.

35. Walsh G. Biopharmaceutical benchmarks. *Nature Biotechnology*. 2010;28:917-24.
36. Aviv H, Leder P. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc Natl Acad Sci USA*. 1972;69(6):1408-12.
37. Oswald T, Avery MD, Colin M, Macleod MD, Maclyn McCarty MD. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *Journal of Experimental Medicine*. 1944;79(2):137-58.
38. Birnboim HC, Doly J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res*. 1979;7(6):1513-23.
39. Birnboim HC. A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods in Enzymology*. 1983;100:243-55.