



Centro Universitário de Brasília

Faculdade de Ciências da Saúde

MECANISMOS DE FUNCIONAMENTO
DA EXPRESSÃO GÊNICA EM PROCARIOTOS DE
INTERESSE NA BIOTECNOLOGIA

Rafaella Christina Lima da Costa Almeida

Brasília - 2003



Centro Universitário de Brasília
Faculdade de Ciências da Saúde
Licenciatura em Ciências Biológicas

MECANISMOS DE FUNCIONAMENTO DA EXPRESSÃO GÊNICA EM PROCARIOTOS DE INTERESSE NA BIOTECNOLOGIA

Rafaella Christina Lima da Costa Almeida

Monografia apresentada à Faculdade de
Ciências da Saúde do Centro Universitário de Brasília como parte
dos requisitos para a obtenção do grau de Licenciado em Biologia.

Orientação: Prof. Paulo Queiroz (FACS – UniCEUB)
Prof. Cláudio Henrique Cerri e Silva (FACS - UniCEUB)

Brasília - 2003

Dedico esse trabalho exclusivamente a Deus,
o verdadeiro autor da vida e da diferenciação celular.

Pelo seu amor, cuidado e proteção para com
a minha pessoa, me ensinou a amar e receber amor.

Não tenho palavras para agradecer Sua bondade.

AGRADECIMENTOS

O meu agradecimento especial a minha mãe pelo apoio e paciência dedicados a mim ao longo deste curso. O meu empenho em tentar recompensa-la é inevitável.

Agradeço a minha grande família, e principalmente aos meus avós que mesmo indiretamente me proporcionaram alegria e esforço para concluir um sonho. Em especial as importantes orações da minha prima Adriana, que foram fundamentais para que obtivesse uma nova chance para vencer.

A toda minha liderança espiritual, Bispos Robson e Lúcia Rodovalho, Pastores Geraldo e Denise Alcântara, Lúcia Helena Damasceno e Roberta Campos pelo incentivo para me tornar uma líder.

Ao bom humor e orientação do Prof. Dr. Paulo Roberto Queiroz pelo atendimento dado e pela nova visão da Biologia Molecular.

As colegas de curso Andréia e Denise, pelo empréstimo de livros necessários para realização desse trabalho.

A professora Tânia Andrade pela importante colaboração com o empréstimo de livro e da tese de doutorado da professora Maria Beatriz Nunes, material fundamental para obtenção de dados desse trabalho.

Aos professores Marcelo Ximenes, Cláudio Cerri e Adrienne Fernandes por terem despertado o valioso interesse pelo curso de Biologia.

RESUMO

Os diversos mecanismos funcionais utilizados pelos procariotos para sua expressão gênica têm sido objeto de estudo importante na busca da otimização dos problemas relacionados aos processos de obtenção de proteínas heterólogas. É importante destacar que esse processo inteiro deriva de pesquisas desenvolvidas desde a época de Mendel (1881) com a descoberta da hereditariedade, passando por Watson e Crick (1953) que desenvolveram o traçado da molécula de DNA e sua função que deram origem ao dogma central da biologia molecular. A trajetória unidirecional da informação que caracteriza o dogma se refere à informação contida no DNA em forma de gene que será transcrito em uma molécula única de mRNA e esta será traduzida em proteína. Essas informações estão presentes em forma de código genético que por sua vez é universal, ou seja, é idêntico para todos os seres incluindo os procariotos. Pesquisas mostraram importante viabilidade desses seres nas pesquisas relacionadas à engenharia genética. O avanço da biotecnologia e a disposição de grandes informações tornaram as bactérias mais úteis do que se imaginava. Devido a sua grande capacidade de regular a expressão de seus genes através de operons, facilidade de manutenção, facilidade de purificação, amplificação da produção de proteínas de interesse, entre outras, esses microrganismos tornaram-se importantes agentes no que se refere à recombinação de DNA. A utilização dos procariotos como células hospedeiras de genes exógenos, associado ao que muito se conhece da sua genética tem revolucionado a biologia molecular. Isso tem levando ao futuro grandes perspectivas para obtenção de forma otimizada de importantes proteínas para o homem como a insulina, hormônio de crescimento, interferon, antígenos, etc.

Palavras-chave: Procariotos, regulação gênica, biotecnologia, proteína.

SUMÁRIO

Introdução.....	1
1. Dogma Central da Biologia Molecular.....	3
1.1. Transcrição	3
1.2. A regulação Ocorre Majoritariamente na Transcrição	5
1.3. Tradução	6
2. Regulação Gênica.....	7
2.1. Mecanismos Regulatórios	8
▪ Operon da Lactose.....	9
▪ Operon Triptofano	10
▪ Operon Arabinose.....	10
▪ Operon Histidina	11
2.1.2. Regulação na Tradução	12
2.2. Transdução de Sinais.....	12
3. Biotecnologia.....	13
3.1. Transgenia	14
3.2. Vetores.....	15
4. Aplicações e Importância da Utilização de Genes Expressos em Bactérias	16
4.1. Insulina Humana.....	17
4.2. Antígenos.....	18
4.3. Outras pesquisas	19
Conclusão	20
Referências Bibliográficas.....	21

INTRODUÇÃO

Após muitos estudos iniciados com a descoberta da hereditariedade por Mendel, chegou-se à informação da existência dos genes e que eles constituíam a estrutura química dos cromossomos, funcionando como portadores de informações. A chamada genética molecular surgiu no início da década de 40, para mostrar que a maioria dos genes codifica algum tipo de proteína. Oswald Avery, C. M. McLeod e M. McCarthy em 1944 demonstraram que o DNA de uma bactéria poderia ser alterado pela adição de um DNA exógeno - transgênia (GRIFFITHS *et al* 2001).

Em 1953 uma descoberta importante e que serve como referencial até os dias de hoje, foi a proposição de James Watson e Francis Crick, do formato estrutural da dupla hélice de DNA. Além disso, eles provaram o papel biológico do DNA e sua capacidade de replicação onde se pudesse carregar suas vastas informações de célula para célula e de geração para geração. Essas informações são passadas através de códigos representados por nucleotídeos onde são conhecidos quatro tipos. Notou-se então, que a dupla hélice era interligada por esses nucleotídeos através de pontes de hidrogênio e que uma fita complementa a outra. O fluxo unidirecional de informação do gene para a molécula de mRNA e desta p/ a proteína define o que é hoje chamado de dogma central da biologia molecular. Essa descoberta foi um marco para ciência o que rendeu aos pesquisadores em 1962 o Prêmio Nobel (ALBERTS *et al* 1997).

Um modelo foi traçado em 1961 por François Jacob e Jacques Monod, para demonstrar que a transcrição dos genes pode ser regulada por meio de proteínas que se ligam a regiões adjacentes ao sítio de síntese de RNA. Essa pesquisa foi realizada em genes bacterianos para demonstrar que eles podem ser ativados ou desativados na produção de transcritos de RNA. Nathans e Smith em 1970 descobriram as enzimas de restrição que cortam o DNA em seqüências-alvo, favorecendo a produção de DNA recombinante (GRIFFITHS *et al* 2001).

As descobertas a respeito dos genes, dentre elas que as informações contidas nos genes podem ou não expressar proteínas tem levado os pesquisadores a investigar e caracterizar as funções dos diversos genes. Contudo, o advento das técnicas modernas de engenharia genética tem permitido isolar genes heterólogos de uma espécie e inseri-los em

outra espécie visando expressar de maneira controlada suas proteínas. Células procarióticas, especialmente a bactéria *Escherichia coli*, são muito usadas como hospedeiras desses genes, pois possui genoma conhecido e tem fácil manipulação por apresentar alta eficiência de transformação.

Segundo pesquisas feitas por Souza (1995), para que haja eficiência na taxa de expressão de genes específicos procura-se obter um promotor forte e um eficiente sítio de ligação do mRNA com o ribossomo. Isso se refere a um completo e bem sucedido processo de transcrição e tradução. Visando isso, na tentativa de minimizar os problemas derivados do processo de expressão tem sido construídos vetores que possibilitam regular a síntese de uma proteína específica a ser estudada. Muitas possibilidades favorecem e facilitam a construção de vetores, a principal é a utilização de unidades regulatórias chamadas operons, que são seqüências conhecidas compostas por genes estruturais e elementos que controlam a expressão dos mesmos.

Muito tem sido feito por parte da ciência para encontrar possibilidades mais vantajosas referentes à utilização de genes heterólogos, ou seja, a introdução de genes que uma espécie em uma outra espécie. O objetivo desse trabalho é relacionar os mecanismos de funcionamento da expressão gênica visando aplicações em biotecnologia utilizando genes eucarióticos.

1. O Dogma Central da Biologia Molecular:

Foi descrito na década de 50 por Francis Crick, na tentativa de relacionar o DNA, RNA e as proteínas. O DNA é uma longa fita dupla, composta por nucleotídeos que são representados num código de quatro letras e são interligadas por pontes de hidrogênio. Essas letras nada mais são que abreviações das bases nitrogenadas, purinas (Adeninas e Citosinas) e pirimidinas (Timinas e Guaninas). As informações genéticas estão impressas no DNA, em forma de genes e vão expressar as características na maioria dos seres vivos. O DNA que nos eucariotos está presente no núcleo, nos procariotos está disperso no citoplasma, é o grande guardião do material genético na maioria dos seres vivos (AZEVEDO 1998).

Essa molécula mantém todo o seu patrimônio durante a replicação garantido a passagem das informações genéticas para as células filhas, levando de herança as características das células mãe. Essas informações são imprescindíveis para que a síntese protéica ocorra. Lembrando que a síntese de proteínas é feita nos ribossomos no citoplasma da célula, e a molécula de mRNA atua como intermediária nesse processo. Nos seres procariontes todo o processo ocorre no citoplasma celular. Contudo o dogma central da biologia resume-se na capacidade de replicação do DNA, e sua transcrição em uma fita de RNA que originará na tradução de uma proteína. Todos os passos apresentados desde a replicação do DNA até a síntese protéica representaram o dogma central da biologia molecular (AZEVEDO 1998, GRIFFITHS *et al* 2001).

1.1. Transcrição:

O RNA é a molécula intermediária que será traduzida em uma proteína. O processo de transcrição resulta na produção de uma molécula de RNA mensageiro (mRNA), a partir da cópia de uma seqüência de nucleotídeos de uma das fitas do duplex de DNA, chamada fita molde. A fita molde de DNA gera apenas uma fita de RNA sendo, que neste a timina é substituída pela uracila. Os RNAs podem ser classificados em funcionais, de informação e estruturais, todos são produtos da transcrição.

Existem algumas diferenças na síntese e na estrutura dos mRNA de procariotos e de eucariotos, mesmo ambos exercendo as mesmas funções. Os processos de transcrição e de tradução procarióticos ocorrem simultaneamente na mesma região da célula. Os ribossomos

ligam-se ao mRNA antes da transcrição ser completada. Já nos eucariotos, a síntese e maturação do mRNA ocorrem no núcleo, e as informações são transportadas ao citoplasma, para ser transcrita em proteína (LEWIN 2001). O mRNA dos procariotos praticamente não sofre modificações, devido ao fato de não apresentarem íntrons (regiões não codantes). No entanto, ocorrem interrupções ao longo da leitura no DNA que origina o mRNA, pois diferentes segmentos em tamanho são produzidos. Existem cerca de 4.000 tipos de mRNA diferentes que vão dar origem a proteínas (AZEVEDO 1998).

A transcrição inicia-se quando a enzima RNA polimerase encontra o promotor, para iniciar a síntese do mRNA. Promotores são seqüências específicas no DNA que fazem parte da região reguladora presente numa região adjacente ao gene. Essa informação é extremamente importante para se compreender os mecanismos de expressão gênica em procariotos (GRIFFITHS *et al* 2001). Além do promotor, muitos dos genes bacterianos contêm seqüências regulatórias de DNA que são necessárias para ligar e desligar o gene. O contato da enzima com o promotor resulta na abertura ou separação da dupla hélice deixando as bases de uma das fitas para que haja síntese do RNA. A leitura da fita de DNA ocorre sempre no sentido 3'-5'. Estudos mostram que as bactérias possuem seqüências comuns de bases de região promotora para uma série de genes, as seqüências de consenso. As duas seqüências antes do primeiro nucleotídeo a ser lido, região conhecida como TATA Box, cujas bases são TATAAT e está localizada à 10 pb do início da transcrição, e a região 35 pb com seqüência TTGACA (CAMPBELL 2001).

Em procariotos, a RNA polimerase move-se ao longo da fita molde numa velocidade de aproximadamente 40 pb/s a 37 °C sobre um gene promovendo a sua leitura até que encontre uma seqüência do terminador. A iniciação da leitura do DNA requer a abertura de uma bolha no segmento a ser transcrito, a fita é formada pelo pareamento complementar de bases sintetizadas em outro lugar na célula que vão se alinhando respectivamente: Adenina com Uracila, Timina com Adenina, Citosina com Guanina e Guanina com Citosina (STRYER 1996, ALBERTS *et al* 1997). Esse caminho que se estende do promotor até o terminador formado de um produto direto transcrito, define uma unidade transcricional primária, que em procariotos é rapidamente degradada ou clivada para formar os outros RNAs. Uma unidade transcricional pode conter um ou mais genes (LEWIN 2001).

Enfim, quando a enzima encontra a seqüência terminadora, ela para de adicionar os ribonucleotídeos complementares à cadeia de mRNA sintetizado, liberando o produto completo e desprendendo-se do molde de DNA. Contudo, para que haja interrupção na transcrição, o processo de terminação, tem que haver o reconhecimento de sinais no transcrito pela RNA polimerase ou por fatores auxiliares (LEWIN 2001). As pontes de hidrogênio que mantinham a fita molde de DNA ligada à fita transcrita de RNA são quebradas e a fita mensageira é liberada.

1.2. A regulação Ocorre Majoritariamente na Transcrição

Vários segmentos de DNA, os genes estruturais, podem estar envolvidos na produção de um mesmo polipeptídeo. Estes são sintetizados e transcritos para a mesma fita de mRNA, porém existem mecanismos que acionam ou silenciam os grupos de genes. Este tipo de mRNA que codifica vários polipeptídios é encontrado apenas em procariotos e em vírus de plantas e animais, são chamados de *operons*. Os mecanismos de controle gênico atuam prioritariamente na região de iniciação, e na região de terminação transcricional mediados por proteínas que irão bloquear ou ativar a expressão do gene.

Como já foi dito, a região de iniciação da transcrição contém, além de promotores, sítios de seqüências que regulam no DNA a sua leitura. As seqüências necessárias que vão ligar ou desligar a expressão do gene dependerão de uma série de fatores, incluindo o tipo celular, idade, circunvizinhança e sinais extracelulares. Contudo, para manter seu sistema metabólico ativo é necessário que mecanismos de controle respondam a reações de estresse oriundo do meio ambiente. O mRNA bacteriano é geralmente instável traduzindo em proteínas por apenas alguns minutos (ALBERTS *et al* 1999).

A maneira utilizada pelas células para exprimir seus genes e formar suas derivadas proteínas tem uma relação direta com o tipo celular, e suas diferenciações funcionais e estruturais, tornando-as irreversíveis. Os agentes responsáveis pela ativação de genes específicos são as proteínas regulatórias, elas vão coordenar a estrutura da fita de mRNA. As modificações percebidas nas taxas de expressão gênica em bactérias serão discutidas mais adiante, quando será discutido o controle da síntese de mRNA.

1.3. Tradução

O mRNA que foi sintetizado no núcleo da célula e agora foi transportado para o citoplasma será traduzido em proteína. Os eventos dessa síntese ocorrem em um sítio multienzimático chamado ribossomos, onde efetivamente ocorre a tradução. Eles são compostos por uma subunidade maior que possui três sítios (A, P e E) para a ligação do tRNA, e outra subunidade menor com um sítio de ligação do mRNA. Possuem ainda rRNA que mantém contato com muitas proteínas, conhecidas como proteínas *r*.

Segundo Lewin (2001) logo que a transcrição se inicia nas bactérias, os ribossomos já vão se associando para iniciar a tradução. No sentido 5' os ribossomos (polirribossomos) vão se movendo ao longo da fita de mRNA, lendo sua seqüência de três em três nucleotídeos (códon), que darão origem aos aminoácidos existentes na proteína sintetizada (GRIFFITHS *et al* 2001).

Além do mRNA e ribossomos são necessárias muitas proteínas e a molécula central da síntese, o tRNA. Ele funciona como intermediador que decifra as bases do mRNA, os códon, traduzindo-as em vinte aminoácidos que constituirão as proteínas com seus respectivos anticódon. Uma molécula de tRNA associada a um dos vinte aminoácidos é chamada de tRNA-aminoacil. Ocorre uma reação que vai formar uma ligação peptídica entre a carboxila da cadeia polipeptídica em crescimento e um grupo amino de um aminoácido.

A síntese ocorre em três etapas: início, alongamento e término. Para que o ribossomo inicie a tradução ele tem que encontrar uma seqüência sinal, o código de iniciação AUG, e então se ligar ao DNA. Essas seqüências estão numa região não codificantes, servindo somente para ligação do ribossomo. A iniciação da tradução dá-se quando a subunidade menor do ribossomo encontra a seqüência sinal ou o códon de iniciação do mRNA, no sítio A. O anticódon do tRNA liga-se ao códon de iniciação de forma antiparalela da trajetória do ribossomo pela fita de mRNA, ou seja, o ribossomo move-se no sentido 5'-3' e o tRNA sintetiza de forma 3'-5'. Em seguida a subunidade maior une-se a menor e a tradução é continuada, o aminoácido formado caminha para o sítio P e o sítio A é liberado para chegada de outro tRNA-aminoacil. Essa fase é conhecida como alongamento, onde os dois sítios estão preenchidos e os aminoácidos formarão ligações peptídicas através da enzima peptil transferase. O tRNA deixa o ribossomo pelo

sítio E e poderá ser ativado novamente numa região *pool* de tRNA do citoplasma. Um terceiro tRNA-aminoacil posiciona-se no sítio A e o processo é repetido (STRYER 1996, MADIGAN et al 2000).

A fase de terminação ocorre quando o ribossomo encontra os três códons de terminação UAG, UAA ou UGA. O tRNA não os reconhece, ou seja, não tem anticódon para eles. Fatores de liberação ligam-se ao códon terminal no sítio A e a cadeia polipeptídica inteira é liberada.

2. Regulação Gênica

Os organismos têm a capacidade de selecionar seus genes, para expressar suas características fenotípicas. Num genoma completo apenas uma pequena parte de suas proteínas são expressas e os níveis quantitativos são diferentes. Até mesmo os procariotos com suas reduzidas fontes de informação tem a capacidade de ativar e desativar a expressão dos seus genes para produzir sua proteína necessária dependendo de suas fontes de alimento (ALBERTS *et al* 1999).

Embora os maiores índices regulatórios sejam no nível transcricional, o controle também poderá ser feito no DNA, no momento da ativação ou inativação do gene a ser expresso e na tradução do RNA mensageiro. A expressão gênica pode ser controlada num processo de três estágios: transcrição, alongamento e tradução. Como já dito, o controle da transcrição dá-se mais freqüentemente na fase inicial podendo também ser controlada na fase terminal quando se encontra um gene terminador. Em bactérias não ocorre regulação na fase de processamento de mRNA. Em alguns casos, na tradução a expressão gênica também pode ser controlada similarmente na transcrição (MADIGAN *et al* 2000, LEWIN 2001).

As bactérias possuem várias enzimas biossintéticas responsáveis por suas transformações químicas metabólicas. Essas reações metabólicas podem ser controladas em resposta ao meio em que vivem, por exemplo, o bloqueio da via metabólica levando a obtenção do produto final diretamente do meio extracelular. A possibilidade de escolha entre vários substratos que fornecerão nutrientes como carbono, fósforo, nitrogênio e enxofre de forma que seja mais proveitosa pela célula de um ponto de vista de rendimento e

de custo energético. Outros mecanismos utilizados pelas bactérias estão relacionados a reações químicas e físicas. A resistência dessas células a fatores como falta de oxigênio, lesões genéticas, choque térmico ou osmótico e dessecação, estão relacionadas sistemas de controle (COSTA 1986).

2.1 Mecanismos Regulatórios

Não basta somente afirmar que o mecanismo de síntese de proteínas pode esclarecer como a informação contida no material genético pode determinar uma seqüência exata de aminoácidos que caracteriza um organismo. O controle da expressão gênica, de fato, tem fundamental importância na determinação fenotípica de um organismo. As bactérias apesar de serem seres unicelulares, também sofrem influência regulatória que determinam quais são os genes que serão expressos. Como já se sabe, isto está diretamente associado ao ambiente que ela está e sua disponibilidade nutricional.

Logo após a região do promotor no DNA existe um sítio de ligação de proteínas regulatórias chamado operador. Esses sítios sofrem a ação de proteínas reguladoras indutoras ou repressoras que são peças-chave para que um gene possa ser transcrito. Para alguns genes impedir a ligação de uma proteína repressora ao seu do operon é necessária a ligação do repressor a uma molécula indutora, esse é um pré-requisito para que ele seja transcrito. Esse processo é chamado regulação negativa que também é reconhecido quando a proteína repressora liga-se ao operador impedindo a transcrição. Na regulação positiva a ausência de uma proteína indutora ligada ao operador impede a transcrição ou quando ligada a um operador permite a transcrição. Esse é o principal motivo pelo qual conclui-se que o controle da expressão gênica é feito ao nível da transcrição (GRIFFITHS *et al* 2001).

Genes estruturais que codificam proteínas reguladoras, quando sofrem mutações são impedidos de sintetizar as proteínas reguladoras. A ausência dessas proteínas afetaria a transcrição do seu mRNA correspondente apenas na região atuante ou em outras regiões que expressariam seus mesmos alelos. Esses processos são denominados cis-atuantes, onde a mutação afeta somente a região de DNA contígua a ela, ou trans-atuantes quando a ausência dessa proteína impede a expressão de ambos os alelos. Essa descoberta foi feita por Jacob e Monod em 1961 (LEWIN 2001).

▪ Operon da Lactose

A principal fonte nutricional das bactérias é a glicose. As bactérias podem retirar glicose de um meio contendo lactose somente na presença da enzima codificada pelo gene *lacZ*, a β -galactosidase. Um meio de cultura contendo lactose, induz um grande aumento na quantidade de β -galactosidase, estimulando a síntese de novas moléculas da enzima, ativando o crescimento bacteriano. Outras duas enzimas aumentam em proporções diretas da galactosidase e são necessárias para utilização da glicose pelas bactérias. A β -galactosídeo permease, enzima codificada pelo gene *lacY*, que vai favorecer a entrada da lactose pela membrana da bactéria. E a tiogalactosídeo transacelilase codificada pelo gene *lac A*, que possui um papel ainda incerto e que não afeta o metabolismo da lactose. Portanto o operon da Lactose é um mecanismo indutor da síntese das três enzimas citadas (STRYER 1996).

Além dos três genes estruturais que codificam as enzimas necessárias para a utilização da lactose esse Operon (segmento de DNA) contém um promotor, um operador e um gene regulador chamado I. O gene I sintetiza uma proteína repressora que na ausência de um indutor bloqueia a ligação da RNA polimerase impedindo a síntese das proteínas Lac, a regulação negativa. Isso acontece quando a bactéria está num meio utilizando glicose diretamente, logo seria desnecessário produzir enzimas para hidrolizar lactose em glicose e galactose. Na presença de lactose, isto é de indutor, os genes podem ser expressos porque o repressor é reconhecido pela lactose e são associados, isso incapacita o repressor de se fixar ao operador, regulação positiva. Com isso a RNA polimerase pode fixar-se ao promotor não havendo mais impedimento da expressão dos genes *lacY*, *lacA* e *lacZ* (AZEVEDO 1998, ÉTIENNE 2003).

Um terceiro sistema de controle adicional dentro do sistema da lactose é proposto quando ocorre a presença de glicose e lactose. Entretanto não ocorre a produção da β -galactocidase até que toda a glicose seja consumida. O mecanismo de degradação da glicose que impede a construção do sistema do Operon lac é chamado repressão catabólica. Isso está relacionado com um constituinte celular, o monofosfato de adenosina cíclico (AMPc). Quando os índices de concentração de glicose são altos, um dos seus produtos impede a síntese de adenilato ciclase, o nível de AMPc é baixo porque não é possível converter ATP em AMPc. A alta concentração do AMPc ativa o Operon lac, e quando

ligada ao ativador CAP formando um complexo para permitir a ativação do operon. CAP é uma proteína ativadora da degradação, expressa em altos níveis de AMPc que quando ligada ao promotor sintetiza mRNA_(DARNELL *et al* 1995).

O operon da arginina envolve oito enzimas, desde o ácido glutâmico até a arginina que representa o co-repressor do sistema. Esse sistema é chamado de regulon, pois existe um regulador e vários operadores (AZEVEDO 1998).

▪ **Operon Triptofano**

O triptofano (Trp) é um aminoácido essencial para o homem e deve ser adquirido na alimentação. O Trp é sintetizado por uma série de reações, e cada etapa requer uma enzima específica. Quando não houver mais Trp, a síntese dessas enzimas será ativada e conseqüentemente, a quantidade suficiente de Trp reprime a produção dessas enzimas. O conjunto de genes que codificam as enzimas necessárias para a síntese do Trp constitui o operon do Trp (AZEVEDO 1998).

O operon do Trp também é controlado por uma enzima repressora e por um ponto de terminação interno, o atenuador, e estão localizados entre o promotor e pelo primeiro gene estrutural do operon. O gene regulador sintetiza a enzima repressora (apo-repressora), que não vai ligar-se ao operador, isso permite a ligação da RNA polimerase iniciando a transcrição do aminoácido. Quando o aminoácido está disponível no meio de cultura sua síntese é reprimida. Nesse caso atua como um co-repressor combinando com o repressor inativo ligando-se ao operador impedindo a transcrição. O metabólito final impediu sua própria síntese, até que todo o Trp seja consumido pela célula (ÉTIENNE 2003). Para Stryer (1996), os operons da fenilalanina e da treonina também possuem atenuadores.

▪ **Operon Arabinose**

Os genes codificadores das enzimas desse operon são *araA*, *araB* e *araD*. Possui ainda uma região iniciadora formada por um sítio com dois operadores *araO₁* e *araO₂*, um ponto de ligação para CAP e outro regulador *araI*, um promotor, e uma outra região de controle *araC* para síntese de uma proteína *C* reguladora.

Nesse operon são necessários dois sinais para uma transcrição eficiente, o complexo AMPc-CAP e a arabinose ligada a proteína C. A proteína C funciona como um regulador negativo que na ausência da arabinose (monossacarídeo), muda de conformação reprimindo o operon ara. Outra característica desse operon é que o seu mRNA só é formado quando os níveis de cAMP-CAP é baixo. Quando a proteína C é abundante e o AMPc-CAP não, a transcrição do gene que sintetiza a proteína C cessa pois a proteína liga-se ao centro controlador do operon. A ligação de uma segunda molécula C ao centro controlador bloqueia a síntese dos genes estruturais que sintetizam as proteínas cinase, isomerase e epimerase. Isso faz com que se forme uma alça de DNA bloqueadora da RNA polimerase que vai se ligar a uma região adjacente do promotor (STRYER 1996, GRIFFITHS *et al* 2001).

A presença de dois fatores reguladores positivos, AMPc-CAP abundante e a arabinose está ligada a proteína C, torna a RNA polimerase capaz de se ligar ao promotor para os genes e transcrevê-los. Logo, o mRNA para a proteína c não é formado.

Assim, no operon da arabinose ocorre um controle duplo onde a proteína regula sua própria síntese reprimindo a transcrição. Numa outra situação a ligação de uma molécula alostérica a uma proteína pode mudar sua conformação de inibidora para ativadora da transcrição. As alterações induzidas por moléculas alostéricas são reversíveis, também diferem em suas habilidades em se ligar a um sítio-alvo, o operador. O sistema da arabinose é extremamente sensível às variações nos níveis metabólicos (STRYER 1996).

▪ **Operon Histidina**

A presença ou ausência de histidina no meio não impede a atividade do operon, mas interfere o processo. Se houver histidina no meio, a transcrição do mRNA é parada na etapa que contém os códons para a síntese de poli-histidina, alguns nucleotídeos nessa região formam alças. Quando houver uma série de tRNA carregados com histidina o ribossomo que traduz o peptídeo líder fica perto da RNA polimerase impedindo a formação de uma das alças e permitindo a formação de outras duas. Uma dessas duas alças é um atenuador e quando ela não é formada a síntese ocorre formando as enzimas necessárias para síntese de histidina. Esse sistema não precisa de um gene regulador e o tRNA funciona como co-

repressor (AZEVEDO 1998). O operon da isoleucina-valina é muito semelhante ao da histidina.

2.1.2. Regulação na Tradução

Para evitar o desperdício de energia, as bactérias sintetizam somente seus ribossomos de forma que supra suas necessidades. Considerando que um ribossomo possui 52 r-proteínas e tem rRNAs 16S, 23S e 5S, observa-se que sua fabricação consome muita energia. As proteínas ribossomais em excesso bloqueiam seus próprios rRNAs e de algumas proteínas situadas no mesmo operon. Além disso, as r-proteínas podem se fixar também sobre o mRNA bloqueando sua tradução em ribossomo. A afinidade da r-proteína ao rRNA é maior devido a semelhança da seqüência nucleotídica existente em relação ao mRNA. Em bactérias não ocorre controle na fase de alongamento (ÉTIENNE 2003).

2.2. Transdução de Sinais

Sabe-se que as bactérias regulam seu metabolismo principalmente em resposta a uma extensa variedade de flutuações ambientais tais como: mudanças de temperatura, mudanças de pH, disponibilidade de oxigênio, disponibilidade de nutrientes e até mudanças do número de células presentes no meio. Contudo, as bactérias necessitam de um mecanismo para receber os sinais oriundos do ambiente para transmitir e regular seu objetivo específico. Toda a maquinaria acionada refere-se a transdução de sinais.

Os mecanismos utilizados para resposta regulatória do controle transcricional dependem do sistema descrito. A exemplo do processo que envolve uma proteína quinase que funciona como um receptor do sinal extracelular, nesse caso a histidina, e o transmite para o interior da célula. As quinases fosforilam as cadeias formadas desde o sinal na membrana até o interior da célula. O grupo fosforilado é transferido para uma proteína reguladora na presença de ATP ligando-se ao DNA reprimindo a transcrição. Outros exemplos envolvem enzimas que retiram a fosfatase da proteína repressora não é ativada e a transcrição é iniciada normalmente. Diversos sistemas a exemplo do descrito acima envolvem dois componentes uma proteína de membrana e uma proteína repressora da transcrição. Entretanto, eventualmente alguns sistemas podem envolver mais componentes com várias etapas (MADIGAN *et al* 2000).

Daí a relação entre as mudanças que acontecem no meio extracelular influenciam o metabolismo bacteriano. Todas as informações são transmitidas para o interior da célula alterando o processo da expressão genética do microrganismo. Medidas induzidas artificialmente com objetivo de regular a expressão tem sido bastante utilizadas para obtenção de proteínas heterólogas.

3. Biotecnologia

O termo biotecnologia diz respeito à tecnologia do DNA recombinante, ou seja, a capacidade de isolar, manipular e reintroduzir genes nos seres vivos. A técnica consiste na interferência controlada e intencional onde o cientista pode inserir ou retirar genes de interesse em qualquer organismo. Os avanços da ciência permitiram a modernização da biotecnologia de forma que a obtenção dos resultados tornou-se mais rápida precisa e eficiente.

A nova tecnologia abriu espaço para realização de infinitas possibilidades de compartilhamento de genes entre organismos. Mesmo considerando o código genético uma ferramenta universal, para se obter sucesso nas pesquisas é necessário que haja compatibilidade entre células hospedeiras e a mensagem exógena. Diversas técnicas foram desenvolvidas com intuito de facilitar a obtenção de proteínas de interesse de forma orientada. A mais importante e mais utilizada é a técnica de PCR (“reação em cadeia da polimerase”).

O PCR é utilizado para adquirir grande quantidade de seqüências específicas da molécula de DNA do organismo desejado. Essa amplificação é possível sem a necessidade de separa-lo do restante do DNA da amostra desejada. A reação necessita de um par de primers com cerca de 20-30 nucleotídeos, que são seqüências flaqueadoras, ou seja que indicam a localização da região do DNA desejada. Necessita também das bases desoxirribonucleicas que vão parear complementarmente ao DNA desejado e de uma DNA polimerase termoestável. O procedimento consiste em três etapas de aquecimento e resfriamento da molécula que permitirá a separação seguida da hibridação dos primers ao filamento da molécula e da síntese do DNA onde o alongamento por uma DNA polimerase termoestável derivada de uma bactéria termófila. Todos os novos filamentos de DNA são

servem como moldes nos ciclos sucessivos. A amplificação é de cerca de um milhão de vezes após 20 ciclos e um bilhão de vezes após 30 ciclos que podem ser feitos em menos de uma hora (STRYER 1996, CAMPBELL 2001).

Muitos trabalhos já foram realizados utilizando as técnicas de engenharia genética nas áreas de medicina, agricultura, farmaco-química e em vários países como, Argentina, Estados Unidos, França, Brasil entre outros. Contudo, existe um grande interesse por parte destes na produção em larga escala de proteínas de mamíferos em células bacterianas. Sabe-se porém, que as bactérias não conseguem expressar genes interrompidos, ou seja, genes constituídos de íntrons e éxons, pois elas não possuem a maquinaria necessária para remover esses íntrons. Por isso, para que ela expresse a proteína desejada é necessário introduzir um DNA recombinante complementar ao mRNA (STRYER 1996). Segundo Souza (1995), há alguns fatores como maior grau de estabilidade do mRNA derivados das etapas de transcrição e tradução e seleção do vetor adequado também são indispensáveis na produção de proteínas de interesse, pois essas características são determinantes para expressão de genes de importância biotecnológica.

3.1. Transgenia

Refere-se aos organismos geneticamente modificados, onde há utilização de um gene exógeno e uma célula receptora. Um gene selecionado é adquirido através das técnicas de engenharia genética onde será clonado e multiplicado. Esses fragmentos constituem uma biblioteca de DNA. A disponibilidade de uma grande coleção de clones desafia o pesquisador a achar o gene desejado. Uma das mais importantes contribuições da engenharia de DNA é a possibilidade de adquirir grandes quantidades de qualquer proteína de uma célula (ALBERTS *et al* 1997).

A produção de bactérias transgênicas, isto é contendo gene exógeno de interesse, é possível utilizando as técnicas da engenharia genética. A indústria tem utilizado células bacterianas para fabricar proteínas úteis para o homem. Para Souza (1995), este processo apresenta as vantagens de independência da matéria prima que na maioria das vezes é de difícil disponibilidade, obtenção do produto de interesse em grande quantidade, relativa facilidade de purificação além de reduzir os custos de produção. Nesse sentido muitos

trabalhos já foram realizados como a produção de insulina, do hormônio de crescimento, o interferon, antígenos, etc.

Um gene de interesse pode ser adquirido através da elaboração de uma biblioteca de cDNA. Um cDNA é produzido através do mRNA isolado do total de mRNAs da célula e que dará origem à proteína de interesse. Esse processo é possível graças a ação da enzima transcriptase reversa que produz um duplex de DNA através de um mRNA. Esses cDNA são transportados para as células hospedeiras, as bactérias, pelos vetores onde serão expressos de forma controlada (ÉTIENNE 2003).

3.2. Vetores

Os plasmídios são moléculas constantemente utilizadas como vetores de expressão, que são programados para produzir grandes quantidades de mRNAs estáveis. Plasmídios são moléculas de DNA circular que se replicam independentes e à parte do cromossomo bacteriano. Eles são muito eficientes, desde que haja um promotor e um sítio de ligação compatível a RNA polimerase, para efetuar a tradução de uma proteína desejada. A eficiência de vetores compatíveis e estáveis é um ponto fundamental que requer atenção, pois do contrário acarretaria na perda de material clonado e falta de êxito na obtenção de resultados. Em alguns casos o alto nível de expressão pode vir a ser tóxico para a célula hospedeira, causando a recombinação do plasmídio ou sua perda total, além de retardar o crescimento celular (SOUZA 1995).

As bactérias absorvem DNA do meio e o incorporam naturalmente, transformação de Griffith, ou artificialmente no seu citoplasma (AZEVEDO 1998). As técnicas de transformação artificial são utilizadas para produção de DNA recombinante, mas são menos eficientes e precisa de tratamentos especiais como aumento de permeabilidade de membrana e ação de íons Ca^{2+} (COSTA 1986). Como já foi visto, as bactérias sofrem constantes influências do meio em que vivem, por isso a utilização de eficientes e compatíveis vetores é extremamente importante para expressão de genes. Diante disso, os pesquisadores têm construído vetores que possibilitem a expressão de genes de interesse de forma regulável. Podem ser citados os vetores montados com promotores *lac*, *trp*, citados anteriormente, o operon *tac* e o promotor do bacteriófago lambda (λ). Os promotores *lac* e

trp já foram descritos. O promotor *tac* é um híbrido de seqüências de promotores *lac* e *trp* (SOUZA 1995).

Um outro tipo de vetor utilizado para introduzir genes exógenos em bactérias é o bacteriófago λ . É necessário que haja compatibilidade no tamanho da molécula introduzida com a molécula receptora. Essa é a forma mais conveniente de introduzir um gene estranho em bactéria, pois os plasmídios apenas ocasionalmente conseguem adrentar na célula procariótica intacta. O DNA recombinado do fago é incorporado ao cromossomo bacteriano onde serão replicados e expressos em proteínas (LEHNINGER 1984).

4. Aplicação e Importância da Utilização de Genes Eucariotos Expressos em Bactérias

De acordo com LIMA 2001 existem várias maneiras de obtenção de um gene eucariótico para serem expressos em procariotos. A síntese química desses genes eucariotos oferece as seguintes vantagens:

- Fornece a seqüência exata desejada;
- As seqüências codificadoras e não codificadoras podem ser desenhadas para expressão procariótica;
- Sítios de restrição podem ser removidos ou adicionados;
- Os íntrons podem ser retirados;
- Não há necessidade da etapa de isolamento do mRNA ou do DNA genômico;
- Permite a alteração de gene de forma mais simples.

A elaboração de um vetor para construção de sistema expressão genética em procariotos requer seis princípios básicos:

- Uma região necessária para replicação estável e controle do número de cópias;
- Um marcador seletivo com um gene conferindo resistência a um antibiótico para a hospedeira;
- Um promotor para iniciação da transcrição e seu controle;
- Uma região terminadora da transcrição;
- Um sítio de ligação de ribossomos para iniciação da tradução em uma trinca ATG apropriada;

- Uma região de sítios apropriados para enzimas de restrição para utilização nas clonagens dos genes a serem expressos.

Dessa forma, a possibilidade de obtenção de altos níveis de expressão gênica decorrente da otimização de todo o processo, vem abrindo novas perspectivas para produção de proteínas de interesse utilizando as bactérias. Alguns exemplos serão revisados para simples conferência.

4.1. Insulina Humana

A insulina é um hormônio protéico produzido pelas células β as ilhotas pancreáticas. A insulina é produzida sob a forma de um único polipeptídeo grande denominada pré-pró-insulina, que após passar por processo de clivagem no retículo endoplasmático perde a cadeia de aminoácidos que formam o peptídeo sinal originando a pró-insulina. Essa molécula é precursora da insulina, que é transportada para o complexo de Golgi e é clivada novamente dando origem a insulina na forma ativa. A insulina possui duas cadeias polipeptídicas, a cadeia A composta por 21 aminoácidos e a cadeia B com trinta aminoácidos interligados por pontes dissulfídricas (LIMA *et al* 2001).

No caso da insulina a alternativa para otimizar a expressão do gene em *E. coli*, foi construir um gene sintético da proteína pró-insulina, esse gene possui 18,6% de bases substituídas que não alteram sua conformação. Para facilitar a clonagem foi inserida a enzima de restrição *Eco* RI seguiu do códon para metionina (ATG) no início e *Bam* HI no final. Na montagem da cadeia B e C foram sintetizados quatro oligonucleotídeos distintos e na cadeia A dois que foram anelados em pares para a formação da dupla fita de DNA. Esse gene foi seqüenciado para confirmar a sua autenticidade e clonado num vetor de hiper-expressão (LIMA *et al* 2001).

A construção do vetor de expressão do gene da pró-insulina requereu: uma região para controle do número de cópias modelo pUC8, um marcador com um gene seletivo resistente ao antibiótico tetraciclina para a célula hospedeira, um promotor do fago λ para iniciação da transcrição e seu controle, uma região terminadora da transcrição Rho-independente, um sítio de ligação no fago T7 de ribossomos para iniciar a tradução e uma região múltipla de sítios apropriados para as enzimas de restrição. O promotor é regulado por um repressor codificado pelo gene *cI* que após sofrer mutação tornou-se sensível à

temperatura funcional a 28 °C e não funcional a 42 °C. Com esse mecanismo de controle foi possível produzir o hormônio da insulina humana de forma simples e barata numa escala industrial (LIMA *et al* 2001).

4.2. Antígenos

Há algum tempo tem sido testadas bactérias patogênicas pelas técnicas de engenharia genética como veículo para administração de vacinas. Contudo, considerando os riscos apresentados a crianças e indivíduos imunodeficientes começaram a pesquisar a possibilidade de utilização de bactérias não patogênicas na produção de vacina. Algumas características favorecem esse sistema, pois ocorre facilidade de administração (oral), possibilidade de serem administradas em forma de coquetéis, além de também apresentarem baixo custo. Vários antígenos bacterianos e virais já foram produzidos em bactérias lácticas, amplamente utilizadas na indústria agro alimentícia, entre eles a expressão do gene produtor do fragmento C da toxina tetânica e da peritactina um fragmento da coqueluche. Ambos os genes forma sintetizados a partir do promotor *tac*, uma modificação do promotor *lac*, pois apresentaram maior afinidade com maior taxa de expressão nessas condições (RIBEIRO *et al* 2001).

A produção do antígeno da bactéria causadora da Brucelose, *Brucella abortus*, patogênica para humanos e bovinos foi possível com a utilização do promotor indutível *P_{nisA}*, pois apresentou altos níveis para a produção da proteína L7/L12. Essa proteína foi expressa em três áreas celulares na parede celular, no citoplasma e meio extracelular, todas construídas através das técnicas de engenharia genética. Apesar da competência de expressar a proteína nas três localidades, buscou-se investigar a via de exposição mais adequada para vacinação oral. Os resultados demonstraram que os vetores construídos possuem a capacidade de controlar a expressão nas três áreas testadas otimizando a produção da proteína (RIBEIRO *et al* 2001).

4.3. Outras pesquisas

Enfim, com o advento da construção de vetores muito já foi feito para minimizar os problemas e aumentar o nível de expressão de genes heterólogos e de forma regulável. O gene da somatostatina foi expresso através de um vetor contendo o promotor *lac*. A prolactina humana, hormônio do crescimento e a interleucina 6 foram adquiridas pelo vetor contendo o promotor *trp*. Outros genes também foram expressos tais como o que sintetiza o interferon e o da urogastrona (SOUZA 1995).

CONCLUSÃO

A utilização de microrganismos modificados, ou seja, microrganismos montados com gene exógeno com objetivo de produzir espécies de valor biotecnológicos capazes de sintetizar suas respectivas proteínas estão à frente dos processos convencionais. Devido à relativa dificuldade de obtenção da matéria prima que quando disponíveis são acompanhados de elevado custo, o emprego de novas técnicas como a do DNA recombinante pode ser considerado uma porta principalmente no que diz respeito à área econômica. Nesse ponto de vista não pode ser permitido que o abuso financeiro que é geralmente acompanhado de novas tecnologias cheguem ao consumidor inviabilizando a utilização dos produtos derivados da engenharia genética.

Além disso, as facilidades dispensadas na manipulação dos procariotos como o controle das condições de crescimento, temperatura, pH e nutrientes, reduzido espaço físico para produção, facilidade de manipulação, produção em grande quantidade da proteína de interesse associada a menor risco de contaminação por agentes exógenos aumentam a viabilidade de utilizar as bactérias como fábricas de proteínas eucarióticas.

BIBLIOGRAFIA

ALBERTS, B.; BRAY, D.; LEWIS, J.; RAFF, M.; ROBERTS, K. & WATSON, J. D. *Biologia Molecular da Célula*. 3ª ed. Porto Alegre: Artes Médicas, 1997. 1291p.

ALBERTS, B.; BRAY, D.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K. & WALTER, P. *Fundamentos de Biologia Celular*. Porto Alegre: Artmed, 1999. 757 p.

AZEVEDO, J. L. *Genética de Microrganismos*. Goiânia: UFG, 1998. 490 p.

CAMPBELL, M. K. *Bioquímica*. Porto Alegre: Artmed, 2001. 752p.

COSTA, S. O. P. *Genética Molecular e de Microorganismo*. São Paulo: Manole, 1986. 559p.

DARNELL, J. E.; LODISCH, H.; BALTIMORE, D.; BERK, A.; MATSUADAI, P. & ZIPURSKY, S. *Molecular Cell Biology*. 3ª ed. New Jersey: Scientific American Books, 1995. 1344p.

ÉTIENNE, J. *Bioquímica genética e biología Molecular*. Trad. de Arthur Jardim de Cerqueira. 5ª ed. São Paulo: Santos, 2003. 504p.

GRIFFITHS, A. J. F.; GELBART, W. M.; MILLER, J. H. & LEWONTIN, R. C. *Genética Moderna*. 1ª ed. Rio de Janeiro: Guanabara Koogan, 2001. 589P.

LEHNINGER, A L. *Princípios de bioquímica*. São Paulo: Sarvier, 1984. 725p.

LEWIN, B. *Genes VII*. 1ª ed, editora Artmed, Porto Alegre, RS. 2001. 955p.

LIMA, B. D. A produção de insulina humana por engenharia genética. *Bactéria transgênica produz insulina humana*, Brasília, Biotecnologia, Ciência & Desenvolvimento, nº 23, p. 28-31, 2001.

MADIGAN, M.; T, MARTINKO, J.M. & PARKER, J. *Biology of Microorganisms*. 9ª ed. New Jersey: Prentice Hall, 2000. 991p.

SOUZA, M. B. N. S. *Construção de sistemas de expressão e secreção de proteínas heterólogas em Escherichia coli*. Brasília: UNB, 1995. 152p. Tese de doutorado.

RIBEIRO, L.; PONTES, D.; DORELLA, F.; LANGELLA, P.; LE LOIR, Y.; OLIVEIRA, S. C; AZEVEDO, V. Bactérias lácteas produtoras de antígenos. *Desenvolvimento de uma vacina oral contra brucellose*, Belo Horizonte, Biotecnologia, Ciência & Desenvolvimento, nº 22, p. 36-40, 2001.

STRYER, L. *Bioquímica*. 4ª ed. Rio de Janeiro: Guanabara Koogan, 1996. 1000p.