

EXPRESSÃO HETERÓLOGA DE PROTEÍNAS

Data de aceite: 10/05/2021

Data de submissão: 13/04/2021

Karolinni Bianchi Britto

Universidade Federal do Espírito Santo
Vitória - ES
<http://lattes.cnpq.br/3604508015056223>

Greiciane Gaburro Paneto

Universidade Federal do Espírito Santo
Alegre - ES
<http://lattes.cnpq.br/8176374147579841>

RESUMO: Proteínas são macromoléculas importantes para o funcionamento dos organismos pois desempenham inúmeros papéis centrais, além de serem amplamente utilizadas como reagentes na pesquisa, sendo aplicadas em diversas áreas como indústria farmacêutica, agricultura e diagnóstico de doenças. A expressão heteróloga de proteínas, possível através da tecnologia do DNA recombinante, se destaca por permitir que organismos que não expressam a proteína normalmente possam sintetizá-la em escalas necessárias ao uso que se propõem. Para tal, são necessários vetores de expressão com características adequadas e que possam direcionar corretamente a transcrição e tradução do gene de interesse na célula hospedeira selecionada. Diferentes sistemas de expressão heteróloga de proteínas são conhecidos, podendo-se destacar as bactérias, especialmente *Escherichia coli*, leveduras, fungos filamentosos, microalgas, cultura de células de mamíferos, insetos e plantas. Cada sistema possui vantagens e desvantagens, sendo o melhor escolhido de acordo com a característica da proteína que se

deseja expressar. Quanto à análise e identificação, diversas técnicas são aplicadas para verificar a presença do gene e das proteínas nas células. A purificação das proteínas expressas pelos sistemas ocorre especialmente por cromatografia, podendo ser do tipo cromatografia de troca iônica, de afinidade, de gel filtração, de fase reversa, de imunoafinidade e cromatografia líquida de alta resolução. Sequências específicas de aminoácidos, conhecidas como 'tags', geralmente são adicionadas às proteínas para facilitar principalmente o processo de purificação, apesar de influenciar também em outras etapas do processo. Após a obtenção da proteína purificada segue-se com o processamento de acordo com o objetivo proposto com sua síntese.

PALAVRAS-CHAVE: Expressão heteróloga, proteína, DNA recombinante, vetores, purificação de proteína.

HETEROLOGOUS PROTEIN EXPRESSION

ABSTRACT: Proteins are important macromolecules for the functioning of organisms because they play countless central roles, besides being widely used as reagents in research, being applied in several areas such as pharmaceutical industry, agriculture and disease diagnosis. The heterologous expression of proteins, possible through recombinant DNA technology, stands out by allowing organisms that do not normally express the protein to synthesize it in scales necessary for its proposed use. For this, expression vectors with adequate characteristics and that can correctly direct the transcription and translation of the gene of interest in the selected host cell are needed. Different systems of heterologous expression of proteins are known, especially bacteria *Escherichia coli*, yeast, filamentous fungi, microalgae, mammalian cell culture, insects and

plants. Each system has advantages and disadvantages, and the best one is chosen according to the characteristic of the protein to be expressed. As for the analysis and identification, several techniques are applied to verify the presence of the gene and proteins in the cells. The purification of the proteins expressed by the systems occurs especially by chromatography, and may be of the ion exchange, affinity, gel filtration, reverse phase, immunoaffinity, and high performance liquid chromatography types. Specific amino acid sequences, known as 'tags', are usually added to proteins to facilitate mainly the purification process, although they also influence other steps in the process. Once the purified protein is obtained, it is processed according to the proposed objective of its synthesis.

KEYWORDS: Heterologous expression, protein, recombinant DNA, vectors, protein purification.

1 | INTRODUÇÃO

Proteínas são macromoléculas formadas por um ou mais polipeptídeos, os quais são constituídos por uma cadeia de aminoácidos unidos por ligações peptídicas. A sequência exata dos aminoácidos é determinada pelo gene que codifica para esse polipeptídeo específico. Quando sintetizada, uma cadeia polipeptídica se dobra, assumindo uma conformação específica (WALSH, 2007). Assim, cada tipo de proteína possui uma estrutura tridimensional única e essa estrutura confere também uma função única. Portanto, é evidente que a sequência de aminoácidos desempenha um papel fundamental tanto na determinação da estrutura tridimensional da proteína quanto na sua função (NELSON; COX, 2011).

A conformação adotada pela proteína é amplamente estabilizada por múltiplas interações não covalentes fracas. Um processo de desnaturação pode ocorrer quando há alguma influência, como calor e produtos químicos por exemplo, que perturbe essas interações fracas resultando na interrupção da conformação nativa do polipeptídeo, o que geralmente leva a perda de atividade funcional devido a perda da estrutura. A conformação da proteína pode ser determinada por técnicas de difração de raios X e ressonância magnética nuclear (RMN), por exemplo (WALSH, 2007).

Para que as informações genéticas presentes no DNA da célula sejam expressas na forma de uma proteína, enzima ou anticorpo, um intermediário entre a proteína e o DNA é requerido: RNA mensageiro (RNAm), conforme discutido anteriormente.

Relembrando: o fluxo geral de informações, do DNA às proteínas, foi resumido por Francis Crick como o "Dogma Central" da biologia molecular em que o DNA de fita dupla é transcrito para o RNAm (em eucariotos, com processamento da transcrição), que por sua vez é traduzido pelo ribossomo na cadeia de aminoácidos que formam uma proteína. A síntese de proteínas possui três estágios principais: iniciação, alongamento e terminação (GROVES, 2006).

As proteínas possuem papéis importantes em inúmeros processos celulares, sendo sintetizadas como parte do metabolismo de todas as formas de vida. Tanto as proteínas nativas quanto as recombinantes são muito utilizadas como reagentes na pesquisa científica, sejam em processos industriais mediados por enzimas, diagnóstico de doenças, na indústria farmacêutica e em setores da agricultura (EGELKROUT; RAJAN; HOWARD,

2012).

Uma proteína heteróloga é definida como aquela produzida em organismos diferentes dos de sua origem, podendo ser expressas em sistemas procariotos ou eucariotos (CANÇADO, 2002; RAI; PADH, 2001). Assim, é denominada produção heteróloga de proteínas a expressão de proteínas recombinantes em células diferentes de sua ocorrência natural (WALSH, 2007).

A expressão heteróloga, de uso da tecnologia do DNA recombinante, permite a obtenção de produtos biofarmacêuticos como hormônios, fatores de coagulação, citocinas, anticorpos, vacinas, enzimas, imunossuppressores, entre outros, em uma elevada quantidade quando comparados aos seus produtores nativos (RAMANA; XAVIER; SHARMA, 2017). Existem diferentes sistemas de expressão de proteínas heterólogas. Dentre os procariotos, estão os que utilizam células de bactérias e são os mais comuns nas indústrias biotecnológicas na produção de proteínas recombinantes; e, dentre os eucariotos, encontram-se os de leveduras, cultura de células de mamíferos e insetos, plantas e animais transgênicos (CANÇADO, 2002; WALSH, 2007), que serão discutidos posteriormente.

2 | TECNOLOGIA DO DNA RECOMBINANTE

Por meio de uma tecnologia de manipulação do DNA que surgiu na década de 1970, foi possível o isolamento e caracterização de genes criando uma forma de recombinar segmentos de DNA de diversas fontes em novas moléculas compostas. Essas técnicas de manipulação do DNA são conhecidas como tecnologia de DNA recombinante ou engenharia genética (GROVES, 2006). Normalmente envolve o isolamento, a manipulação e a reintrodução de trechos de DNA nas células, conferindo à célula receptora a capacidade de produzir uma proteína específica. O pedaço de DNA criado artificialmente *in vitro* que contém DNA obtido de duas ou mais fontes é chamado de DNA recombinante (DNAr) (WALSH, 2007).

No caso do desenvolvimento de uma proteína heteróloga, realiza-se inicialmente a identificação e o isolamento do gene responsável por codificar a proteína alvo, a geração de um pedaço apropriado de DNAr contendo a sequência de codificação da proteína e a introdução deste DNAr em uma célula hospedeira apropriada, de modo que a proteína alvo seja produzida em grandes quantidades por essa célula manipulada (WALSH, 2007).

2.1 Clonagem

A inserção e multiplicação de moléculas idênticas de DNA em um organismo hospedeiro refere-se à etapa de clonagem na técnica de DNAr (GROVES, 2006). Para clonar um fragmento de DNA de interesse é necessário vinculá-lo a uma molécula de DNA vetorial, que pode se replicar dentro de uma célula hospedeira. Então, após uma única molécula de DNAr, composta por um vetor e um fragmento de DNA inserido, ser introduzida em uma célula hospedeira, o DNA inserido é replicado junto com o vetor, gerando um grande número de moléculas de DNA idênticas (ALBERTS et al., 2007).

As enzimas de restrição (ERs), endonucleases produzidas por bactérias que normalmente reconhecem sequências específicas de 4 a 8 pb nos chamados locais de

restrição, facilitam a produção das moléculas de DNAr. Tais enzimas cortam o DNA de uma fonte específica em um conjunto reproduzível de fragmentos de restrição para serem introduzidos no vetor. Como exemplos podemos citar: BamHI, EcoRI, HindIII, KpnI e PstI. A inserção dos fragmentos de DNA gerados em vetores recebe o auxílio de DNA ligases (ALBERTS et al., 2007).

Graças às informações sobre diversas sequências de genes disponíveis em bancos de dados, é possível obter grandes quantidade do gene de interesse usando a técnica de Reação em Cadeia da Polimerase (PCR). Primeiramente, o DNA genômico da fonte de interesse é extraído e depois realizada a PCR, gerando cópias do gene de interesse. Locais de reconhecimento para ERs podem ser incorporados nos oligonucleotídeos para permitir a clonagem do gene amplificado (WALSH, 2007).

O corte do DNA pode ser também realizado sem o uso de ERs, no caso de cisalhamento mecânico da solução de DNA. Pode-se ainda partir de um extrato de RNA, que após algumas etapas geram o DNA complementar (cDNA) apto a ser clonado da mesma maneira que o DNA genômico (GROVES, 2006).

2.2 Vetores

O fragmento de DNA que contém o gene de interesse será transferido para uma célula hospedeira através dos chamados vetores. Uma característica essencial do vetor de clonagem utilizado é que ele deve ser capaz de se autorreplicar na célula em que será introduzido, além de possuir marcador de seleção e locais de restrições únicos para inserção das sequências de interesse (FERRIER, 2017). Esses vetores, que podem ser plasmídeos, bacteriófagos λ , cosmídeos, cromossomos artificiais bacterianos (BACs) ou cromossomos artificiais de leveduras (YACs) por exemplo, devem possuir pelo menos um local de clonagem e uma sequência que pode ser clivada por uma ER para permitir sua ligação a um fragmento de DNA clivado de maneira semelhante (GROVES, 2006; WALSH, 2007). Alguns vetores de clonagem são conhecidos como: pBR322, pUC8, pEMBL8 e λ gt10 (BROWN, 2013).

Os plasmídeos, vetores muito utilizados, são moléculas circulares de DNA de fita dupla que são separadas do DNA cromossômico de uma célula. Ocorrem naturalmente em bactérias ou em células eucarióticas inferiores, seu DNA é duplicado antes de cada divisão celular e suas cópias compõem cada célula filha. Os plasmídeos frequentemente carregam um ou mais genes para resistência a antibióticos e os mais comumente usados no DNAr são os que se replicam em *Escherichia coli* (ALBERTS et al., 2007; GROVES, 2006).

O vetor escolhido deve ser aberto com a mesma ER utilizada para cortar o fragmento de DNA contendo a sequência gênica, seguido por uma incubação para promover a adesão das extremidades do fragmento de DNA com o vetor. Após, o DNAr deverá ser transferido para as células hospedeiras, em um processo conhecido como transformação (WALSH, 2007), que pode acontecer por diversas técnicas, entre as quais se destacam: a eletroporação como método de transfecção, o qual consiste em sujeitar as células a um campo elétrico forte, curto e pulsado, abrindo poros na célula e permitindo que a molécula exógena seja inserida (LI; LIN, 2011); microinjeção (SAMPATH KUMAR; PUTTARAJU, 2012); bombardeamento de partículas ou biobalística (CARNEIRO; CARNEIRO; PAIVA, 2004); e, também através de vetores, como bactérias e vírus (CARNEIRO; CARNEIRO;

PAIVA, 2004; FELBERBAUM, 2015; STUDART-GUIMARÃES; LACORTE; BRASILEIRO, 2003).

2.3 Células hospedeiras

Diversas células podem ser usadas na técnica do DNA recombinante. Entre elas encontram-se comumente células bacterianas, leveduras e mamíferas. De acordo com o tipo de produto e a quantidade desejada baseia-se a escolha da célula e a cepa específica (GROVES, 2006).

A replicação do plasmídeo dentro da célula hospedeira começa na origem de replicação (ORI), uma sequência específica de DNA de 50 a 100 pares de bases, e continua em torno do plasmídeo circular. Dessa forma, ocorre a replicação da sequência de DNA inserida no plasmídeo, seja ela qual for (ALBERTS et al., 2007).

Para identificar se as células hospedeiras, como exemplo *E. coli*, absorveram corretamente o plasmídeo, as mesmas são incubadas em condições apropriadas para que se possa identificar qual colônia abriga o fragmento de DNA que contém o gene de interesse. Dentre as estratégias para identificação do DNA em células de *E. coli* usando o vetor pUC18, destaca-se uma em que as células são espalhadas em placas de ágar contendo o antibiótico ampicilina e um produto químico chamado X-Gal. Devido as células de *E. coli* não possuírem o gene de resistência à ampicilina (*ampR*), as células que não foram transformadas não crescerão nesse meio. Já as células que contém o plasmídeo com o gene *ampR*, mesmo sem o fragmento de DNA inserido a ele, crescerão no meio. Porém, devido a essa célula produzir a enzima β -galactosidase haverá a quebra do X-Gal liberando um produto de cor azul, colorindo essas colônias. Finalmente, as células contendo o plasmídeo no qual um fragmento de DNA foi inserido corretamente no gene *lacZ* não produzirão β -galactosidase e, assim, colônias derivadas dessas células crescerão no meio e apresentarão uma cor branca normal, as quais serão identificadas e selecionadas (WALSH, 2007).

Segundo Groves (2006), para escolher a melhor célula hospedeira deve-se observar alguns aspectos, entre eles: se o crescimento da célula é lento ou rápido; o custo; o nível de expressão de produto alcançado pela célula; a facilidade da purificação do produto levando em conta a secreção dele no meio; se o produto, no caso uma proteína, será dobrada corretamente e modificada para atingir sua atividade; e, se os vetores são adequados para ela.

3 | EXPRESSÃO HETERÓLOGA DE PROTEÍNAS

Devido às técnicas de DNA recombinante, que transformam células em fábricas para sintetizar proteínas, a produção comercial de insulina, hormônio do crescimento, fator estimulador de colônias de granulócitos, vacinas e outras proteínas humanas com usos terapêuticos tornou-se possível (ALBERTS et al., 2007; RAMANA; XAVIER; SHARMA, 2017). A expressão heteróloga de proteínas provenientes da manipulação de genes específicos representa a última etapa da tecnologia do DNA recombinante (CARUSO, 2007).

Inicialmente, obtém-se um clone do gene que codifica a proteína de interesse, após, os vetores que expressarão as proteínas são projetados com promotores para a transcrição do DNA e inseridos nas células hospedeiras. Para facilitar a purificação da proteína expressa, entre outros benefícios, normalmente implanta-se uma sequência nucleotídica curta (podendo ser seis resíduos de histidina no carbono terminal, por exemplo) ao final do DNA complementar, conhecida também como 'tag' (ALBERTS et al., 2007), que será discutida posteriormente.

Os vetores citados anteriormente foram projetados para isolamento e propagação de fragmentos de DNA com genes que codificam uma proteína específica, porém, para se alcançar altos níveis de expressão da proteína requerida, o vetor em questão deve suportar alto nível de transcrição e tradução. Os vetores de expressão devem possuir além de alguns elementos como origem de replicação, marcador selecionável com um gene de resistência a antibiótico e locais de restrição para clonagem, também região com promotor da transcrição, região de início da tradução, bem como terminadores da transcrição e da tradução, sequências reguladoras, sequência Shine-Dalgarno (RBS - *Ribosome Binding Site*) no caso de procariotos, proteínas de fusão para facilitar dobramento e purificação (opcional), ou seja, deve conter todas as sequências de DNA específicas que irão direcionar a transcrição e tradução do gene de interesse. Esses vetores estão disponíveis comercialmente e cada um é adaptado para funcionar melhor em um tipo específico de célula hospedeira (KAUR; KUMAR; KAUR, 2018; SØRENSEN; MORTENSEN, 2005; WALSH, 2007).

A clonagem de genes para expressão de proteínas geralmente começa com um cDNA, pois a maioria dos genes eucarióticos contém íntrons. Esse cDNA é usado como modelo na PCR onde o gene de interesse será amplificado, clonado, sequenciado e subclonado em um ou mais vetores de expressão. Alguns motivos levam a preferência da clonagem inicial em um vetor de não expressão antes do uso do vetor de expressão, como por exemplo o tamanho maior do vetor de expressão e seu baixo número de cópias complicando um pouco mais a clonagem e sequenciamento, a ocorrência de erros de sequência no gene ou polimorfismos de nucleotídeo único e também erros nos iniciadores de PCR. Assim, para melhorar a expressão e posterior purificação das proteínas esse processo pode ser realizado, confirmando antes a sequência do gene em questão e corrigindo algum erro que possa ter ocorrido (HARTLEY, 2006).

Muitos sistemas de expressão comerciais desenvolvidos para inúmeras aplicações e compatibilidades estão disponíveis, como exemplos pode-se citar o sistema de expressão pET baseado no promotor do fago T7 e também sistemas que utilizam os promotores *trc*, *lac*, *tac* e promotor PL do fago λ (RESENDE, 2015; SØRENSEN; MORTENSEN, 2005). Os promotores são elementos de um vetor que estão relacionados diretamente ao rendimento proteico, devido seu efeito na força e duração da transcrição. A RNA polimerase se liga a sequência promotora, localizada ao lado do gene alvo, iniciando a síntese de RNAm. Um promotor efetivo para expressão heteróloga de proteínas deve ser forte, apresentar baixa atividade de transcrição basal e sua indução deve ser simples e de baixo custo (FRANCIS; PAGE, 2010; RESENDE, 2015).

Espera-se de uma expressão heteróloga que a proteína de interesse seja estável, solúvel, expressa em grande quantidade, não tóxica para a célula hospedeira e facilmente

purificada. Não existe um sistema ideal para todas as proteínas possíveis, cada um possui características favoráveis e limitações (WALSH, 2007).

3.1 Sistemas de expressão heteróloga de proteínas

O uso de bactérias como *E. coli* como sistema de expressão heteróloga de proteínas é vantajoso devido sua biologia molecular ser bem caracterizada, possuir rápido crescimento, baixos custos de produção (GOMES et al., 2016; WALSH, 2007), escalonamento direto para maiores produções, além de fácil manipulação. Como desvantagens têm-se a falta de processamentos pós-traducionais, como a glicosilação; acúmulo intracelular de corpos de inclusão, que são agregados constituídos de proteínas desnaturadas ou incorretamente enoveladas; o processamento inadequado de transcritos de genes eucarióticos que contêm sequências não codificantes ou íntrons; a presença de lipopolissacarídeo em sua superfície; capacidade limitada de formar pontes dissulfídicas; viés de códon; acúmulo de endotoxinas e a falta de um sistema para secretar as proteínas ao meio extracelular, sendo que essa última desvantagem faz com que seja necessária uma purificação mais extensa para separar a proteína de interesse das demais proteínas homólogas adicionais produzidas pelas células de *E. coli* (BROWN, 2013; GOMES et al., 2016; GROVES, 2006; WALSH, 2007).

Alternativas para contornar algumas desvantagens desse sistema têm sido implantadas, como o uso de proteínas de fusão, uso de peptídeo sinal para direcionamento da proteína, condições especiais de crescimento, uso de aditivos, co-expressão com chaperonas e escolha da linhagem mais adequada (KAUR; KUMAR; KAUR, 2018). Deve-se selecionar a cepa de *E. coli* mais interessante e viável para clonagem e expressão de determinada proteína (SØRENSEN; MORTENSEN, 2005), pois existem diversas disponíveis comercialmente, como exemplo BL21 (Novagen), BL21 (DE3) (Novagen), BLR (DE3) (Addgene), Origami B (Novagen), T7 Express (Novagen), entre outras (KAUR; KUMAR; KAUR, 2018). A insulina humana recombinante foi o primeiro produto biofarmacêutico feito por engenharia genética, sendo produzida em *E. coli* com aprovação de marketing em 1982 (WALSH, 2007).

Um sistema de expressão eucariótico interessante e antigo é o de leveduras, representado na maioria das vezes por *Saccharomyces cerevisiae*, mas também por *Hansenulla polymorpha* e *Pichia pastoris* (BALAMURUGAN et al., 2006). Os sistemas eucarióticos permitem a modificação e o processamento de proteínas ao contrário dos bacterianos, o que constitui vantagem (CANÇADO, 2002; GOMES et al., 2016). Como vantagens também caracterizam-se o crescimento rápido em meios relativamente de baixo custo, não há produção de endotoxinas, parede externa resistente a danos físicos, possibilidade de secretar produtos de proteína para o meio mais rapidamente do que a *E. coli*, organismos listados como seguros pela grande aplicação industrial já utilizada, capacidade de processar transcritos de RNA para remover os íntrons e promover glicosilação de proteínas, esta última pode representar também uma desvantagem devido ao padrão de glicosilação variar um pouco das proteínas nativas podendo ocorrer também hiperglicosilação (GROVES, 2006; WALSH, 2007). Outra desvantagem observada é em relação aos níveis de expressão de proteínas heterólogas que situam-se abaixo dos alcançados por *E. coli* (WALSH, 2007) e também a possibilidade de viés do códon (GOMES

et al., 2016). Os principais biofarmacêuticos obtidos através de leveduras são a albumina sérica humana, insulina e seus análogos, e vacinas contra papilomavírus e hepatite (NIELSEN, 2013).

Fungos filamentosos também possuem capacidade de produzir grandes quantidades de proteína e são muito utilizados na indústria de alimentos principalmente como produtores de enzimas, como por exemplo *Aspergillus niger*, *Aspergillus oryzae* e *Trichoderma reesei* (GOMES et al., 2016; LANDOWSKI et al., 2016). Possuem vantagens em relação ao crescimento rápido e robusto, secreção das proteínas nos meios extracelulares facilitando a purificação e também a capacidade de realizar modificações pós-traducionais como os demais sistemas eucarióticos (HAVLIK et al., 2017; WALSH, 2007). Porém, como desvantagem tem-se a maior complexidade desse sistema e a relativa falta de compreensão de sua fisiologia em comparação às bactérias (SU et al., 2012).

As microalgas, principalmente da espécie *Chlamydomonas reinhardtii*, representam outro sistema para a produção de proteínas recombinantes, devido ao baixo custo e tecnologia simples (DYO; PURTON, 2018). *C. reinhardtii* é um organismo modelo cujos três genomas foram sequenciados, sendo possível então direcionar a modificação genética (TAUNT; STOFFELS; PURTON, 2018). A introdução do gene de interesse no genoma de pequenos cloroplastos é desejada, pois garante uma expressão estável e de alto nível (DYO; PURTON, 2018). Outras vantagens apresentadas são a formação de ligações dissulfeto e dobragem adequada. Muitas proteínas recombinantes, incluindo anticorpos monoclonais, fatores de crescimento e vacinas são produzidas em *C. reinhardtii* (TAUNT; STOFFELS; PURTON, 2018).

Outras células eucarióticas usadas incluem linhagens celulares derivadas de tumor, células de ovário de hamster chinês e células de rim de hamster bebê. O crescimento dessas células é muito mais lento do que leveduras ou bactérias além de serem mais sensíveis às mudanças no pH, temperatura, nível de oxigênio, metabólitos e às forças de cisalhamento. O cultivo para produção em larga escala é mais complicado e apresentam níveis variados de expressão das proteínas heterólogas (GOMES et al., 2016; GROVES, 2006), além de geralmente exigirem a suplementação de fatores de crescimento, aminoácidos, agentes redutores ou vitaminas (ZHU, 2012). Porém, dentre as vantagens encontram-se a capacidade de secretar o produto proteico no meio, de processar adequadamente os transcritos de RNA para expressão gênica e modificar adequadamente a proteína expressa por clivagem, glicosilação e redobramento, por exemplo (GROVES, 2006). A cultura de células de mamíferos geralmente é usada apenas na fabricação de proteínas terapêuticas que mostram modificações pós-traducionais extensas e essenciais, incluindo gonadotrofinas, algumas citocinas, anticorpos monoclonais, fatores de coagulação, enzimas, entre outros (DUMONT et al., 2016; WALSH, 2007). Os promotores do citomegalovírus são amplamente utilizados na construção de cassetes de expressão para células de mamíferos (KHAN, 2014).

Sobre os sistemas de cultura de células de insetos envolve-se a infecção dessas células cultivadas com o vetor de expressão baculovírus geneticamente modificado, que carrega integrado ao seu genoma o gene que codifica a proteína desejada controlado por um potente promotor viral e pelo seu terminador correspondente (FELBERBAUM, 2015; STÖGER et al., 2005). Alguns insetos utilizados são *Spodoptera frugiperda*, *Drosophila*

melanogaster e *Autographa californica* (GECHELE et al., 2015). São células que crescem mais rápido comparadas a de mamíferos, não há risco de contaminação por príons e DNA oncogênico, garantem alto rendimento (CONTRERAS-GÓMEZ et al., 2014; GECHELE et al., 2015) e servem de ferramenta para produção de glicoproteína recombinante. Porém, como desvantagens encontram-se as modificações pós-traducionais que podem ser incompletas, podem diferir dos padrões associados às glicoproteínas humanas nativas e culturas com condições mais exigentes. O antígeno de superfície da hepatite B e IFN- γ são exemplos de proteínas terapêuticas produzidas com sucesso em escala laboratorial em linhas celulares de insetos (WALSH, 2007).

A produção de proteínas heterólogas em animais transgênicos também ganhou um certo destaque, pois como cada célula do animal transgênico resultante abrigará uma cópia do DNA transferido, essa nova informação genética introduzida pode ser transmitida de uma geração para a seguinte, produzindo a proteína de interesse continuamente. Para facilitar a liberação da proteína, uma alternativa foi direcionar a produção de proteínas para a glândula mamária, na qual cabras e ovelhas provaram ser interessantes devido a características como altas capacidades de produção de leite e facilidade de manuseio e criação (WALSH, 2007). A primeira proteína recombinante aprovada em 2006 foi a antitrombina, secretada no leite de cabras (MAKSIMENKO et al., 2013). Além do leite, anticorpos e outras proteínas já foram produzidos no sangue de porcos e coelhos transgênicos. Porém, várias razões tornam esse sistema inviável industrialmente, entre elas a possibilidade de coletar pequenos volumes de sangue, a complexidade do soro que contém uma variedade de proteínas nativas tornando a purificação mais difícil, a falta de estabilidade das proteínas e os efeitos colaterais fisiológicos negativos que podem ser causados pela proteína no animal produtor (WALSH, 2007).

Também foi demonstrada a produção de β -lactamase na clara de ovos de galinhas transgênicas (HARVEY et al., 2002). Algumas vantagens como dobragem adequada das proteínas, glicosilação correta e modificações pós-traducionais adequadas são destacadas (GOMES et al., 2016), porém, o custo mais alto, tempo de produção elevado e baixo rendimento representam problemas, além ainda do risco de contaminação por vírus (CANÇADO, 2002; GOMES et al., 2016; LARRICK; THOMAS, 2001).

O uso de plantas como sistema de expressão de proteínas heterólogas, em comparação aos demais eucariotos, apresenta-se com maior segurança, menor tempo, baixo custo e é superior em termos de armazenamento e distribuição. Constituem um sistema ideal para expressar proteínas heterólogas que requerem modificações pós-traducionais um pouco mais complicadas. Tais proteínas expressas podem ser localizadas em diferentes órgãos da planta e existe a possibilidade de manipular o tempo de expressão para estágios de crescimento específicos (ŁOJEWSKA et al., 2016). Oferecem uma fonte de material menos custosa para seu cultivo em comparação a outros microrganismos e devido a facilidade pode-se cultivar um número maior para aumentar o rendimento total das proteínas (ESPÍRITO SANTO, 2013).

Em relação ao uso de plantas transgênicas, uma gama de proteínas terapêuticas podem ser expressas pelo tecido da planta, pois a planta pode receber a transferência dos genes de interesse através de vetores principalmente pelo patógeno vegetal *Agrobacterium*, no qual uma porção do plasmídeo é translocada e integrada ao genoma da célula vegetal,

ou também através de vetores virais de plantas (BUYEL; TWYMAN; FISCHER, 2017; WALSH, 2007). No primeiro método, plantas transgênicas com uma expressão estável são obtidas, enquanto no segundo caso a expressão transitória do gene é obtido (YAO et al., 2015). Algumas desvantagens do sistema de plantas transgênicas são padrões de glicosilação diferentes da proteína humana nativa (WALSH, 2007), presença de metabólitos que podem contaminar o produto bruto, natureza geográfica/sazonal do crescimento da planta (BALAMURUGAN et al., 2006), a variação dos níveis de expressão de acordo com o destino e a necessidade de melhorar o desenvolvimento dos ensaios (GOMES et al., 2016) invasion and metastasis. Tissue and serum levels of MMP2 and MMP9 correlate with disease prognosis. Real-time PCR is emerging as an alternative or supplementary technique to immunohistochemistry (IHC).

3.2 Análise e seleção dos recombinantes

Quanto ao processo de seleção, o mesmo pode envolver a resistência a antibióticos ou a habilidade das células hospedeiras de crescerem na ausência de certos suplementos nutricionais, capacidade gerada pelos vetores inseridos, como previamente discutido a respeito da inserção do fragmento de interesse no gene lacZ, o que determinará quais colônias foram transformadas pela cor apresentada (WALSH, 2007). Pode-se realizar também uma PCR seguida de sequenciamento para confirmar a presença do gene de interesse.

Além disso, técnicas imunológicas podem ser usadas para detectar a presença da proteína que foi produzida pelos clones (GROVES, 2006), como a verificação da expressão da proteína pela técnica de *Western blotting* acompanhada por eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS-PAGE) (LIU; YANG, 2012), e ainda, dependendo do produto em questão, existe a possibilidade de encontrar o clone adequadamente transformado que expressou uma enzima, por exemplo, procurando pela atividade dela na célula (GROVES, 2006).

3.3 Purificação de proteínas

Antes da purificação é necessário que a proteína seja retirada da célula hospedeira e o processo vai depender se o sistema de expressão libera o produto no meio intracelular ou extracelular. Nos dois casos, as células podem ser coletadas por centrifugação ou microfiltração (WALSH, 2007). Ainda, no caso de produto intracelular, realiza-se a ruptura celular para obter a proteína de interesse. Essa ruptura é frequentemente alcançada por métodos mecânicos como homogeneização ou agitação vigorosa com abrasivos, mas também pode-se recorrer ao uso de produtos químicos como detergentes, solventes como tolueno, agentes caotrópicos como ureia, através de enzimas como lisozima, exposição a condições alcalinas e sonicação (AHMAD et al., 2014). Após a ruptura das células, ocorre a concentração do produto bruto, podendo ocorrer principalmente por ultrafiltração ou também através de sais que induzem a precipitação (WALSH, 2007).

No caso do produto intracelular produzido em procariotos, a secreção no espaço extra citoplasmático pode ser alcançada fundindo a proteína recombinante a um peptídeo sinal adequado. Alguns exemplos de peptídeos são: OmpA (proteína A da membrana externa), PhoE (proteína E dos poros da membrana externa), Lpp (lipoproteína mureína) e OmpT

(protease VII) (CHOI; LEE, 2004). O peptídeo sinal é clivado e a proteína recombinante se dobra no periplasma com a ajuda de chaperonas e formam-se ligações dissulfeto gerando a conformação correta (KAUR; KUMAR; KAUR, 2018). Geralmente a proteína é expressa em maiores quantidades no citoplasma bacteriano, porém na forma de corpos de inclusão. Então, é necessário um processamento adequado desses agregados para se obter a forma solúvel da proteína, passando pelo isolamento, purificação, solubilização, renaturação e dobramento, através de reagentes específicos e técnicas diversas (CLARK, 2001).

Para purificar proteínas recombinantes são utilizados os mesmos meios de purificação de proteínas tradicionais, porém, a purificação pode ser até mais direta devido a capacidade de alcançar altos níveis de expressão da proteína desejada. Técnicas de cromatografia que as separam considerando suas propriedades específicas e diferenças são amplamente utilizadas, como exemplo cita-se a cromatografia de troca iônica, de afinidade, de imunoafinidade e cromatografia líquida de alta resolução (HPLC), sendo que essa última poderá ser utilizada desde que a alta pressão aplicada não afete o produto proteico (WALSH, 2007).

O procedimento de purificação consiste na separação da proteína alvo, mantendo sua estrutura química e atividade biológica (LABROU, 2014), e representam entre 45 a 92% dos custos totais da fabricação de proteínas recombinantes (SARASWAT et al., 2013). Como princípio básico da separação cromatográfica tem-se uma fase móvel fluida que transporta espécies em direção a uma fase estacionária sólida, baseando-se nas diferenças de afinidade. Ao longo da coluna, alguns componentes da amostra transportada pelo eluente apresentarão interações mais fortes com a fase estacionária do que outros, gerando perfis de concentração e diferentes velocidades de eluição. Assim, espécies mais retidas eluirão por último, permitindo a coleta do produto de interesse com maior pureza (FARIA; RODRIGUES, 2015).

Cada metodologia de purificação utiliza uma característica particular das proteínas, por exemplo, quando se trata da especificidade do ligante, a cromatografia de afinidade é usada, tendo como base a interação bioespecífica entre uma proteína e um ligante apropriado, sendo o método mais popular; quando as proteínas possuem diferenças na carga a um determinado pH a cromatografia de troca iônica é preferida; quando as proteínas variam em tamanho é utilizada a cromatografia por exclusão de tamanho ou gel filtração; e, quando se trata de hidrofobicidade, são utilizadas a cromatografia de fase reversa e a cromatografia de interação hidrofóbica (OWCZAREK; GERSZBERG; HNATUSZKO-KONKA, 2019). Dependendo da proteína em questão, várias etapas de cromatografia podem ser requeridas (WALSH, 2007).

Algumas considerações importantes sobre o processo de purificação de proteínas são: deve ser simples e com o mínimo de etapas; uso de reagentes mais baratos; deve haver monitoramento constante; o produto final deve ser altamente purificado; gastar pouco tempo; permitir alta reprodutibilidade e sempre utilizar aparelhagem confiável (KHAN, 2014).

A incorporação de 'tags' ou proteínas de fusão podem facilitar o processo de purificação. Utiliza-se da engenharia genética para incorporação de tags específicas de peptídeos ou proteínas à proteína de interesse, que ligam-se às proteínas alvo por

um sítio de reconhecimento de uma protease específica (KOSOBOKOVA; SKRYPNIK; KOSORUKOV, 2016). Foram projetadas com sucesso tags que permitem a purificação rápida e direta da proteína híbrida por técnicas como cromatografia de afinidade, troca iônica, interação hidrofóbica ou imunoafinidade. Dependendo do marcador adicionado pode ser gerada uma carga positiva na proteína o que facilita sua purificação por troca catiônica, por exemplo (WALSH, 2007).

Após a purificação da proteína híbrida, é necessária a remoção do marcador, que pode ser realizada por meios químicos ou enzimáticos (WALSH, 2007), adicionando outra etapa e custo ao processo de produção de proteína recombinante (LI, 2011), porém, dependendo da sequência adicionada pode não afetar a estrutura final da proteína (OWCZAREK; GERSZBERG; HNATUSZKO-KONKA, 2019). Essas proteínas de fusão que se ligam às proteínas recombinantes facilitam não só a purificação, como também podem proteger as proteínas da proteólise, aumentar a solubilidade, facilitar o reconhecimento e melhorar a expressão. Podem ser tanto sequências curtas de aminoácidos como resíduo de histidina (poli-His), resíduo de arginina (poli-Arg) e epítipo FLAG, quanto proteínas como tioredoxina (TRX), proteína ligante da maltose (MBP), glutationa S-transferase (GST) e proteína verde fluorescente (GFP) (BUCHER; EVDOKIMOV; WAUGH, 2002).

Para acompanhar o progresso da purificação de proteínas ensaios analíticos são essenciais, entre eles estão métodos físicos e químicos, para determinar a qualidade da preparação final e a caracterização da proteína (VEDADI et al., 2010). Um método de quantificação do produto final é a medição da absorbância da proteína utilizando a ligação com corantes, e, para avaliar a pureza utiliza-se a técnica de SDS-PAGE, podendo ser acompanhada de *Western blotting*, além também de outras técnicas como eletroforese capilar, HPLC e espectrometria de massas (WALSH, 2007).

Adicionalmente, podem ser realizadas diferentes análises que não somente detectam impurezas proteicas, mas verificam se a substância está em total conformidade com o que foi proposto inicialmente, como mapeamento de peptídeos, análise de aminoácidos, sequenciamento N-terminal e análises espectrofotométricas. O processamento seguinte vai depender do objetivo com a proteína obtida, o que pode envolver a adição de excipientes, filtragem estéril, liofilização ou preparo de solução, por exemplo (WALSH, 2007).

4 | CONSIDERAÇÕES FINAIS

A expressão heteróloga de proteínas, que surgiu através da tecnologia do DNA recombinante, permitiu um avanço na produção em larga escala de muitos produtos importantes como hormônios, vacinas, entre outros. No presente capítulo foram descritos de forma geral o histórico, aplicação e todo processo que envolve desde a manipulação do gene de interesse, escolha de vetores, características dos sistemas de expressão heteróloga de proteínas, métodos de análise e identificação até a completa purificação das proteínas.

REFERÊNCIAS

- AHMAD, M. *et al.* **Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production.** Applied Microbiology and Biotechnology, v. 98, p. 5301–5317, 2014.
- ALBERTS, B. *et al.* **Molecular Biology of the Cell.** 5. ed. New York: Garland Science, 2007.
- BALAMURUGAN, V.; SEN, A.; SARAVANAN, P.; SINGH, R. K. **Biotechnology in the Production of Vaccine or Antigen for animal health.** Journal of Animal and Veterinary Advances, v. 5, n. 6, p. 487–495, 2006.
- BROWN, T. A. **Gene Cloning and DNA Analysis: An Introduction.** 6. ed. Hong Kong: John Wiley & Sons, 2013.
- BUCHER, M. H.; EVDOKIMOV, A. G.; WAUGH, D. S. **Differential effects of short affinity tags on the crystallization of Pyrococcus furiosus maltodextrinbinding protein.** Acta Crystallographica Section D Biological Crystallography, v. 58, p. 392–397, 2002.
- BUYEL, J. F.; TWYMAN, R. M.; FISCHER, R. **Very-large-scale production of antibodies in plants: The biologization of manufacturing.** Biotechnology Advances, v. 35, n. 4, p. 458–465, 2017.
- CANÇADO, L. J. **Utilização de Sementes de Tabaco Transgênico como Biorreatores para Produção de um Fragmento scFv de um Anticorpo Monoclonal.** [s.l.] Universidade Estadual de Campinas UNICAMP, 2002.
- CARNEIRO, A. A.; CARNEIRO, N. P.; PAIVA, E. **Transformação genética de milho utilizando o bombardeamento de partículas.** Sete Lagoas: Embrapa Milho e Sorgo, 2004.
- CARUSO, C. S. **Clonagem, expressão e caracterização de proteínas recombinantes de Xylella fastidiosa.** [s.l.] Universidade de São Paulo, 2007.
- CHOI, J. H.; LEE, S. Y. **Secretory and extracellular production of recombinant proteins using *Escherichia coli*.** Applied Microbiology and Biotechnology, v. 64, n. 5, p. 625–635, 2004.
- CLARK, E. D. B. **Protein refolding for industrial processes.** Current Opinion in Biotechnology, v. 12, n. 2, p. 202–207, 2001.
- CONTRERAS-GÓMEZ, A. *et al.* **Protein production using the baculovirus-insect cell expression system.** Biotechnology Progress, v. 30, n. 1, p. 1–18, 2014.
- DUMONT, J. *et al.* **Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives.** Critical Reviews in Biotechnology, v. 36, n. 6, p. 1110–1122, 2016.
- DYO, Y. M.; PURTON, S. **The algal chloroplast as a synthetic biology platform for production of therapeutic proteins.** Microbiology, v. 164, n. 2, p. 113–121, 2018.
- EGELKROUT, E.; RAJAN, V.; HOWARD, J. A. **Overproduction of recombinant proteins in plants.** Plant Science, v. 184, p. 83–101, 2012.
- ESPÍRITO SANTO, P. F. R. **Pesquisa de estirpes de *Saccharomyces cerevisiae* para expressão de proteína recombinante.** [s.l.] Universidade de Coimbra, 2013.

FARIA, R. P. V.; RODRIGUES, A. E. **Instrumental aspects of Simulated Moving Bed chromatography**. Journal of Chromatography A, v. 1421, p. 82–102, 2015.

FELBERBAUM, R. S. **The baculovirus expression vector system: A commercial manufacturing platform for viral vaccines and gene therapy vectors**. Biotechnology Journal, v. 10, n. 5, p. 702–714, 2015.

FERRIER, D. R. **Bioquímica Ilustrada**. 7. ed. Porto Alegre: Artmed, 2017.

FRANCIS, D. M.; PAGE, R. **Strategies to optimize protein expression in *E. coli***. Current Protocols in Protein Science, n. SUPPL. 61, 2010.

GECCHELE, E. *et al.* **A comparative analysis of recombinant protein expression in different biofactories: Bacteria, insect cells and plant systems**. Journal of Visualized Experiments, n. 97, 2015.

GOMES, A. R. *et al.* **An Overview of Heterologous Expression Host Systems for the Production of Recombinant Proteins**. Advances in Animal and Veterinary Sciences, v. 4, n. 7, p. 346–356, 2016.

GROVES, M. J. **Pharmaceutical Biotechnology**. 2. ed. New York: Taylor & Francis Group, LLC, 2006.

HARTLEY, J. L. **Cloning technologies for protein expression and purification**. Current Opinion in Biotechnology, v. 17, n. 4, p. 359–366, 2006.

HARVEY, A. J. *et al.* **Expression of exogenous protein in the egg white of transgenic chickens**. Nature Biotechnology, v. 20, n. 4, p. 396–399, 2002.

HAVLIK, D. *et al.* **Establishment of *Neurospora crassa* as a host for heterologous protein production using a human antibody fragment as a model product**. Microbial Cell Factories, v. 16, n. 1, p. 128, 2017.

KAUR, J.; KUMAR, A.; KAUR, J. **Strategies for optimization of heterologous protein expression in *E. coli* : Roadblocks and reinforcements**. International Journal of Biological Macromolecules, v. 106, p. 803–822, 2018.

KHAN, K. H. **Gene Expression Systems and Recombinant Protein Purification**. Research Journal of Pharmaceutical, Biological and Chemical Sciences, v. 5, n. 6, p. 450–463, 2014.

KOSOBOKOVA, E. N.; SKRYPNIK, K. A.; KOSORUKOV, V. S. **Overview of fusion tags for recombinant proteins**. Biochemistry (Moscow), v. 81, n. 3, p. 187–200, 2016.

LABROU, N. E. **Protein purification: An overview**. Methods in Molecular Biology, v. 1129, p. 3–10, 2014.

LANDOWSKI, C. P. *et al.* **Enabling low cost biopharmaceuticals: high level interferon alpha-2b production in *Trichoderma reesei***. Microbial Cell Factories, v. 15, n. 104, 2016.

LARRICK, J. W.; THOMAS, D. W. **Producing proteins in transgenic plants and animals**. Current Opinion in Biotechnology, v. 12, n. 4, p. 411–418, 2001.

LI, J.; LIN, H. **Numerical simulation of molecular uptake via electroporation**. Bioelectrochemistry, v. 82, n. 1, p. 10–21, 2011.

LI, Y. **Self-cleaving fusion tags for recombinant protein production**. *Biotechnology Letters*, v. 33, n. 5, p. 869–881, 2011.

LIU, Z. Q.; YANG, P. C. **Construction of pET-32 α (+) vector for protein expression and purification**. *North American Journal of Medical Sciences*, v. 4, n. 12, p. 651–655, 2012.

ŁOJEWSKA, E. *et al.* **Extraction and purification methods in downstream processing of plant-based recombinant proteins**. *Protein Expression and Purification*, v. 120, p. 110–117, 2016.

MAKSIMENKO, O. G. *et al.* **Use of Transgenic Animals in Biotechnology: Prospects and Problems**. *Acta Naturae*, v. 5, n. 1, p. 33–46, 2013.

NELSON, D. L.; COX, M. M. **Princípios de Bioquímica de Lehninger**. 5. ed. Porto Alegre: Artmed, 2011.

NIELSEN, J. **Production of biopharmaceutical proteins by yeast: Advances through metabolic engineering**. *Bioengineered*, v. 4, n. 4, p. 207–211, 2013.

OWCZAREK, B.; GERSZBERG, A.; HNATUSZKO-KONKA, K. **A Brief Reminder of Systems of Production and Chromatography-Based Recovery of Recombinant Protein Biopharmaceuticals**. *BioMed Research International*, v. 2019, 2019.

RAMANA, K. V.; XAVIER, J. R.; SHARMA, R. K. **Recent Trends in Pharmaceutical Biotechnology**. *Pharmaceutical Biotechnology: Current Research*, v. 1, n. 1, p. 1–10, 2017.

RAI, M.; PADH, H. **Expression systems for production of heterologous proteins**. *Current Science*, v. 80, n. 9, p. 1121–1128, 2001.

RESENDE, R. R. **Biotecnologia aplicada à saúde - vol. 2: fundamentos e aplicações**. São Paulo: Blucher, 2015.

SAMPATH KUMAR, S.; PUTTARAJU, H. P. **Improved microinjection technique for mosquito vectors**. *Indian Journal of Medical Research*, v. 136, n. 6, p. 971–978, 2012.

SARASWAT, M. *et al.* **Preparative Purification of Recombinant Proteins: Current Status and Future Trends**. *BioMed Research International*, v. 2013, p. 1–18, 2013.

SØRENSEN, H. P.; MORTENSEN, K. K. **Advanced genetic strategies for recombinant protein expression in *Escherichia coli***. *Journal of Biotechnology*, v. 115, p. 113–128, 2005.

STOGER, E. *et al.* **Sowing the seeds of success: Pharmaceutical proteins from plants**. *Current Opinion in Biotechnology*, v. 16, n. 2, p. 167–173, 2005.

STUDART-GUIMARÃES, C.; LACORTE, C.; BRASILEIRO, A. C. M. **Transformação Genética em Espécies Florestais**. *Ciência Florestal*, v. 13, n. 1, p. 167–178, 2003.

SU, X.; SCHMITZ, G.; ZHANG, M.; MACKIE, R. I.; CANN, I. K. **Heterologous gene expression in filamentous fungi**. In: GADD, G. M.; SARIASLANI, S. *Advances in Applied Microbiology*. USA: Elsevier, 2012. p 1-61.

TAUNT, H. N.; STOFFELS, L.; PURTON, S. **Green biologics: The algal chloroplast as a platform for making biopharmaceuticals**. *Bioengineered*, v. 9, n. 1, p. 48–54, 2018.

VEDADI, M. *et al.* **Biophysical characterization of recombinant proteins: A key to higher structural genomics success.** Journal of Structural Biology, v. 172, n. 1, p. 107–119, 2010.

WALSH, G. **Pharmaceutical Biotechnology Concepts and Applications.** England: John Wiley & Sons Ltd, 2007.

YAO, J. *et al.* **Plants as factories for human pharmaceuticals: Applications and challenges.** International Journal of Molecular Sciences, v. 16, n. 12, p. 28549–28565, 2015.

ZHU, J. **Mammalian cell protein expression for biopharmaceutical production.** Biotechnology Advances, v. 30, n. 5, p. 1158–1170, 2012.