

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE FARMÁCIA  
DISCIPLINA DE TRABALHO DE CONCLUSÃO DE CURSO DE FARMÁCIA**

**Clonagem, Expressão e Produção da DNA Polimerase de Alta  
Fidelidade de *Pyrococcus furiosus***

**Jordana Beatris Ludwig**

**Porto Alegre, junho de 2010.**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE FARMÁCIA  
DISCIPLINA DE TRABALHO DE CONCLUSÃO DE CURSO DE FARMÁCIA**

**Clonagem, Expressão e Produção da DNA Polimerase de Alta  
Fidelidade de *Pyrococcus furiosus***

**Jordana Beatris Ludwig**

**Trabalho de Conclusão de Curso**

**Disciplina de Estágio Curricular em Farmácia**

**Prof. Dr. Giancarlo Pasquali  
Orientador**

**M.Sc. Carolina Pereira Silveira  
Co-orientadora**

**Porto Alegre, junho de 2010.**

## AGRADECIMENTOS

Ao meu orientador Prof. Dr. Giancarlo Pasquali, um profissional exemplar, pelos ensinamentos, conselhos e comprometimento com este trabalho.

À minha co-orientadora Carolina Pereira Silveira, pela paciência, pelos inúmeros esclarecimentos e por toda ajuda nos experimentos, me orientando sem poupar esforços.

À Profa. Dra. Marilene Henning Vainstein, por aceitar o convite como banca, mas principalmente por disponibilizar o seu laboratório para o desenvolvimento prático deste estudo.

Ao Prof. Dr. Rogério Margis, pelo empréstimo da coluna de  $\text{Ni}^{2+}$ , além de aceitar o convite para banca deste trabalho.

Aos amigos do laboratório 212 pela companhia durante toda a graduação e pelo incentivo nesta etapa final.

Aos meus pais, Danilo e Seloi, pela educação dada e por sempre me apoiarem em todos os momentos da minha vida. Principalmente à minha mãe por estar sempre presente, muitas vezes deixando de lado os seus sonhos para a realização dos meus.

Aos meus irmãos por acreditarem em mim. Em especial ao Joel, pela confiança e pela oportunidade de trabalhar durante a graduação.

Ao Deividi, pelo amor, amizade, incentivo e compreensão durante os dias difíceis dessa jornada. Agradeço muito o seu companheirismo.

## RESUMO

A reação em cadeia da DNA polimerase (PCR) é uma das técnicas de biologia molecular mais importantes da atualidade e amplamente utilizada em laboratórios clínicos e de pesquisa. Considerando a alta fidelidade requerida para algumas PCRs e o alto custo para a aquisição de DNA polimerases termoestáveis de alta fidelidade, o objetivo que norteou o desenvolvimento deste trabalho foi a clonagem, a expressão e a produção da proteína *Pfu* DNA polimerase para fins comerciais. A sequência completa do gene de *Pyrococcus furiosus* codificadora da *Pfu* DNA polimerase foi projetada e adaptada para a frequência de códons de *Escherichia coli*. A síntese da sequência gênica foi solicitada à empresa GenScript. A sequência recombinante, denominada *rPfu*, foi transferida do vetor de clonagem pUC57 da GenScript aos plasmídeos de expressão de *E. coli* pET-23a(+) e pGEX-4T-1. Células competentes de *E. coli* da linhagem BL21(DE3)pLysS foram transformadas com os plasmídeos de expressão por eletroporação. A análise de extratos proteicos totais de ambas as linhagens recombinantes em gel de poliacrilamida mostrou a eficiência dos testes de expressão. Utilizando-se o vetor pET-23a(+)-*rPfu*, obteve-se a expressão de uma proteína de aproximadamente 95 kDa, sendo que 5 kDa correspondem à cauda de histidinas originária do plasmídeo. A *rPfu* produzida a partir da expressão de pET-23a(+)-*rPfu* foi purificada utilizando-se o método de cromatografia imobilizada de íons metálicos (IMAC). Utilizando-se o vetor pGEX-4T-1-*rPfu*, obteve-se a expressão de uma proteína de aproximadamente 116 kDa, sendo 26 kDa referentes à proteína de fusão, glutathione S-transferase (GST) originária do plasmídeo. Procedimentos de purificação desta versão da *rPfu*, de avaliação da atividade enzimática e de *scale-up* para a produção de quantidades comerciais da proteínas estão sendo conduzidos.

Palavras-chave: DNA polimerase, fidelidade, PCR, *Pyrococcus furiosus*, termoestável.

## LISTA DE TABELAS

<b>Tabela 1:</b> DNA polimerases termoestáveis e de alta fidelidade.....	19
<b>Tabela 2:</b> Custo para a aquisição de DNA polimerases termoestáveis e de alta fidelidade na apresentação de 500 U no mercado brasileiro .....	21
<b>Tabela 3:</b> Resultados da busca de patentes sobre DNA polimerases de alta fidelidade no INPI.....	23
<b>Tabela 4:</b> Patentes depositadas e acessíveis no EPO e no USPTO referentes à <i>Pfu</i> DNA polimerase.....	23
<b>Tabela 5:</b> Modificações de códons realizadas na sequência original da DNA polimerase de <i>P. furiosus</i> para o uso preferencial de códons de <i>E. coli</i> .....	35
<b>Tabela 6:</b> Exclusão de sítios de restrição pré-existentes na sequência original da DNA polimerase de <i>P. furiosus</i> para a adaptação a plasmídeos de expressão em <i>E. coli</i> .....	36

## LISTA DE FIGURAS

<b>Figura 1:</b> Comparação gráfica de <i>codon usage</i> de <i>P. furiosus</i> (barras vermelhas) com <i>E. coli</i> K-12 (barras pretas) .....	34
<b>Figura 2:</b> Alinhamento das sequências nucleotídicas da <i>rPfu</i> e da <i>Pfu</i> DNA polimerase original.....	37
<b>Figura 3:</b> Alinhamento das sequências peptídicas deduzidas da <i>rPfu</i> e <i>Pfu</i> DNA polimerase....	39
<b>Figura 4:</b> Análise das clivagens dos plasmídeos pET-23a(+), pGEX-4T-1 e pUC57- <i>rPfu</i> com <i>EcoRI</i> e <i>XhoI</i> por eletroforese em gel de agarose a 0,8% .....	41
<b>Figura 5:</b> Representação esquemática da ligação do fragmento da <i>rPfu</i> DNA polimerase ao vetor pET-23a(+)......	42
<b>Figura 6:</b> Representação esquemática da ligação do fragmento da <i>rPfu</i> DNA polimerase ao vetor pGEX-4T-1.....	43
<b>Figura 7:</b> Análise de minipreparações plasmidiais derivadas de colônias de <i>E. coli</i> XL1-Blue transformadas com os plasmídeos pET-23a(+) e pGEX-4T-1 ligados ao fragmento da <i>rPfu</i> DNA polimerase.....	44
<b>Figura 8:</b> Análise da expressão de pET-23a(+)- <i>rPfu</i> por SDS-PAGE.....	46
<b>Figura 9:</b> Análise da expressão de pGEX-4T-1- <i>rPfu</i> por SDS-PAGE.....	46
<b>Figura 10:</b> Análise da purificação da <i>rPfu</i> por SDS-PAGE .....	47

## SUMÁRIO

1. INTRODUÇÃO.....	9
2. OBJETIVOS GERAL E ESPECÍFICOS.....	10
3. REVISÃO BIBLIOGRÁFICA .....	11
3.1 Reação em cadeia da DNA polimerase.....	11
3.1.1 Princípio da PCR.....	11
3.1.2 Otimização do procedimento de PCR.....	12
3.1.3 Contaminações .....	13
3.1.4 Variações da técnica de PCR.....	14
3.1.5 Aplicações da PCR .....	15
3.2 DNA polimerases termoestáveis .....	16
3.3 DNA polimerases termoestáveis e de alta fidelidade para PCRs .....	18
3.4 Custos de comercialização de DNA polimerases termoestáveis e de alta fidelidade e propriedade intelectual.....	21
3.4.1 Custos de comercialização .....	21
3.4.2 Propriedade intelectual .....	22
3.5 Produção de proteínas recombinantes em <i>Escherichia coli</i> .....	25
4. MATERIAIS E MÉTODOS .....	29
4.1 Identificação de gene de <i>P. furiosus</i> potencialmente codificador da <i>Pfu</i> DNA polimerase e análises <i>in silico</i> .....	29
4.2 Projeção, adaptação e síntese da sequência nucleotídica da <i>Pfu</i> DNA polimerase.....	29
4.3 Clonagem da sequência da <i>rPfu</i> DNA polimerase em plasmídeos de expressão e transformação de <i>E. coli</i> .....	30
4.4 Extração de plasmídeos.....	31
4.5 Expressão da <i>rPfu</i> DNA polimerase.....	31
4.6 Purificação da <i>rPfu</i> DNA polimerase.....	31
4.7 Eletroforese de DNA em gel de agarose.....	32
4.8 Eletroforese de extratos proteicos em gel de poliacrilamida.....	32
4.9 Sequenciamento de DNA.....	32
5. RESULTADOS E DISCUSSÃO .....	33
5.1 Seleção e definição de sequência nucleotídica para a produção de uma DNA polimerase termoestável e de alta fidelidade para PCR .....	33

5.2 Alinhamento das sequências nucleotídicas e peptídicas da <i>rPfu</i> e <i>Pfu</i> DNA polimerases.....	35
5.3 Síntese e clonagem da sequência da <i>rPfu</i> DNA polimerase em pET-23a(+) e pGEX-4T-1 .....	40
5.4 Expressão da <i>rPfu</i> DNA polimerase em <i>E. coli</i> .....	45
5.5 Purificação da <i>rPfu</i> DNA polimerase por cromatografia em colunas de níquel.....	47
6. CONCLUSÕES E PERSPECTIVAS.....	48



## 1. INTRODUÇÃO

A reação em cadeia da DNA polimerase (PCR) é uma das mais importantes técnicas da biologia molecular, a qual permite a amplificação *in vitro* de moléculas de DNA ou cDNA de forma extremamente rápida. Os componentes básicos da reação são molécula de DNA-molde, oligonucleotídeos iniciadores ou *primers*, desoxirribonucleosídeos trifosfatados (dNTP), DNA polimerase, íon magnésio ( $Mg^{2+}$ ) cofator da enzima e tampão de reação. Este é um processo que decorre em três passos: desnaturação, anelamento e extensão, que em conjunto designam um ciclo da PCR e que se repetem em um número específico de vezes (Uhlmann *et al.*, 1998).

As DNA polimerases termoestáveis são enzimas fundamentais para a realização de PCRs, pois se mantêm ativas em altas temperaturas, permitindo a realização da reação sem a necessidade de adição de mais enzima após cada ciclo de desnaturação (Haki & Rakshit, 2003). Entre as DNA polimerases termoestáveis de uso na PCR estão aquelas de alta fidelidade, como a obtida de *Pyrococcus furiosus* (*Pfu*) e que possui atividade exonucleásica 3'-5', o que lhe permite corrigir erros introduzidos durante a polimerização com uma taxa de erro de apenas  $1,3 \times 10^{-6}$  mutação/pb/duplicação (Cline *et al.*, 1996). O alto custo para a aquisição de DNA polimerases termoestáveis de alta fidelidade disponíveis comercialmente foi o principal fator motivador para o desenvolvimento do presente trabalho.

## 2. OBJETIVO GERAL E ESPECÍFICO

O OBJETIVO GERAL proposto para este trabalho é a clonagem, a expressão e a produção da proteína *Pfu* DNA polimerase recombinante para fins comerciais.

Os seguintes OBJETIVOS ESPECÍFICOS foram propostos para o desenvolvimento desta monografia:

1. Identificação de gene(s) de *P. furiosus* potencialmente codificadores da *Pfu* DNA polimerase disponíveis em bancos de dados de sequências gênicas;
2. Projeção e síntese da sequência da *Pfu* DNA polimerase ideal para expressão em *Escherichia coli*;
3. Adaptação da sequência da *Pfu* DNA polimerase a plasmídeos de expressão de *E. coli*;
4. Expressão em *E. coli* e caracterização bioquímica e molecular da *Pfu* DNA polimerase;
5. *Scale-up* da produção da *Pfu* DNA polimerase.

### 3. REVISÃO BIBLIOGRÁFICA

#### 3.1 Reação em cadeia da DNA polimerase

A técnica da reação em cadeia da DNA polimerase (PCR, do inglês, *polymerase chain reaction*) foi desenvolvida e nomeada por Kary Mullis e colegas da Cetus Corporation, em 1985. Rapidamente, esta técnica tornou-se um procedimento de rotina em todos os laboratórios de biologia molecular para a identificação e manipulação de material genético, de clonagem, sequenciamento, mutagênese, diagnóstico e análises genéticas (Mullis, 1990).

##### 3.1.1 Princípio da PCR

A amplificação pela PCR requer uma quantidade muito pequena de DNA. Este DNA-molde é desnaturado e a solução contendo os *primers* é adicionada a ele e, pelo pareamento específico (A-T/C-G) por ligações de hidrogênio, estes *primers* anelam-se às sequências de DNA que lhe são complementares. Na presença de uma DNA polimerase termoestável e de precursores de DNA (os quatro desoxirribonucleosídeos trifosfatados: dATP, dCTP, dGTP e dTTP), sob condições ideais de pH, salinidade e o cofator da enzima [íons magnésio bivalentes ( $Mg^{2+}$ )], a síntese de novas fitas de DNA inicia, as quais serão complementares a cada uma das fitas do DNA-molde. O volume total da reação pode variar dependendo do objetivo a ser atingido. Em geral, é realizada em volumes de 10 a 50  $\mu$ l (Uhlmann *et al.*, 1998).

A PCR consiste em ciclos repetitivos de 3 etapas, e é realizada em termocicladores, equipamentos que automaticamente controlam e alternam as temperaturas durante períodos programados de tempo para o número apropriado de ciclos de PCR. Na primeira etapa, as moléculas de DNA de fita dupla separam-se completamente (**desnaturação**), formando duas fitas simples que serão usadas como moldes para os *primers* e para a DNA polimerase. Normalmente, a fita de DNA é desnaturada por 1-2 min a 94 °C. As duas cadeias de DNA são mantidas unidas por ligações de hidrogênio que, por serem relativamente fracas, quebram-se a altas temperaturas, ao passo que as ligações entre as moléculas de fosfato e desoxirribose, por serem ligações covalentes mais fortes, permanecem intactas (Uhlmann *et al.*, 1998).

Na segunda etapa, a temperatura é diminuída para permitir a hibridização ou ligação dos *primers* (**anelamento**) às sequências complementares na molécula de DNA. A temperatura de anelamento, entre 50 a 60 °C, dependendo da temperatura média de desnaturação ( $T_m$ , do inglês, *melting temperature*) dos *primers*, é a variável mais importante para determinar a especificidade da PCR, de forma que as temperaturas e os tempos de

anelamento variam de acordo com a sequência a ser amplificada. *Primers* exclusivos para cada extremidade da sequência alvo são necessários para amplificar somente a sequência de interesse. O comprimento do *primer* é geralmente de 15-30 nucleotídeos (Kim *et al.*, 2002).

Na terceira etapa (**extensão** ou **polimerização**), ocorre a síntese de uma segunda fita complementar de DNA que ocorre pela adição de dNTP por uma enzima DNA polimerase na região de anelamento de cada *primer*, numa temperatura de aproximadamente 72 °C. A extensão deve ocorrer na temperatura recomendada para atividade da enzima utilizada por um período de 30 s a 5 min, dependendo do tamanho da sequência-alvo a ser amplificada. (Kim *et al.*, 2002).

Após esse período, a temperatura é elevada mais uma vez a aproximadamente 95 °C, para que se separem as fitas duplas de curta extensão do DNA. As fitas simples tornam-se moldes para outro ciclo de síntese de DNA. Assim, o ciclo de desnaturação, anelamento e extensão é repetido por 25 a 35 vezes, produzindo-se de 100 ng a mais de 1 µg de DNA. É aconselhável executar o número mínimo de ciclos necessários para produzir uma quantidade viável de DNA para o uso pretendido, minimizando o número de mutações e obtendo-se, assim, o produto específico desejado (Eckert & Kunkel, 1991).

O resultado da PCR (amplicon) é geralmente analisado por eletroforese em gel de agarose ou de poliacrilamida. Géis de agarose a 2% são geralmente suficientes para a visualização de fragmentos de 50 a 500 pb, sendo que a resolução e a sensibilidade são aumentados com a utilização de géis de poliacrilamida numa concentração entre 6 e 10%. Métodos de eletroforese capilar também têm tornado-se frequentes na análise de amplicons (Kramer & Coen, 2001).

### 3.1.2 Otimização do procedimento de PCR

A padronização de um protocolo de PCR pode significar uma série de testes, onde se pode alterar a concentração dos reagentes básicos, adicionar-se algum potencializador, assim como alterar as condições de amplificação, principalmente a temperatura de anelamento e o tempo de alongamento da nova fita de DNA formada. Atualmente, programas de computação ou até mesmo sites de pesquisa *on-line* da internet, tais como *Basic Local Alignment Tool* (BLAST), no sítio do *National Center for Biotechnology Information* (NCBI, [www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)), podem ser usados para ajudar na análise das sequências escolhidas. A projeção dos oligonucleotídeos iniciadores é geralmente realizada observando-se algumas recomendações gerais que permitam aumentar a probabilidade de obtenção de melhores resultados como, por exemplo, conteúdo G+C similar entre os *primers*, e a projeção

de um número mínimo de estruturas secundárias (baixa complementaridade de um *primer* com o outro, particularmente na região 3').

A  $T_m$  de um *primer* para PCR pode ser calculada pela equação simples e genérica: (número de A+T) x 2 °C + (número de G+C) x 4 °C . Cálculos mais exatos levam em conta a variação de energia livre ( $\Delta G^0$ ) de cada ligação de hidrogênio entre os pares A+T e G+C, além do número de nucleotídeos. Muitos laboratórios utilizam temperaturas de anelamento de 3 °C a 5 °C abaixo da  $T_m$  dos *primers* calculada usando-se a fórmula mais simples como início para os experimentos de otimização da PCR.

A concentração de  $Mg^{2+}$  também deve ser controlada. O excesso de  $Mg^{2+}$  livre provoca a redução da fidelidade enzimática e aumento de produtos não específicos. A ausência ou pouca quantidade de  $Mg^{2+}$  mantém inativas as DNA polimerases termoestáveis (Kramer & Coen, 2001).

Alguns cossolventes podem ser utilizados para aumentar a especificidade da reação como glicerol, dimetilsulfóxido (DMSO), polietilenoglicol (PEG), betaína e formamida. Algumas amidas de baixo peso molecular também têm potencial para aumentar a especificidade da PCR (Chakrabarti & Schutt, 2001).

### 3.1.3 Contaminações

A alta sensibilidade da PCR pode constituir-se um grave problema, caso não sejam tomadas algumas medidas de segurança e procedimentos especiais com o objetivo de evitar contaminações, conforme listagem abaixo:

- (a) Preparar a reação em um laboratório isolado ou, pelo menos, em área especialmente designada do laboratório para o preparo da PCR;
- (b) Manter equipamentos e reagentes para uso na PCR separadamente do equipamento geral do laboratório e, principalmente, dos reagentes usados para a análise de produtos de PCR;
- (c) Manter estéreis todos os reagentes que podem ser autoclavados e dividir as soluções-estoque em alíquotas para armazenamento em quantidades reduzidas;
- (d) Preparar uma pré-mistura contendo todos os componentes da reação com exceção da DNA polimerase e do DNA-molde (que devem ser adicionados por último);
- (e) Usar luvas descartáveis durante o preparo da PCR;
- (f) Incluir controles negativos, para verificar possíveis contaminações, e controles positivos com amostras usadas previamente para confirmar a funcionalidade da mistura de reação previamente preparada (Erlich *et al.*, 1991).

### 3.1.4 Variações da técnica de PCR

A técnica da PCR idealizada por Kary Mullis tinha como objetivo a amplificação de uma única região específica do DNA. Entretanto, poucos anos depois começaram a ser desenvolvidas variações do protocolo básico da PCR com uma variedade de objetivos. Algumas destas técnicas estão citadas abaixo.

A “**PCR multiplex**” foi desenvolvida com o objetivo de permitir a amplificação simultânea de mais de uma região de DNA de tamanhos distintos, incluindo mais de um par de *primers* durante os ciclos de amplificação. Esta vantagem pode simplificar alguns experimentos, reduzindo tempo e trabalho. Desde a sua introdução, a “PCR multiplex” tem sido aplicada com sucesso em muitas áreas de diagnóstico, incluindo a análise de supressão de genes, a mutação e a análise de polimorfismos, a análise quantitativa e detecção de RNA (Elnifro, *et al.*, 2000).

A PCR convencional não permite a estimativa exata de valores quantitativos. Por esta razão, foi desenvolvida a **PCR em tempo real** ou PCR quantitativa (qPCR), que permite realizar a avaliação da quantidade de produto formado durante o curso da reação pelo monitoramento da fluorescência de corantes ou sondas introduzidas na reação. A fluorescência é proporcional à quantidade de produto formado, e permite registrar o número de ciclos de amplificação necessários para obter uma determinada quantidade de moléculas de DNA. Os usos típicos da PCR em tempo real incluem a detecção de patógenos, análises da expressão gênica, análises de polimorfismo de base única (SNP), análises de aberrações cromossômicas e, mais recentemente, a detecção de proteínas pela imuno-PCR em tempo real (Kubista *et al.*, 2006).

A técnica de PCR precedida de transcrição reversa (**RT-PCR**, do inglês, **Reverse Transcription-PCR**) parte de um RNA-molde que é convertido em DNA complementar (cDNA) usando uma transcriptase reversa (RT, do inglês, *reverse transcriptase*). O produto da RT, isto é, o cDNA, é a seguir submetido à PCR como DNA-molde. A RT-PCR é uma ferramenta muito sensível para estudos de expressão gênica em nível de RNA, especialmente na quantificação de RNA mensageiro (mRNA) ou de níveis de RNA viral (Uhlmann *et al.*, 1998).

A “**hot-start PCR**” foi desenvolvida ao observar-se que a mistura de todos os reagentes da PCR antes da etapa de desnaturação inicial proporcionava maiores oportunidades de ligações não específicas dos *primers* com regiões do DNA-molde. A utilização de uma

DNA polimerase ligada a um anticorpo termosensível, isto é, que se desnatura no primeiro ciclo de desnaturação e permite a ativação da enzima, reduz essa possibilidade de amplificar produtos indesejados. Como alternativa à DNA polimerase conjugada a anticorpo, pode-se realizar a adição de enzima após a desnaturação inicial. Embora permita resultados semelhantes à “*hot-start PCR*”, esta prática requer maior manipulação e, conseqüentemente, consumo de materiais e tempo (Arezi *et al.*, 2003).

A “*nested PCR*” é realizada com a utilização de *primers* internos ao primeiro par de *primers*. Os produtos de uma reação de amplificação inicial são diluídos e utilizados como molde para uma segunda PCR, na qual é utilizado outro conjunto de *primers*, correspondente às sequências localizadas próximas, mas internas àquelas utilizadas na primeira reação (Elnifro, *et al.*, 2000).

### 3.1.5 Aplicações da PCR

A PCR tem sido amplamente aplicada em várias áreas de análises da biologia molecular. Ela pode ser usada para clonar diretamente quaisquer fragmentos de DNA para os quais seja possível projetar *primers* flanqueadores. O molde para uma reação pode ser tanto DNA quanto RNA, de modo que tanto DNA genômico como cDNAs podem servir de molde para PCRs. Como conclusão, além da clonagem de fragmentos específicos ou aleatórios de DNA, a PCR permite estimar a presença ou a ausência de determinadas sequências em nível de genoma ou de transcritos, o número de cópias gênicas e o nível de expressão gênica em quantidades de mRNA originais (Erlich *et al.*, 1991).

Várias doenças de origem genética, como distrofia muscular de Duchene,  $\beta$ -talassemia, doença de Huntington, doença de Tay-Sachs, fibrose cística, hemocromatose hereditária, hemofilia, síndrome de Lesch-Nylon, têm sido diagnosticadas pela PCR (Kim *et al.*, 2002).

A técnica também possibilita o diagnóstico de infecções virais, bacterianas e aquelas causadas por protozoários circulantes de maneira rápida e precisa. Alguns patógenos que podem ser identificados incluem *Chlamydia trachomates*, Citomegalovírus, Enterovírus, *Haemophylus influenza*, *Herpes simplex*, *Mycobacterium tuberculosis*, *Neisseria gonorrhoeae*, Papiloma Vírus (HPV), *Plasmodium* sp., *Toxoplasma gondii*, *Varicella zooster*, Vírus da Hepatite B, Vírus da Imunodeficiência Humana (HIV), entre outros (Yang 2004; Speers, 2006).

A sensibilidade extrema da PCR tem grande uso na medicina forense, pois permite a determinação de traços mínimos de evidências biológicas (manchas de sangue, sêmen, cabelos,

etc.) que são frequentemente encontradas em cenas de crimes. A partir destas amostras, é possível gerar-se uma “impressão digital de DNA” (do inglês, *DNA fingerprinting* ou dactiloscopia de DNA) do indivíduo da qual a amostra proveio (Walsh, 2004).

Outra aplicação da PCR é no diagnóstico precoce do câncer, que apesar dos avanços na tecnologia de diagnóstico por imagem, tratamento cirúrgico e nas modalidades terapêuticas, continua sendo uma importante causa de mortalidade no mundo. A técnica de PCR em tempo real permite o diagnóstico de vários tipos de câncer, entre eles: de mama, de útero, do colo-retal, de bexiga, de próstata e vários tipos de leucemia e linfomas. Um diagnóstico acurado na fase inicial do câncer permite um excelente prognóstico para o paciente (Deepak *et al.*, 2007).

### 3.2 DNA polimerases termoestáveis

A introdução de DNA polimerases termoestáveis na PCR foi um dos mais importantes desenvolvimentos em pesquisa genética, e que tornou a PCR uma técnica amplamente utilizada em laboratórios clínicos e de pesquisa (Kim *et al.*, 2007). Antes da introdução destas enzimas termoestáveis, DNA polimerases como o fragmento Klenow da DNA polimerase I de *E. coli* e a T4 DNA polimerase, as quais são termolábeis, eram as enzimas utilizadas originalmente por Kary Mullis e seguidores iniciais da técnica. Isto significava que uma nova alíquota da enzima tinha de ser adicionada após cada ciclo de desnaturação e etapa de hibridização dos *primers*, o que tornava o processo demorado, trabalhoso e mais caro, além de proporcionar resultados de baixa qualidade. Com a descoberta de DNA polimerases termoestáveis como a originária da bactéria termofílica *Thermus aquaticus* (*Taq* DNA polimerase), foi possível a automatização do processo nos equipamentos termocicladores (Haki & Rakshit, 2003).

A enzima *Taq* DNA polimerase tem temperatura ótima de ação a 72 °C e é razoavelmente estável a 94 °C. Esta enzima pode ser adicionada apenas uma vez, no início da reação, e permanecerá ativa durante a série completa de ciclos de amplificação. É a enzima mais utilizada em PCRs ao redor do mundo desde sua introdução comercial (Haki & Rakshit, 2003). Desprovida de atividade exonucleásica de revisão 3'-5', a *Taq* DNA polimerase sintetiza DNA mais rápido do que as enzimas com atividade de revisão 3'-5', porém com uma maior taxa de erro. Estima-se que a *Taq* DNA polimerase incorpore um nucleotídeo errado a cada 100 replicados. Devido à sua alta processividade, versões desta enzima são escolhidas



para o sequenciamento de DNA e para procedimentos de detecção. Porém, quando a fidelidade é necessária, outras enzimas são requeridas (Vieille & Zeikus, 2001).

Outras DNA polimerases termoestáveis foram isoladas de outros organismos e estão comercialmente disponíveis para realização da PCR. Entre elas destacam-se a *Tbr*, isolada de *Thermus brocianus*; a *Tth*, isolada de *Thermus thermophilus* (Kramer & Coen, 2001); e a *BstI*, isolada de *Bacillus stearothermophilus* (Haki & Rakshit, 2003).

Algumas propriedades das DNA polimerases são essenciais para a amplificação de DNA e sequenciamento, sendo elas a termoestabilidade, a taxa de extensão, a processividade e a fidelidade. A **termoestabilidade** das DNA polimerases é essencial para sua utilização em procedimentos de amplificação. A estabilidade térmica de uma DNA polimerase é geralmente caracterizada por sua meia-vida em uma temperatura específica. A sobrevivência de uma enzima em um determinado procedimento difere significativamente, no entanto, dependendo do meio e de outras condições de reação como, por exemplo, quantidade de detergentes, betaína ou outros açúcares, pode-se aumentar a estabilidade das DNA polimerases. Assim, apesar dos dados de termoestabilidade de algumas polimerases já publicados, é importante assegurar que a polimerase permaneça ativa nas condições operacionais selecionadas (Pavlov *et al.*, 2004).

A **taxa de extensão** específica representa o número de nucleotídeos polimerizados por segundo por molécula de DNA polimerase. Esta taxa depende fortemente dos meios de reação e do DNA-molde. Algumas moléculas de DNA podem até mesmo formar estruturas que impedem o alongamento das cadeias a partir dos *primers* (Pavlov *et al.*, 2004).

O conceito de **processividade** foi introduzido como a probabilidade de que, após a incorporação de um nucleotídeo, a DNA polimerase não se dissocie do DNA quando se transloca para a próxima posição. A processividade caracteriza o número médio de nucleotídeos incorporados antes da dissociação da enzima do DNA-molde. Assim como a taxa de extensão específica, a processividade também depende dos componentes do meio da reação e da sequência de DNA-molde. Muitas vezes, aumentar a concentração de uma DNA polimerase, mesmo com baixa processividade, aumenta o comprimento médio dos produtos sintetizados a partir de um longo DNA-molde. A processividade da *Taq* DNA polimerase, a qual pertence à família A de DNA polimerases, é relativamente alta, com cerca de 40 nucleotídeos. Já as DNA polimerases de *P. furiosus* e *Thermococcus litoralis* (“Vent”), entre outras, estas pertencentes à grande família B, têm uma processividade baixa de 4-30 nucleotídeos (Pavlov *et al.*, 2004).

A **fidelidade** é uma propriedade intrínseca das DNA polimerases e é definida como a capacidade de inserir o nucleotídeo correto à fita em extensão, mantendo fielmente o pareamento Watson-Crick (Biles & Connolly, 2004). A eficiência com que DNA polimerases inserem um nucleotídeo correto varia  $10^7$  vezes, com efeitos correspondentes na especificidade. Assim, DNA polimerases ineficientes, ou seja, aquelas com uma baixa eficiência para inserção do nucleotídeo correto, apresentam baixa fidelidade, ao passo que as DNA polimerases eficientes possuem uma alta fidelidade (Pavlov *et al.*, 2004).

A reduzida capacidade de polimerases pouco fiéis de inserir nucleotídeos corretos significa que essas enzimas são controladas cineticamente pela virtude de sua eficiência catalítica para inserção de nucleotídeos corretos. Isto leva à interessante conclusão de que polimerases utilizadas em PCR de alto desempenho, ou seja, aquelas que produzem um rendimento elevado de produtos, criarão um grande número de pareamentos incorretos durante a amplificação, devido à grande quantidade de DNA sintetizado. Felizmente, muitas DNA polimerases termoestáveis têm uma atividade exonucleásica 3'-5' de revisão associada. O equilíbrio de nucleotídeos, a concentração, o pH, as temperaturas da reação, o sal, a concentração de metais bivalentes e a sequência do DNA-molde afetam significativamente a fidelidade da reação (Pavlov *et al.*, 2004).

Embora o processo da PCR desenvolveu-se rapidamente por intermédio da invenção de melhores instrumentos, a necessidade de fidelidade manteve-se um sério desafio. Sabe-se que as cinco atividades distintas em que erros durante a replicação podem ocorrer são (i) a taxa de formação de ligação fosfodiéster; (ii) a ligação do dNTP pela polimerase; (iii) a taxa de liberação de pirofosfato; (iv) a contaminação após a incorporação errônea; e (v) a capacidade de revisão 3'-5' (Haki & Rakshit, 2003).

### **3.3 DNA polimerases termoestáveis e de alta fidelidade para PCRs**

Considerando a alta fidelidade requerida para algumas PCRs, várias enzimas foram desenvolvidas para suprir esta necessidade. As mais conhecidas comercialmente estão apresentadas na Tabela 1. Mais de 50 genes de DNA polimerases foram clonados a partir de vários organismos, incluindo os termofílicos e outras arqueobactérias (Kim *et al.*, 2007).

**Tabela 1:** DNA polimerases termoestáveis e de alta fidelidade.

DNA polimerase	Organismo	Referência
<i>Psp</i> “Deep Vent”	<i>Pyrococcus</i> sp. GB-D	Cline <i>et al.</i> (1996)
<i>Pfu</i>	<i>Pyrococcus furiosus</i>	Lundberg <i>et al.</i> (1991)
<i>Pwo</i>	<i>Pyrococcus woesei</i>	Frey & Suppman (1995)
<i>Tma</i>	<i>Thermotoga maritima</i>	Bost <i>et al.</i> (1994)
<i>Tli</i> “Vent”	<i>Thermococcus litoralis</i>	Perler <i>et al.</i> (1992)

Fonte: Haki & Rakshit, 2003

DNA polimerases derivadas de duas arqueobactérias hipertermofílicas, *Pyrococcus* e *Thermococcus*, apresentaram uma recente popularidade para uso em PCR, porque oferecem maior fidelidade do que a *Taq* DNA polimerase. Em geral, a família B de DNA polimerases tem atividade exonucleásica 3'-5' e permite amplificar o DNA-molde com alta fidelidade. No entanto, as enzimas desta família apresentam uma baixa taxa de alongamento em comparação com a família A. Este problema tem sido superado pela otimização dos tampões de reação e a construção de versões mutantes com diminuição da atividade exonucleásica 3'-5' (Kim *et al.*, 2007).

A DNA polimerase isolada de *P. furiosus* (*Pfu*) é termoestável e termofílica, otimamente ativa a 75 °C. A *Pfu* DNA polimerase apresenta as seguintes características de DNA polimerases termoestáveis: possui termoestabilidade de 6 h a 95 °C; uma taxa de extensão de 25 bases/s; e uma processividade menor que 20 bases (Takagi *et al.*, 1997). Esta enzima possui atividade polimerásica 5'-3' e atividade exonucleásica de correção no sentido 3'-5', sendo útil nas reações que exigem alta fidelidade de amplificação de DNA de até 25 kb (Cline *et al.*, 1996).

A *Pfu* DNA polimerase apresenta maior fidelidade na PCR quando comparada com outras DNA polimerases de alta fidelidade, com uma taxa média de erro de  $1,3 \times 10^{-6}$  mutação/pb/duplicação, ou seja, cerca de oito vezes mais precisa do que *Taq* DNA polimerase (Kim *et al.*, 2008). As taxas médias de erro das DNA polimerases termoestáveis foram encontradas na seguinte ordem: *Pfu* ( $1,3 \times 10^{-6}$ ) < *Deep Vent* ( $2,7 \times 10^{-6}$ ) < *Vent* ( $2,8 \times 10^{-6}$ ) < *Taq* ( $8,0 \times 10^{-6}$ ) << *UITma* ( $5,5 \times 10^{-5}$ ). Estes resultados, obtidos por Cline *et al.* (1996) estão de acordo com as taxas médias de erro obtidas por Flaman *et al.* (1994).

Cline *et al.* (1996) tentaram melhorar ainda mais a fidelidade da *Pfu*, otimizando as condições da PCR. As menores taxas de erro para a *Pfu* foram observadas quando

amplificações de PCR foram realizadas na presença de 2-3 mM de sulfato de magnésio ( $\text{MgSO}_4$ ), 100-300 mM de cada dNTP e uma faixa de pH entre 8,5 e 9,1.

Atualmente, muitas estruturas de DNA polimerases foram determinadas, demonstrando-se a alta conservação entre elas. Os componentes estruturais de DNA polimerases têm sido geralmente divididos em três domínios distintos, denominados dedos, palma e polegar. Vários aspectos da replicação do DNA, como o substrato de ligação, a transferência de nucleotídeos, a fidelidade e a processabilidade, têm sido propostos a partir dos complexos das estruturas binárias/terciárias de DNA polimerases com o DNA de diversos substratos. Os domínios dedos e polegar mudam de posição dependendo da ligação da polimerase ao seu substrato. A ligação da DNA polimerase abre as conformações dos domínios dedos e polegar. Quando o substrato é ligado, os dois domínios avançam em direção ao domínio palma para manter mais fortemente o pareamento entre as fitas *primer*/molde.

Na *Pfu* DNA polimerase existem cinco domínios, denominados dedos, palma, polegar, exonuclease e domínio N-terminal. O domínio exonuclease também muda sua conformação dependendo da replicação do DNA ou edição. Quando um nucleotídeo incompatível é incorporado nas fitas recém sintetizadas, as fitas *primer*/molde ligam a polimerase mais fracamente, ou são desalinhadas com o respectivo sítio ativo da polimerase. Eventualmente, a dupla hélice desenrola-se e nucleotídeos incompatíveis são movidos para o sítio ativo do domínio exonuclease e são extirpados. Na conformação fechada do DNA ligado no domínio polegar, a parte 3' final da fita do *primer* não pode ligar no domínio exonuclease devido ao impedimento estérico causado principalmente pela borda do domínio polegar. Apenas as conformações abertas do domínio polegar permitem a ligação de fitas únicas de DNA *primer* (DNA fita-simples ou ssDNA, do inglês, *single-stranded DNA*) no sítio ativo exonuclease. Em geral, as conformações de cada domínio são bem coordenadas com os outros para realizar o processo de replicação do DNA (Kim *et al.*, 2008).

Algumas questões que ainda hoje estimulam cientistas na área, portanto, são as melhorias exigidas na fidelidade da PCR, o desenvolvimento de estratégias para o uso mais econômico das enzimas, assegurando a rápida multiplicação de fitas menores. Uma DNA polimerase termoestável tendo estas características certamente melhorará os resultados obtidos pela PCR (Haki & Rakshit, 2003).

### 3.4 Custos de comercialização de DNA polimerases termoestáveis e de alta fidelidade e propriedade intelectual

#### 3.4.1 Custos de comercialização

Considerando o alto custo para a aquisição de DNA polimerases termoestáveis e de alta fidelidade disponíveis no mercado, foi realizada uma pesquisa de preços das principais enzimas de alta fidelidade comercializadas no mercado brasileiro. Na Tabela 2 estão apresentadas estas enzimas e seus custos para aquisição. Como conclusão, podemos observar que os valores de 500 unidades enzimáticas (U) destas enzimas no mercado brasileiro variam de R\$ 555,00 a aproximadamente R\$ 2.000,00.

**Tabela 2:** Custo para a aquisição de DNA polimerases termoestáveis e de alta fidelidade na apresentação de 500 U no mercado brasileiro.

Descrição	Concentração	Marca	Preço R\$
Cloned <i>Pfu</i> DNA Polymerase	2,5 U/ $\mu$ l	Stratagene	1.997,00
Deep Vent <sub>R</sub> <sup>TM</sup> DNA Polymerase	2 U/ $\mu$ l	New England	797,00
Fideli <i>Taq</i> <sup>TM</sup> DNA Polymerase	5 U/ $\mu$ l	GE Healthcare	1.232,00
Gene Amp <sup>®</sup> High Fidelity PCR System	5 U/ $\mu$ l	Applied Biosystems	1.411,00
KOD Hi-Fi DNA Polymerase	5 U/ $\mu$ l	Novagen	Descontinuada
<i>Pfu</i> DNA Polymerase	2-3 U/ $\mu$ l	Promega	Descontinuada
<i>Pfu</i> DNA Polymerase (recombinant)	5 U/ $\mu$ l	Fermentas	1.120,00
<i>Pfx</i> 50 <sup>TM</sup> DNA Polymerase	5 U/ $\mu$ l	Invitrogen	678,00
Phusion <sup>®</sup> High Fidelity DNA Polymerase	2 U/ $\mu$ l	New England	1.927,00
Platinum <sup>®</sup> <i>Taq</i> DNA Polymerase High Fidelity	5 U/ $\mu$ l	Invitrogen	785,00
Platinum <sup>®</sup> <i>Pfx</i> DNA	2,5 U/ $\mu$ l	Invitrogen	877,00

Polymerase			
Vent <sub>R</sub> <sup>®</sup> DNA Polymerase	2 U/μl	New England	555,00

### 3.4.2 Propriedade intelectual

A aplicação em escala industrial e empresarial dos avanços científicos e tecnológicos advindos da pesquisa biológica constitui o chamado setor de biotecnologia. O mesmo representa um dos setores mais promissores da economia e os avanços biotecnológicos têm potencial comprovado de impactar diretamente a qualidade de vida da população e gerar desenvolvimento econômico e social. Como exemplos deste fantástico impacto pode-se citar o advento de animais e vegetais transgênicos, o desenvolvimento de novos medicamentos, vacinas e terapias, bem como a produção de adjuvantes alimentares e nutracêuticos. O uso estratégico da propriedade intelectual é importante para se assegurar maior competitividade à biotecnologia nacional, e que os benefícios dos investimentos em biotecnologia sejam revertidos em prol do desenvolvimento nacional.

Uma busca por patentes sobre sequências e proteínas recombinantes de DNA polimerases realizada no sítio do Instituto Nacional de Propriedade Intelectual (INPI) na internet ([www.inpi.gov.br](http://www.inpi.gov.br)), acessado dia 01/05/2010, revelou a inexistência de documentos de proteção intelectual sobre estes temas. Nesta busca, foram pesquisados nos títulos e nos resumos das obras depositadas os seguintes termos: DNA polimerase, *Pfu*, *Pyrococcus furiosus*, termoestável. O resultado da pesquisa está mostrado na Tabela 3.

Realizou-se também uma busca, com os mesmos termos pesquisados no INPI, no sítio da base de dados com informações bibliográficas de documentos de patentes de vários países da América Latina (LATIPAT, <http://lp.espacenet.com/>) acessado no dia 01/05/2010. Nesta base de patentes, também não foi encontrada qualquer patente da sequência em estudo.

**Tabela 3:** Resultados da busca de patentes sobre DNA polimerases de alta fidelidade no INPI.

Termo pesquisado	Local pesquisado	Resultado (Número de processos)
<i>Pfu</i>	Título e resumo	0
<i>Pyrococcus furiosus</i>	Título e resumo	0
DNA polimerase	Título	14
	Resumo	2
Termoestável	Título	21
	Resumo	37

Finalmente, para a verificação de patentes internacionais sobre os temas, uma busca foi realizada com a sequência peptídica da *Pfu* DNA polimerase no sítio conjunto do *European Patent Office* (EPO) e do *United States Patent and Trademark Office* (USPTO, <http://www.ebi.ac.uk/Tools/fasta33/index.html>). Como resultado foram encontradas 20 patentes, conforme mostrado na Tabela 4.

**Tabela 4:** Patentes depositadas e acessíveis no EPO e no USPTO referentes à *Pfu* DNA polimerase.

Base de dados	Empresa	Número da patente	Data	Resumo
EPO	Applera Corporation (US)	WO2006074288	13/07/2006	<i>Polypeptides having nucleic acid binding activity</i>
EPO	Applera Corporation (US)	WO2006074233	13/07/2006	<i>Polypeptides having nucleic acid binding activity and compositions and methods for nucleic acid amplification</i>
EPO	Bio-Rad Laboratories, Inc (US)	EP1832652	12/07/2007	<i>Improved nucleic acid modifying enzymes</i>
EPO	Genesys Ltd (GB)	WO2009087394	16/07/2009	<i>CREN7 Chimeric Protein</i>
EPO	Medical Research Council (US)	WO2008029085	13/03/2009	<i>Polymerase</i>
EPO	Roche Diagnostics	EP1350841	08/10/2003	<i>Thermostable or thermoactive DNA</i>

	GmnH (DE)			<i>polymerase with attenuated 3'-5' exonuclease activity</i>
EPO	Stratagene California (US)	EP2110432	21/10/2009	<i>High fidelity DNA polymerase compositions and uses therefor</i>
EPO	University of Newcastle Upon Tyne (GB)	WO03089637	30/10/2003	<i>Mutation of DNA Polymerase from Archaeobacteria</i>
EPO	University of Newcastle Upon Tyne (GB)	WO2006030174	23/03/2006	<i>DNA Polymerase</i>
USPTO	Amersham Life Science, Inc	US5827716	27/10/1998	<i>Modified pol-II type DNA polymerases</i>
USPTO	Bio-Rad Laboratories, Inc	US7541170	02/06/2009	<i>Nucleic acid modifyind enzymes</i>
USPTO	Bio-Rad Laboratories, Inc	US7560260	14/07/2009	<i>Compositions with polymerases activity</i>
USPTO	MJ Bioworks, Inc.	US6627424	30/07/2003	<i>Nucleic acid modifyind enzymes</i>
USPTO	Roche Molecular Systems, Inc.	US7148049	12/12/2006	<i>Thermostable or thermoactive DNA polymerase molecules with attenuated 3'-5' exonuclease activity</i>
USPTO	Stratagene	US5489523	06/02/1996	<i>Exonuclease-deficient thermostable Pyrococcus furiosus DNA polymerase I</i>
USPTO	Stratagene	US5545552	13/08/1996	<i>Purified thermostable P. furiosus DNA pol I</i>
USPTO	Stratagene	US5866395	02/02/1999	<i>Purified thermostable P. furiosus DNA pol I</i>
USPTO	Stratagene	US5948663	07/09/1999	<i>Purified thermostable P. furiosus DNA pol I</i>
USPTO	Stratagene	US6489150	03/12/2002	<i>Purified thermostable P. furiosus DNA pol I</i>
USPTO	Stratagene	US7045328	16/05/2006	<i>Purified thermostable P. furiosus DNA pol I</i>



Como conclusões das buscas realizadas para proteção intelectual sobre sequências e proteínas recombinantes da *Pfu* DNA polimerase, a empresa Stratagene é a principal titular de patentes depositadas no EPO e no USPTO. Porém, no âmbito do mercado brasileiro, não há patentes equivalentes depositadas. Segundo nossa compreensão da legislação brasileira sobre propriedade intelectual (Lei da Propriedade Industrial 9.279/96, em vigor desde 15 de maio de 1997 em substituição à Lei 5772/71, e atualizada de acordo com a Lei 10.196/01, disponível em [http://www.inpi.gov.br/menu-esquerdo/patente/pasta\\_legislacao](http://www.inpi.gov.br/menu-esquerdo/patente/pasta_legislacao), a produção comercial de qualquer versão de DNA polimerases no Brasil, incluindo-se a *Pfu* DNA polimerase, é permitida sem a necessidade de coleta de *royalties*, uma vez que tais invenções não estão protegidas no país. Com respeito aos produtos do desenvolvimento do presente projeto de pesquisa, caso obtenha-se sucesso na produção de uma *Pfu* DNA polimerase funcional e com potencial comercial, uma consulta a órgãos e profissionais especializados em proteção intelectual será realizada antes da comercialização.

### **3.5 Produção de proteínas recombinantes em *Escherichia coli***

A bactéria *E. coli* foi o primeiro organismo utilizado como célula hospedeira e é frequentemente a primeira escolha para a produção de proteínas recombinantes e engenharia metabólica. Além de ser o organismo padrão em laboratórios de biologia molecular no mundo inteiro para expressar proteínas, uma parcela significativa de proteínas comercializadas para fins terapêuticos, aproximadamente 40%, é feita em sistemas de expressão em *E. coli* (Ni & Chen, 2009).

Essa bactéria é um sistema de expressão altamente versátil devido à rápida multiplicação celular, a alta velocidade de produção de proteínas e ao extenso conhecimento fisiológico. Porém, cada proteína a ser expressa representa um desafio, independentemente do organismo hospedeiro, e o alto nível de síntese deve ser otimizado em cada situação por meio de variações de diferentes parâmetros de cultura, indução e estabilização dos produtos (Swartz, 2001).

De fato, segundo Sivashanmugam *et al.* (2009), as vantagens de rápida multiplicação, com densidade elevada em um meio de baixo custo, a genética bem caracterizada e a disponibilidade de um grande número de vetores de clonagem e cepas mutantes, permitiram que *E. coli* oferecesse um meio rápido, com alto rendimento econômico para produção de proteínas recombinantes.

Aproximadamente 80% das proteínas produzidas para resolução de suas estruturas tridimensionais, e cujos resultados são submetidos ao Banco de Dados de Proteínas (PDB, do

inglês, *Protein Data Bank*, <http://www.pdb.org/pdb/home/home.do>), são oriundas de sistemas de expressão de *E. coli*. (Sorensen & Mortensen, 2005). A expressão de proteínas recombinantes em *E. coli* depende, naturalmente, de vários fatores para garantir a obtenção de alta produtividade com baixo custo agregado, tais como as cepas bacterianas, um forte promotor, a estabilidade de mRNA e do plasmídeo, a frequência de códons do gene de interesse, a condição adequada de cultura como temperatura, aeração e composição dos meios (Jana & Deb, 2005). A cepa de *E. coli* para a expressão recombinante é altamente importante. Cepas para expressão devem ser deficientes de proteases prejudiciais, devem manter a expressão do plasmídeo estável e conferir os elementos genéticos relevantes para o sistema de expressão. *E. coli* BL21 é o hospedeiro mais comum no atendimento a estes requisitos e tem se revelado um padrão para aplicações de expressão recombinante. *E. coli* BL21 é uma cepa robusta, capaz de multiplicar-se vigorosamente em meios mínimos. Adicionalmente, é não-patogênica e improvável de sobreviver em tecidos humanos ou animais ou causar doenças e intoxicações (Sorensen & Mortensen, 2005).

Um promotor gênico eficiente deve apresentar várias características desejáveis. Entre elas, deve ser forte, ter um baixo nível de expressão basal, ou seja, ser fortemente regulado, ser facilmente transferível para outras cepas de *E. coli* para facilitar os ensaios de rendimento de proteína com um grande número de linhagens, e sua indução deve ser realizada de forma simples e rentável (Jana & Deb, 2005). O promotor gênico T7, derivado de bacteriófagos lambda T7 e presente no sistema de expressão pET (comercializado pela Novagen) é, de longe, o mais utilizado na preparação de proteínas recombinantes. Em 2003, o sistema pET representou mais de 90% dos protocolos de preparação de proteínas cujas estruturas foram depositada no PDB. Os sistemas que utilizam o promotor PL/repressor cI como, por exemplo, pLEX (Invitrogen), o promotor TRC como, por exemplo, pTrc (GE Healthcare), o promotor Tac como, por exemplo, pGEX (GE Healthcare) e híbridos lac/T5 como, por exemplo, pQE (Qiagen) são também comumente utilizados em laboratórios de biologia molecular para a produção de proteínas recombinantes em *E. coli* (Sorensen & Mortensen, 2005).

Os promotores comumente empregados para a expressão heteróloga de proteínas exigem a adição de uma molécula indutora, o esgotamento ou a adição de um nutriente ou, ainda, a mudança de um fator físico ou físico-químico como o pH ou a temperatura. O isopropil- $\beta$ -D-tiogalactopiranosídeo (IPTG) é um forte indutor de genes sob o controle do promotor *lac*. É, na realidade, um inibidor da atividade da proteína repressora LacI nos sistemas de expressão *lac*. No entanto, devido à sua toxicidade e alto custo, ele é indesejável na produção de proteínas recombinantes em grande escala ou para finalidades terapêuticas.

Uma alternativa é o uso da molécula análoga e original lactose, pelo seu baixo custo e por não ser tóxica às células (Jana & Deb, 2005).

Os sistemas de termo regulação são particularmente adequados para a produção de proteínas em grande escala, pois são altamente produtivos, podem ser facilmente regulados, minimizam a manipulação das culturas, e evitam o uso de produtos químicos indutores. Apesar destas vantagens, estes sistemas ainda apresentam muitos problemas relacionados com as respostas fisiológicas durante a indução de proteínas recombinantes, o que causa um maior consumo de energia e carga metabólica, afetando a produtividade e a qualidade do produto final (Cruz *et al.*, 2010)

A estabilidade do mRNA heterólogo é muitas vezes comprometida por várias diferentes formas de ribonucleases (RNases) celulares que participam da degradação do RNA em *E. coli*, incluindo endonucleases (RNase E, RNase K e RNase III) e 39 exonucleases, principalmente RNase II e polinucleotídeo fosforilase (PNPase). A proteção de mRNAs das diversas RNases bacterianas depende da flexibilidade do RNA, da proteção por ribossomos e da modulação da estabilidade do mRNA por poliadenilação (Jana & Deb, 2005; Sorensen & Mortensen, 2005).

A estabilidade do plasmídeo é influenciada por suas propriedades, incluindo tamanho e seqüência de nucleotídeos, pela origem genética do hospedeiro, bem como pela temperatura da cultura, da taxa de multiplicação e das concentrações dos substratos. Uma estratégia para aumentar a estabilidade dos plasmídeos e tornar a expressão mais eficiente é o emprego de um plasmídeo “fora de controle” (do inglês, *runaway*), no qual a replicação pode ser induzida de forma independente à multiplicação bacteriana e em fase específica da fermentação, permitindo mais de 1.000 cópias do plasmídeo por célula hospedeira na fase de expressão (Schmidt, 2005).

Nos genes endógenos de *E. coli*, assim como nos genes de todos os demais organismos vivos, há naturalmente combinações de códons mais frequentes do que outras. Assim, existem os denominados códons principais e os códons raros. Os códons AGG/AGA, AUA, CUA, CGA e CCC geralmente ocasionam problemas de erros de tradução, pois são encontrados em baixa frequência em *E. coli* (Jana & Deb, 2005). Quando os produtos da transcrição contêm códons raros, geralmente ocorrem problemas na tradução durante a fase de produção, levando à redução na quantidade e na qualidade da proteína recombinante. Alguns destes erros incluem a substituição de aminoácidos, a troca de fase de leitura em que um ou alguns pares de nucleotídeos são eliminados (“deletados”) ou inseridos, ou a terminação prematura da tradução (Sorensen & Mortensen, 2005). Algumas estratégias para contornar o problema do

uso preferencial de códons (do inglês, *codon usage*) é a cotransformação da célula hospedeira com um plasmídeo contendo a sequência do tRNA do anticódon raro ou projetar um gene sintético com o uso de códons otimizados (Jana & Deb, 2005).

Por último, mas não menos importante, a composição do meio deve ser cuidadosamente formulada e monitorada, porque pode ter importantes efeitos metabólicos sobre a produção de células e proteína recombinante. Em geral, a multiplicação de *E. coli* é inibida quando os seguintes nutrientes estão presentes acima de determinadas concentrações: glicose (50 g/l), amônia (3 g/l), ferro (1,15 g/l), magnésio (8,7 g/l), fósforo (10 g/l) e zinco (0,038 g/l). Portanto, estes e outros componentes dos meios, sejam eles substratos ou produtos do metabolismo celular, devem ser controlados (Lee, 1996).

A temperatura também é uma importante variável que pode ser utilizada no controle do metabolismo celular. Ao se reduzir a temperatura do meio de cultivo de *E. coli* de 37 °C para 26-30 °C, uma redução na captação de nutrientes e na velocidade de multiplicação podem ocorrer. Em contrapartida, a redução da temperatura de cultivo promove a redução da formação de subprodutos tóxicos, muitas vezes incrementando a quantidade e a qualidade da proteína recombinante produzida (Lee, 1996).

Várias técnicas de cultura de *E. coli* de alta densidade celular têm sido desenvolvidas para melhorar a produtividade e também para proporcionar vantagens como a redução do volume de cultura, a redução no investimento em equipamentos e nos custos de produção. A obtenção do cultivo de alta densidade celular é iniciada pelo estudo em biorreator, que frequentemente opera em regime de bateladas alimentadas para atingir elevadas concentrações celulares. Culturas celulares de alta densidade devem, portanto, ser consideradas como estratégia para melhorar o rendimento e a qualidade da proteína recombinante produzida, conforme sua natureza (Lee, 1996).

## 4. MATERIAIS E MÉTODOS

### 4.1 Identificação de gene de *P. furiosus* potencialmente codificador da *Pfu* DNA polimerase e análises *in silico*

A sequência nucleotídica de *P. furiosus* codificadora da DNA polimerase foi obtida do banco de dados de sequências gênicas do NCBI pelo emprego das palavras-chave “*Pyrococcus furiosus*” e “*polymerase*”.

Com o objetivo de atestar a funcionalidade da enzima codificada, foram realizadas análises *in silico* no banco de dados Pfam 24.0 (Instituto Sanger, [www.sanger.ac.uk/search/sequence](http://www.sanger.ac.uk/search/sequence)) e nos bancos de dados COG e CGAP Kegg disponíveis pelas páginas do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais ([www.biotec.icb.ufmg.br/pct/proc.php](http://www.biotec.icb.ufmg.br/pct/proc.php)).

### 4.2 Projeção, adaptação e síntese da sequência nucleotídica da *Pfu* DNA polimerase

Os códons da sequência nucleotídica original codificadora da DNA polimerase de *P. furiosus* foram comparados com a frequência de uso de códons (*codon usage*) de *E. coli* K-12 por intermédio de ferramenta *Graphical Codon Usage Analyser 2.0* ([www.gcua.de](http://www.gcua.de)). Para verificar a correspondência dos códons com os aminoácidos foi utilizado o programa Expasy (<http://expasy.org/tools/dna.html>). Após a otimização manual dos códons de *P. furiosus* para *E. coli*, a sequência otimizada da *Pfu* DNA polimerase para expressão em *E. coli* passou a se chamar *rPfu* (do inglês, *recombinant Pfu*).

Os plasmídeos escolhidos para a adaptação da sequência e expressão da *rPfu* em *E. coli* foram pET-23a(+) (Novagen) e pGEX-4T-1 (GE Healthcare). Estes plasmídeos permitem a adição de uma cauda de histidinas (*His-tag*) e da glutatona S-transferase (GST), respectivamente, permitindo que a proteína de interesse seja mais facilmente purificada. Para tanto, as enzimas de restrição escolhidas foram *EcoRI* e *XhoI* (Promega), permitindo que a sequência mantivesse a fase de leitura das sequências codificadoras da *His-tag* e da GST de ambos os plasmídeos. A confirmação das fusões transduccionais foi realizada manualmente e pelo emprego do programa Vector NTI Advanced<sup>TM</sup> versão 9 (Invitrogen).

Para comparação e alinhamentos múltiplos das sequências de nucleotídeos e peptídica de *Pfu* e *rPfu*, utilizou-se o programa ClustalW ([www.ebi.ac.uk/clustalw](http://www.ebi.ac.uk/clustalw)).

A síntese da sequência final projetada e otimizada (*rPfu* DNA polimerase) foi solicitada à empresa GenScript ([www.genscript.com](http://www.genscript.com)). Uma massa total de 4 µg do plasmídeo

de clonagem pUC57 contendo a sequência de interesse foram fornecidos pela empresa GenScript.

### **4.3 Clonagem da sequência da *rPfu* DNA polimerase em plasmídeos de expressão e transformação de *E. coli***

As reações de clivagem de DNA plasmidial foram realizadas utilizando-se o tampão adequado à melhor eficiência para clivagem de cada enzima, conforme recomendações dos fabricantes e as regras gerais para reações com endonucleases segundo Sambrook & Russel (2001). Primeiramente foi realizada a clivagem de 3 µg dos plasmídeos pUC57-*rPfu*, pET-23a(+) e pGEX-4T-1 com 5 U da enzima de restrição *EcoRI*, incubando-se as reações a 37 °C por 1 h e 30 min e precipitando-se com acetato de amônio e etanol por 2 h a -20 °C, seguindo-se centrifugação a 14.000 rpm por 15 min. Os sobrenadantes foram descartados e os precipitados foram secos por 1 h à temperatura ambiente. Após estas etapas, procedeu-se as reações dos plasmídeos com a enzima de restrição *XhoI*. Os produtos das reações foram analisados por eletroforese em gel de agarose. A banda de DNA correspondente ao gene da *rPfu* DNA polimerase foi purificada de gel a partir do produto de clivagem de pUC57-*rPfu* utilizando-se o “*GFX PCR DNA and Purification Band Gel Kit*” (GE Healthcare) segundo orientações do fabricante. Reações de ligação do fragmento da *rPfu* DNA polimerase com os plasmídeos pET-23a(+) e pGEX-4T-1, ambos clivados com *EcoRI* e *XhoI*, foram realizadas com a T4 DNA ligase (Promega) segundo orientações do fabricante. As reações de ligação de DNA foram procedidas utilizando-se uma proporção molar vetor: fragmento de 3:1. As reações foram incubadas a 22 °C por 1 h e inativadas a 65 °C por 10 min.

Aproximadamente 100 ng de plasmídeos ou 2 µl dos produtos das reações de ligação foram misturados a 50 µl de células eletrocompetentes de *E. coli* XL1-Blue. As células foram submetidas à eletroporação em eletroporador (Eppendorf 2510) a 2.500 V durante 5 ms. A seguir, as suspensões de células foram inoculadas em 1 ml de meio líquido Luria-Bertani (LB) e incubadas a 37 °C a 200 rpm. Após 1 h, 100 µl de células foram distribuídas em placas de Petri contendo meio LB sólido acrescido de ampicilina a 100 µg/ml. As placas foram incubadas a 37 °C por 16 h. Plasmídeos foram purificados a partir de 12 colônias formadas, resistentes ao antibiótico (recombinantes), derivadas de células transformadas com pET-23a(+)-*rPfu* ou pGEX-4T-1-*rPfu*.

#### 4.4 Extração de plasmídeos

Plasmídeos foram purificados pelo método da lise alcalina a partir de 1,5 ml de culturas bacterianas incubadas por 16 h a 37 °C e agitação a 200 rpm em meio LB líquido contendo 100 µg/ml de ampicilina essencialmente como descrito em Sambrook & Russel (2001). A quantificação dos plasmídeos e do fragmento de DNA da *rPfu* foi realizada no aparelho Qubit<sup>®</sup> (Invitrogen), conforme orientações do fabricante.

#### 4.5 Expressão da *rPfu* DNA polimerase

Os plasmídeos pET-23a(+)-*rPfu* e pGEX-4T-1-*rPfu* foram utilizados para transformar por eletroporação linhagens de *E.coli* BL21(DE3)pLysS. Uma colônia transformada foi inoculada em frasco Erlenmeyer de 125 ml contendo 10 ml de meio LB líquido suplementado com 100 µg/ml de ampicilina. A cultura foi incubada a 37 °C com agitação de 200 rpm por 16 h. Após este período, 1 ml da cultura foi inoculado em frasco Erlenmeyer de 500 ml contendo 50 ml de meio LB e ampicilina a 100 µg/ml. A cultura foi incubada a 37 °C em incubador giratório com agitação de 200 rpm até atingir uma densidade óptica a 600 nm (OD<sub>600</sub>) de 0,5. Em seguida, um volume de IPTG foi adicionado à cultura para obter-se uma concentração final de 1 mM. A indução foi testada em temperatura de 37 °C por 4 h, sendo retiradas alíquotas a cada hora. Após a indução, as células foram recuperadas por centrifugação a 3.000 g, a 4 °C durante 15 min. Os sedimentos celulares foram congelados e armazenados a -20 °C.

#### 4.6 Purificação da *rPfu* DNA polimerase

Os sedimentos celulares foram ressuspensos em tampão LEW, utilizando-se 5 µl de tampão para cada mg de sedimento obtido do cultivo, acrescidos de 1 mg/ml de lisozima. A lise celular foi realizada em desmembrador sônico VibraCell com 3 ciclos a 30 W por 30 s com 1 min de intervalo. O lisado foi centrifugado a 10.000 g por 30 min a 4 °C. A purificação da *rPfu* DNA polimerase foi realizada em coluna Protina Ni-TED 1000 (Macherey-Nagel), conforme recomendações do fabricante. A fração solúvel foi aplicada na coluna e a enzima foi eluída com 1,5 ml de imidazol em quatro diferentes concentrações de 40, 60, 100 e 250 mM. A pureza e a massa molecular aparente foram determinadas por eletroforese em gel de poliacrilamida.

#### 4.7 Eletroforese de DNA em gel de agarose

Para as análises de DNA em géis de agarose, foi utilizado o sistema de eletroforese horizontal com gel submerso, sendo o preparo dos géis realizado como descrito em Sambrook & Russel (2001). A agarose, na concentração de 0,8%, foi fundida em tampão TAE 1x (Tris-acetato a 40 mM, EDTA a 1 mM, pH 8,2), adicionada de 2,5 µl de *Gel Red* (Biotium). A eletroforese foi realizada em tampão TAE 1x e o tamanho dos fragmentos de DNA foi estimado com base no marcador de tamanhos de fragmentos de 1 kb (Invitrogen). Para as migrações de DNA nos géis de agarose foi utilizada a fonte Power PAC (Bio-Rad) ajustada a 100 V. Após a eletroforese, o DNA foi visualizado em transiluminador ultravioleta a 254 nm (Bio-Rad). As imagens dos géis foram captadas pelo uso do equipamento Gel-Doc (Bio-Rad).

#### 4.8 Eletroforese de extratos proteicos em gel de poliacrilamida

A separação de proteínas por eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio (SDS-PAGE) foi executada utilizando-se o método de Laemmli (1970), usando-se gel de empacotamento a 5% e gel de corrida a 12%, em uma cuba de eletroforese vertical (Bio-Rad). As amostras foram desnaturadas por fervura (3 min a 100 °C) em tampão de amostra (Tris.HCl a 125 mM, pH 6,8; SDS a 4%; glicerol a 40%; β-mercaptoetanol a 400 mM e azul de bromofenol a 0,002%). Um volume de 15 µl de cada amostra desnaturada foi aplicada nos géis. Como proteínas padrão de massa molecular, 12 proteínas variando de 14,4 a 220 kDa foram utilizadas (GE Healthcare). Os géis foram corados com *Coomassie Brilliant Blue R-250*.

#### 4.9 Sequenciamento de DNA

O sequenciamento do DNA foi realizado no sequenciador automático *ABI-PRISM 3100 Genetic Analyzer* (Applied Biosystems), seguindo os protocolos que acompanham o equipamento e o kit *BigDye Terminator v3.1 Cycle Sequencing* (Applied Biosystems). A quantidade de DNA-molde utilizada para o sequenciamento foi de 60 ng e 4,5 pmol de *primer* T7 e PTac para pET-23a(+) e pGEX-4T-1, respectivamente. A qualidade dos resultados foi analisada pelo programa *Sequencing Analysis* (Applied Biosystems).



## 5. RESULTADOS E DISCUSSÃO

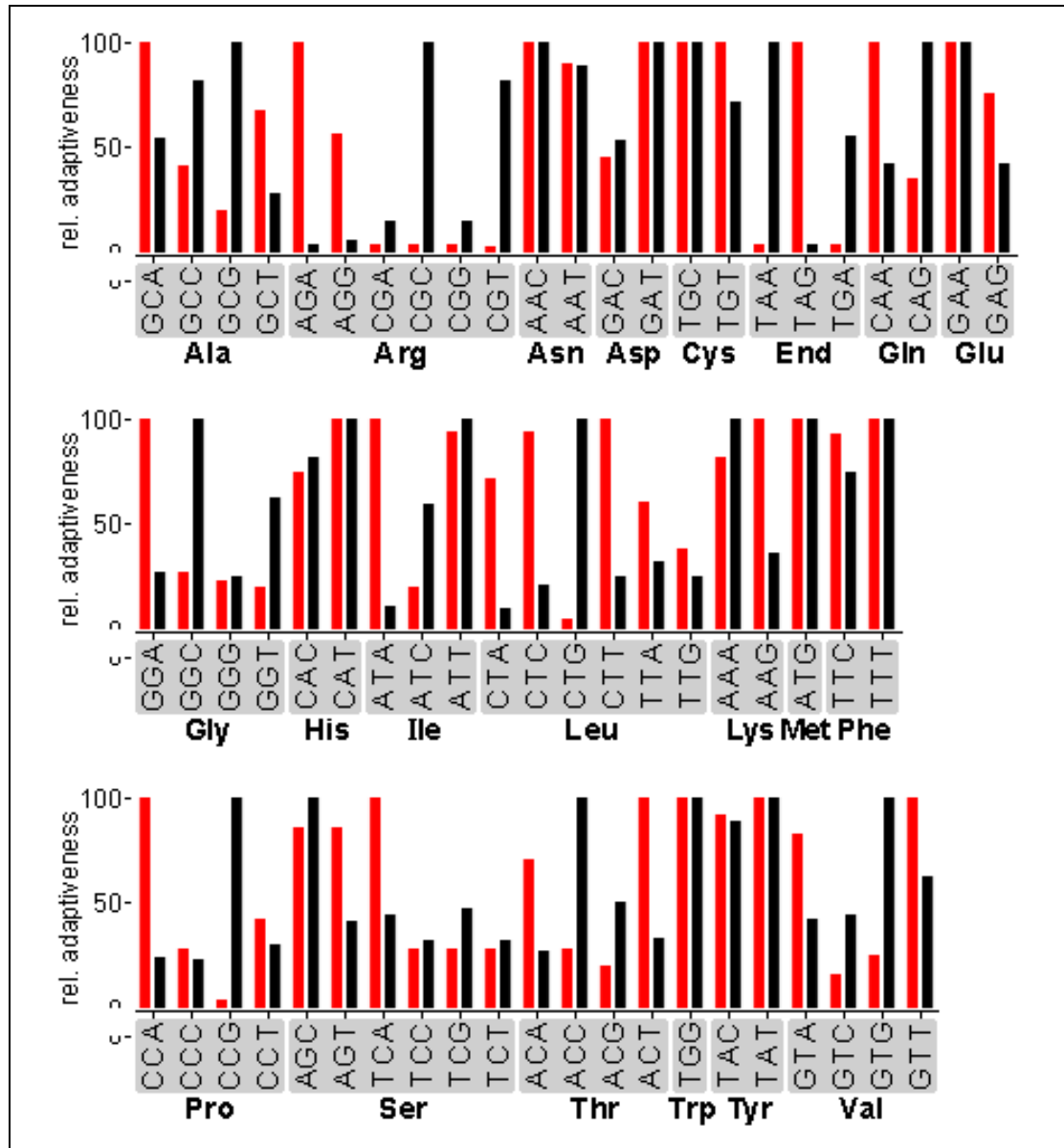
### 5.1 Seleção e definição de sequência nucleotídica para a produção de uma DNA polimerase termoestável e de alta fidelidade para PCR

Considerando as DNA polimerases existentes e com a sequência nucleotídica completa disponível em bancos de dados, a *Pfu* DNA polimerase isolada de *Pyrococcus furiosus* possui uma das mais baixas taxas de erro entre as DNA polimerases de alta fidelidade. Ela também possui uma das menores sequências nucleotídicas, gerando consequentemente um menor custo para a síntese da sequência.

A partir das palavras-chave “*Pyrococcus furiosus*” e “*DNA polymerase*”, foi possível identificar a sequência completa da *Pfu* DNA polimerase deste organismo no banco de sequências nucleotídicas do NCBI. A sequência completa está identificada pelo código de acesso D12983 e possui 2.785 pb. Somente a região codificadora (CDS, do inglês, *coding sequence*), que se estende da posição 237 da sequência nucleotídica até o códon de terminação, na posição 2.564, foi utilizada. A CDS codifica um total de 775 aminoácidos, com uma massa molecular estimada de 90 kDa.

As análises *in silico* realizadas por alinhamento de sequências, comparando-se a sequência deduzida de aminoácidos da *Pfu* DNA polimerase com sequências de proteínas curadas depositadas no Instituto Sanger permitiram confirmar a classificação da enzima como uma DNA polimerase da família B, com função de replicação, recombinação e reparação do DNA em bactérias.

A sequência original da DNA polimerase de *P. furiosus* foi adaptada para a frequência de uso de códons (*codon usage*) de *E. coli* K-12, para que a proteína pudesse ser expressa sem problemas na tradução durante a fase de produção, o que poderia levar a uma redução na quantidade e qualidade da proteína recombinante. Na Figura 1 estão representadas as diferenças entre o uso preferencial de códons de *P. furiosus* e *E. coli*. Esta diferença atinge o valor bastante significativo de 39,03 % entre os microrganismos. Portanto, a conversão dos códons de *P. furiosus* para os códons preferenciais de *E. coli* é possivelmente fundamental para a expressão da proteína neste último microrganismo.



**Figura 1:** Comparação gráfica de *codon usage* de *P. furiosus* (barras vermelhas) com *E. coli* K-12 (barras pretas). Na ordenada está representado o percentual de adaptabilidade relativa (em inglês, *rel. adaptiveness*). Na abcissa estão representados os códons e aminoácidos correspondentes. Os gráficos foram gerados com o uso da ferramenta *Graphical Codon Usage Analyser 2.0* ([www.gcu.de](http://www.gcu.de)).

A partir da análise dos gráficos representados na Figura 1, realizaram-se modificações nos códons que apresentavam diferenças maiores do que 50% entre os microrganismos. Portanto, foram realizadas modificações nos códons que codificam Alanina (Ala, A), Arginina (Arg, R), Glicina (Gly, G), Glutamato (Glu, E), Glutamina (Gln, Q), Isoleucina (Ile, I), Leucina (Leu, L), Lisina (Lys, K), Prolina (Pro, P), Serina (Ser, S), Treonina (Thr, T) e Valina (Val, V) conforme representadas na Tabela 5. Em geral, somente o último nucleotídeo do códon foi modificado. Para alguns aminoácidos, no entanto, foi necessária a troca de dois nucleotídeos.

Após a modificação manual da sequência nucleotídica para a otimização do uso de códons por *E. coli*, foram buscados na sequência final os sítios de clivagem para as principais enzimas de restrição utilizadas em clonagem molecular, como *Bgl*III, *Bam*HI, *Eco*RI, *Eco*RV, *Hind*III, *Hpa*I, *Nco*I, *Nde*I, *Pst*I, *Pvu*I, *Not*I, *Kpn*I, *Sac*I, *Sal*I, *Sma*I, *Xba*I e *Xho*I. Foram encontrados sítios de restrição para as enzimas *Bam*HI, *Eco*RI, *Eco*RV, *Nco*I, *Sma*I, *Xho*I, conforme indicados na Figura 2. Procedeu-se a substituição dos nucleotídeos sem, no entanto, alterar-se os aminoácidos codificados. A retirada dos sítios de restrição foi fundamental para permitir a ligação da sequência aos plasmídeos de expressão de *E. coli* e posteriores clonagens. As modificações realizadas estão detalhadas na Tabela 6.

Por último, realizou-se a inclusão de sítios de restrição para as enzimas *Eco*RI (porção 5`-terminal) e *Xho*I (porção 3`-terminal), a fim de possibilitar a inserção da sequência nucleotídica sintetizada nos plasmídeos pET-23a(+) e pGEX-4T-1, os quais também possuem sítios de restrição para estas enzimas. A versão final da sequência gênica projetada foi denominada *rPfu* DNA polimerase, designando-se a versão recombinante da DNA polimerase de *P. furiosus*. Os resultados de todas as modificações estão representados na Figura 2.

**Tabela 5:** Modificações de códons realizadas na sequência original da DNA polimerase de *P. furiosus* para o uso preferencial de códons de *E. coli*.

Aminoácido	Códon Original	Códon Final
Alanina	GCA, GCT	GCG
Arginina	AGA, AGG	CGC
Glicina	GGA, GGG	GGC
Glutamato	GAG	GAA
Glutamina	CAA	CAG
Isoleucina	ATA	ATT
Leucina	CTA, CTC, CTT TTA, TTG	CTG
Lisina	AAG	AAA
Prolina	CCA, CCC, CCT	CCG
Serina	AGT TCA, TCC, TCT	AGC TCG
Treonina	ACA, ACG, ACT	ACC
Valina	GTA, GTC, GTT	GTG

**Tabela 6:** Exclusão de sítios de restrição pré-existent na sequência original da DNA polimerase de *P. furiosus* para a adaptação a plasmídeos de expressão em *E. coli*.

Enzima de restrição	Região de localização (pb)	Códon Inicial	Códon Final
<i>Bam</i> HI	1.417	GAT	GAC
<i>Eco</i> RI	974 1.473	TTC AAT	TTT AAC
<i>Eco</i> RV	1.297	GAT	GAC
<i>Nco</i> I	174	CAT	CAC
<i>Sma</i> I	1.338 1.647	ATC ATC	ATT ATT
<i>Xho</i> I	1.038	CGC	CGG

## 5.2 Alinhamento das sequências nucleotídicas e peptídicas da *rPfu* e *Pfu* DNA polimerase

Na Figura 2 está representada a comparação das sequências nucleotídicas da *rPfu* DNA polimerase e da *Pfu* DNA polimerase original, conforme depositada no NCBI sob o código de acesso D12983. Também na Figura 2 estão salientados os nucleotídeos inalterados (indicados por um asterisco) e os alterados (ausência de asteriscos) para a otimização do uso de códons por *E. coli*. O códon de terminação da *Pfu* DNA polimerase foi retirado na versão da *rPfu* para não interromper a leitura da sequência quando esta estiver ligada aos plasmídeos pET-23a(+) e pGEX-4T-1.

A análise comparativa das sequências peptídicas dos genes da *rPfu* DNA polimerase com a sequência original da *Pfu* está mostrada na Figura 3. Por ela, está demonstrada que as sequências peptídicas são idênticas, como esperado, satisfazendo o objetivo inicial de se alterar os códons sem, no entanto, alterar qualquer aminoácido da sequência peptídica original.

<i>rPfu</i>	GAATTCATGATTCTGGATGTGGATTACATTACCGAAGAAGCAAACCGGTGATTCGCCTGTTCAAA	60
<i>Pfu</i>	ATGATTTTAGATGTGGATTACATAACTGAAGAAGGAAAACCTGTTATTAGGCTATTCAA	60
	*****	
<i>rPfu</i>	AAAGAAAACGGCAAATTTAAAATTGAACATGATCGCACCTTCGCCGTACATTTACGCG	120
<i>Pfu</i>	AAAGAGAACGGAAAATTTAAGATAGAGCATGATAGAACTTTAGACCATAACATTTACGCT	120
	*****	
<i>rPfu</i>	CTGCTGCGGATGATTCGAAAATTGAAGAAGTGAAAAAATTACCGGCGAACGCCACGGC	180
<i>Pfu</i>	CTTCTCAGGGATGATTCAAAGATTGAAGAAGTTAAGAAAATAACGGGGAAAGGCATGGA	180
	** * * *	
<i>rPfu</i>	AAAATTGTGCGCATGTGGATGTGAAAAAGTGGAAAAAATTTCTGGGCAAACCGATT	240
<i>Pfu</i>	AAGATTGTGAGAATTTGTTGATGTAGAGAAGTTGAGAAAAGTTTCTCGGCAAGCCTATT	240
	** *****	
<i>rPfu</i>	ACCGTGTGAAACTGTATCTGGAACATCCCGAGGATGTGCCGACCATTCGCGAAAAAGTG	300
<i>Pfu</i>	ACCGTGTGAAACTTTATTTGGAACATCCCAAGATGTTCCCACTATTAGAGAAAAAGTT	300
	*****	

rPfu	CGCGAACATCCGGCGGTGGTGGACATCTTCGAATACGATATTCGGTTTGCAGAACGCTAC	360
Pfu	AGAGAACATCCAGCAGTTGTGGACATCTTCGAATACGATATTCATTTCAGAAAGAGATAC	360
	* ***** ** * ***** ***** ** * **	
rPfu	CTGATCGACAAAGGCTGATTCGGATGGAAGGCGAAGAAGAACTGAAAATCTGGCCTTC	420
Pfu	CTCATCGACAAAGGCTAATACCAATGGAGGGGAAGAAGACTAAAGATCTTGGCCTTC	420
	** ***** ** * ***** ** * ***** ** * ***** ** * *****	
rPfu	GATATTGAAACCCTGTATCACGAAGGCGAAGAATTTGGCAAAGGCCGATTATTATGATT	480
Pfu	GATATAGAAACCCTCTATCACGAAGGAGAGAGTTTGGAAAAGGCCAATTATAATGATT	480
	***** ***** ***** ***** ***** ***** ***** *****	
rPfu	AGCTATGCGGATGAAAATGAAGCGAAAGTGATTACCTGGAAAAACATTGATCTGCCGTAC	540
Pfu	AGTTATGCAGATGAAAATGAAGCAAGGTGATTACTTGGAAAAACATAGATCTTCCATAC	540
	** ***** ***** ** * ***** ***** ***** ***** ** * **	
rPfu	GTGGAAGTGGTGTGCGAGCGAACGCGAAATGATTAACGCTTCTGCGCATTATCCGCGAA	600
Pfu	GTTGAGGTTGTATCAAGCGAGAGAGATGATAAAGAGATTTCTCAGGATTATCAGGGAG	600
	** *	
rPfu	AAAGATCCGGACATTATTGTGACCTATAATGGCGACTCGTTCGACTTCCCGTATCTGGCG	660
Pfu	AAAGATCCTGACATTATAGTTACTTATAATGGAGACTCATTCGACTTCCATATTTAGCG	660
	***** ***** ** * ***** ***** ***** ***** ** * **	
rPfu	AAACGCGCGAAAAAAGTGGGCATTAACCTGACCATTGGCCGCGATGGCAGCGAACCGAAA	720
Pfu	AAAAGGGCAGAAAAAAGTGGGATTAATTAACCATGGAAGAGATGGAAGCGAGCCCAAG	720
	*** *	
rPfu	ATGCAGCGCATTGGCGATATGACCGCGGTGGAAGTGAAAGGCCGCATTCATTTGACCTG	780
Pfu	ATGCAGAGAATAGGCGATATGACGCGTGTAGAAGTCAAGGAAGAATACATTTGACTTG	780
	***** *	
rPfu	TATCATGTGATTACCCGCACCATTAATCTGCCGACTACACCCTGGAAGCGGTGTATGAA	840
Pfu	TATCATGTAATAACAAGGACAATAAATCTCCAACATACACACTAGAGGCTGTATATGAA	840
	***** ** *	
rPfu	GCGATTTTTGGCAAACCGAAAAGTGTACGCCGACGAAATTCGCAAAGCCTGGGAA	900
Pfu	GCAATTTTTGGAAAGCCAAAGGAGAAGGTATACGCCGACGAGATAGCAAAGCCTGGGAA	900
	** ***** ** *	
rPfu	AGCGGCGAAAACCTGGAACGCGTGGCCAAATACTCGATGGAAGATGCGAAAGCGACCTAT	960
Pfu	AGTGAGAGAACCCTGAGAGAGTTGCCAAATACTCGATGGAAGATGCAAAGGCAACTTAT	960
	** *	
rPfu	GAATGGGCAAAGAATTTCTGCCGATGGAATTCAGCTGTGCGCCCTGGTGGGCCAGCGG	1020
Pfu	GAATCGGGAAGAATTCCTTCCAATGGAATTCAGCTTCAAGATTAGTTGGACAACCT	1020
	***** ** *	
rPfu	CTGTGGATGTGTGCGGTCGAGCACCGGCAACCTGGTGAATGGTTCCTGCTGCGCAAA	1080
Pfu	TTATGGGATGTTTCAAGGTCAGCACAGGCAACCTGTAGAGTGGTTCCTACTTAGGAAA	1080
	* ***** ** *	
rPfu	GCCTACGAACGCAACGAAGTGGCGCCGAACAACCGAGCGAAGAAGAATATCAGCGCCGC	1140
Pfu	GCCTACGAAAGAAACGAAGTAGCTCCAAACAAGCCAAGTGAAGAGGAGTATCAAAGAAG	1140
	***** *	
rPfu	CTGCGCAAAGCTACACCGGTGGCTTCGTGAAAGAACCGAAAAAGCCTGTGGGAAAAC	1200
Pfu	CTCAGGAGAGCTACACAGGTGGATTCGTTAAAGAGCCAGAAAAGGGTGTGGGAAAAC	1200
	** *	
rPfu	ATTGTGTACCTGGATTTTCGCGCCCTGTATCCGTCGATTATTATTACCCACAATGTGTCG	1260
Pfu	ATAGTATACCTAGATTTTAGAGCCCTATATCCCTCGATTATAATTACCCACAATGTTTCT	1260
	** *	
rPfu	CCGATACCTGAATCTGGAAGGTGCAAAAACCTATGACATCGCGCCGAGGTGGGCCAC	1320
Pfu	CCCATACTCTAAATCTTGGGGATGCAAGAACCTATGATATCGCTCCTCAAGTAGGCCAC	1320
	** ***** ** *	
rPfu	AAATCTGCAAAGACATTCGGGTTTTATTCGAGCCTGCTGGGCCATCTGCTGGAAGAA	1380
Pfu	AAGTTCTGCAAGGACATCCCTGGTTTTATACCAAGTCTCTTGGGACATTTGTTAGAGGAA	1380
	** ***** ***** ** * ***** ** * * * * * * * * * * * * * * *	
rPfu	CGCCAGAAAATTAAAACCAAAATGAAAGAACCAGGACCCGATTGAAAAAATTCGCTG	1440
Pfu	AGACAAAAGATTAAGACAAAAATGAAGGAACTCAAGATCCTATAGAAAAAATACTCCTT	1440
	* *	
rPfu	GACTATCGCCAGAAAGCGATTAACTGCTGGCGAACTCGTTCACGGCTATTATGGCTAT	1500

<i>Pfu</i>	GACTATAGACAAAAAGCGATAAACTCTTAGCAAATTCCTTTCTACGGATATTATGGCTAT	1500
	***** * ** ***** * ** * ** * ** ***** *****	
<i>rPfu</i>	GCGAAAGCGCGCTGGTACTGTAAAGAATGTGCGGAAAGCGTGACCGCTGGGGCCGCAAA	1560
<i>Pfu</i>	GCAAAAGCAAGATGGTACTGTAAAGAGTGTGCTGAGAGCGTTACTGCCTGGGGAAGAAAG	1560
	** ***** * ***** ** ***** * ** ***** * ** ***** * **	
<i>rPfu</i>	TACATCGAAGTGGTGTGGAAGAAGTGAAGAAAAATTTGGCTTTAAAGTCTGTACATT	1620
<i>Pfu</i>	TACATCGAGTTAGTATGGAAGGAGCTCGAAGAAAAGTTTGGATTAAAGTCCTCTACATT	1620
	***** * ** ***** * ** ***** ***** ***** * ** *****	
<i>rPfu</i>	GACACCGATGGTCTGTATGCGACCATTCCGGGCGCGAAAGCGAAGAAATTAAGAAAAAA	1680
<i>Pfu</i>	GACTACTGATGGTCTCTATGCAACTATCCAGGAGGAGAAAGTGAAGAAATTAAGAAAAAA	1680
	***** ***** ***** * ** * ** * ** * ** ***** * ** *****	
<i>rPfu</i>	GCGTGGAATTTGTGAAATACATTAATTCGAAACTGCCGGCCTGCTGGAAGTGAATAT	1740
<i>Pfu</i>	GCCCTAGAATTTGTAAATACATAAATTCAGAAAGCTCCCTGGACTGCTAGAGCTTGAATAT	1740
	** * ** ***** ***** ***** * ** * ** * ** * ** ***** * ** *****	
<i>rPfu</i>	GAAGGCTTTTATAAACGCGGCTTCTTCGTGACCAAAAAACGCTATGCGGTGATTGATGAA	1800
<i>Pfu</i>	GAAGGGTTTATAAGAGGGGATTCTTCGTACGAAAGAGGTATGCAGTAATAGATGAA	1800
	***** ***** * ** ***** * ** * ** * ** ***** * ** *****	
<i>rPfu</i>	GAAGGCAAAGTATTACCGTGGTCTGGAATTTGTGCGCGCGATTGGAGCGAAATTCGG	1860
<i>Pfu</i>	GAAGGAAAAGTCACTACTCGTGGTTTAGAGATAGTTAGGAGAGATTGGAGTGAATTCGA	1860
	***** ***** ***** ***** * ** * ** * ** * ** ***** *****	
<i>rPfu</i>	AAAGAAACCAGGCGCGCTGCTGGAACCAATTCGAAACACGGCGATGTGAAGAAGCG	1920
<i>Pfu</i>	AAAGAACTCAAGCTAGAGTTTGGAGACAATACTAAAACACGGAGATGTGAAGAAGCT	1920
	***** * ** * ** ***** * ** * ** * ** ***** ***** *****	
<i>rPfu</i>	GTGGCATTGTGAAGAAGTGATTACGAAACTGGCCAATTATGAAATTCGCCCGAAAAA	1980
<i>Pfu</i>	GTGAGAATAGTAAAAGAAGTAATACAAAACCTTGCCAATTATGAAATTCACCAGAGAAG	1980
	*** * ** * ** ***** * ** * ** ***** ***** ***** * ** * **	
<i>rPfu</i>	CTGGCGATTTATGAACAGATTACCGCCCGCTGCATGAATATAAGCGATTGGTCCGCAC	2040
<i>Pfu</i>	CTCGCAATATATGAGCAGATAACAAGACATTACATGAGTATAAGCGATAGGTCTCTCAC	2040
	** * ** * ** ***** ***** * ** * ** * ** ***** ***** ***** * **	
<i>rPfu</i>	GTGGCGGTGGCGAAAAAAGTGGCGGCAAGGCGTGAAAAATTAACCGGGCATGGTGATT	2100
<i>Pfu</i>	GTAGCTGTGCAAAAGAACTAGCTGCTAAGGAGTAAATAAAGCCAGGAATGTAATT	2100
	** * ** * ** * ** ***** * ** * ** ***** * ** * ** ***** * **	
<i>rPfu</i>	GGCTACATTGTGCTGCGCGCGGATGGTCCGATTAGCAATCGCGGATTCTGGCGGAAGAA	2160
<i>Pfu</i>	GGATACATAGTACTTAGAGCGGATGGTCCAATTAGCAATAGGGCAATTCTAGCTGAGGAA	2160
	** ***** * ** * ** ***** ***** ***** * ** ***** * ** * **	
<i>rPfu</i>	TACGATCCGAAAAAACACAATATGACGCGGAATATTACATTGAAAACCAGGTGCTGCCG	2220
<i>Pfu</i>	TACGATCCCAAAAAGCACAGTATGACGCGAATATTACATTGAGAACCAGGTTCTTCCA	2220
	***** ***** ***** ***** ***** ***** ***** ***** * **	
<i>rPfu</i>	GCGTGCTGCGCATTCGGAAGGCTTTGGCTACCGCAAAGAAGACCTGCGCTACCAGAAA	2280
<i>Pfu</i>	GCGTACTTAGGATATTGGAGGGATTTGGATACAGAAAGGAAGACCTCAGATACCAAAAG	2280
	***** * ** * ** ***** * ** ***** * ** * ** ***** * ** *****	
<i>rPfu</i>	ACCGCCAGGTGGGCTGACCTCGTGGCTGAACATTAAGAAATCGTAGCTCGAG	2334
<i>Pfu</i>	ACAAGACAAGTCGGCTAACTTCCTGGCTTAACATTAAGAAATCTAA	2325
	** * ** * ** ***** * ** ***** ***** ***** ***** * **	

**Figura 2:** Alinhamento das sequências nucleotídicas da *rPfu* e da *Pfu* DNA polimerase original. Os nucleotídeos sem asteriscos foram alterados intencionalmente para otimização do uso de códons por *E. coli*. Em cores verde e amarela estão indicados os códons de início e término da tradução, respectivamente. Os hexanucleotídeos sublinhados na sequência da *Pfu* DNA polimerase representam os sítios de restrição para as enzimas *Bam*HI, *Eco*RI, *Eco*RV, *Nco*I, *Sma*I, *Xho*I, respectivamente, que foram eliminados da sequência da *rPfu* DNA polimerase, para efeitos da manipulação da mesma. Os hexanucleotídeos sublinhados nas posições 5'-terminal e 3'-terminal da sequência da *rPfu* DNA polimerase representam os sítios de *Eco*RI e *Xho*I artificialmente projetados para as clonagens posteriores em plasmídeos de expressão. O alinhamento das sequências nucleotídicas foi realizado empregando-se o programa ClustalW ([www.ebi.ac.uk/clustalw](http://www.ebi.ac.uk/clustalw)).

	10	20	30	40	50	60
<i>rPfu</i>	MILDVDYITEEGKPVIRLFKKENGKFKIEHDRTRFRPYIYALLRDDSKEEVKKITGERHG					
	.....					
<i>Pfu</i>	MILDVDYITEEGKPVIRLFKKENGKFKIEHDRTRFRPYIYALLRDDSKEEVKKITGERHG					
	10	20	30	40	50	60
	70	80	90	100	110	120
<i>rPfu</i>	KIVRIVDVEKVEKKFLGKPITVWKLYLEHPQDVPTIREKVVREHPAVVDIFEYDIPFAKRY					
	.....					
<i>Pfu</i>	KIVRIVDVEKVEKKFLGKPITVWKLYLEHPQDVPTIREKVVREHPAVVDIFEYDIPFAKRY					
	70	80	90	100	110	120
	130	140	150	160	170	180
<i>rPfu</i>	LIDKGLIPMEGEEELKILAFDIETLYHEGEEFGKGPIIMISYADENEAKVITWKNIDLPI					
	.....					
<i>Pfu</i>	LIDKGLIPMEGEEELKILAFDIETLYHEGEEFGKGPIIMISYADENEAKVITWKNIDLPI					
	130	140	150	160	170	180
	190	200	210	220	230	240
<i>rPfu</i>	VEVVSSEREMIKRFLRIIREKDPDIVTYNGDSFDFPYLAKRAEKLGIKLTIGRDGSEPK					
	.....					
<i>Pfu</i>	VEVVSSEREMIKRFLRIIREKDPDIVTYNGDSFDFPYLAKRAEKLGIKLTIGRDGSEPK					
	190	200	210	220	230	240
	250	260	270	280	290	300
<i>rPfu</i>	MQRIGDMTAVEVKGRIHFDLYHVITRTINLPTYTLEAVYEAI FGKPKVKYADEIAKAWE					
	.....					
<i>Pfu</i>	MQRIGDMTAVEVKGRIHFDLYHVITRTINLPTYTLEAVYEAI FGKPKVKYADEIAKAWE					
	250	260	270	280	290	300
	310	320	330	340	350	360
<i>rPfu</i>	SGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVGPPLWDVSRSS TGNLVEWFLLRK					
	.....					
<i>Pfu</i>	SGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVGPPLWDVSRSS TGNLVEWFLLRK					
	310	320	330	340	350	360
	370	380	390	400	410	420
<i>rPfu</i>	AYERNEVAPNKPSEEEYQRRRESYTGGFVKEPEKGLWENIVYLDFRALYPSIIITHNVS					
	.....					
<i>Pfu</i>	AYERNEVAPNKPSEEEYQRRRESYTGGFVKEPEKGLWENIVYLDFRALYPSIIITHNVS					
	370	380	390	400	410	420
	430	440	450	460	470	480
<i>rPfu</i>	PDTLNLEGCKNYDIAPQVGHKFC KDIPGFIPSL LGHLEERQKIKTKMKETQDPIEKILL					
	.....					
<i>Pfu</i>	PDTLNLEGCKNYDIAPQVGHKFC KDIPGFIPSL LGHLEERQKIKTKMKETQDPIEKILL					
	430	440	450	460	470	480
	490	500	510	520	530	540
<i>rPfu</i>	DYRQKAIKLLANSFYGYGYAKARWYCKECAESVTAWGRKYIELVWKELEEKFGFKVLYI					
	.....					
<i>Pfu</i>	DYRQKAIKLLANSFYGYGYAKARWYCKECAESVTAWGRKYIELVWKELEEKFGFKVLYI					
	490	500	510	520	530	540
	550	560	570	580	590	600
<i>rPfu</i>	DTDGLYATIPGGESEEI KKKALEFVKY INSKLPGLLELEYEGFYKRGFFVTKKRYAVIDE					
	.....					
<i>Pfu</i>	DTDGLYATIPGGESEEI KKKALEFVKY INSKLPGLLELEYEGFYKRGFFVTKKRYAVIDE					
	550	560	570	580	590	600
	610	620	630	640	650	660
<i>rPfu</i>	EGKVITRGLIIVRRDWSEIAKETQARVLETILKHGDVEEAVRIVKEVIQKLANYEIPPEK					
	.....					
<i>Pfu</i>	EGKVITRGLIIVRRDWSEIAKETQARVLETILKHGDVEEAVRIVKEVIQKLANYEIPPEK					
	610	620	630	640	650	660

	670	680	690	700	710	720
<i>rPfu</i>	LAIYEQITRPLHEYKAIGPHVAVAKKLAAGVKIKPGMVIQYIVLRGDGPISNRAILAE					
	.....	.....	.....	.....	.....	.....
<i>Pfu</i>	LAIYEQITRPLHEYKAIGPHVAVAKKLAAGVKIKPGMVIQYIVLRGDGPISNRAILAE					
	670	680	690	700	710	720
	730	740	750	760	770	
<i>rPfu</i>	YDPKHKYDAEYYIENQVLPVLRILEGFGYRKEDLRYQKTRQVGLTSWLKKS					
	.....	.....	.....	.....	.....	.....
<i>Pfu</i>	YDPKHKYDAEYYIENQVLPVLRILEGFGYRKEDLRYQKTRQVGLTSWLNKKS					
	730	740	750	760	770	

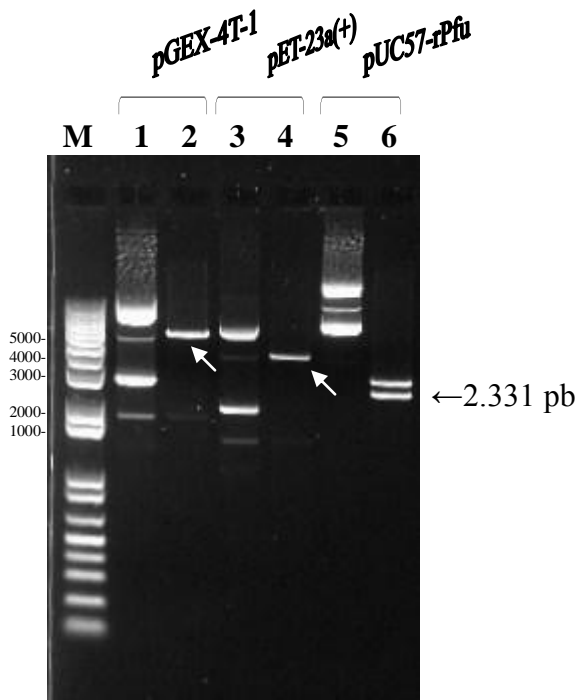
**Figura 3:** Alinhamento das seqüências peptídicas deduzidas das *rPfu* e *Pfu* DNA polimerase. O alinhamento das seqüências peptídicas foi realizado empregando-se o programa ClustalW ([www.ebi.ac.uk/clustalw](http://www.ebi.ac.uk/clustalw)).

### 5.3 Síntese e clonagem da seqüência da *rPfu* DNA polimerase em pET-23a(+) e pGEX-4T-1

A seqüência nucleotídica projetada para a *rPfu* DNA polimerase foi encaminhada via *on-line* para a empresa GenScript, junto a qual a síntese foi solicitada sob condições padrões. Tais condições determinaram o encaminhamento do fragmento sintetizado ligado ao plasmídeo de clonagem pUC57, em uma massa liofilizada de plasmídeo de 4 µg. O plasmídeo foi denominado pUC57-*rPfu*.

O plasmídeo pUC57-*rPfu* e os plasmídeos de expressão em *E. coli* pET-23a(+) e pGEX-4T-1 foram clivados com as enzimas de restrição *EcoRI* e *XhoI*. Os produtos das clivagens foram confirmados por eletroforese em gel de agarose, conforme mostrado na Figura 4. À direita da Figura 4 está salientada a banda de 2.331 pb correspondente ao fragmento da *rPfu* DNA polimerase. Este fragmento foi devidamente purificado para ser ligado aos plasmídeos pET-23a(+) e pGEX-4T-1.

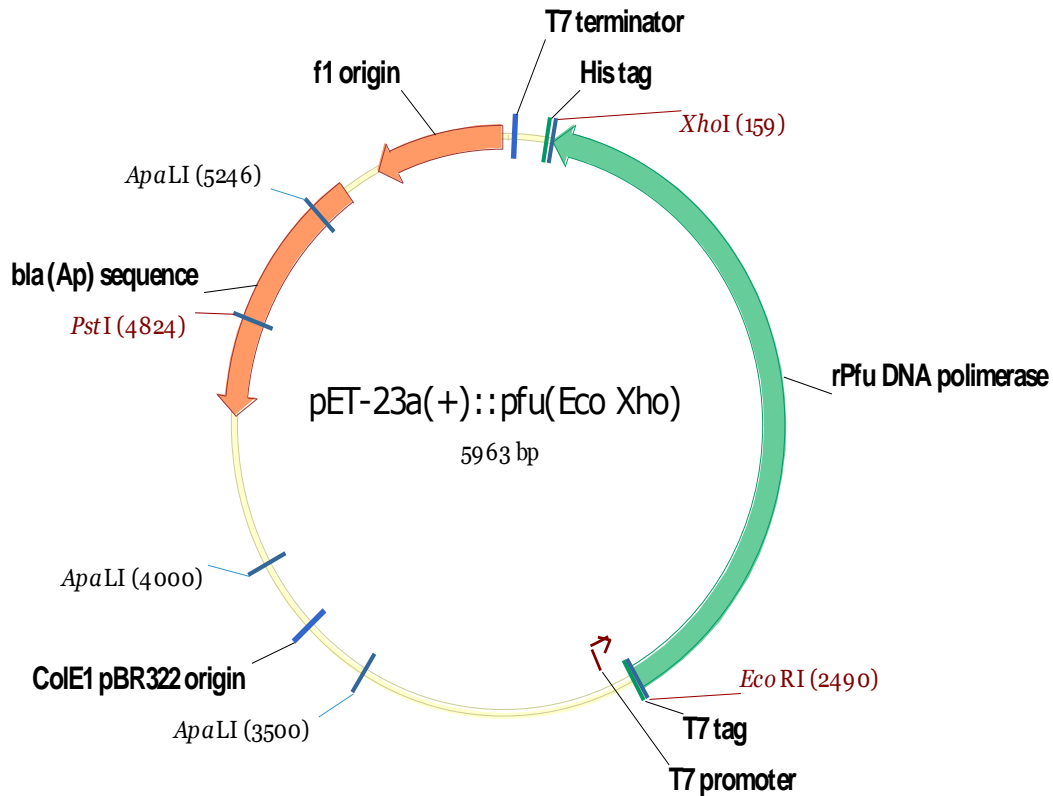




**Figura 4:** Análise das clivagens dos plasmídeos pET-23a(+), pGEX-4T-1 e pUC57-*rPfu* com *EcoRI* e *XhoI* por eletroforese em gel de agarose a 0,8%. Canaletas: (M) Marcador de tamanhos de fragmentos de DNA de 1 kb (Invitrogen); (1) pGEX-4T-1 íntegro; (2) pGEX-4T-1 clivado, seta indicando a banda de 4.900 pb; (3) pET-23a(+) íntegro; (4) pET-23a(+) clivado, seta indicando a banda de 3.666 pb; (5) pUC57-*rPfu* íntegro; (6) pUC57-*rPfu* clivado.

O fragmento de 2.331 pb da *rPfu* DNA polimerase purificado de gel de agarose foi ligado às versões de pET-23a(+) e pGEX-4T-1 devidamente clivados com *EcoRI* e *XhoI*. A representação esquemática da *rPfu* DNA polimerase ligada a pET-23a(+) está mostrada na Figura 5, demonstrando-se a correta fusão, em fase de leitura, da sequência de códons do fragmento com a sequência de códons do plasmídeo que codificam a cauda de histidinas. O vetor resultante desta reação de ligação foi denominado pET-23a(+)-*rPfu*. A ligação a pGEX-4T-1 está representada na Figura 6, demonstrando-se a fusão correta, em fase de leitura, das sequências nucleotídicas do fragmento com a codificadora da GST do plasmídeo. O vetor resultante desta reação de ligação foi denominado pGEX-4T-1-*rPfu*.

A

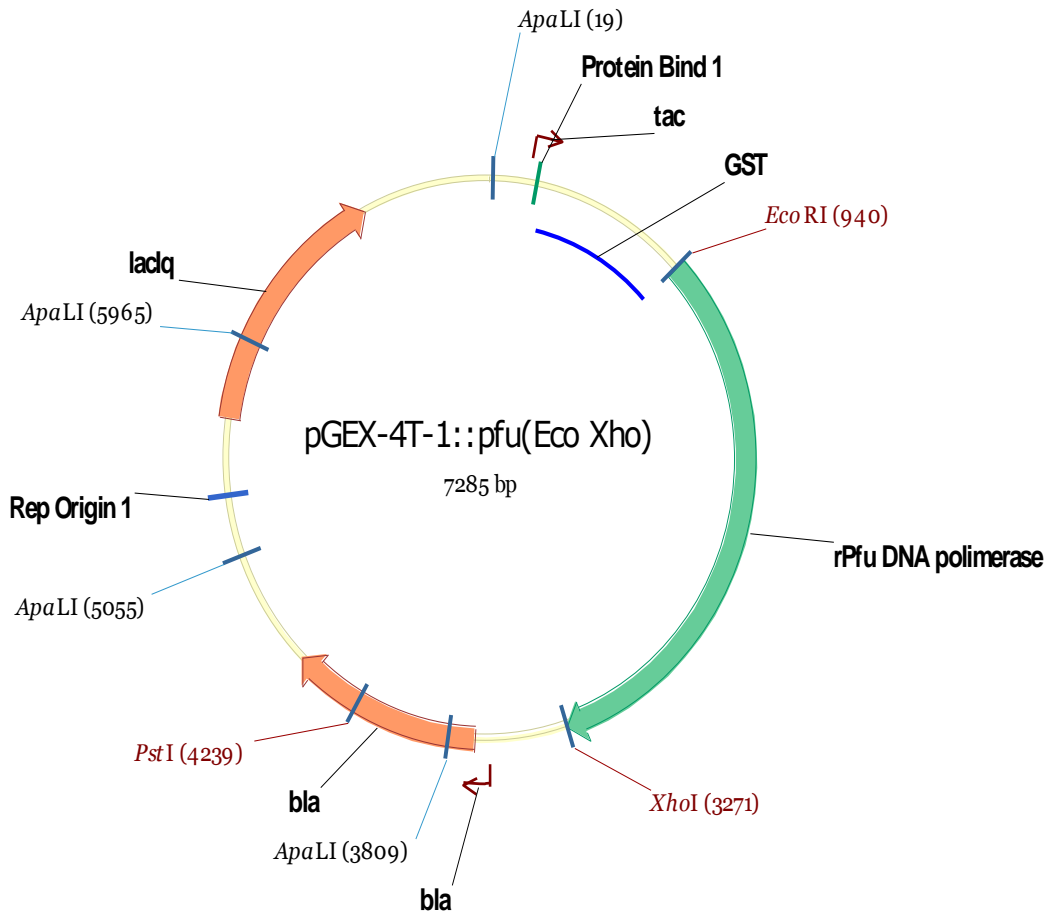


B

pET-23a(+) (5'-end)	<b>EcoRI</b>		<b>rPfu</b>
GGA CAG CAA ATG GGT CGC GGA TCC	GAA TTC	ATG ATT CTG GAT GTG GAT ...	
CCT GTC GTT TAC CCA GCG CCT AGG	CTT AAG	TAC TAA GAC CTA CAC GTA ...	
Gly Gln Gln Met Gly Arg Gly Ser	Glu Phe	Met Ile Leu Asp Val Asp ...	
	<b>XhoI</b>		pET-23a(+) (3-end)
<b>rPfu</b>	CTC GAG	CAC CAC CAC CAC CAC CAC TGA	
...TCG TGG CTG AAC ATT AAA AAA TCG	GAG CTC	GTG GTG GTG GTG GTG GTG ACT	
...AGC ACC GAC TTG TAA TTT TTT AGC	Leu Glu	His His His His His His Stop codon	
Ser Trp Leu Asn Ile Lys Lys Ser			

**Figura 5:** Representação esquemática da ligação do fragmento da *rPfu* DNA polimerase ao vetor pET-23a(+). A: Mapa do vetor pET-23a(+) (Novagen) contendo o inserto da *rPfu* DNA polimerase. B: Detalhes das regiões 5'- e 3'-terminais de clonagem do vetor pET-23a(+)*rPfu*. Para a expressão do gene da *rPfu* DNA Polimerase foram utilizados os sítios *EcoRI* (5') e *XhoI* (3').

A



B

pGEX-4T-1 (5'-end) EcoRI rPfu

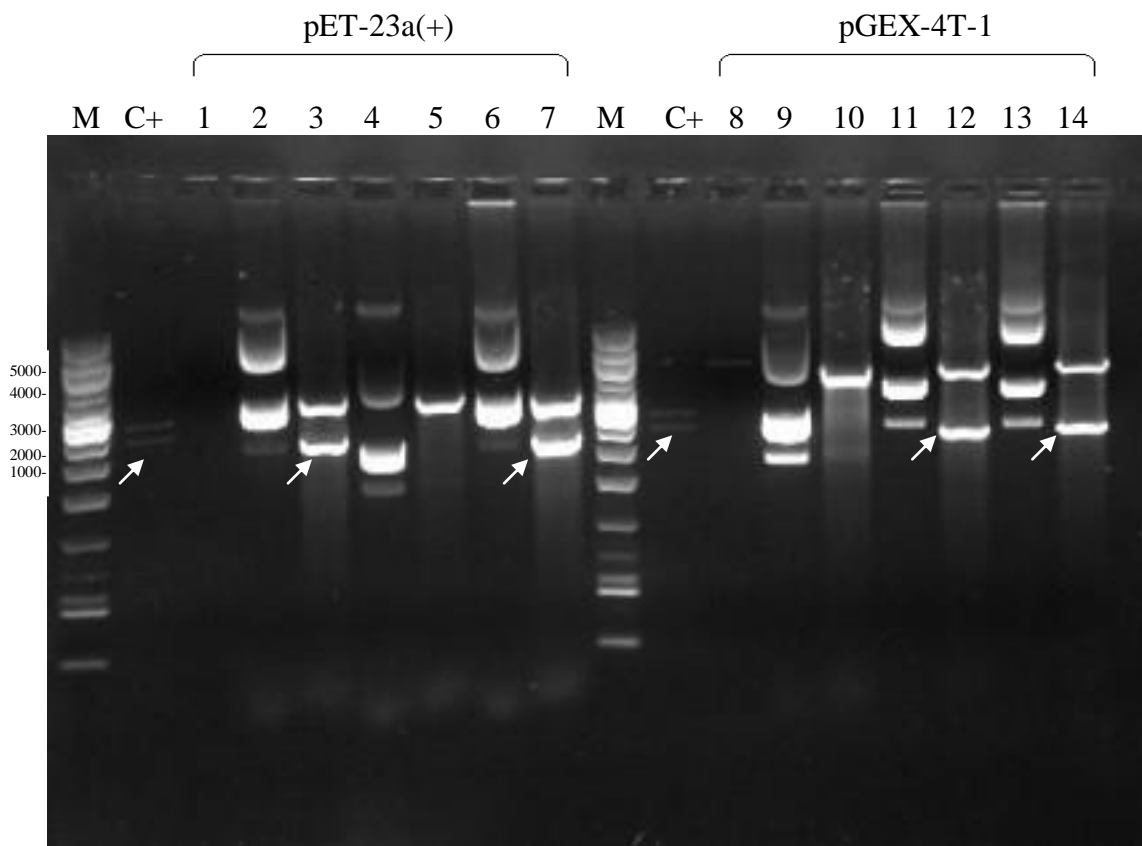
CTG GTT CCG CGT GGA TCC CCG	GAA TTC	ATG ATT CTG GAT GTG GAT ...
GAC CAA GGC GCA CCT AGG GGC	CTT AAG	TAC TAA GAC CTA CAC GTA ...
Leu Val Pro Arg Gly Ser Pro	Glu Phe	Met Ile Leu Asp Val Asp ...

rPfu	XhoI	pGEX-4T-1 (3-end)
...TCG TGG CTG AAC ATT AAA AAA TCG	CTC GAG	CGG CCG CAT CGT GAC TGA
...AGC ACC GAC TTG TAA TTT TTT AGC	GAG CTC	GTC GGC GTA GCA CTG ACT
Ser Trp Leu Asn Ile Lys Lys Ser	Leu Glu	Arg Pro His Arg Asp Stop codon

**Figura 6:** Representação esquemática da ligação do fragmento da *rPfu* DNA polimerase ao vetor pGEX-4T-1. A: Mapa do vetor pGEX-4T-1 (GE Healthcare) contendo o inserto da *rPfu* DNA polimerase. B: Detalhes das regiões 5'- e 3'-terminais de clonagem do vetor pGEX-4T-1-*rPfu*. Para a expressão do gene da *rPfu* DNA Polimerase foram utilizados os sítios *EcoRI* (5') e *XhoI* (3').

Como ao menos cinco produtos distintos podem ser formados a partir da ligação do inserto (*rPfu* DNA polimerase) com o plasmídeo (pET-23a(+) ou pGEX-4T-1), a análise dos produtos das ligações foi feita por clivagem com enzimas de restrição e confirmada por sequenciamento. Após a transformação genética de *E. coli* XL1-Blue por eletroporação utilizando-se os produtos das reações de ligação, um *screening* de 12 colônias de cada grupo foi realizado para verificar quais delas seriam positivas (fragmentos da *rPfu* DNA polimerase ligado aos plasmídeos). Das 12 minipreparações plasmidiais realizadas aleatoriamente, duas colônias de cada tipo demonstraram resultado positivo após a clivagem dos plasmídeos com as enzimas flanqueadoras do inserto *EcoRI* e *XhoI*. Os produtos das clivagens dos plasmídeos derivados destas colônias, juntamente com uma amostra negativa para cada vetor, estão mostrados na Figura 7. Junto às amostras positivas, podem ser observadas bandas referentes aos fragmentos da *rPfu* DNA polimerase, de 2.331 pb, também observado no plasmídeo pUC57-*rPfu* clivado (controle positivo). Junto às amostras positivas também é possível observar as bandas de 3.666 pb ou 4.900 pb, correspondentes aos plasmídeos pET-23a(+) e pGEX-4T-1, respectivamente.



**Figura 7:** Análise de minipreparações plasmidiais derivadas de colônias de *E. coli* XL1-Blue transformadas com os plasmídeos pET-23a(+) e pGEX-4T-1 ligados ao fragmento da *rPfu* DNA polimerase. Os plasmídeos foram clivados com *EcoRI* e *XhoI* e os produtos das clivagens foram resolvidos por eletroforese em gel de agarose a 0,8%. Canaletas: (M) Marcador de tamanhos de fragmentos de DNA de 1 kb (Invitrogen); (C+) pUC57-*rPfu*

clivado; (1) pET-23a(+) clivado; (8) pGEX-4T-1 clivado. Amostras de pET-23a(+)-*rPfu* não clivadas (2, 4 e 6) e clivadas (3, 5 e 7), sendo 3 e 7 positivas e 5 negativa; Amostras de pGEX-4T-1-*rPfu* não clivadas (9, 11 e 13) e clivadas (10, 12 e 14) sendo 12 e 14 positivas e 10 negativa. As setas indicam a posição dos fragmentos referentes à *rPfu* DNA polimerase.

O sequenciamento de DNA confirmou a ligação em fase de leitura da *rPfu* DNA polimerase com o plasmídeo pET-23a(+). Porém o resultado do sequenciamento com o plasmídeo pGEX-4T-1 não foi satisfatório.

#### 5.4 Expressão da *rPfu* DNA polimerase em *E. coli*

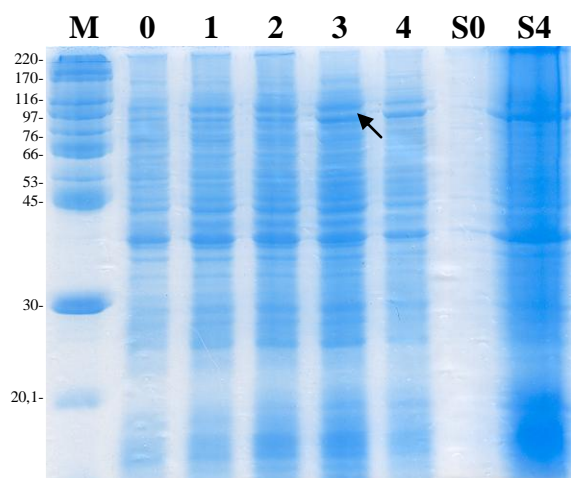
Para a expressão da *rPfu* DNA polimerase em bactérias, linhagens de *E. coli* BL21(DE3)pLysS foram utilizadas. Lu & Erickson (1997), ao clonarem a DNA polimerase de *P. furiosus* utilizando o plasmídeo pET11, observaram que esta construção era tóxica em células de *E. coli* BL21(DE3). Entretanto, ao utilizarem a cepa BL21(DE3)pLysS de *E. coli*, os autores conseguiram produzir e caracterizar a enzima. Esta linhagem possui o plasmídeo pLysS, o qual codifica a lisozima T7. Esta lisozima liga a RNA polimerase T7 e promove a diminuição da expressão basal da proteína de interesse. Assim, a proteína de interesse é expressa apenas após a indução por IPTG (Dubendorff & Studier, 1991).

Os plasmídeos positivos pET-23a(+)-*rPfu* e pGEX-4T-1-*rPfu* foram introduzidos por eletroporação em células de *E. coli* BL21(DE3)pLysS para expressar a *rPfu*. Colônias positivas foram, a seguir, inoculadas em meio de cultivo para a indução da expressão proteica. Após culturas, indução com IPTG e extrações protéicas totais, a melhor expressão pode ser observada pela maior presença de proteína recombinante nas frações sobrenadantes, produzidas em culturas à temperatura de 37 °C e sob indução de 1 mM de IPTG por 3 h, como demonstrado nas Figuras 8 e 9. Foi observado que a maior proporção da proteína recombinante se encontrava na fração solúvel, conforme análises por SDS-PAGE das frações solúveis e insolúveis (resultados não apresentados) e, também, baseados em estudos realizados por Dabrowski & Kur (1998).

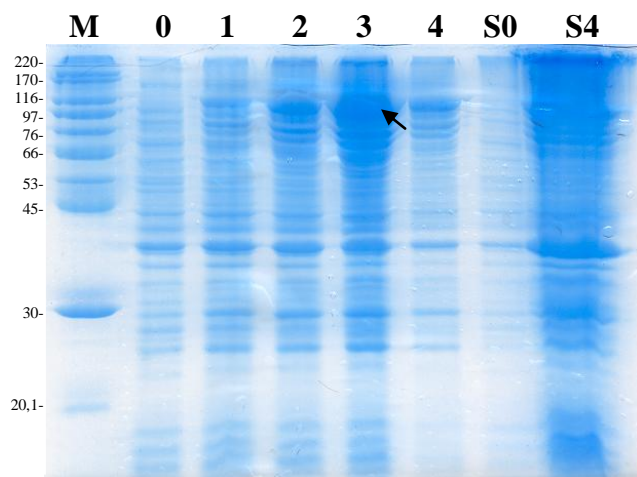
Para a *rPfu* DNA polimerase produzida a partir de pET-23a(+)-*rPfu*, a confirmação de sua produção pode ser observada pela presença de uma banda mais intensa no gel de poliacrilamida com aproximadamente 95 kDa (Figura 8), sendo 90 kDa correspondentes à *rPfu* DNA polimerase e 5 kDa referentes à a cauda de histidinas fusionada no vetor pET-23a(+).

Para a *rPfu* DNA polimerase produzida a partir de pGEX-4T-1-*rPfu*, uma banda mais intensa de aproximadamente 116 kDa é observável (Figura 9), sendo 90 kDa correspondentes

à *rPfu* DNA polimerase e 26 kDa referentes à proteína de fusão GST, conferida pelo vetor pGEX-4T-1.



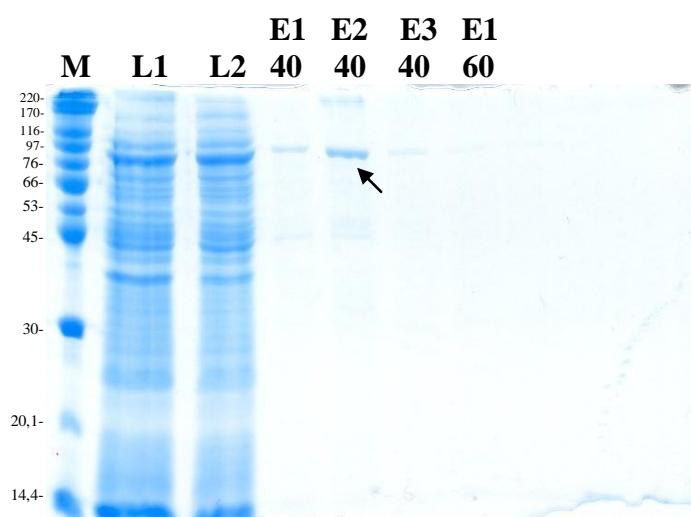
**Figura 8:** Análise da expressão de pET-23a(+)-*rPfu* por SDS-PAGE. Canaletas: (M) Padrão de massa molecular de proteínas HMW-SDS *Marker Kit* e LMW-SDS *Marker Kit* (GE Healthcare). Os valores indicados à esquerda das bandas representam as massas em kDa; Amostras de proteínas totais extraídas de células de *E. coli* BL21(DE3)pLys transformadas com pET-23a(+)-*rPfu* resultantes de diferentes tempos de cultura, indução e frações: (0) cultivo não induzido; (1) cultivo induzido por 1 h; (2) cultivo induzido por 2 h; (3) cultivo induzido por 3 h; (4) cultivo induzido por 4 h; (S0) cultivo não induzido da fração sedimento; (S4) cultivo induzido por 4 h da fração sedimento. A seta indica a banda de 95 kDa referente à *rPfu* DNA polimerase.



**Figura 9:** Análise da expressão de pGEX-4T-1-*rPfu* por SDS-PAGE. Canaletas: (M) Padrão de massa molecular de proteínas HMW-SDS *Marker Kit* e LMW-SDS *Marker Kit* (GE Healthcare). Os valores indicados à esquerda das bandas representam as massas em kDa; Amostras de proteínas totais extraídas de células de *E. coli* BL21(DE3)pLys transformadas com pGEX-4T-1-*rPfu* resultantes de diferentes tempos de cultura, indução e frações: (0) cultivo não induzido; (1) cultivo induzido por 1 h; (2) cultivo induzido por 2 h; (3) cultivo induzido por 3 h; (4) cultivo induzido por 4 h; (S0) cultivo não induzido da fração sedimento; (S4) cultivo induzido por 4 h da fração sedimento. A seta indica a banda de 116 kDa referente à *rPfu* DNA polimerase.

### 5.5 Purificação da *rPfu* DNA polimerase por cromatografia em colunas de níquel

A purificação da *rPfu* DNA polimerase foi realizada apenas a partir da fração solúvel derivada de extratos de proteínas totais de *E. coli* BL21(DE3)pLys transformadas com pET-23a(+)-*rPfu*. Até o momento, ainda não foram realizadas tentativas de purificação da versão da *rPfu* DNA polimerase fusionada à GST, codificada pelo plasmídeo pGEX-4T-1-*rPfu*. Assim, a purificação foi realizada a partir do sobrenadante resultante da centrifugação executada após o rompimento das células com lisozima e por ultrassonicação. A histidina, por meio de seu anel imidazólico, possui afinidade por íons metálicos e a presença de seis resíduos de histidinas na região N-terminal da *rPfu* DNA polimerase possibilitou sua purificação pelo método de cromatografia de afinidade por íons metálicos imobilizados (IMAC, do inglês, *Immobilized Metal ion Affinity Chromatography*). Assim, uma coluna de Protina Ni-TED 1000 (Macherey-Nagel) foi utilizada para a purificação final. Após a aplicação do sobrenadante de centrifugação e um bom volume de tampão de lavagem da matriz, as proteínas contendo sequências poli-histidinas puderam ser facilmente eluídas da coluna por um gradiente de imidazol de 40 a 250 mM. Uma banda mais intensa de 90 kDa pode ser observada neste eluído, conforme representado na Figura 9, o que está de acordo com o peso molecular desta proteína descrita na literatura (Uemori *et al.*, 1993; Dabrowski & Kur, 1998).



**Figura 10:** Análise da purificação da *rPfu* por SDS-PAGE. Canaletas: (M) Padrão de massa molecular de proteínas HMW-SDS *Marker Kit* e LMW-SDS *Marker Kit* (GE Healthcare). Os valores indicados à esquerda das bandas representam as massas em kDa; (L1) Lavagem 1; (L2) Lavagem 2; (E1 40) Eluição 1 com imidazol 40 mM; (E2 40) Eluição 2 com imidazol 40 mM; (E3 40) Eluição 3 com imidazol 40 mM; (E1 60) Eluição 1 com imidazol 60 mM;

## 6. CONCLUSÕES E PERSPECTIVAS

Pelo presente estudo, o gene da DNA polimerase de *P. furiosus* foi adaptado e sintetizado, sendo otimizado para a frequência de códons de *E. coli* e para os sítios de restrição de dois plasmídeos de expressão, simultaneamente. A nova versão do gene, denominada *rPfu* DNA polimerase, foi clonada em pET-23a(+) e pGEX-4T-1 e a proteína recombinante foi expressa em *E. coli* BL21(DE3)pLysS com sucesso. Até a conclusão da presente monografia e por questões temporais, apenas a versão da *rPfu* DNA polimerase codificada por pET-23a(+)-*rPfu* foi purificada por cromatografia de afinidade em colunas de Ni<sup>2+</sup>. A purificação da versão codificada por pGEX-4T-1-*rPfu* constitui-se na primeira atividade a ser realizada futuramente. Para tanto, extratos protéicos serão submetidos à cromatografia em colunas onde a GST está imobilizada. Os resultados em quantidade e qualidade das proteínas recombinantes produzidas empregando-se ambos os métodos indicarão qual sistema de purificação apresenta melhor rendimento e menores custos.

Até o momento, não pudemos realizar ensaios para determinar a atividade de DNA polimerase das proteínas recombinantes produzidas, o que é de fundamental importância a todo o Projeto. Estes ensaios também serão realizados de forma comparativa a outras DNA polimerases termoestáveis e de alta fidelidade disponíveis no mercado. A comprovação de que as versões da *rPfu* DNA polimerase sejam tão ou mais ativas do que as versões comerciais indicará a possibilidade de torná-las produtos comerciais. Assim, é objetivo futuro realizar consulta mais segura sobre os direitos legais de produção comercial da enzima no Brasil e a proposição de produção em escala industrial da mesma, ou *scale-up* do procedimento de produção.



## REFERÊNCIAS BIBLIOGRÁFICAS

- AREZI, B.; XING, W.; SORGE, J. A.; HOGREFE, H. H. **Amplification efficiency of thermostable DNA polymerases.** Analytical Biochemistry, v. 321, p. 226–235, 2003.
- BILES, B. D.; CONNOLLY, B. A. **Low-fidelity *Pyrococcus furiosus* DNA polymerase mutants useful in error-prone PCR.** Nucleic Acids Research, v. 32, n. 22, p. 176-182, 2004.
- CLINE, J.; BRAMAN, J. C.; HOGREFE, H. H. **PCR fidelity of *Pfu* DNA Polymerases and other thermostable DNA polymerases.** Nucleic Acids Research, v. 24, n. 18, p. 3546-3551, 1996.
- CRUZ, N. A. V.; CASPETA, L.; PÉREZ, N. O.; RAMÍREZ, O. T.; ROLDÁN, M. A. T. **Production of recombinant proteins in *E. coli* by the heat inducible expression system based on the phage lambda pL and/or pR promoters.** Microbial Cell Factories, v. 9, n. 18, 2010.
- DABROWSKI, S.; KUR, J. **Cloning and Expression in *Escherichia coli* of the Recombinant His-Tagged DNA Polymerases from *Pyrococcus furiosus* and *Pyrococcus woesei*.** Protein Expression and Purification, v. 14, p. 131-138, 1998.
- DEEPAK, S. A.; KOTTAPALLI, K. R.; RAKWAL, R.; OROS, G.; RANGAPPA, K. S.; IWAHASHI, H.; MASUO, Y.; AGRAWAL, G. K. **Real-Time PCR: Revolutionizing Detection and Expression Analysis of Genes.** Current Genomics, v. 8, p. 234-251, 2007.
- DUBENDORFF, J. W.; STUDIER, F. W. **Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor.** J. Mol. Biol. v. 219, n. 1, p. 45-59, 1991
- ECKERT, K. A.; KUNKEL, T. A. **DNA Polymerase Fidelity and the Polymerase Chain Reaction.** Cold Spring Harbor laboratory Press, v. 1, p. 17-24, 1991.
- ELNIFRO, E. M.; ASHSHI, A. M.; COOPER, R. J.; KLAPPER, P. E. **Multiplex PCR: Optimization and Application in Diagnostic Virology.** Clinical Microbiology Reviews, v. 13, n. 4, p. 559-570, 2000.
- ERLICH, H. A.; GELFRAND, D.; SNINSKY, J. J., **Recent Advances in the polymerase Chain Reaction.** Science, v. 292, p. 1643-1651, 1991.

HAKI, G. D.; RAKSHIT, S. R. **Developments in industrially important thermostable enzymes: a review.** *Bioresource Technology*, v. 89, p. 17-34, 2003.

JANA, S.; DEB, J. K. **Strategies for efficient production of heterologous proteins in *Escherichia coli*.** *Appl. Microbiol Biotechnol*, v. 67, p. 289-298, 2005.

KIM, S. W.; KIM, D. U.; KIM, J. K.; KANG, L. W.; CHO, H. S. **Crystal structure of *Pfu*, the high fidelity DNA polymerase from *Pyrococcus furiosus*.** *International Journal of Biological Macromolecules*, v. 42, p. 356-361, 2008.

KIM, Y.; FLYNN, T. R.; DONOFF, R. B.; WONG, D. T. W.; TODD, R. **The Gene: The Polymerase Chain Reaction and Its Clinical Application.** *J Oral Maxillofac Surg*, v. 60, p. 808-815, 2002.

KIM, Y. J.; LEE, H.S.; BAE, S.S.; JEON, J. H.; LIM, J. K.; CHO, Y.; NAM, K. H.; KANG, S. G.; KIM, S. J.; KWON, S. T.; LEE, J. H. **Cloning, Purification, and Characterization of a New DNA Polymerase from a Hyperthermophilic Archaeon, *Thermococcus* sp. NA1.** *J. Microbiol. Biotechnol*, v. 17, n. 7, p. 1090-1097, 2007.

KRAMER, M. F.; DONALD, M. C. **Enzymatic Amplification of DNA by PCR: Standard Procedures and Optimization.** *Current Protocols in Molecular Biology*, 2001.

KUBISTA, M.; ANDRADE, J. M.; BENGTSSON, M.; FOROSTAN, A.; JONÁK, J.; LIN, K.; SINDELKA, R.; SJOBACK, R.; SJOGREEN, B.; STROMBOM, L.; STAHLBERG, A.; ZORIC, N. **The real-time polymerase chain reaction.** *Molecular Aspects of Medicine*, v. 27, p. 95-125, 2006.

LAEMMLI, U. K. **Cleavage of structural proteins during the assembly of the head of bacteriophage T4.** *Nature*, v. 227, p. 680-685, 1970.

LEE, S. Y. **High cell-density culture of *Escherichia coli*.** *Trends Biotechnology*, v. 14, p. 98-105, 1996.

LU, C.; ERICKSON, H. P. **Expression in *Escherichia coli* of the Thermostable DNA Polymerases from *Pyrococcus furiosus*.** *Protein Expression and Purification*, v. 11, p. 179-184, 1997.

MULLIS, K. **The unusual origin of the polymerase chain reaction.** *Scientific American*, v. 262, n. 4, p. 56-65, 1990.

NI, Y.; CHEN, R. **Extracellular recombinant protein production from *Escherichia coli***. Biotechnol Lett, v. 31, p. 1661–1670, 2009.

PAVLOV, A. R.; PAVLOVA, N. V.; KOZYAVKIN, S. A.; SLESAREV, A. I. **Recent developments in the optimization of thermostable DNA polymerases for efficient applications**. Trends in Biotechnology, v. 22, n. 5, p. 253-260, 2004.

SAMBROOK, J.; RUSSELL, D. W. **Molecular cloning: a Laboratory manual**. 3<sup>rd</sup> ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York: CSHL, 2001.

SCHMIDT, F. R. **Optimization and scale up of industrial fermentation processes**. Appl Microbiol Biotechnol, v. 68, p. 425-435, 2005.

SIVASHANMUGAM, A.; MURRAY, V.; CUI, C.; ZHANG, Y.; WANG, J.; LI, Q. **Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli***. Protein Science, v. 18, p. 936-948, 2009.

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

SPEERS, D. J. **Clinical Applications of Molecular Biology for Infectious Diseases**. Clin Biochem Rev, v. 27, p. 39-51, 2006.

SWARTZ, J. R. **Advances in *Escherichia coli* production of therapeutic proteins**. Current Opinion in Biotechnology, v. 12, p. 195-201, 2001.

TAKAGI, M.; NISHIOKA, M.; KAKIHARA, H.; KITABAYASHI, M.; INOUE, H.; KAWAKAMI, B.; OKA, M.; IMANAKA, T. **Characterization of DNA Polymerase from *Pyrococcus* sp. Strain KOD1 and Its Application to PCR**. Applied and Environmental Microbiology, v. 63, n. 11, p. 4504-4510, 1997.

UEMORI, T.; ISHINO, Y.; TOH, H.; ASADA, K.; KATO, I. **Organization and nucleotide sequence of the DNA polymerase gene from the archaeon *Pyrococcus furiosus***. Nucleic Acids Research, v. 21, n. 2, p. 259-265, 1993.

UHLMANN, V.; SILVA, I.; LUTTICH, K.; PICTON, S.; O'LEARY, J. J. **In cell amplification**. J Clin Pathol: Mol Pathol, v. 51, p. 119–130, 1998.

VIEILLE, C.; ZEIKUS, G. J. **Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability.** Microbiology and Molecular Biology Reviews, v. 65, n. 1, p. 1-43, 2001.

WALSH, S. J. **Recent advances in forensic genetics.** Expert Rev Mol Diagn, v. 4, n. 1, p. 31-40, 2004.

YANG, S.; ROTHMAN, R. E. **PCR-based diagnostics for infectious diseases: uses, limitations, and future applications in acute-care settings.** Lancet Infect Dis, v. 4, p. 337-348, 2004.