



Pontifícia Universidade Católica do Rio grande do Sul  
Faculdade de Biociências  
Programa de Pós-Graduação em Biologia Celular e Molecular

**Clonagem, Superexpressão, Purificação e Caracterização da  
Proteína Recombinante Humana Fator Estimulador de Colônias de  
Granulócitos**

Autor

Ana Letícia de Souza Vanz

Orientador

Prof. Dr. Luiz Augusto Basso

Co-orientador

Prof. Dr. Diógenes Santiago Santos

Porto Alegre

Fevereiro de 2008

Pontifícia Universidade Católica do Rio Grande do Sul  
Faculdade de Biociências  
Programa de Pós-Graduação em Biologia Celular e Molecular

**Clonagem, Superexpressão, Purificação e Caracterização da  
Proteína Recombinante Humana Fator Estimulador de Colônias de  
Granulócitos**

Dissertação apresentada ao  
Programa de Pós-Graduação  
em Biologia Celular e Molecular  
como requisito para obtenção do  
grau de Mestre.

Autor  
Ana Letícia de Souza Vanz

Orientador  
Prof. Dr. Luiz Augusto Basso

Co-orientador  
Prof. Dr. Diógenes Santiago Santos

Porto Alegre  
Fevereiro de 2008

## **AGRADECIMENTOS**

*Aos professores Prof. Dr. Diógenes Santiago Santos e Prof. Dr. Luiz Augusto Basso pela possibilidade de desenvolver este trabalho de mestrado, o qual me trouxe enorme crescimento pessoal e científico.*

*Aos doutores Gaby Renard, Cláudia Paiva Nunes, Eraldo Batista Júnior, Maria Martha Campos e Jocelei Maria Chies pela ajuda essencial no desenvolvimento desta pesquisa através de compreensão, conhecimento e enriquecedor convívio.*

*A todos os colegas do laboratório que de alguma maneira me ajudaram a concluir este trabalho. Certamente são verdadeiros amigos que estarão sempre presentes na minha vida.*

*Ao Programa de Pós-Graduação em Biologia Celular e Molecular da PUCRS.*

*Ao Laboratório Americano de Farmacoterapia S.A. (FARMASA) pela bolsa de estudos para realizar o curso de Mestrado.*

## SUMÁRIO

AGRADECIMENTOS.....	i
LISTA DE ABREVIATURAS.....	iv
RESUMO.....	vi
ABSTRACT.....	vii
1. INTRODUÇÃO.....	1
1.1 ATIVIDADE BIOLÓGICA DO G-CSF.....	2
1.2 CARACTERIZAÇÃO MOLECULAR E BIOLÓGICA.....	3
1.2.1 Interação com receptor.....	6
1.3 IMPORTÂNCIA CLÍNICA.....	8
1.4 CLONAGEM E EXPRESSÃO DO G-CSF.....	11
1.4.1 Corpos de inclusão: solubilização e renaturação.....	12
1.5 IMPORTÂNCIA SOCIAL E ECONÔMICA.....	15
2. OBJETIVOS.....	19
2.1 Objetivo geral.....	19
2.2 Objetivos específicos.....	19
3. ARTIGO CIENTÍFICO.....	20
Abstract.....	23
Background.....	25
Results and discussion.....	27
Conclusion.....	32
Methods.....	33
List of abbreviations.....	39
Competing interests.....	39

Authors' contributions.....	39
Acknowledgments.....	40
References.....	40
Figure legends.....	46
4. CONSIDERAÇÕES FINAIS.....	56
REFERÊNCIAS BIBLIOGRÁFICAS.....	58
ANEXO I - Testes de superexpressão.....	64
ANEXO II - Testes de solubilização.....	67
ANEXO III - Testes de purificação.....	71
ANEXO IV - Comprovante de aceite para publicação do artigo científico.....	73

## LISTA DE ABREVIATURAS

CFU - Unidades formadoras de colônias  
CHO – Linhagem celular derivada do ovário de hamster chinês  
CSFs - Fatores estimuladores de colônias  
*E. coli* – *Escherichia coli*  
FDA - Departamento de controle de drogas e alimentos dos EUA  
FPLC - Cromatografia líquida de rápida performance  
hG-CSF- Fator estimulador de colônias de granulócitos humano  
rhG-CSF- Fator estimulador de colônias de granulócitos recombinante humano  
G-CSF- Fator estimulador de colônias de granulócitos  
GM-CSF- Fator estimulador de colônias de granulócitos e macrófagos  
GMP – Boas práticas de manipulação  
IL-1 – Interleucina 1  
IL-3 – Interleucina 3  
IL- 4 – Interleucina 4  
INF- $\gamma$  – Interferon  $\gamma$   
IPTG - isopropil  $\beta$ -D-tiogalactopiranosídeo  
JAKS - *Janus Kinases*  
kDa – Quilo Dalton  
LB – meio de cultura Luria Bertani  
LMA - Leucemia Mielóide Aguda  
M-CSF- Fator estimulador de colônias de macrófagos  
NK - *Natural killer*  
OD<sub>600</sub> – Densidade ótica à 600 nm  
PCR – Reação em cadeia pela polimerase  
PMSF- Fenilmetilsulfonil fluoreto  
SDS-PAGE – Eletroforese em gel de poliacrilamida com dodecil sulfato de sódio  
SDS - Dodecil sulfato de sódio  
AIDS – Síndrome da imunodeficiência adquirida  
STATs – Transdutores de sinal e ativadores da transcrição  
SUS – Sistema Único de Saúde  
TNF-  $\alpha$  – Fator de necrose tumoral  $\alpha$

SEC-HPLC – Cromatografia líquida de exclusão molecular

RP-HPLC – Cromatografia líquida de fase reversa

## RESUMO

O fator estimulador de colônias de granulócitos (G-CSF) é uma citocina hematopoiética que age sobre a linhagem de neutrófilos promovendo proliferação e diferenciação de seus precursores e ativação dos neutrófilos maduros. O G-CSF é uma proteína com 18,8 kDa, constituída por 174 aminoácidos possuindo duas pontes dissulfeto intra-moleculares e uma cisteína livre no resíduo 17. Este biofármaco tem sido empregado com sucesso em pacientes com câncer que recebem altas doses de quimioterapia. Além disso, também tem sido usado como tratamento ou profilaticamente a fim de reforçar o sistema imune em pacientes com HIV, pneumonia, pacientes diabéticos com infecções nos pés, leucemia e neutropenia febril. Em função desta ampla aplicação clínica, hG-CSF recombinante tem sido produzido por engenharia genética em *Escherichia coli* e foi aprovado para uso em neutropenia provocada por quimioterapia pelo FDA em 1991. Filgrastima (nome genérico do G-CSF) teve sua patente expirada em 2006, tornando-se alvo das indústrias farmacêuticas. Atualmente, o Brasil é totalmente dependente da importação deste biofármaco. Logo, o objetivo deste estudo é desenvolver uma metodologia para a posterior produção de uma Filgrastima nacional. Neste trabalho, o gene para hG-CSF foi construído por PCR, clonado no vetor de expressão pET23a(+), usando as enzimas de restrição *NdeI* e *BamHI*. Os testes de superexpressão foram realizados em diferentes cepas de *E. coli*, mostrando a melhor condição de expressão na fração insolúvel na cepa BL21(DE3). Na tentativa de solubilizar os corpos de inclusão e purificar a proteína, inúmeros protocolos foram testados. Por fim, foi descrito um eficiente protocolo de isolamento e solubilização dos corpos de inclusão por um processo de múltiplos passos de lavagem e um método de purificação usando somente uma coluna cromatográfica de troca catiônica. A identidade da proteína foi confirmada por seqüenciamento N-terminal e *Western blotting*. A caracterização do rhG-CSF homogêneo, através de SEC-HPLC e RP-HPLC, mostrou resultados similares aos do padrão internacional. O teste de atividade biológica, *in vivo*, demonstrou que o rhG-CSF produzido tem potencial equivalente ao padrão internacional utilizado. A proteína foi produzida por um processo simples e econômico, sendo de extrema importância em um processo industrial, podendo trazer benefícios para a saúde da população.



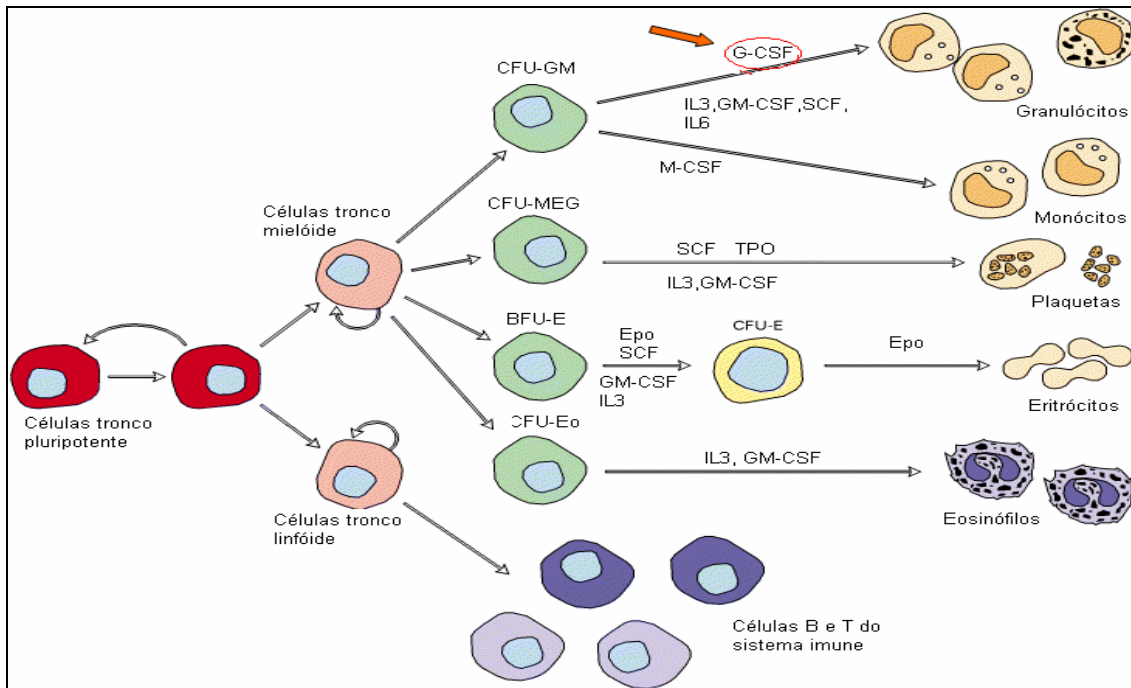
## ABSTRACT

The granulocyte colony stimulating factor (G-CSF) is a hematopoietic cytokine that acts on cells of the neutrophil lineage causing proliferation and differentiation of committed precursor and activation of mature neutrophils. The hG-CSF is an 18.8 kDa protein consisting of 174 amino acid polypeptide chain with two intra-molecular disulphide bonds and one free cysteine at residue 17. This biopharmaceutical has been widely used with success in oncology patients who receive high-dose chemotherapy. It is also used as treatment and prophylactically to improve the immune system in patients with HIV, pneumonia, diabetic foot infections, leukemia and febrile neutropenia. Based on this large clinical application, the recombinant hG-CSF has been produced in genetically engineered *Escherichia coli* and was approved for use in chemotherapy-induced neutropenia by the U.S Food and Drug Administration in 1991. Filgrastim (generic name of G-CSF) lost its patent protection in 2006 becoming a target of Brazilian pharmaceutical industries. Currently, Brazil is totally dependent of the importation of this biopharmaceutical. Therefore, the aim of this work is to develop a methodology for subsequent production of a national Filgrastim. In this work the granulocyte colony-stimulating factor gene was assembled by PCR, Its amplicon was cloned into pET23a(+) expression vector using *NdeI* and *BamHI*, restriction enzymes. The overexpression was tested in different strains of *E. coli* cells and the best condition for expression of the protein was in the BL21(DE3) strain. In order to solubilize the inclusion bodies and purify the protein, many protocols have been tested. Finally, it was described an efficient protocol of isolation of inclusion bodies through a multi-step washing procedure and purification method of the recombinant protein from inclusion bodies using only a cationic exchange column. The immunoassay and N-terminal sequencing confirmed the identity of rhG-CSF. Characterization of homogeneous rhG-CSF using SEC-HPLC and RP-HPLC has shown similarity to the international standard. The biological activity assay, *in vivo*, has shown an equivalent biological effect to those obtained with the standard reference rhG-CSF. The protein rhG-CSF was produced through simple, cost effective and economically feasible process of rhG-CSF, which has extreme importance to the industrial process and healthcare community.

## 1. INTRODUÇÃO

A produção de células hematopoiéticas é mantida pelas células-tronco pluripotentes comuns na medula óssea (1), as quais representam em torno de 0,01% das células nucleadas (2). Estima-se que para a manutenção do *pool* de células hematopoiéticas de um adulto que pese 70 kg, a medula óssea deva produzir aproximadamente um trilhão de células ao dia (2).

A diferenciação celular inicia-se a partir das células-tronco pluripotentes que se dividem para gerar mais células-tronco (auto-renovação) e várias células progenitoras comprometidas (geralmente designadas como Unidades Formadoras de Colônias - CFU), que são capazes de originar apenas alguns tipos celulares. As células progenitoras comprometidas diferenciam-se em células sanguíneas maduras sob a influência de moléculas sinalizadoras chamadas Fatores Estimulantes de Colônias (CSFs), entre elas está o Fator Estimulador de Colônias de Granulócitos, G-CSF (do inglês, *granulocyte colony stimulating factor*) (Figura 1) (1).



**Figura 1.** Diferenciação celular.

Fonte: Lodish H. et al. (3)

Estas citocinas são secretadas pelo estroma, que é o espaço medular constituído por fibroblastos, adipócitos, macrófagos, linfócitos, células endoteliais e matriz extracelular. O estroma constitui o microambiente que possibilita o crescimento e a diferenciação das células hematopoiéticas (2,4).

### 1.1 ATIVIDADE BIOLÓGICA DO G-CSF

A primeira linha de defesa celular contra agentes infecciosos compreende: granulócitos polimorfonucleares, macrófagos, células NK (do inglês, *Natural Killer*) e linfócitos citotóxicos (5,6). Assim, o fator estimulador de colônias de granulócitos (G-CSF) torna-se uma das principais citocinas envolvidas na estratégia de defesa do sistema imune (5), pois seu papel é estimular a proliferação, diferenciação e

atividade dos granulócitos polimorfonucleares (7,8). Estas células exercem sua função de defesa por quimiotaxia, fagocitose, morte intracelular e liberação de vários produtos, incluindo enzimas lisossomais (5).

Os granulócitos recebem esta denominação por possuírem grânulos densamente coráveis em seu citoplasma, e se dividem, segundo Janeway e colaboradores (9) em:

-neutrófilos: que são as células fagocitárias mais numerosas e importantes da resposta inata;

-eosinófilos: importantes na defesa contra infecções por parasitas e em respostas alérgicas;

-basófilos: participam de reações de hipersensibilidade imediata e têm provavelmente uma função similar e complementar a dos eosinófilos e mastócitos.

O G-CSF atua principalmente promovendo a maturação dos neutrófilos e estimulando sua atividade fagocítica e quimiotática, além de estar envolvido com o processo de segmentação nuclear dos neutrófilos maduros. O G-CSF também atua em outras linhagens celulares, apresentando um importante papel na mobilização de células-tronco hematopoiéticas da medula óssea para a circulação (10). Também é capaz de modular a resposta inflamatória, reduzindo a liberação as citocinas pró-inflamatórias por monócitos e macrófagos ativados (11).

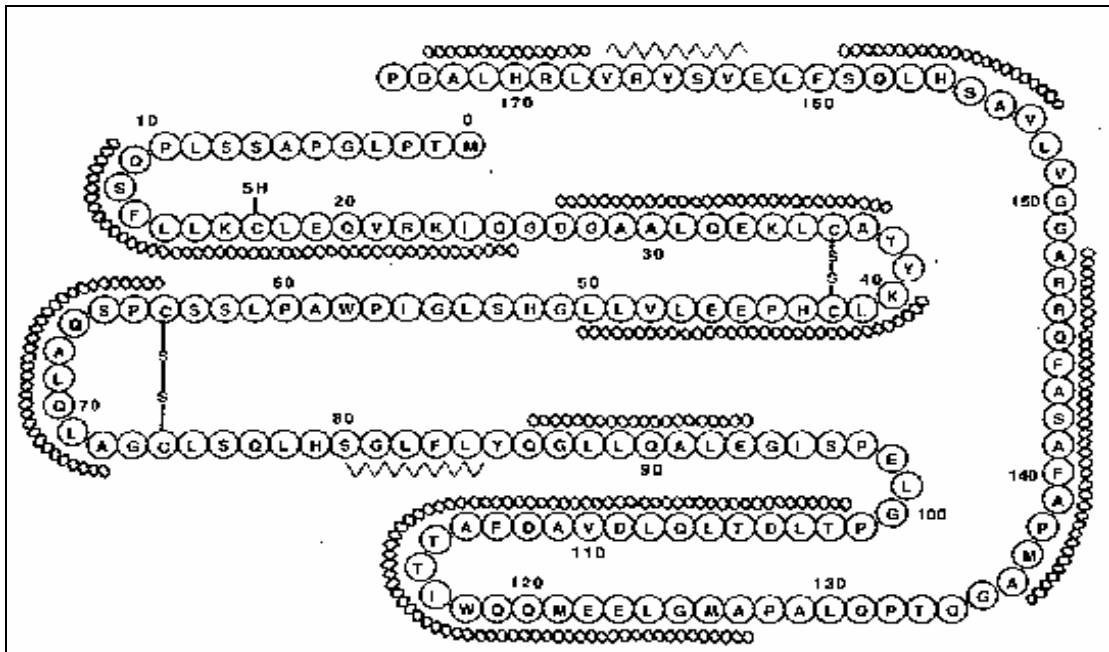
## 1.2 CARACTERIZAÇÃO MOLECULAR E BIOLÓGICA

O G-CSF é produzido por uma variedade de células, das quais monócitos e macrófagos são considerados os maiores produtores. O aumento da produção e liberação desta proteína tem sido observado após estimulação com:

- lipopolissacarídeos (LPS): um componente da parede celular de bactérias Gram-negativas;
- IL-1 e TNF- $\alpha$ : citocinas pró-inflamatórias que coordenam respostas do organismo contra infecções;
- IL-3: interleucina relacionada com hematopoiese;
- IL-4: interleucina envolvida na ativação de células B e síntese de eosinófilos;
- GM-CSF: fator estimulador de granulócitos e macrófagos;
- M-CSF: fator estimulador de macrófagos;
- INF- $\gamma$ : possui ação antiviral;
- *Mycobacterium avium*.

Além disso, outras células como fibroblastos, células endoteliais, linfócitos T e granulócitos polimorfonucleares estão relacionados com a produção de G-CSF (9,10).

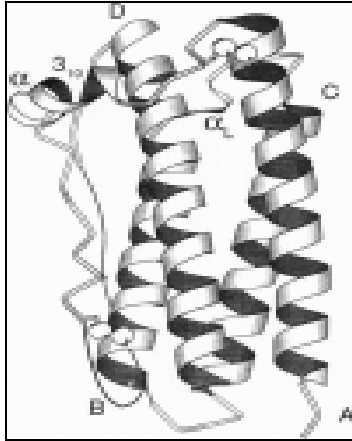
A Filgrastima (nome genérico do G-CSF) é uma molécula monomérica em solução que possui uma seqüência de 174 aminoácidos e tem peso molecular aproximado de 18,8 kDa. A molécula contém um resíduo de cisteína livre na posição 17 e duas pontes dissulfeto, Cys<sup>(36)</sup>-Cys<sup>(42)</sup> e Cys<sup>(64)</sup>-Cys<sup>(74)</sup>, que são importantes para a estabilidade estrutural da proteína e dobramento correto desta longa cadeia de aminoácidos (Figura 2) (8).



**Figura 2.** Sequência de aminoácidos do G-CSF. Duas pontes dissulfeto.  $\alpha\alpha\alpha\alpha$  segmentos em  $\alpha$ -hélice,  $\sim\sim\sim$  segmentos em folhas- $\beta$ .

Fonte: Lu S. *et al* (8)

A estrutura secundária da molécula contém aproximadamente 69% de  $\alpha$ -hélice, 4% de folha- $\beta$ , 5% de  $\beta$ -curvatura, sendo que o restante da molécula ainda não foi caracterizado (Figura 2) (8). É composta por quatro hélices, denominadas A, B, C e D, como mostra a Figura 3. A hélice A forma-se entre os resíduos 11-39, a B entre 71-91, a C entre 100-123 e a D entre 143-172 (12).



**Figura 3.** Estrutura do G-CSF

Fonte: Hill C. *et al.* (12)

O gene que codifica para o G-CSF em humano está localizado no cromossomo 17 q21-22 (13). A forma nativa produzida naturalmente é glicosilada, o que confere resistência à degradação por proteases e pode influenciar toda a estrutura da proteína. A glicosilação também parece prevenir a agregação da proteína pelo aumento da solubilidade e estabilidade desta molécula que é altamente hidrofóbica em pH neutro. Porém, tanto a forma nativa quanto a não glicosilada apresentam atividade biológica (14).

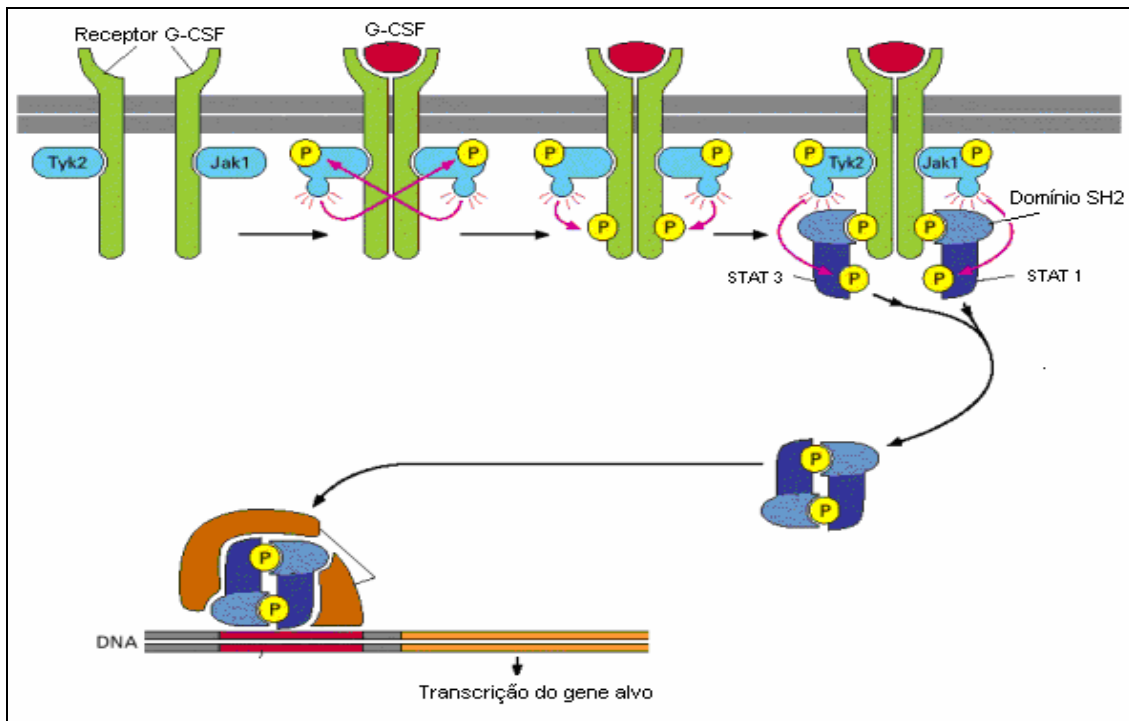
### 1.2.1 Interação com receptor

Todos os efeitos biológicos do G-CSF são mediados via ativação do receptor G-CSF (G-CSFR), um membro da superfamília de receptores de citocinas tipo I. O receptor de G-CSF humano é conhecido por CD114. Este receptor é primeiramente expresso em progenitores neutrofilicos e neutrófilos maduros, transmitindo principalmente sinais de proliferação, diferenciação e sobrevivência destas células. Pode também ser encontrado em monócitos, com a função de atenuar a liberação de citocinas pró-inflamatórias quando estas células se encontram ativadas. Foi ainda

identificado em plaquetas maduras e várias outras células não hematopoiéticas e tecidos, incluindo células endoteliais, placenta e células de leucemia mielóide. A maturação de neutrófilos permite um aumento da expressão dos receptores, sendo que cada célula pode apresentar aproximadamente 200-1000 receptores (10).

A maioria das citocinas, incluindo o G-CSF, liga-se a receptores que podem ativar as proteínas de regulação gênica mantidas em estado latente na membrana plasmática. Os receptores de G-CSF estão associados à tirosinas quinases citoplasmáticas chamadas Jaks (Janus quinases) - JAK1, JAK2 e TYK2 -, que são ativadas quando esta citocina liga-se ao receptor. Após a ligação, as duas cadeias do receptor ficam próximas e as Jaks se fosforilam reciprocamente ativando-se que, por sua vez, fosforilam as tirosinas do receptor criando um local para ligação das proteínas reguladoras gênicas citoplasmáticas STATs ( do inglês *Signal Transducers and Activators of Transcription*) - STAT1, STAT 3 e STAT 5. Estas possuem o domínio SH2, que se liga ao receptor. Depois da ligação, as Jaks fosforilam e ativam estas proteínas regulatórias causando dissociação do receptor. Por meio do domínio SH2 ocorre a formação de dímeros pelas duas STATs. Então, estas proteínas direcionam-se ao núcleo, onde estimulam a transcrição de genes-alvo específicos (Figura 4) (1).





**Figura 4.** Rota de sinalização Jak-STAT ativado por citocina.

Fonte: Alberts et al. (1)

A resposta mediada por STATs é frequentemente regulada por mecanismo de retroalimentação negativa. Os dímeros de STAT também podem ativar genes que codificam proteínas inibitórias. Estas se ligam tanto no receptor de citocina como na própria STAT bloqueando a cascata de sinalização. Entretanto, este mecanismo pode não ser suficiente para bloquear totalmente a cascata, ocorrendo uma inativação das Jaks e STATs por desfosforilação, através de enzimas específicas (1).

### 1.3 IMPORTÂNCIA CLÍNICA

Baseado nas atividades biológicas, o G-CSF tem sido usado no tratamento de diversas patologias, sobretudo em neutropenias provocadas pela quimioterapia

usada no tratamento de tumores, pela radioterapia e pelo uso de drogas que suprimem a produção de células mielóides (5,15,16). Desde 1991, o uso deste biofármaco é liberado pela agência reguladora dos Estados Unidos *U.S. Food & Drug Administration* (FDA), com a finalidade de diminuir a incidência de infecções associadas com neutropenia induzida por quimioterapia em pacientes com câncer (17).

O quadro de neutropenia se caracteriza pela diminuição da contagem normal de neutrófilos no sangue, sendo que para um adulto estes valores são de 1500 a 7000/mm<sup>3</sup> (1,5 – 7,0 x 10<sup>9</sup>/L). Segundo Kufe e colaboradores (18), os níveis de neutropenia se dividem em:

- neutropenia média, quando a contagem de neutrófilos fica entre 1000/mm<sup>3</sup> (1,0 x 10<sup>9</sup> /L) e 1500/mm<sup>3</sup> (1,5 x 10<sup>9</sup> /L);
- neutropenia moderada, quando a contagem de neutrófilos fica entre 500/mm<sup>3</sup> (0,5 x 10<sup>9</sup> /L) e 1000/mm<sup>3</sup> (1,0 x 10<sup>9</sup> /L);
- neutropenia severa, quando a contagem de neutrófilos fica abaixo de 500/mm<sup>3</sup> (0,5 x 10<sup>9</sup> /L).

A associação de neutropenia e infecções continua sendo a maior causa de morbidade e mortalidade em pacientes com câncer que recebem tratamento de quimioterapia (15). Inúmeros estudos vêm apresentando resultados positivos com o uso de G-CSF nestes pacientes. Crianças com câncer que recebem doses elevadas de quimioterapia se tornam suscetíveis a infecções, e a severidade destas infecções está diretamente relacionada com a duração da neutropenia causada pela quimioterapia. Nos últimos 10-15 anos, a terapia profilática com G-CSF tem apresentado bastante sucesso tanto em pacientes adultos como crianças. Análises adicionais mostraram que a incidência de neutropenia febril, bem como a duração da neutropenia, os dias de hospitalização e a terapia com antibióticos diminuem

expressivamente com o uso profilático de fatores de crescimento, entre eles o G-CSF (16). Um estudo realizado por Carbonero e colaboradores (15) em pacientes neutropênicos portadores de tumor sólido tratados com quimioterapia mostrou uma melhora clinicamente relevante e a conseqüente diminuição dos custos hospitalares após o uso de G-CSF.

Embora ainda existam controvérsias em relação ao uso de G-CSF em pacientes com Leucemia Mielóide Aguda (LMA), um estudo de Harrousseau e colaboradores (19) realizado com pacientes em tratamento quimioterápico, demonstrou que a administração de G-CSF rotineiramente após as sessões de quimioterapia reduz não somente o tempo de neutropenia, mas também diminui significativamente os dias de hospitalização e terapia com outras drogas como antibióticos e/ou antifúngicos. Nesta análise, o autor ressalta que o impacto destes resultados na diminuição dos gastos hospitalares deveria ser analisado.

Nos últimos anos, outros alvos para o uso de G-CSF vêm sendo apresentados, como o tratamento de infecções em pacientes não neutropênicos, a fim de melhorar sua defesa imune, bem como medida preventiva (5,6). Entre eles, há relatos de benefícios em caso de peritonites e infecção tecidual em ratos e camundongos (5).

Babalola e colaboradores (20) propuseram, em um estudo com camundongos, avaliar a eficiência da administração conjunta de G-CSF e do antibiótico ceftazidima no tratamento de pneumonia causada por *Pseudomonas aeruginosa*, tanto em hospedeiros neutropênicos como não-neutropênicos, após a indução de infecção. Este estudo apresentou resultados satisfatórios, sugerindo uma nova opção de tratamento para pneumonia, devido aos resultados positivos obtidos em pacientes neutropênicos.

Análises adicionais realizadas em pacientes portadores de HIV em estágio avançado demonstraram que a administração de G-CSF foi efetiva na prevenção de neutropenia severa, levando a uma redução na incidência e duração de infecções bacterianas e diminuição dos dias de tratamento com antibióticos por via intravenosa e dias de hospitalização. Estes resultados sugerem que o uso de Filgrastima permite a manutenção de uma contagem normal de neutrófilos, promovendo um benefício clínico significativo em pacientes portadores de HIV (21).

Outro estudo foi realizado em pacientes com diabetes dependentes de insulina com infecção nos pés, cuja ocorrência é muito comum neste tipo de paciente. Foi demonstrada uma significativa melhora na erradicação de patógenos, encurtando o tempo de internação no hospital e diminuindo o uso de antibióticos (22).

Devido ao aumento do número de pacientes com problemas no coração nos últimos 40 anos, se fazem necessárias novas estratégias de tratamentos. Um estudo feito por Harada e colaboradores (23) sugere que citocinas hematopoiéticas, como G-CSF, podem ser capazes de melhorar a função cardíaca e reduzir a taxa de mortalidade após infarto no miocárdio em camundongos. Essa ação se dá possivelmente pelo efeito protetor sobre os cardiomiócitos, a promoção da angiogênese e a prevenção da remodelação cardíaca do ventrículo esquerdo após infarto.

#### 1.4 CLONAGEM E EXPRESSÃO DO G-CSF

A clonagem molecular e expressão do gene do G-CSF humano foram descritas primeiramente por dois grupos de pesquisa e publicados por: Nagata, 1986 e Souza, 1986 (24,25). O grupo de pesquisa liderado por Larry Souza clonou o gene

do G-CSF e produziu a proteína recombinante. Em 1989, junto com o laboratório americano AMGEN, patenteou o processo (Patente US 4,810,643) e lançou no mercado o produto nomeado Neupogem<sup>®</sup> (Filgrastima) (26).

A proteína recombinante pode ser expressa na forma glicosilada em células eucarióticas (linhagem celular derivada de ovário de Hamster chinês - CHO) e na forma não-glicosilada, com uma metionina extra na posição N-terminal em procariotos (*E. coli*). Entre muitos organismos, a enterobactéria *E. coli* é um dos procariotos mais utilizados para manipulação genética e produção industrial de proteínas heterólogas para fins terapêuticos. *E. coli* oferece várias vantagens, incluindo sua estrutura molecular bem caracterizada, altos níveis de expressão de proteínas heterólogas, crescimento em meio de cultura de baixo custo, acúmulo rápido de biomassa e simples processo de escalonamento (27,28).

#### 1.4.1 Corpos de inclusão: solubilização e renaturação

A superexpressão de proteínas eucarióticas recombinantes, por exemplo, G-CSF, no citoplasma de *E. coli* frequentemente demanda a formação de pontes dissulfeto e modificações pós-traducionais, como a glicosilação, para alcançar uma conformação igual à nativa e biologicamente ativa. Isto pode ser um desafio, pois a maquinaria de ligação da proteína e a condição conformacional em procariotos são diferentes dos eucariotos, muitas vezes levando à deposição de proteínas na forma inativa de agregados insolúveis, conhecidos como corpos de inclusão (27,29). Além disso, o uso de promotores fortes e altas concentrações de indutores permite uma super produção de proteína, que algumas vezes pode exceder 50% do total das proteínas bacterianas. Nestas condições, o dobramento correto da proteína é prejudicado (27).

Os corpos de inclusão podem ser observados em microscópio óptico na forma cilíndrica ou ovóide, com cerca de  $2 \mu\text{m}^3$ . Apesar de ser uma partícula densa, é altamente hidratada e apresenta poros em sua superfície (30).

O correto dobramento da proteína e sua forma biologicamente ativa podem ser obtidos a partir dos corpos de inclusão por completa desnaturação, seguida de renaturação *in vitro*. A solubilização pode ser alcançada com ajuda de agentes desnaturantes como: uréia, cloreto de guanidina e *N-lauroyl-sarcosine* (29,31), os quais provocam um completo ou parcial desdobramento na estrutura das proteínas (32). A renaturação pode ser alcançada por diminuição gradativa da concentração do desnaturante utilizado por troca de tampão, empregando diálise e cromatografia de gel filtração. Por consumir muito tempo e freqüentemente produzir baixas quantidades da proteína ativa, este processo todo não é tão simples de ser realizado, podendo resultar em passos adicionais no processo industrial, muitas vezes não sendo conveniente (32).

Uma das estratégias para evitar a formação destes agregados é a expressão do gene em baixas temperaturas, sendo que a vantagem é a diminuição do grau de transcrição, tradução e redução da força de interação hidrofóbica a qual contribui para o dobramento errôneo da proteína (27,33,34). Outra alternativa é reduzir a taxa de síntese da proteína alvo, promovendo um dobramento adequado desta. Isto pode ser alcançado usando promotores fracos ou diminuindo a concentração de indutores (27).

Uma das estratégias também empregadas na prevenção da formação de corpos de inclusão é a coexpressão de chaperonas moleculares (35). As chaperonas são proteínas que se ligam a proteínas não dobradas ou a cadeias polipeptídicas parcialmente dobradas para evitar a associação imprópria de segmentos hidrofóbicos expostos, os quais poderiam levar a um dobramento não-nativo, além

de agregação de polipeptídeos e à precipitação (36). Embora seja uma idéia atrativa, não garante o correto dobramento de proteínas recombinantes, visto que estas moléculas podem não estar em uma quantidade suficiente para atingir esse objetivo nas células bacterianas. Estudos mostram que diferentes chaperonas podem ser limitadas a diferentes polipeptídeos (35).

No entanto, a expressão de proteínas na forma de corpos de inclusão pode ser vantajosa. Inicialmente, os corpos de inclusão podem se acumular no citoplasma em altos níveis, muitas vezes ultrapassando 25% do total de proteínas. Além disso, eles contêm uma grande quantidade de proteína praticamente pura. Como eles não têm atividade biológica, facilitam assim a produção de proteínas tóxicas em *E. coli*. Outro grande benefício é que esses agregados insolúveis são, na maior parte, protegidos de degradação proteolítica, permitindo alta produção protéica. E finalmente, a produção de proteína pode ter um aumento significativo quando produzida por alta densidade de cultura celular (37). O desafio é tirar vantagem desses altos níveis de expressão, sendo capaz de converter proteínas na forma de corpos de inclusão em produtos solúveis e bioativos (32).

Embora existam descritos na literatura diversos protocolos, a estratégia geral usada envolve basicamente três passos: isolamento e lavagem dos corpos de inclusão, solubilização das proteínas agregadas e *refolding*.

Desta maneira, as células que contêm corpos de inclusão são rompidas e estes podem ser separados por centrifugação. Os passos de lavagem são então utilizados para remover proteínas de membrana e outros contaminantes. Geralmente são usados: agentes quelantes como EDTA, para impedir a oxidação de cisteínas; baixas concentrações de desnaturantes e detergentes fracos como Triton X-100, deoxicolato e octilglucosídeo e inibidores de enzimas proteolíticas como PMSF (Fenilmetilsulfonil fluoreto). Os procedimentos de solubilização são normalmente

baseados em agentes desnaturantes como uréia e cloreto de guanidina, detergentes e/ou exposição destes corpos de inclusão a pH extremos. O último passo é a renaturação da proteína pela remoção do excesso de desnaturante, pelo uso das técnicas de diálise e cromatografia de gel filtração (30,32).

A recuperação de proteínas ativas a partir de corpos de inclusão é um processo complexo, embora existam diferentes protocolos que possam guiar esses procedimentos de acordo com o produto final desejado. Assim, a produção econômica do biofármaco Filgrastima em escala industrial é um grande desafio em relação a sua biossíntese e otimização.

## 1.5 IMPORTÂNCIA SOCIAL E ECONÔMICA

A garantia de acesso a medicamentos é parte integrante e essencial de uma adequada política assistencial em saúde. Em 1993, o Estado criou o Programa de Medicamentos Excepcionais (38). Os medicamentos incluídos no programa são aqueles de alto valor unitário e/ou usados para tratamento crônico, se tornando excessivamente caros para serem adquiridos pela população. Utilizados no nível ambulatorial, a maioria deles é de uso prolongado e parte deles integra tratamentos que duram a vida toda de um paciente. Essa política tem enorme alcance em todas as classes sociais, uma vez que, se não fossem distribuídos gratuitamente, tais medicamentos seriam acessíveis a poucas pessoas em função do alto custo.

Em termos operacionais, os recursos para aquisição destes medicamentos são transferidos pelo Ministério da Saúde para as Unidades Federativas. Desde o início do programa, muitas inclusões e exclusões foram realizadas, sendo que em 2006 estavam na lista 226 medicamentos ditos excepcionais. Do total de recursos aplicados pelo governo federal na assistência farmacêutica em 2006 (mais de R\$4,2



bilhões), um terço (R\$1,4 bilhão) foi investido no financiamento de medicamentos excepcionais (38,39,40).

Dentre estes medicamentos, se encontra o biofármaco Filgrastima. Em novembro de 2002, o Ministério da Saúde redigiu um protocolo clínico considerando os seguintes critérios de inclusão de pacientes para o tratamento com G-CSF:

- mobilização de células progenitoras para transplante de medula óssea;
- neutropenia associada ao transplante de medula óssea (contagem de neutrófilos igual ou abaixo de  $500/\text{mm}^3$ , 21 dias após o transplante);
- neutropenia induzida por quimioterapia (período mínimo de neutropenia de 10 -14 dias com contagem de neutrófilos abaixo de  $500/\text{mm}^3$ );
- neutropenia crônica grave (contagem de neutrófilos igual ou abaixo de  $500/\text{mm}^3$ );
- Aids com neutropenia (com contagem de neutrófilos abaixo de  $1000/\text{mm}^3$ );
- anemia aplástica grave em terapia com tripla imunossupressão (contagem de neutrófilos igual ou abaixo de  $200/\text{mm}^3$ );
- mielodisplasia com neutropenia grave e infecção de repetição (adultos com anemia refratária com ou sem sideroblastos em anel com contagem de neutrófilos abaixo de  $500/\text{mm}^3$  e infecções de repetição necessitando de hospitalizações (41).

Os avanços científicos que se iniciaram na primeira metade do século passado e que culminaram com o conhecimento da estrutura do DNA por Watson e Crick (42) propiciaram as bases moleculares para permitir o emprego industrial de microorganismos ou células modificadas geneticamente, objetivando a produção de proteínas de interesse em diversas áreas e, em especial, na saúde humana. A engenharia genética permite produzir proteínas terapêuticas, assim como produtos baseados em ácidos nucleicos, as quais são denominadas biofármacos (43).

Uma segunda geração de produtos, denominados “biossimilares” começou a ser desenvolvida amparada pela Lei de Patentes que regulamenta o tempo de exclusividade dos dados de processo de produção dos produtos biológicos inovadores. Em 2005, *European Medicines Agency Home* (EMA) emitiu uma série de esboços no marco da lei estabelecida pela União Europeia que estabelecem os padrões científicos para assistir aos laboratórios no processo de aprovação de biossimilares (EMA/CHMP/437/04) (44). O Hormônio de crescimento (Omnitrope) foi o primeiro biossimilar a ser liberado para uso pelas agências reguladoras da Europa, em abril de 2006 (45).

Esta segunda geração de produtos biológicos não pode ser considerada como “genéricos”, pois os testes exigidos para desenvolver estes produtos são muito mais rigorosos que os realizados para um genérico tradicional produzido por síntese química. Por esta razão, o termo usado para se referir ao biofármaco não inovador na Europa é “biossimilar” e nos EUA “*Follow-on proteins*” (46). Diferentemente da Europa, nos Estados Unidos as autoridades estudam o assunto de forma extensiva junto à comunidade científica, mas, até agora (Janeiro de 2008) ainda não foi aprovada nenhuma regulamentação para produção de biossimilares (47,48).

A Filgrastima, produzida pelo laboratório Amgen Inc, perdeu sua proteção por patente em 2006. Esse biofármaco representa um mercado global de milhões de dólares como apresentado na Tabela 1.

**Tabela1-Vendas Globais de Biofármacos**

<b>Produto</b>	<b>Substância ativa</b>	<b>Laboratório</b>	<b>Expiração das Patentes</b>	<b>Vendas globais (2005 em \$milhões)</b>
Cerezyme	imiglucerase	Genzyme	2001	933
Humulin	insulina humana	Eli Lilly	2001	1005
Novolin	insulin humana	Novo Nordisk	2001	1618
Intron-A	interferon alpha2b	Schering-Plough	2002	287
Avonex	interferon-beta	Biogen Idec	2003	1543
Humatrope	somatropina	Eli Lilly	2003	414
Nutropin	somatropina	Genentech	2003	370
Procrit	eritropoietina	J&J	2004	3324
Epogen	eritropoietina	Amgen	2004	2455
<b>Neupogen</b>	<b>filgrastima</b>	<b>Amgen</b>	<b>2006</b>	<b>1216</b>
			Total	13,165

Fonte: Adaptado de Pisani and Bonduelle. Opportunities and barriers in the biosimilar market: evolution or revolution for generics companies? (49)

O uso da engenharia genética é crescente em todo mundo e tem se tornado uma questão estratégica para o desenvolvimento científico dos principais países (50).

Logo, o objetivo deste trabalho foi desenvolver uma metodologia para a posterior produção de uma Filgrastima nacional. Este trabalho pretende contribuir para a sensível redução de custos com importação e o fornecimento do fármaco a um número maior de pacientes que dele necessita para tratamentos de saúde.

## 2. OBJETIVOS

### 2.1 OBJETIVO GERAL

Este estudo faz parte de um projeto maior que visa o escalonamento da produção de biofármacos adquiridos pelo Sistema Único de Saúde (SUS) federal ou estadual usados no tratamento de diversos tipos de doenças. O objetivo deste estudo é a determinação de um eficiente protocolo de superexpressão e purificação da proteína recombinante humana Fator Estimulador de Colônias de Granulócitos (G-CSF), que possa ser empregado para produção deste biofármaco em escala industrial.

### 2.2 OBJETIVOS ESPECÍFICOS

De modo mais concreto, esta proposta abrange os seguintes objetivos específicos:

- 1) Construção do gene que codifica para proteína G-CSF;
- 2) Subclonagem dos produtos da PCR no plasmídeo pCR-Blunt® e posterior subclonagem em vetor de expressão pET23a(+);
- 3) Determinar protocolo de superexpressão da proteína recombinante em *E. coli*;
- 4) Determinar o protocolo de purificação da proteína recombinante pelo emprego de Cromatografia líquida de rápida performance (FPLC);
- 5) Seqüenciamento dos aminoácidos da porção N-terminal da proteína purificada;
- 6) Análises físico-químicas, de identidade e de atividade biológica da proteína.

### **3. ARTIGO CIENTÍFICO**

Title of the article:

Human granulocyte colony stimulating factor (hG-CSF): cloning, overexpression, purification and characterization

Periodic chosen for submission:

Microbiol Cell Factories

# **Human granulocyte colony stimulating factor (hG-CSF): cloning, overexpression, purification and characterization.**

Ana Leticia S Vanz<sup>1,2,3</sup>, Gaby Renard<sup>3</sup>, Mario S Palma<sup>4</sup>, Jocelei M Chies<sup>3</sup>, Sérgio L Dalmora<sup>5</sup>,  
Luiz A Basso<sup>1,2\*</sup>, Diógenes S Santos<sup>1,2\*</sup>

<sup>1</sup>Programa de Pós-Graduação em Biologia Celular e Molecular - PUCRS. Av. Ipiranga, 6690 –  
Partenon – Porto Alegre, ZIP CODE 90610000 / Phone: + 55 51 33203318 – Brazil.

<sup>2</sup>Centro de Pesquisas em Biologia Molecular e Funcional, Instituto de Pesquisas Biomédicas - PUCRS.  
Av. Ipiranga, 6681 – Tecnopuc - Prédio 92A - Partenon - Porto Alegre, ZIP CODE 90619900 / Phone:  
+55 51 33203629 – Brazil.

<sup>3</sup>Quatro G Pesquisa e Desenvolvimento LTDA. Av. Ipiranga, 6681 – Tecnopuc - Prédio 92A -  
Partenon - Porto Alegre, ZIP CODE 90619900 / Phone: +55 51 33526560 – Brazil.

<sup>4</sup>Laboratório de Biologia Estrutural e Zooquímica, Centro de Estudos de Insetos Sociais,  
Departamento de Biologia, Instituto de Biociências, Universidade Estadual Paulista - Rio Claro, ZIP  
CODE 13506-900 / Phone: +55 19 35264163 – Brazil.

<sup>5</sup>Departamento de Farmácia Industrial, Centro de Ciências da Saúde, Universidade Federal de Santa  
Maria - Santa Maria, ZIP CODE 97105-900 / Phone: +55 55 33208952 – Brazil.

\*Corresponding authors

E-mails addresses:

ALSV: [anavanz@hotmail.com](mailto:anavanz@hotmail.com)

GR: [gaby.renard@pucrs.br](mailto:gaby.renard@pucrs.br)

MSP: [mspalma@rc.unesp.br](mailto:mspalma@rc.unesp.br)

JMC: [jocleleichies@terra.com.br](mailto:jocleleichies@terra.com.br)

SLD: [sdalmora@ccs.ufsm.br](mailto:sdalmora@ccs.ufsm.br)

LAB: [luiz.basso@puers.br](mailto:luiz.basso@puers.br)

DSS: [diogenes@puers.br](mailto:diogenes@puers.br)

# Abstract

## Background

Biopharmaceutical drugs are mainly recombinant proteins produced by biotechnological tools. The patents of many biopharmaceuticals have expired, and biosimilars are thus currently being developed. Human granulocyte colony stimulating factor (hG-CSF) is a hematopoietic cytokine that acts on cells of the neutrophil lineage causing proliferation and differentiation of committed precursor cells and activation of mature neutrophils.

Recombinant hG-CSF has been produced in genetically engineered *Escherichia coli* (Filgrastim) and successfully used to treat cancer patients suffering from chemotherapy-induced neutropenia. Filgrastim is a 175 amino acid protein, containing an extra N-terminal methionine, which is needed for expression in *E. coli*. Here we describe a simple and low-cost process that is amenable to scaling-up for the production and purification of homogeneous and active recombinant hG-CSF expressed in *E. coli* cells.

## Results

Here we describe cloning of the human granulocyte colony-stimulating factor coding DNA sequence, protein expression in *E. coli* BL21(DE3) host cells in the absence of isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) induction, efficient isolation and solubilization of inclusion bodies by a multi-step washing procedure, and a purification protocol using a single cationic exchange column. Characterization of homogeneous rhG-CSF by size exclusion and reverse phase chromatography showed similar yields to the standard. The immunoassay and N-terminal sequencing confirmed the identity of rhG-CSF. The biological activity assay, *in vivo*, showed an equivalent biological effect (109.4 %) to the standard reference rhG-CSF. The homogeneous rhG-CSF protein yield was 3.2 mg of bioactive protein per liter of cell culture.



## **Conclusion**

The recombinant protein expression in the absence of IPTG induction is advantageous since cost is reduced, and the protein purification protocol using a single chromatographic step should reduce cost even further for large scale production. The physicochemical, immunological and biological analyses showed that this protocol can be useful to develop therapeutic bioproducts. In summary, the combination of different experimental strategies presented here allowed an efficient and cost-effective protocol for rhG-CSF production. These data may be of interest to biopharmaceutical companies interested in developing biosimilars and healthcare community.

## Background

Biopharmaceuticals are medicinal products comprising biotechnology-derived recombinant proteins as active substances, according to the European Agency for the Evaluation of Medicinal Products [1]. Biopharmaceuticals have revolutionized the treatment of many diseases. Some of the biopharmaceutical patents have already expired, thereby allowing production of follow-on or biosimilar products. Public and private sectors of a number of countries have been encouraged to share staff, funding and facilities to increase technology transfers between universities and the industry [2]. It is due to an emerging multibillion-dollar market for companies willing to produce biosimilar versions of these products and to potential savings for healthcare payers and consumers that represents a driver of demand for biosimilars. Filgrastim (recombinant human granulocyte colony stimulating factor, rhG-CSF), produced by Amgen, had its patent expired in 2006. This biopharmaceutical generated global sales of \$5.6 billion (June 2005 to June 2006) and its market in Europe and USA has the potential to generate sales of approximately \$605 million in 2010 [3].

The granulocyte colony stimulating factor (G-CSF) is central to neutrophil-based immune defenses due to its regulatory role in the growth, differentiation, survival, and activation of neutrophils and their precursors [4]. Cancer chemotherapy can suppress production of these white blood cells, leaving patients vulnerable to potentially life-threatening infections and sepsis. G-CSF has thus been widely used with success in cancer patients whose treatment requires high-dose chemotherapy [5]. Furthermore, G-CSF can be used to reinforce the immune system in patients with HIV, pneumonia, diabetic foot infections, leukemia and febrile neutropenia [6–10]. Based on this ample clinical application, the recombinant human G-CSF has been produced in genetically engineered *Escherichia coli* and was approved for use in chemotherapy-induced neutropenia by the U.S Food and Drug Administration in 1991 [11]. It should be pointed out that two types of G-CSF are clinically available: a glycosylated form (lenograstim) which is produced by using the expression in

mammalian cells, and nonglycosylated form (filgrastim) which is produced by using the expression in *E. coli*.

Molecular cloning and expression of cDNA for hG-CSF have been described [12, 13]. The mature human G-CSF is a 18.8 kDa protein of 174 amino-acid polypeptide chain with two intra-molecular disulphide bonds between residues Cys<sup>36</sup>-Cys<sup>42</sup> and Cys<sup>64</sup>-Cys<sup>74</sup> and one free cysteine at residue 17 [14]. Native hG-CSF has a single *O*-glycosylation site at Thr<sup>133</sup>, which protects the protein from aggregation but is not crucial for biological activity [15]. The recombinant hG-CSF produced by *E.coli*, has identical biological activity to that of endogenous protein, but differs in that it contains an N-terminal methionine residue and is not glycosylated.

Recent publications describe various protocols of cloning, expression and purification of the rhG-CSF. These protocols involve use of several chromatography columns, high amount of detergents for the purification of G-CSF [16, 17, 18] and some of them were not applicable to recombinant G-CSF expressed in *E. coli* as inclusion bodies [19]. In industry, production of biopharmaceuticals employing a simple and cost-effective process involving fewer steps and yielding high levels of active protein is an essential prerequisite.

Here we describe the cloning of recombinant human granulocyte colony-stimulating factor gene, protein expression in *E. coli* cells, a straightforward purification protocol of the active recombinant protein from inclusion bodies and characterization of rhG-CSF by analytical methods. We believe that the combination of the different experimental strategies presented here provides an efficient protocol that may be useful in the industrial process of rhG-CSF protein production.

## Results and discussion

### Sequence synthesis, cloning and expression

Synthesis of G-CSF coding DNA sequence was carried out by a method developed in our laboratory, to which a patent has been filed [20]. We used the native sequence of hG-CSF, without making any modification such as replacement of GC rich regions with AT rich regions or replacement of rare codons, as have been reported by others [16, 21, 22]. A PCR amplification fragment consistent with that expected for hG-CSF (522 bp) was detected on agarose gel (Fig. 1), and its DNA sequence confirmed by automatic sequencing. Recombinant hG-CSF protein was expressed in insoluble form in *E. coli* BL21(DE3) host cells. The best conditions for expression of the rhG-CSF protein in the BL21(DE3) strain were reached at 24 hours in the absence of IPTG induction. In the pET system, target genes are positioned downstream of bacteriophage T7 late promoter. Typically, production hosts contain a prophage ( $\lambda$ DE3) encoding the highly processive T7 RNA polymerase under control of the IPTG-inducible *lacUV5* promoter that would ensure tight control of recombinant gene basal expression [23]. In agreement with the results presented here, high levels of protein expression in the absence of inducer have been shown to occur in the pET system. It has been proposed that protein expression in the absence of induction is a property of *lac*-controlled system when cells approach stationary phase in complex medium and that cyclic AMP, acetate, and low pH are required to achieve high-level expression in the absence of IPTG induction, which may be part of a general cellular response to nutrition limitation [24]. We have previously described this particular feature of the pET system for a number of recombinant proteins [25, 26, 27]. The growth rates were the same for *Escherichia coli* BL21(DE3) host cells harboring either pET 23a(+) (control) or the recombinant pET 23a(+):rhG-CSF plasmid. The absence of IPTG induction to obtain protein expression is thus advantageous since cost is reduced. The presence of the rhG-CSF protein was not observed in

the soluble fraction of BL21(DE3) host cells (data not shown). SDS-PAGE showed expression of an insoluble protein with a molecular mass value consistent with that expected for rhG-CSF (18.8 kDa) (Fig. 2). Interestingly, IPTG induction appears to abolish expression of rhG-CSF (Fig. 2, lanes: 4, 8, 10, 14). As pointed out above, T7 RNA polymerase in the pET system is under control of the IPTG-inducible *lacUV5* promoter, a system that has been used to produce substantial amounts of target protein from a wide variety of genes. However, a few proteins are made in disappointingly small amounts, for reasons that are obvious in some cases and obscure in others. Some likely reasons for obtaining low levels of expression are as follows (pET System Manual, Invitrogen): The target protein itself may interfere with gene expression or with the integrity of the cell, sometimes pulse labeling shows a gradual or rapid decrease in the rate of protein synthesis as target protein accumulates, or sometimes all protein synthesis stops before any target protein can be detected, and occasionally, considerable lysis of a culture is observed. More recently, it has been shown that the predominant contributory factor for decreased production of target proteins from pET plasmids is chromosomal mutation in the prophage of BL21( $\lambda$ DE3) host cells, resulting in decreased levels of functional T7 RNA Polymerase and ensuing reduction in target protein expression [28]. However, the reason for the lack of rhG-CSF protein expression described here is not known yet.

### **Purification and characterization of the recombinant G-CSF**

The appropriate genetics for soluble protein production has not yet been established and thus prevention of inclusion body formation is essentially a trial-and-error process. Aggregation may be minimized by controlling process parameters such as temperature, reduction of recombinant protein expression rate, adjusting the codon usage, or by co-expression of plasmid-encoded chaperones [29]. However, these strategies do not necessarily result in the same degree of success for different polypeptide chains. Notwithstanding,

inclusion bodies can be easily purified and these protein aggregates have also been observed as a source of relatively pure polypeptide [29]. The recovery of biologically active protein from inclusion bodies has several advantages: large amounts of highly enriched protein can be expressed as inclusion bodies; these aggregates are resistant to proteolysis by *E. coli* proteases, allowing high-yield protein production; facilitate production of protein that can be toxic or lethal to the host cell, because inclusion bodies have no biological activity; and, finally, it can be isolated and purified, decreasing the downstream process [30, 31]. We have thus employed a strategy to recover active protein from inclusion bodies as described by others [32]. This strategy involves three steps: inclusion body isolation and washing; solubilization of the aggregated protein, and refolding of the solubilized protein. Non-ionic detergents such as Triton X were used to solubilize the bacterial cell wall components that contaminate the inclusion body preparation. EDTA was added to chelate divalent metal ions, which maintain the structural integrity of the cell membrane. A second wash procedure incorporated sodium deoxycholate to remove any residual cell debris particles, especially lipopolysaccharides units that contribute to the unacceptable levels of endotoxins in protein preparations from *E. coli*. The third step using sodium chloride helps displace nucleic acids or any other contaminants that are non-specifically bound to the G-CSF protein in the inclusion bodies by ionic interactions.

rhG-CSF was further purified from inclusion bodies by a cationic exchange column, yielding approximately 3.2 mg of recombinant protein per liter of cell culture. Analysis by SDS-PAGE from the final step of the purification showed a single band of 18.8 kDa, similar to reference standard (Fig. 3). The same general pattern was revealed when the sample shown in Figure 3 was transferred from SDS-PAGE to a nitrocellulose membrane and immunodetected with G-CSF-specific antiserum, confirming a typical profile (Fig. 4) as described by others [33, 34]. The first 19 amino-terminal amino acid residues of purified recombinant protein were determined to be MTPLGPASSLPQSFLKCL by Edman

degradation. These results unambiguously demonstrate purification of rhG-CSF protein and that the N-terminal methionine was not removed by post-translational processing.

The solubilization of inclusion bodies pellet method described here employed a combination of small quantity of denaturing agent (2 M urea) and high alkaline pH, different from the majority of the published protocols that used high concentration of solubilizing agent (6 - 8 M). Key to the development of an efficient and economic denaturant-based solubilization step is the determination of the minimum amount of denaturant needed to solubilize the protein and to allow for full bioactivity recovery in the refolding step. In addition, the purification protocol of hG-CSF utilizes only one cation exchange column, which is an improvement upon previous protocols that utilized multiple chromatographic steps [16, 17, 18, 35]. A method has been described for G-CSF extraction from inclusion bodies produced in *E. coli* that used high amount of detergent and chaotropic agent, and hence additional steps had to be employed to remove these agents, including G-25, CM-cellulose and G-75 chromatographic columns to obtain pure protein [16]. A G-CSF purification protocol has been described that employed immobilized metal affinity chromatography (IMAC) matrix, cation exchange and gel filtration chromatographic steps, the final yields of the protein are not clearly evident [18]. On a commercial scale, reducing the number of protein purification steps is practical and economical because each purification step not only adds to the cost of the final product but also causes successive yield losses of the recombinant protein.

### **Physicochemical properties**

Characterization of a biological product (which includes the determination of physicochemical properties, biological activity, immunochemical properties, purity and impurities) by appropriate techniques is necessary to confirm an efficient expression and purification protocols. Protein instability encompasses many complicated and interrelated

physical and chemical processes. Any of these can occur during the production, isolation and purification of proteins. Chemical degradations that include, at least, deamidation of Asparagine and Glutamine side chain and oxidation of Methionine, as well as dimeric forms and related substances of higher molecular mass, have been recognized to be an important cause of partial loss in activity of therapeutic proteins [36, 37]. According to Massiero and coworkers [38], rhG-CSF that undergoes some form of degradation retains only 15 % of biological activity. The rhG-CSF here described was analysed by HPLC to detect any of these alterations, including chemical degradation, dimers and related substances of higher molecular mass.

Figure 5 shows a typical reverse-phase chromatogram of rhG-CSF demonstrating the resolution of the symmetrical peak that corresponds to hG-CSF with the retention time of 31.6min similar to standard. The deamidates and sulfoxides were analyzed (Fig. 5) and estimated to be 2.09 % and 2.20 %, respectively. These values are lower than what is routinely regarded as safe level of 5 % for many biopharmaceuticals [39]. Acidic size-exclusion chromatography of rhG-CSF (Fig. 6) demonstrates the presence of a small amount of high molecular weight aggregates (0.26 %) and absence of dimers. This level is lower than what is routinely regarded as safe (0.5 %) for many biopharmaceuticals [39]. The main peak represents monomers eluted at a retention time similar to that of the standard. Glycosylated G-CSF is freely soluble in phosphate-buffered saline. The absence of carbohydrate however results in a more hydrophobic protein that does not tolerate high salt concentration or neutral pH. It is thus necessary to run size-exclusion chromatography under acidic conditions. At higher pH, recovery of monomers is poor and multimers are irreversibly bound to the resin [40]. The analytical results of recombinant hG-CSF protein reported here are similar to the reference standard. The characterization of rhG-CSF plays a vital role in its development as useful therapeutic agent. Physicochemical techniques can produce information about the



structure and composition of this therapeutic protein, but cannot yet predict their biological activity, for which biological assays are required.

### **Biological potency evaluation**

In order to determine the biological potency of rhG-CSF in the present study, we employed an *in vivo* model of neutropenia, by treating mice with a single dose of ifosfamine. The percentage of neutrophils in serum was monitored 6 h after the last rhG-CSF injection or vehicle injection. The data (Fig. 7) clearly demonstrate that treatment with rh-GCSF in the sample group produced a significant and dose-dependent increase of neutrophil counts ( $P < 0.001$ ), when compared to the control group ( $P < 0.001$ ). The neutrophil counts in the standard group were not significantly different from that observed in the sample group when comparing the same doses of rhG-CSF ( $P > 0.05$ , for the three doses).

Statistical analysis by the Finney test [41] revealed a biological activity of 109.4 % for the rhG-CSF in the sample group. It appears to be worth pointing out that the potency of Filgrastim (rhG-CSF) itself is not described in any Pharmacopoeia, although values ranging from 90 to 110 % have been suggested for biological medicines, according to the guidelines from the European Pharmacopoeia [42]. The data presented clearly indicate that the rhG-CSF produced by the combination of experimental strategies described here can produce biological effects that are comparable or even superior to those obtained with the standard reference rhG-CSF. It is thus tempting to infer that generation of rhG-CSF on a large scale using the protocols described here may represent an interesting option for producing biosimilars.

## **Conclusions**

In this report, we describe an efficient protocol for cloning, expression and purification of rhG-CSF. The recombinant protein expression in the absence of IPTG may be advantageous since cost is reduced. In addition, the protein purification protocol by liquid

chromatography using a single chromatographic step may be a valuable and cost-effective approach to large scale production. The physicochemical, immunological and biological analyses showed that this protocol can be useful to develop therapeutic bioproducts. The data presented here may be of interest to biopharmaceutical companies interested in developing biosimilars, which offer a great opportunity to scientific, economical and industrial growth.

## Methods

### Amplification, cloning and overexpression of hG-CSF coding DNA sequence

Oligonucleotides were manually designed and synthesized, in even numbers, corresponding to the double stranded DNA, in a sequential manner. The G-CSF nucleotide sequence (Accession number NM\_000759) of 522 bp, was divided in 12 short sequences of approximately 50 bp each. Oligonucleotides contained overlapping regions of about 10 bases at their 5'- and 3'-ends. The pairs of oligonucleotides were assembled and PCR amplified to yield the hG-CSF coding DNA sequence in a step-wise fashion. PCR-amplified DNA fragments with the expected sizes were detected by 2 % agarose gel electrophoresis stained with 0.5 µg mL<sup>-1</sup> of Ethidium Bromide, purified and another round of PCR amplification carried out. A thorough description of this method has been presented elsewhere [20]. The primers employed in the present study are given in Table 1. The primers P1 and P12 contained the restriction sites for, respectively, *Nde* I and *Bam* HI. The 522 bp final PCR product was agarose-gel purified and cloned into the pCR<sup>®</sup>-Blunt vector (Invitrogen) and subcloned into the pET 23a(+) expression vector (Novagen) using the *Nde* I and *Bam* HI restriction enzymes (Boehringer Mannheim). The nucleotide sequence of rhG-CSF was determined in order to confirm the identity, integrity, and absence of PCR-introduced mutations in the cloned gene. *Escherichia coli* BL21(DE3) (Novagen) electrocompetent cells

were transformed with the recombinant pET 23a(+):rhG-CSF plasmid and grown on LB agar plates containing  $50 \mu\text{g mL}^{-1}$  of ampicillin. The protocol for recombinant protein expression described here represents a choice from a number of tests carried out at different experimental conditions. Soluble and insoluble fractions were analyzed by 12 % SDS-PAGE [43]. To express and purify the recombinant protein, two liters of 4YT medium (32 g Bacto tryptone, 20 g yeast extract and 5 g NaCl per litre, pH 7.2) containing ampicillin were inoculated with a single colony and grown in shaker flasks for 24 h at 180 rpm and 37 °C without IPTG induction. Cells were harvested by centrifugation at  $15,900 \times g$  for 30 min at 4°C and stored at -20°C.

#### **Isolation of inclusion bodies**

The frozen cell paste was resuspended in lysis buffer (50 mM Tris buffer, at pH 8.0, 1 mM EDTA and 1 mM PMSF) at a pellet:buffer ratio of 1:10 (w/v). The inclusion bodies containing rhG-CSF were recovered from the cells using a French press (Constant Systems LTD) under 137.9 MPa. Prior to centrifugation, a 1:2 dilution of the lysate was carried out to reduce viscosity and to obtain a better yield of inclusion bodies. The resulting lysate solution was centrifuged at  $48,000 \times g$  for 30 minutes at 4°C to pellet the inclusion bodies. The supernatant was discarded, and the fraction containing the inclusion bodies was subjected to a three-step wash procedure to eliminate endotoxins, proteins and DNA of the host cells. In all steps, the pellet was suspended in the specific buffer at 1:40 (w/v) ratio at room temperature, stirred for 30 min and re-pellet by centrifugation. The first buffer was comprised of: 50 mM Tris, pH 8.0, 5 mM EDTA and 2 % Triton X-100. The composition of the second buffer was 50 mM Tris, pH 8.0, 5 mM EDTA, 1% sodium deoxycholate. In the third wash step, the wash buffer contained 50 mM Tris buffer, pH 8.0, 5 mM EDTA, and 1 M NaCl.

## **Solubilization and purification**

The protocol of solubilization and denaturation of rhG-CSF in inclusion bodies employed a combination of solubilization agent (2 M urea) and high alkaline pH. The solubilized pellet was stirred for 30 min at room temperature, after diluting the protein to a concentration of 2 mg mL<sup>-1</sup>. Acetic acid was added to bring the pH to a value of 8.0. Refolding of 190 mL of the solubilized protein was carried out by dialysis (MWCO 6,000-8,000 Da) in two steps for 12-16 hours. The first one was carried out against 4 L 50 mM Tris HCl pH 8.0 and the second step against 4 L 25 mM sodium acetate buffer pH 4.5 (buffer A). The purification protocol was performed using the ÄKTA System (GE Healthcare). The soluble protein was purified on a cation exchange column, HiPrep Resource S column (GE Healthcare), preequilibrated with buffer A, and the adsorbed material was washed with buffer A (4 column volumes), eluted with a linear gradient of 0.2 to 0.26 M Tris HCl pH 8.0 (10 column volumes), 0.26 – 1.0 M Tris HCl pH 8.0 (5 column volumes), followed by an isocratic elution with 1.0 M Tris HCl pH 8.0 (5 column volumes) at a flow rate of 1 mL min<sup>-1</sup>. The recombinant hG-CSF eluted between 0.207 – 0.215 M Tris HCl pH 8.0. Protein expression and all protein purification steps were analyzed by 12 % SDS-PAGE, and protein concentrations were determined by the method of Bradford *et al.* [44] (Bio-Rad Laboratories).

## **N-terminal amino acid sequencing**

The N-terminal amino acid residues of homogeneous recombinant hG-CSF were identified by automated Edman degradation sequencing as we have described elsewhere [45].

## **Western Blotting**

Recombinant rhG-CSF on 12 % SDS-PAGE was transferred to a nitrocellulose membrane by electrophoresis at a constant voltage of 70V in 25 mM Tris pH 8.8, 192 mM glycine, containing methanol 20 %, for 1h using a Trans-blot apparatus (Bio-Rad, USA). The

membrane was blocked for 1 h at room temperature by 50 mM sodium phosphate, 150 mM sodium chloride pH 7.0, containing 5 % (w/v) dried skim milk powder. The membrane was incubated for 1 h at room temperature with 1.5 µg G-CSF polyclonal antibody (Santa Cruz Biotechnology), in milk-containing phosphate buffered saline as above. The blot was then incubated with a secondary antibody conjugated with horseradish peroxidase (HRPO) (1:1000), and diaminobenzidine (DAB) was used for the development of signal. The size (molecular mass) of the protein was compared to pre-loaded standards (Santa Cruz™ molecular weight standards).

### **Reverse phase liquid chromatography (RP-HPLC) analysis**

The sulphoxides and deamidates were analyzed as we have described elsewhere [46]. A HPLC system (Shimadzu, Kyoto, Japan) equipped with an SCL-10A<sub>vp</sub> system controller, LC-10AD<sub>vp</sub> pump, DGU-14A degasser, SIL-10AD<sub>vp</sub> autosampler, and an SPD-M10A<sub>vp</sub> photodiode array (PDA) detector was used. The detector was set at 280 nm and peak areas were integrated automatically by computational analysis, using a Shimadzu Class VP® software program. The experiments were carried out on reversed phase Phenomenex (Torrance, USA) Jupiter C4 column (250 x 4.6 mm I.D, with a pore size of 30 nm). The liquid chromatography system was operated at controlled room temperature (25 °C). The elution was performed by a fast gradient at a constant flow rate of 0.5 mL min<sup>-1</sup>. Mobile phase A consisted of water: acetonitrile (90:10 v/v) containing 0.1% trifluoroacetic acid (TFA) and mobile phase B consisted of water: acetonitrile (20:80, v/v) containing 0.1% TFA. The injection volume was 50 µL for both standard and sample. The Filgrastim substance reference used was from NIBISC (National Institute for Biological Standards and Control).

### **Size exclusion chromatography (SEC-HPLC) analysis**

The detection of dimers and monomers was performed as we have described elsewhere [33]. The HPLC system (Shimadzu, Kyoto, Japan) equipped with an SCL-10A<sub>vp</sub> system controller, LC-10AD<sub>vp</sub> pump, DGU-14A degasser, SIL-10AD<sub>vp</sub> autosampler, and an SPD-M10A<sub>vp</sub> photodiode array (PDA) detector was used. The detector was set at 214 nm and peak areas were integrated automatically by computational analysis, using a Shimadzu Class VP<sup>®</sup> software program. The experiments were carried out on size exclusion Phenomenex (Torrance, USA) BioSep-SEC 2000 column (300 x 7.8 mm I.D., with a particle size of 5 µm and pore size of 14.5 nm). The HPLC system was operated at room temperature, using mobile phase of phosphoric acid (pH 2.5; 0.1 M) adjusted by the addition of sodium hydroxide (10 N). At the beginning of the run, 50 µL bovine albumin (1 mg mL<sup>-1</sup>) was injected to reduce non-specific adsorption. The injection volume was the same for both, standards and samples, and all determinations were carried out in triplicate. Flow rate was maintained at 1 mL min<sup>-1</sup>. The Filgrastim substance reference used was from NIBISC (National Institute for Biological Standards and Control).

### **Biological Assay**

Male 7-8 week-old BALB-c mice (19 - 24 g) obtained from the Department of Industrial Pharmacy, Health Science Centre, Universidade Federal de Santa Maria (Brazil), were used in this study. Animals were housed under conditions of optimum light, temperature and humidity (12 h light-dark cycle, 22 ± 2 °C, 65 % humidity), with food and water provided *ad libitum*. All procedures used in the present study were approved by “Principles of Laboratory Animal Care” from NIH publication No. 85-23 and were approved by the local Animal Ethics Committee with the registration number 19/2005. The number of animals used was the minimum necessary to demonstrate the consistent effects.

The animals were allocated to sample, standard, and control groups in a fully randomized order and identified by color code for assay (6 mice per group). Standard and test sample were diluted to the concentrations of 4, 12, and 36  $\mu\text{g mL}^{-1}$ , in phosphate buffered saline (PBS), containing 0.1% bovine serum albumin (BSA). To induce neutropenia, all mice received a single dose of ifosfamine (4.3 mg/0.5 mL per animal) by intraperitoneal route, on day 0, according to the methodology we have described elsewhere [46]. Animals in the treated groups received multiple injections of 0.5 mL rhG-CSF (standard or sample groups), from day 1 to day 4. The control group was treated with vehicle (0.1 % BSA in PBS) at the same schedule of administration. Six hours after the last rhG-CSF injection or vehicle, peripheral blood was collected from the orbital venus sinus. Smears were prepared on glass slides and stained by the May-Grünwald-Giemsa method, and the white cells were counted and expressed as the total number of neutrophils. The Filgrastim standard used was from NIBISC (National Institute for Biological Standards and Control).

### **Statistical analysis**

The results (% of neutrophils) are presented as the mean  $\pm$  S.D. of 6 animals. Statistical comparison of data was performed by one-way analysis of variance (ANOVA) followed by Bonferroni's post-test, by means of GraphPad Prism Program (Version 4.0). *P*-values less than 0.05 ( $P < 0.05$  or less) were considered significant.

For determining the possible differences in the biological activity between the standard and the sample groups, the test described by Finney [41] was adopted, through the parallel line method (3 x 3), using a PLA 1.2 program (Stegmann Systemberatung, Rodgau, Germany). Analysis of variance was performed for each assay, and the assumption of linearity and parallelism of the log dose-log response lines was tested ( $P < 0.05$ ). Statistical weights were computed as the reciprocal of the variance of the log potency. Estimates of log potency

were examined for heterogeneity using a  $\chi^2$  test ( $P = 0.05$ ) and were combined as weight geometric means of homogeneous estimates ( $P > 0.05$ ) [47].

## List of abbreviations

BSA: bovine serum albumin, DAB: diaminobenzidine, HRPO: horseradish peroxidase, IPTG: isopropyl- $\beta$ -D-thiogalactopyranoside, kDa: kilodalton, PBS: phosphate buffered saline containing, PCR: polymerase chain reaction, PMSF: phenyl methyl sulfonyl fluoride, rhG-CSF: recombinant human granulocyte colony stimulating factor, RP-HPLC: reverse phase liquid chromatography, SDS-PAGE: sodium dodecyl sulfate-polyacrylamide gel electrophoresis, SEC-HPLC: Size exclusion liquid chromatography.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

ALSV performed most of the experiments and drafted the manuscript. GR supervised cloning, expression and purification experiments and helped draft the manuscript. MSP carried N-terminal protein sequencing analysis. JMC participated in the analysis and interpretation of protein chemistry results. SLD supervised physicochemical and biological activity analysis and helped draft the manuscript. The study was conceived and coordinated by DS and LB, who have also revised the manuscript. All authors read and approved the final manuscript.



## Acknowledgements

Financial support for this work was provided by FARMASA (Laboratório Americano de Farmacoterapia S.A.) to Quatro G P & D. Financial support was also provided by Millennium Initiative Program MCT-CNPq, Ministry of Health-Department of Science and Technology and PRONEX/CNPq/FAPERGS (Brazil) to D.S.S. and L.A.B. DSS and LAB. DSS (CNPq, 304051/1975-06), LAB (CNPq, 5201182/99-5) and MSP (CNPq, 500079/90-0) are research career awardees of the National Council for Scientific and Technological Development of Brazil.

## References

1. EMEA Guideline EMEA/CHMP/BWP/49348/2005 **Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues** [[www.emea.eu.int](http://www.emea.eu.int)]
2. Riche M: **Opportunities for UK Companies in Biopharmaceuticals and Biotechnology in Brazil.** *UK Trade and Investment*, 2006  
[[www.uktradeinvest.gov.uk/ukti/biotechnology](http://www.uktradeinvest.gov.uk/ukti/biotechnology)]
3. Pisani J, Bonduelle Y: **Opportunities and barriers in the biosimilar market: evolution or revolution for generics companies?**  
[<http://www.ableindia.org/biosimilars.pdf>]
4. Barreda D, Hanington P, Belosevic M: **Regulation of myeloid development and function by colony stimulating factors.** *Dev Comp Immunol* 2004, **28**: 509-554.
5. Wittman B.; Horan J.; Lyman GH. **Prophylactic colony-stimulating factor in children receiving myelosuppressive chemotherapy: a meta-analysis of randomized controlled trials.** *Cancer Treat Rev* 2006, **32**:289-303.

6. Kuritzkes D, Parenti D, Ward D, Rachlis A, Wong R, Mallon K, Rich W, Jacobson M: **Filgrastim prevents severe neutropenia and reduces infective morbidity in patients with advanced HIV infection: results of a randomized, multicenter, controlled trial.** *Aids* 1998, **12**:65-74.
7. Babalola C, Nightingale C, Nicolau D: **Adjunctive efficacy of granulocyte colony-stimulating factor on treatment of *Pseudomonas aeruginosa* pneumonia in neutropenic and non-neutropenic hosts.** *J Antimicrob Chemother* 2004, **53**:1098-100.
8. Gough A, Clapperton M, Rolando N, Foster A, Howard J, Edmonds M: **Randomized placebo-controlled trial of granulocyte colony-stimulating factor in diabetic foot infection.** *The Lancet* 1997, **350**:855-859.
9. Harousseau JL, Witz B, Lioure B, Berger M, Desablens B, Delain M, Guilhot F, Le Prise PY, Abgrall JF, Deconick E, Guyotat D, Vilque JP, Casassus P, Tournilhac O, Audhuy B, Solary E: **Granulocyte colony-stimulating factor after intensive consolidation chemotherapy in acute myeloid leukemia: Results of a randomized trial of the group Ouest-Est leucémies aiguës myéloblastiques.** *J Clin Oncol* 2000, **18**:780-787.
10. Carbonero R, Mayordomo J, Tornamira M, Brea M, Rueda A, Guillem V, Arcediano A, Yubero A, Ribera F, Gomes C, Tres A, Peres-Gracia J, Lumbreras R, Hornedo J, Funes H, Paz-Ares L: **Granulocyte colony-stimulating factor in the treatment of high-risk febrile neutropenia: a multicenter randomized trial.** *J Natl Cancer Inst* 2001, **93**:31-38.
11. **U.S Food & Drug Administration. U.S Department of Health and Human Services.** Therapeutic Biological Products Approvals  
[<http://www.accessdata.fda.gov/scripts/cder/onctools/druglist.cfm>]

12. Nagata S, Tsuchiya M, Asano S, Kaziro Y, Yamazaki T, Yamamoto O, Hirata Y, Kubota N, Oheda M, Nomura H, Ono M: **Molecular cloning and expression of cDNA for human granulocyte colony –stimulating factor.** *Nature* 1986, **319**:415-418.
13. Souza L: **Recombinant human granulocyte colony-stimulating factor: effects on normal and leukemic myeloid cell.** *Science* 1986, **232**:61-65.
14. Lu S, Boone T, Souza L, Lai P: **Disulfide and secondary structure of recombinant human granulocyte colony-stimulating factor.** *Arch Biochem Biophys* 1989, **1**:81-92.
15. Oh-eda M, Hasegawa M, Hattori K, Kuboniwa H, Kojima T, Orita T, Tomonou K, Yamazaki T, Ochi N: **O-linked sugar chain of human granulocyte colony-stimulating factor protects it against polymerization and denaturation allowing it to retain its biological activity.** *J Biol Chem* 1990, **265**:11432-11435.
16. Souza L: **Production of pluripotent granulocyte colony-stimulating factor.** 1989. Patent US 4,810,643.
17. Boone T, Miller, Allan L, Andresen, Jeffrey W: **Method for purifying granulocyte colony stimulating factor.** 1998. Patent US 5,849,883.
18. Porekar VG, Menart V: **Process for the purification and/or isolation of biologically active granulocyte colony stimulating factor.** 2005. Patent US 2005/01595589
19. Sassenfeld H: **Purification of G-CSF.** 1991. Patent US5,055,555.
20. Santos DS, Chies JM, Basso LA, Renard G, Fonseca IO: **Method for the obtention of chimeric nucleotide sequences and chimeric nucleodite sequences.** 2006. Patent application WO2007/068068 A2.
21. Jevsevar S, Menart V: **Synthetic gene coding for human granulocyte-colony stimulating factor for the expression in *E coli*.** 2006. Patent US2006/0228781.

22. Jeong K, Lee S: **Secretory production of human granulocyte-colony stimulating factor in *Escherichia coli***. *Protein Expr Purif* 2001, **23**:311-318.
23. Dubendorff JW, Studier FW. **Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor**. *J Mol Biol* 1991, **219**:45-59.
24. Grossman TH, Kawasaki E, Punreddy SR, Osburne M: **Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability**. *Gene* 1998, **209**:95-103.
25. Rizzi C, Frazzon J, Ely F, Weber P, Fonseca I, Gallas M, Oliveira J, Mendes M, Souza B, Palma M, Santos D, Basso L: **DAHP synthase from *Mycobacterium tuberculosis* H37Rv: cloning, expression, and purification of functional enzyme**. *Protein Expr Purif* 2005, **40**:23-30.
26. Mendonça J, Ely F, Palma M, Frazzon J, Basso L, Santos D: **Functional characterization by genetic complementation of *aroB*-encoded dehydroquinate synthase from *Mycobacterium tuberculosis* H37Rv and its heterologous expression and purification**. *J Bacteriol* 2007, **189**:6246-6252.
27. Magalhaes M, Pereira C, Basso L, Santos D: **Cloning and expression of functional shikimate dehydrogenase (EC.1.1.25) from *Mycobacterium tuberculosis* H37Rv**. *Protein Expr Purif* 2002, **29**:59-64.
28. Venhanayagam JGG, Flower AM. Decreased gene expression from T7 promoters may be due to impaired production of active T7 RNA polymerase. *Microb Cell Fact* 2005, **4**:3.
29. Villaverde A, Carrió MM: **Protein aggregation in recombinant bacteria: biological role of inclusion bodies**. *Biotechnol Lett* 2003, **2**:1385-1395.
30. Choi J, Keum K, Lee S: **Production of recombinant proteins by high cell density culture of *Escherichia coli***. *Chemical Engineering Science* 2006, **61**:876-885.

31. Clark E: **Protein refolding for industrial process**. *Current Opinion in Biotechnology Engineering* 2001, **12**:202-207.
32. Komath UD, Lodha S, Chigurupati J: **Process for preparing G-CSF**. 2004. Patent WO2004/001056.
33. Codevilla C, Brum L, Oliveira P, Dolman C, Rafferty B, Dalmora S: **Validation of an SEC-HPLC method for the analysis of rhG-CSF in pharmaceutical formulation**. *Journal of liquid chromatography & related technologies* 2004, **27**:2681-2690.
34. Hammerling U, Kroon R, Sjodin L: **In vitro bioassay with enhanced sensitivity for human granulocyte colony-stimulating factor**. *J Pharm Biomed Anal* 1995, **13**:9-20.
35. Bae CS, Yang DS, Chang KR, Seong BL, Lee J: **Enhanced secretion of human granulocyte colony-stimulating factor directed by a novel hybrid fusion peptide from recombinant *Saccharomyces cerevisiae* at high cell concentration**. *Biotechnol. Bioeng.* 1998,**57**: 600-609.
36. Manning M, Patel K, Borchardt R: **Stability of protein pharmaceuticals**. *Pharm Res* 1989, **6**:903-918.
37. Lu H, Fausset P, Narhi L, Horan T, Shinagawa K, Shimamoto G, Boone T: **Chemical modification and site-directed mutagenesis of methionine residues in recombinant human granulocyte colony-stimulating factor: effect on stability and biological activity**. *Arch Biochem Biophys* 1999, **362**:1-11.
38. Massiero S: **Study of chromatography methods for the potency evaluation of filgrastim and correlation with biological assay**. *Master's degree dissertation*. Federal University of Santa Maria, Brazil; 2006.
39. **U.S Food & Drug Administration** 2004N-0355: Scientific considerations related to developing follow-on protein products.  
  
[<http://www.fda.gov/ohrms/dockets/DOCKETS/04n0355/04n-0355-c000005-01-vol1.pdf>]

40. Pearlman R.; Wang J. **Formulation, characterization, and stability of protein drugs.** In *Pharmaceutical biotechnology*. Volume 9. Edited by Plenum Press. New York: 1996:303-325.
41. Finney JD: **Statistical methods in biological assay.** Charles Griffin. London: **1978.**
42. European Pharmacopoeia. 4<sup>th</sup> ed. Strasbourg: EDQM, 2002.
43. Laemmli U: **Cleavage of structural proteins during the assembly of the head of bacteriophage T4.** *Nature*, **227**:680-685.
44. Bradford MM, Mcrorie RA, Williams WL. **A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding.** *Anal. Biochem.* 1976, **72**: 248-254.
45. Fonseca IO, Magalhaes MLB, Oliveira JS, Silva RG, Mendes MA, Palma MS, Santos DS, Basso LA. **Functional shikimate dehydrogenase from *Mycobacterium tuberculosis* H37Rv: purification and characterization.** *Protein Expr. Purif.* 2006, **46**: 429-437.
46. Dalmora SL, Massiero S, Oliveira P, Sangoi M, Brum L: **Validation of an RP-HPLC method and assessment of rhG-CSF in pharmaceutical formulation by liquid chromatography and biological assay.** *Journal of liquid chromatography & related technologies* 2006, **29**:1753-1767.
47. Gaines D, Tydeman, M: **Iterative weighted regression analysis of logit responses. A computer program for the analysis of bioassay and immunoassays.** *Comput. Programs. Biomed.* 1980, **15**:13-22.

## Figure legends

**Figure 1. Agarose gel electrophoresis of PCR product.** Agarose gel (2%) electrophoresis of PCR amplified hG-CSF coding DNA sequence. Lane 1: DNA markers (1 Kb Plus DNA Ladder™, Life Technologies, Gibco BRL); lane 2: PCR amplification of hG-CSF (522bp).

**Figure 2. Analysis of recombinant hG-CSF protein expression by SDS-PAGE.** SDS-PAGE (12 %) analysis showing the insoluble fractions of heterologous protein expression in *E. coli* BL21 (DE3) host cells transformed with either pET 23a(+) (control) or pET 23a(+):hG-CSF. The best expression conditions were 24 h at 37°C without IPTG addition. Lane 1: Molecular mass markers (BenchMark™ Protein Ladder, Invitrogen), lanes 2,6,12: control induced with 0.4 mM IPTG and further growth for 3, 8 and 24h respectively; lanes 3,7,13: control without induction and further growth for 3, 8 and 24h respectively; lanes 4,8,10,14: host cells containing the recombinant pET 23a(+):hG-CSF plasmid induced with 0.4 mM IPTG and further growth for 3, 8, 17 and 24 h respectively; and lanes 5,9,11,15: host cells containing the recombinant pET 23a(+):hG-CSF plasmid without IPTG induction and further growth for 3, 8,17 and 24 h respectively. The predicted molecular mass of rhG-CSF is 18.8 kDa.

**Figure 3. SDS-PAGE of homogeneous rhG-CSF.** SDS-PAGE (12 %) analysis of purified rhG-CSF showing a single protein band. Lane 1: Protein molecular mass markers (BenchMark™ Protein Ladder, Invitrogen), lane 2: homogeneous rhG-CSF (18.8kDa) after elution from a cation exchange column, lane 3: reference standard. Lanes 2 and 3 contain 1.5µg of total protein each.

**Figure 4. Western blot of homogeneous rhG-CSF.** A single band was observed for rhG-CSF transferred to a nitrocellulose membrane and detected using a polyclonal antiserum. Lane 1: crude extract of host cells transformed with pET 23a(+):hG-CSF; lane 2: crude extract of host cells transformed with pET 23a(+) (control); lane 3: insoluble fractions of host cells transformed with pET 23a(+):hG-CSF; lane 4: insoluble fractions of pET 23a(+); lane 5: homogeneous hG-CSF; lane 6: reference standard; and lane 7: protein molecular mass markers (Cruz Marker™).

**Figure 5. Reverse-phase HPLC analysis of hG-CSF.** Sulphoxides and deamidates in hG-CSF were analyzed by Reverse-phase HPLC. 200µg mL<sup>-1</sup> of either sample or standard was loaded for analysis. (A): homogeneous hG-CSF, peak 1- main peak. (B): reference standard, peak 1- main peak. Absorbance is in milliabsorbance units (mAU).

**Figure 6. Size exclusion HPLC analysis of hG-CSF.** Size-exclusion HPLC chromatography was employed to detect dimeric forms of hG-CSF. 200µg mL<sup>-1</sup> of either sample or standard was loaded for analysis. (A): homogeneous hG-CSF, peak 1- monomer and peak 2- excipient. (B): reference standard, peak 1- monomer and peak 2- excipient. Absorbance is in milliabsorbance units (mAU).

**Figure 7. Biological assay for hG-CSF.** Serum neutrophil counts after treatment with rhG-CSF. Mice in the standard and the sample groups were treated with different doses of rhG-CSF (4, 12, and 36 µg/0.5ml/mouse). The control group was treated with vehicle (0.1 % BSA in PBS). The asterisks represent significant differences in comparison to the control group (\*\*  $P < 0.001$ ).  $N = 6$  mice per group.



## Tables

**Table 1. Oligonucleotides employed to obtain the coding DNA sequence of hG-CSF.**

---

P1- 5' AAC <u>CATATG</u> ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTC AAGTG 3'
P2- 5' GCGCTGCGCCATCGCCCTGGATCTTCCTCACTTGCTCTAAGCACTTGAGCA 3'
P3- 5' GGCAGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAGCT 3'
P4- 5' CCAGAGAGTGTCCGAGCAGCACCAGCTCCTCGGGGTGGCACAGCTTGTAGG 3'
P5- 5' CACTCTCTGGGCATCCCCTGGGCTCCCCTGAGCAGCTGCCCCAGCCAGGCCCTG CAG 3'
P6- 5' GGTAGAGGAAAAGGCCGCTATGGAGTTGGCTCAAGCAGCCTGCCAGCTGCA GGGCC 3'
P7- 5' TTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT C 3'
P8- 5' TGGTGGCAAAGTCGGCGACGTCCAGCTGCAGTGTGTCCAAGGTGGGACCCAACT C 3'
P9'- 5' TTTGCCACCACCATCTGGCAGCAGATGGAAGAAGTGGGAATGGCCCCT GCC CT 3'
P10- 5' AAAGCAGAGGCGAAGGCCGGCATGGCACCCCTGGGTGGGCTGCAGGGCAGGGG 3'
P11- 5' CCTCTGCTTTCAGCGCCGGGCAGGAGGGTCTGGTTGCCTCCCATCTGC AGAGCTTCC 3'
P12- 5' A <u>AGGATCCT</u> CAGGGCTGGGCAAGGTGGCGTAGAACGCGGTACGACACCTCCAGGAAGCTCTG 3'

---

Primer sequences for construction of hG-CSF. The G-CSF nucleotide sequence (Accession number NM\_000759) of 522 bp, was divided into 12 DNA sequences of approximately 50 bp each, containing 10-bp overlapping regions at their 5'- and 3'-ends for adjacent oligonucleotides. The restriction sites, *Nde* I and *Bam* HI, are underlined in primers P1 and P12, respectively.

Figure 1

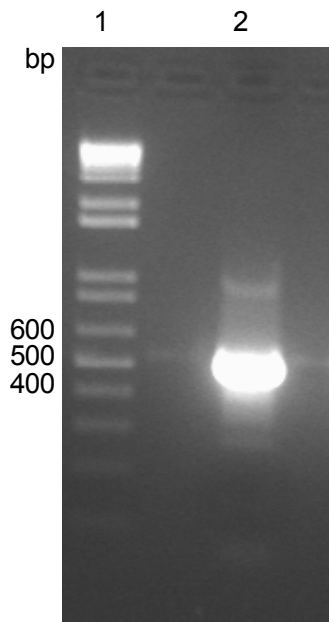


Figure 2

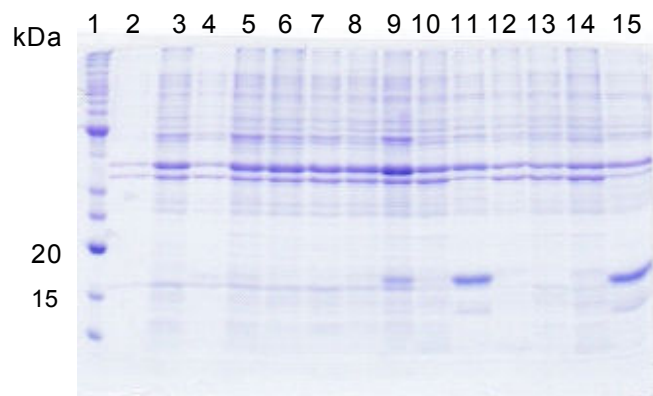


Figure 3

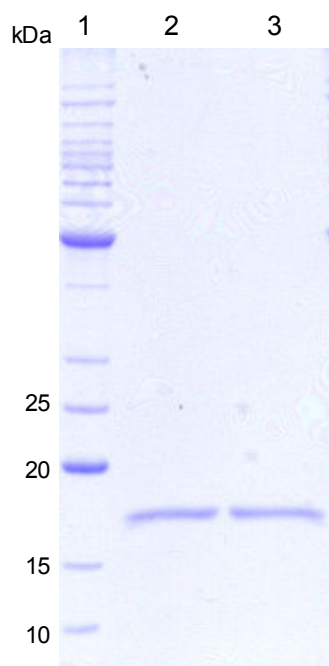


Figure 4

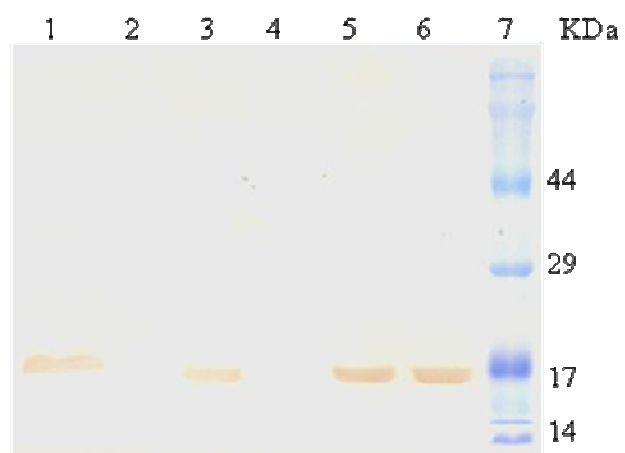


Figure 5

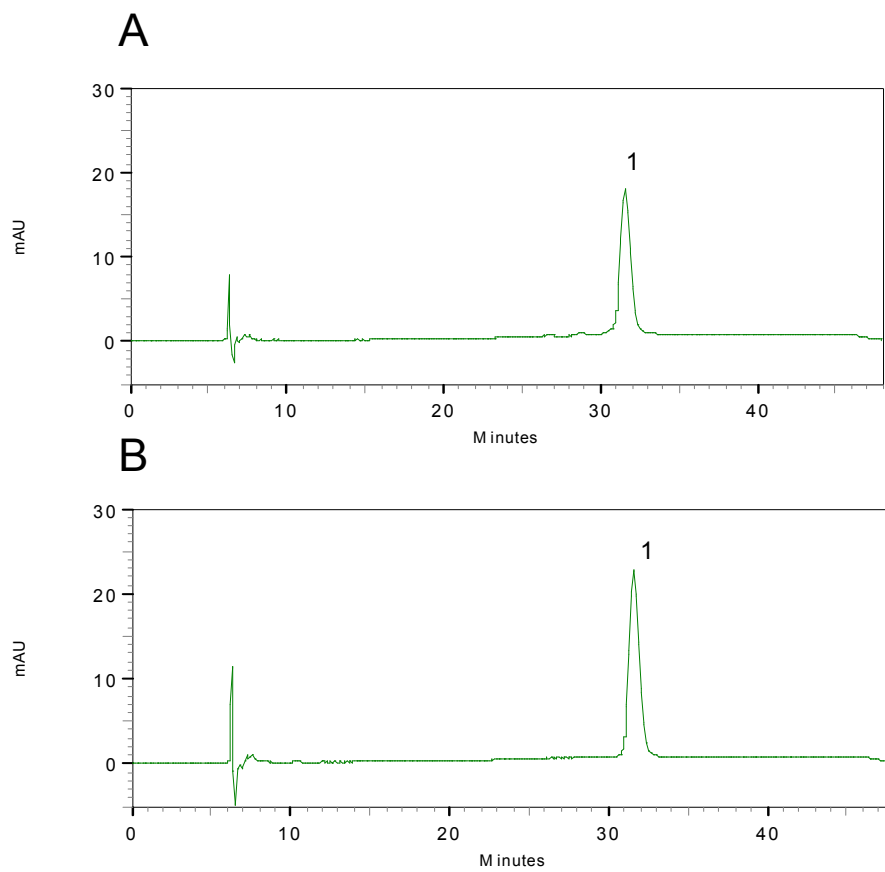


Figure 6

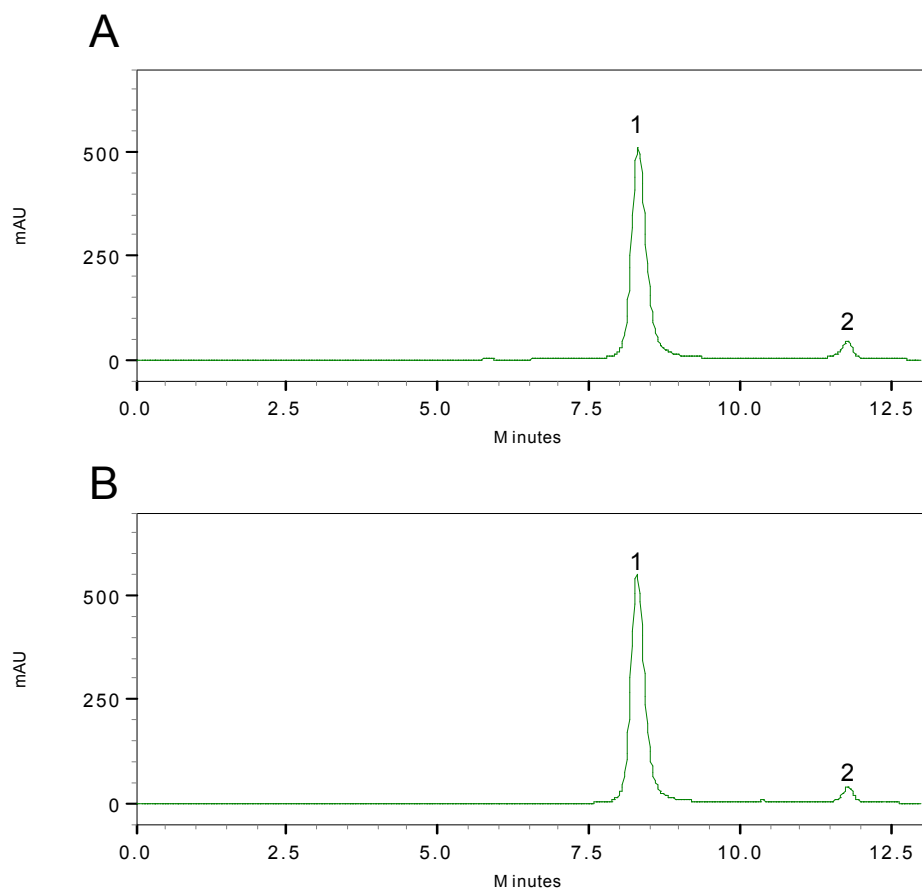
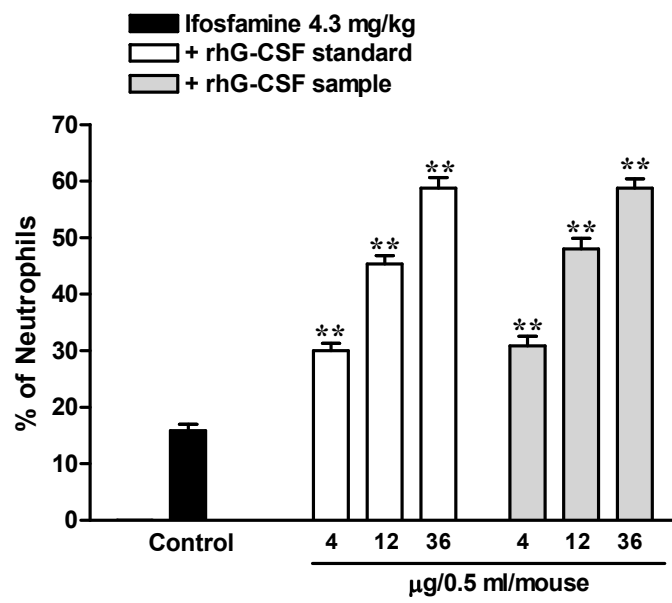


Figure 7





#### 4. CONSIDERAÇÕES FINAIS

O nascimento da indústria da biotecnologia, na década de 70, foi promovido pela descoberta das tecnologias de DNA recombinante e anticorpos monoclonais. Desde então, a indústria da biotecnologia se tornou, não só uma das principais fontes de novos medicamentos como também um motor de crescimento econômico e social. O desenvolvimento de biofármacos permitiu encontrar opções de tratamento para doenças crônicas e pouco freqüentes que não possuíam terapia ou cujas terapias não eram suficientemente bem-sucedidas para todo tipo de paciente. Atualmente, cerca de 300 milhões de pacientes no mundo são beneficiados com estes novos medicamentos (44).

Proteínas terapêuticas têm muitas vantagens sobre medicamentos obtidos por síntese química. Primeiro, devido à natureza complexa dessas proteínas e à alta especificidade quanto à sua função, a imitação por pequenas moléculas químicas é extremamente difícil. Além disso, provocam menos efeitos adversos devido a sua alta especificidade. Por serem moléculas naturais do nosso organismo são bem toleradas por induzirem pouca resposta imune. Por fim, podem ser usadas no tratamento de doenças na qual houve deleção ou mutação genética (51).

Com a expiração das patentes dos produtos biológicos inovadores, o desenvolvimento e produção de biossimilares tornaram-se um mercado possível e promissor. Neste trabalho foi desenvolvido um eficiente protocolo de expressão e purificação da proteína recombinante G-CSF em *E. coli*.

Após inúmeros testes (descritos no Anexo I), a melhor condição para expressão da proteína foi observada na cepa BL21(DE3) sem a indução por IPTG. Na tentativa de solubilizar a proteína a partir de corpos de inclusão foram utilizados não só diferentes agentes solubilizantes, como também um método para romper os

corpos de inclusão por diferença de pressão (Anexo II). O protocolo de purificação é baseado na lavagem dos corpos de inclusão, solubilização e desnaturação destes, *refolding* e uso de apenas uma coluna cromatográfica (troca catiônica) para finalizar a purificação da proteína. Para definir este protocolo, diferentes colunas cromatográficas foram testadas. O Anexo III apresenta os testes mais relevantes.

As análises físico-químicas, imunológicas e de atividade biológica *in vivo*, mostraram-se compatíveis quando comparadas com o padrão internacional. Portanto, os dados demonstraram que esse protocolo pode ser empregado para produção desta proteína terapêutica.

Este protocolo apresenta vantagens como a expressão da proteína sem a necessidade de indução por IPTG, o que favorece a diminuição dos custos de produção. Além disso, o protocolo de purificação do rhG-CSF utiliza somente uma coluna cromatográfica, mostrando-se mais atraente que outros protocolos que empregam múltiplas colunas cromatográficas (52,53,54). Em uma escala comercial, a redução no número de passos durante a purificação da proteína tem importância econômica e prática, por que cada passo não somente adiciona custos no produto final como também pode provocar perda na produção da proteína recombinante. Este é o primeiro passo no desenvolvimento de um biofármaco, o qual está sendo direcionado para uma escala piloto.

A metodologia estabelecida é econômica e simples, podendo ser reproduzida em escala industrial. Este é um importante projeto de desenvolvimento tecnológico concebido no Brasil para produção de biossimilares. Um dos principais impactos previstos será a diminuição das importações, dado que esse medicamento é totalmente importado, além do impacto social aumentando o acesso da população a esses medicamentos.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *Biologia molecular da célula*. 4.ed. Porto Alegre: Artes Médicas, 2002.
2. Zago M, Falcão R, Pasquini R. *Hematologia: fundamentos e prática*. São Paulo: Atheneu, 2001.
3. Lodish H, Berk A, Matsudaira P, Kaiser C, Krieger M, Scott M et al. *Biología Celular e Molecular*. 5ª ed. Artmed: Porto Alegre, 2005.
4. Hoffbrand A, Petti T. *Essential Haematology*. 3. ed. Oxford: Blackwell Science; 1993.
5. Hartung T, Aulock S, Wendel A. Role of granulocyte colony –stimulating factor in infection and inflammation. *Med Microbiol Immunol*. 1998 Oct;187(2):61-9.
6. Hubel K, Engert A. Clinical applications of granulocyte colony-stimulating factor: an update and summary. *Ann Hematol*. 2003 Apr;82(4):207-13.
7. Metcalf D. The Molecular biology and function of the granulocyte macrophage colony stimulating factors. *Blood*. 1986; 67(2):257-67.
8. Lu HS, Boone TC, Souza LM, Lai PH.. Disulfide and secondary structure of recombinant human granulocyte colony-stimulating factor. *Arch Biochem Biophys*. 1989 Jan;268(1):81-92.
9. Janeway C, Travers P, Walport M, Shlomchik M. *Imunobiologia: o sistema imunológico na saúde e na doença*. 4. ed. Porto Alegre: Artmed; 2000.
10. Barreda D, Hanington P, Belosevic M. Regulation of myeloid development and function by colony stimulating factors. *Dev Comp Immunol*. 2004 May 3;28(5):509-54.

11. Boneberg E, Hartug T. Molecular aspects of anti-inflammatory action of G-CSF. *Inflamm Res*. 2002 Mar;51(3):119-28.
12. Hill C, Osslund T, Eisenberg D. The structure of granulocyte colony-stimulating factor and its relationship to other growth factors. *Biochemistry*. 1993 June 90:5167-5171.
13. Simmers RN, Webber LM, Shannon MF, Garson OM, Wong G, Vadas MA, et al. Localization of the G-CSF gene on chromosome 17 proximal to the breakpoint in the t(15;17) in acute promyelocytic leukemia. *Blood*. 1987 Jul;70(1):330-2.
14. Oh-eda M, Hasegawa M, Hattori K, Kuboniwa H, Kojima T, Orita T, Tomonou K, Yamazaki T, Ochi N: O-linked sugar chain of human granulocyte colony-stimulating factor protects it against polymerization and denaturation allowing it to retain its biological activity. *J Biol Chem* 1990 July, 15(20):11432-11435.
15. García-Carbonero R, Mayordomo JI, Tornamira MV, López-Brea M, Rueda A, Guillem V, et al. Granulocyte colony-stimulating factor in the treatment of high-risk febrile neutropenia: a multicenter randomized trial. *J Natl Cancer Inst*. 2001 Jan 3;93(1):31-8.
16. Wittman B, Horan J, Lyman GH. Prophylactic colony-stimulating factor in children receiving myelosuppressive chemotherapy: a meta-analysis of randomized controlled trials. *Cancer Treat Rev*. 2006 Jun;32(4):289-303.
17. U.S FOOD & DRUG ADMINISTRATION. U.S Department of Health and Human Services. Therapeutic Biological Products Approvals. Disponível em: <[http://www.fda.gov/cder/biologics/biologics\\_table.htm](http://www.fda.gov/cder/biologics/biologics_table.htm)>. Acesso em: 4 jul. 2006. <http://www.accessdata.fda.gov/scripts/cder/onctools/druglist.cfm>
18. Kufe, Donald W, Pollock, Raphael E, Weichselbaum, Ralph R, et al. *Cancer Medicine*. 6.ed. Hamilton: Emil editora, 2003.
19. Harousseau JL, Witz B, Lioure B, Hunault-Berger M, Desablens B, Delain M, et al. Granulocyte colony-stimulating factor after intensive consolidation chemotherapy

in acute myeloid leukemia: results of a randomized trial of the Groupe Ouest-Est Leucémies Aigues Myeloblastiques.. *Journal of Clinical Oncology*. J Clin Oncol. 2000 Feb;18(4):780-7.

20. Babalola C, Nightingale C, Nicolau D. Adjunctive efficacy of granulocyte colony-stimulating factor on treatment of *Pseudomonas aeruginosa* pneumonia in neutropenic and non-neutropenic hosts. *J Antimicrob Chemother*. 2004 Jun;53(6):1098-100.

21. Kuritzkes DR, Parenti D, Ward DJ, Rachlis A, Wong RJ, Mallon KP. Filgrastim prevents severe neutropenia and reduces infective morbidity in patients with advanced HIV infection: results of a randomized, multicenter, controlled trial. *AIDS*. 1998 Jan 1;12(1):65-74.

22. Gough A, Clapperton M, Rolando N, Foster AV, Philpott-Howard J, Edmonds ME. Randomised placebo-controlled trial of granulocyte colony-stimulating factor in diabetic foot infection. *Lancet*. 1997 Sep 20;350(9081):855-9.

23. Harada M, Qin Y, Takano H, Minamino T, Zou Y, Toko H, et al. G-CSF prevents cardiac remodeling after myocardial infarction by activating the JAK-Stat pathway in cardiomyocytes. *Nat Med*. 2005 Mar;11(3):305-11.

24. Nagata S, Tsuchiya M, Asano S, Kaziro Y, Yamazaki T, Yamamoto O, et al. Molecular cloning and expression of cDNA for human granulocyte colony-stimulating factor. *Nature*. 1986 Jan 30-Feb 5;319(6052):415-8.

25. Souza LM, Boone TC, Gabilove J, Lai PH, Zsebo KM, Murdock DC, et al. Recombinant human granulocyte colony-stimulating factor: effects on normal and leukemic myeloid cells. *Science*. 1986 Apr 4;232(4746):61-5.

26. Souza L. Production of pluripotent granulocyte colony-stimulating factor. 1989. Patente US 4,810,643.

27. Baneyx F, Mujacic M. Recombinant protein folding and misfolding in *Escherichia Coli*. *Nat Biotechnol*. 2004 Nov;22(11):1399-408.

28. Walsh G. Biopharmaceuticals: biochemistry and Biotechnology. 2.ed. England: John Wiley & Sons; 2004.
29. Jevsevar S, Gaberc-Porekar V, Fonda I, Podobnik B, Grdadolnik J, Menart V. Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. *Biotechnol Prog.* 2005 Mar-Apr;21(2):632-9.
30. Villaverde A, Mar Carrio M: Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol Lett* 2003, 2:1385-1395.
31. Tsumoto K, Ejima D, Kumagai I, Arakawa T. Practical considerations in refolding proteins from inclusion bodies. *Protein Expr Purif.* 2003 Mar;28(1):1-8.
32. Clark E. Protein refolding for industrial process. *Curr Opin Biotechnol.* 2001 Apr;12(2):202-7.
33. Chalmers JJ, Kim E, Telford JN, Wong EY, Tacon WC, Shuler ML, et al. Effect of temperature on *Escherichia coli* overproducing  $\beta$ -lactamase or human epidermal growth factor. *Appl Environ Microbiol.* 1990 Jan;56(1):104-11.
34. Vasina J, Baneyx F. Expression of aggregation-prone recombinant proteins at low temperature: a comparative study of the *Escherichia coli* *cspA* and *tac* promoter systems. *Protein Expression and Purification.* 1997; 9:211-218.
35. Sorensen HP, Mortensen KK. Soluble expression of recombinant protein in the cytoplasm of *Escherichia coli*. *Microbiol Cell Factories* 2005 Jan; 4(1): 1-8.
36. Voet D, Voet J. Bioquímica. Porto Alegre: Artes médicas; 2006.
37. Choi J, Keum K, Lee S. Production of recombinant proteins by high cell density culture of *Escherichia coli*. *Chemical Engineering Science.* 2006 61:876-885.

38. BRASIL. Ministério da Saúde. Programa de Medicamentos Excepcionais. Disponível em: [http://www.opas.org.br/medicamentos/docs/pcdt/04\\_programa.pdf](http://www.opas.org.br/medicamentos/docs/pcdt/04_programa.pdf)
39. BRASIL. Ministério da Saúde. Política Federal de Assistência Farmacêutica 1990 a 2002. Brasília, 2002. Disponível em: [http://bvsmms.saude.gov.br/bvs/publicacoes/polit\\_fed\\_assist\\_farm.pdf](http://bvsmms.saude.gov.br/bvs/publicacoes/polit_fed_assist_farm.pdf)
40. BRASIL. Ministério da Saúde. Esclarecimento sobre medicamentos excepcionais. Disponível em: [http://portal.saude.gov.br/portal/aplicacoes/noticias/noticias\\_detalhe.cfm?co\\_seq\\_noticia=27873](http://portal.saude.gov.br/portal/aplicacoes/noticias/noticias_detalhe.cfm?co_seq_noticia=27873)
41. BRASIL. Ministério da Saúde. Portaria SAS/MS nº 862, de 04 de novembro de 2002. Protocolo Clínico e Diretrizes Terapêuticas. Neutropenia: tratamento com estimuladores de colônias das células progenitoras da medula óssea. Disponível em: [http://www.opas.org.br/medicamentos/docs/pcdt/do\\_m27\\_01.pdf](http://www.opas.org.br/medicamentos/docs/pcdt/do_m27_01.pdf)
42. Watson JD, Crick FH. A structure for deoxyribose nucleic acid. *Nature*. 1953; 171(4356):737-738.
43. Spada S. Directory of approved biopharmaceutical products. Florida: CRC Press LLC, 2005.
44. Roche. Medicamentos biológicos. Disponível em: <http://www.roche.com.br/NR/rdonlyres/27273F4A-1D0B-4FE4-9DD2-8951FB89EE6E/6488/medicamentosbiologicos1.pdf>
45. Teare I. Biosimilar warfare: the arrival of generic biopharmaceuticals the omnitrope decision. Lawtext Publishing Limited. Disponível em: [www.lawtext.com](http://www.lawtext.com)
46. Posição da Roche a respeito dos produtos biológicos similares. Disponível em: <http://www.roche.com.br/NR/rdonlyres/27273F4A-1D0B-4FE4-9DD2-8951FB89EE6E/6482/posicaodarochearespeitodosprodutosbiologicossimila.pdf>

47. Roger S, Mikhail A. Biosimilars: opportunity or cause for concern?. *J Pharm Pharmaceut Sci.* 2007; 10(3):405-410.
48. Covic A, Kuhlmann. Biosimilars: recent developments. *Int Urol Nephrol.* 2007; 36:261-266.
49. Pisani J, Bonduelle Y: Opportunities and barriers in the biosimilar market: evolution or revolution for generics companies? Disponível em: <http://www.ableindia.org/biosimilars.pdf>
50. A Revolução dos Biofármacos. *Revista de Manguinhos. Fiocruz.* 2005 Out; 24-27.
51. Leader B, Baca Q, Golan D. Protein therapeutics: a summary and pharmacological classification. *Nature Reviews.* 2008, Jan; 7:21-39.
52. Souza L: Production of pluripotent granulocyte colony-stimulating factor. 1989. Patente US 4,810,643.
53. Boone T, Miller, Allan L, Andresen, Jeffrey W: Method for purifying granulocyte colony stimulating factor. 1998. Patente US 5,849,883.
54. Porekar VG, Menart V: Process for the purification and/or isolation of biologically active granulocyte colony stimulating factor. 2005. Patente US 2005/01595589



## ANEXO I

### Testes de Superexpressão

A superexpressão da proteína recombinante humana G-CSF (rhG-CSF) foi comparada em diferentes cepas de *E. coli*: BL21(DE3), Rosetta(DE3), C41(DE3), C43(DE3), C41(DE3)pLys, C43(DE3)pLys. As distintas cepas de *E. coli* foram transformadas por eletroporação com o plasmídeo pET-23a(+) contendo o gene de interesse que codifica para a proteína G-CSF. O mesmo procedimento foi realizado com o vetor pET-23a(+), sem o gene de interesse clonado, para controle negativo de expressão. As colônias foram cultivadas em placas contendo meio LB (Luria Bertani) sólido com os antibióticos apropriados para cada cepa de acordo com a **Tabela 1** e incubadas durante 16h a 37°C.

A colônia selecionada foi inoculada em meio de cultura líquido com os respectivos antibióticos sob agitação de 180 rpm, e crescida a 37°C até chegar na fase logarítmica ( $OD_{600}$  de 0,4-0,6). As culturas foram crescidas com e sem a adição de 0,4 mM de isopropil  $\beta$ -D-tiogalactopiranosídeo (IPTG) e incubadas sob a agitação de 180 rpm em diferentes temperaturas (37, 25, 15 ou 12°C) por mais 24, 48, 54 ou 72 horas. Após a indução foram coletados 3 mL de cada cultura periodicamente e centrifugados a 23.000 g por 1 minuto.

Para análise do resultado da expressão, as células foram ressuspensas em tampão Tris/HCl 10 mM pH 8.0 e posteriormente rompidas por sonicação (3 vezes por 10 segundos a 21% de amplitude). As frações solúvel (sobrenadante) e insolúvel (sedimento) foram separadas e analisadas em gel de poliacrilamida 12% (SDS-PAGE). Os antibióticos utilizados e suas respectivas concentrações finais para o

cultivo de cada cepa de *E.coli* estão listados na **Tabela 1**. A **Tabela 2** apresenta as diferentes condições testadas.

**Tabela 1.** Antibióticos utilizados e suas respectivas concentrações finais.

<b>Cepas de <i>E. coli</i> transformadas com o plasmídeo pET-23a(+)*</b>	<b>Concentração de uso dos antibióticos</b>
BL21(DE3)	50 µg/mL de Carbenicilina
Rosetta(DE3)	50 µg/mL de Carbenicilina e 34 µg/mL de Cloranfenicol
C41(DE3)	50 µg/mL de Carbenicilina
C43(DE3)	50 µg/mL de Carbenicilina
C41(DE3)pLys	50 µg/mL de Carbenicilina e 34 µg/mL de Cloranfenicol
C43(DE3) pLys	50 µg/mL de Carbenicilina e 34 µg/mL de Cloranfenicol

Fonte: Protocolo de células competentes - Novagen® - [www.novagen.com](http://www.novagen.com)

\* O plasmídeo pET-23a(+) contém um gene de resistência a carbenicilina.

**Tabela 2:** Diferentes condições testadas para expressão da proteína rhG-CSF em *E. coli*. As células marcadas com “X” representam as diferentes temperaturas testadas para cada cepa no meio de cultivo 4YT.

	Meio de cultivo 4YT			
Temperatura	37°C	25°C	15°C	12°C
Cepa de <i>E. coli</i>				
BL21(DE3)	X	X	X	X
Rosetta (DE3)	X			
C41(DE3)	X			
C43(DE3)	X			
C41(DE3)pLysS	X			
C43(DE3)pLysS	X			

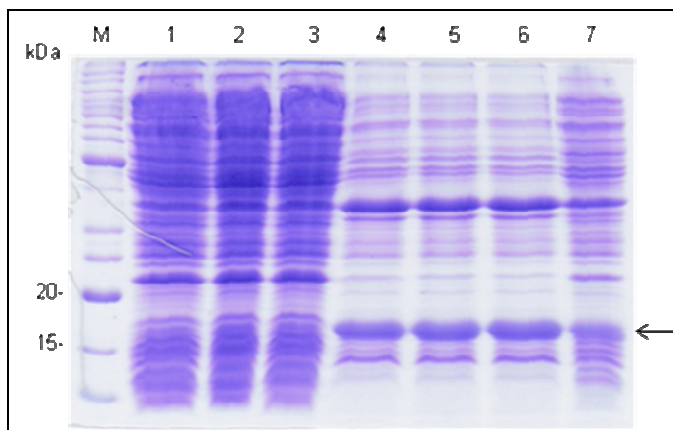
Dentre as condições testadas, apenas as cepas BL21(DE3) e Rosetta (DE3) sob a temperatura de 37°C e não induzida, apresentaram a expressão da proteína, sendo esta na fração insolúvel. Não foi observada sua expressão na fração solúvel em nenhuma das condições testadas (dados não mostrados).

## ANEXO II

### Testes de solubilização

#### 1. Prensa de French

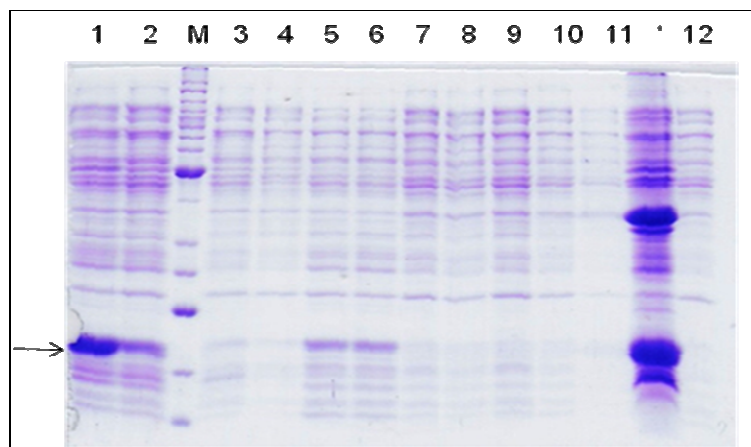
Após determinada a expressão de rhG-CSF em BL21(DE3) a 37°C, foi necessário estabelecer o protocolo de solubilização visto que a proteína é expressa em corpos de inclusão. Após crescimento celular, a amostra foi dividida entre três tubos contendo cada tubo 1g de célula ressuspensa em 5mL de Tris HCl 50mM pH 8,0. A amostra foi rompida em diferentes pressões 20.000, 27.000 e 35.000 psi, sendo que em cada pressão a mesma amostra foi passada 1, 2 e 3 vezes na prensa. As frações foram separadas por centrifugação a 23.000 g a 4°C por 30min e então analisadas. A **Figura 1** apresenta as amostras solúveis e insolúveis após rompimento sob 27.000 psi. As amostras correspondentes a 20.000 e 35.000 psi não estão representadas, pois apresentaram resultados muito similares. Não foi observada a proteína rhG-CSF na fração solúvel e nem diferenças com o número de passagens da mesma amostra na prensa.



**Figura 1:** Análise em SDS-PAGE 12% de rhG-CSF (18,8 kDa) após rompimento na prensa de French sob 27.000 psi. M: marcador de peso molecular (BenchMark™ Protein Ladder, Invitrogen). Canaleta 1,2 e 3: fração solúvel passada 1, 2 e 3 vezes, respectivamente, na prensa; canaletas 4,5 e 6: fração insolúvel passada 1, 2 e 3 vezes, respectivamente, na prensa; canaleta 7: extrato bruto antes do rompimento.

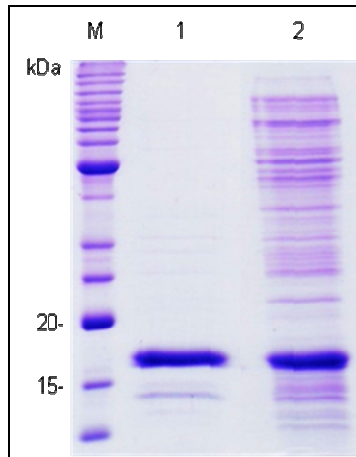
## 2. Agentes solubilizantes

Outros protocolos utilizando diferentes agentes solubilizantes foram testados de acordo com a **Figura 2**. Depois do crescimento em agitador, as células foram rompidas por sonicação e a fração insolúvel separada para posterior tentativa de solubilização. A amostra foi separada em 12 tubos de polipropileno (1,5mL) e ressuspensa com 1mL de cada agente solubilizante. Após 1 hora de agitação em temperatura ambiente, as frações solúvel e insolúvel foram separadas por centrifugação sob 23000 g a 4 °C por 30 min. Os agentes solubilizantes foram retirados da fração solúvel por meio de diálise contra o tampão Tris/HCl 50mM pH 8,0 durante 12 horas. Foram utilizados sacos de diálise de 6.000-8.000 Da. Após a diálise, as frações solúvel e insolúvel foram novamente separadas como descrito acima e analisadas em SDS-PAGE (**Figura 2**).



**Figura 2.** Análise da fração solúvel do rhG-CSF (18,8kDa) em SDS-PAGE 12% após teste de solubilização com diferentes agentes solubilizantes. M: marcador de peso molecular (BenchMark™ Protein Ladder, Invitrogen). Canaleta 1: uréia 8M; canaleta 2: uréia 6M; canaleta 3: uréia 4 M; canaleta 4: uréia 2M; canaleta 5: GndCl 6M; canaleta 6: GndCl 4M; canaleta 7: tween 0,75%; canaleta 8: tween 0,5%; canaleta 9: tween 0,25%; canaleta 10: KCl 0,5M; canaleta 11: KCl 1M; canaleta 12: KCl 2M. (\*) amostra insolúvel após rompimento.

Como pode ser visto na **Figura 2**, a solubilização de rhG-CSF com uréia 8M mostrou-se mais eficiente. Observou-se também, a solubilização de uma grande quantidade de proteínas contaminantes com uréia 2M. A partir destes resultados, o primeiro protocolo de solubilização foi estabelecido sendo: solubilização das proteínas contaminantes com uréia 2M e posterior solubilização do rhG-CSF com uréia 8M. A retirada de toda uréia é realizada por meio de diálise e a fração solúvel é separada para purificação (**Figura 3**).

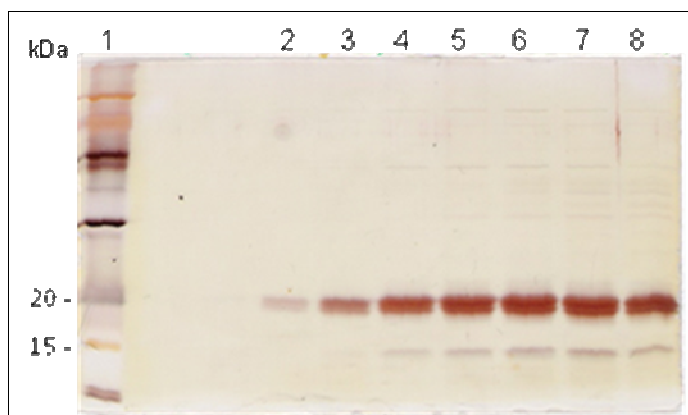


**Figura 3.** Análise da solubilização do rhG-CSF (18,8kDa) em SDS-PAGE 12%. M: marcador de peso molecular (Low Range, Invitrogen); canaleta 1: solubilização das proteínas contaminantes com uréia 2M e posterior solubilização do rhG-CSF com uréia 8M; canaleta 2: solubilização do rhG-CSF com uréia 8M.

## ANEXO III

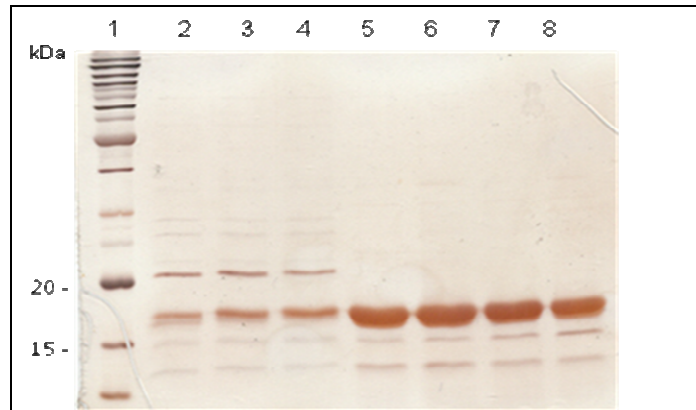
### Testes de purificação

Os testes de purificação foram realizados por meio de Cromatografia Líquida de Alta Performance (FPLC) empregando o equipamento do tipo ÄKTA-Pilot (GE Healthcare) para o controle do fluxo, pressão da coluna cromatográfica, detecção de múltiplos comprimentos de onda, controle de gradiente, aplicação e coleta das amostras. Após o preparo da amostra como descrito no **Anexo 2**, foram utilizadas diferentes colunas de troca catiônica (Hitrap SP-FF, SP-XL e CM-FF) e uma coluna de exclusão molecular (Hitrap Sephacryl S-200 HR) a fim de obter o melhor rendimento da proteína recombinante. Os resultados mais relevantes estão apresentados abaixo nas **Figuras 1 e 2**.



**Figura 1.** Análise da purificação em coluna de troca catiônica Hitrap SP-XL do rhG-CSF em SDS-PAGE 12% corado com prata. Os tampões utilizados foram: tampão A1: ácido cítrico 50mM pH 4,5 e tampão B1: Tp A1 + NaCl 1M. Canaleta 1: marcador de peso molecular (Low Range, Invitrogen); canaletas 2-8: diferentes frações do pico do cromatograma mostrando o rhG-CSF (18,8kDa) e alguns contaminantes.





**Figura 2.** Análise da purificação em coluna de exclusão molecular Hitrap Sephacryl S-200 HR do rhG-CSF em SDS-PAGE 12% corado com prata. O tampão utilizado foi: tampão A1: ácido cítrico 50mM pH 4,5. Canaleta Linha 1: marcador de peso molecular (Low Range, Invitrogen); canaletas 2-8: diferentes frações do pico do cromatograma mostrando o rhG-CSF (18,8kDa ) e alguns contaminantes.

## ANEXO IV

**From:** Antonio Pedro Villaverde Corrales [mailto:Antoni.Villaverde@uab.cat]  
**Sent:** Wed 3/12/2008 10:52 AM  
**To:** Luiz Augusto Basso  
**Subject:** ms 1888263491179891 is acceptable for publication

MS: 1888263491179891

Research

Human granulocyte colony stimulating factor (hG-CSF): cloning, overexpression, purification and characterization

Ana LS Vanz, Gaby Renard, Mario S Palma, Joceli M Chies, Sergio L Dalmora, Luiz A Basso and Diogenes S Santos

Microbial Cell Factories

Dear Dr. Basso,

your manuscript has been transferred to our production department. In-house editors will shortly contact you regarding any format change required prior to the publication of your file.

Best regards,

A. Villaverde

Tel: +44 (0)20 7323 0323

Facsimile: +44 (0)20 7631 9923

e-mail: Antoni.Villaverde@uab.es

Web: <http://www.microbialcellfactories.com/>