Pflügers Arch – Eur J Physiol (2001) 442 [Suppl 1]: R184-R186 © Springer Verlag 2001

Human granulocyte colony stimulating factor (hG-CSF) expressed by methylotrophic yeast *Pichia pastoris*

Marija Anžur Lasnik[1], Vladka Gaberc Porekar[2], Anton Štalc[1]

[1]LEK d.d., R&D, Verovškova 69, 1526 Ljubljana, Slovenia;
[2]National Institute of Chemistry, Ljubljana, Slovenia

Abstract Human granulocyte colony stimulating factor (hG-CSF) was expressed in the methylotrophic yeast *Pichia pastoris*, using two different constructs which resulted in proteins with different N-terminal sequences. In the first construct, a hexa-histidine tag and enterokinase cleavage site were added to the N-terminus of the protein to achieve one-step separation and exact processing. In the second construct, the gene was fused to the α-MF prepro leader at the Lys-Arg processing site (without Glu-Ala spacer). The PCR products were cloned in pPIC9 commercial vector and integrated into the alcohol oxidase region of the host genome. Transformation was done by electroporation or spheroplasting. Selection of good producing clones was performed by immunoblot analyses of the supernatants from shake-flask fermentation. Proper processing of the products was confirmed by N-terminal sequencing of the secreted proteins. With both plasmid constructs, the target proteins, bearing the histidine tag or not, represented majority of the secreted proteins. Although the proteins were present in the soluble form, they were highly aggregated, which interfered with purification. The most efficient way to obtain monomeric, biologically active protein was complete denaturation by guanidine-HCl or urea and subsequent renaturation during gel filtration chromatography.

Key words G-CSF • *ichia pastoris* • methylotrophic yeast

Introduction

Methylotrophic yeasts have some advantages as host organisms for production of pharmaceutically important proteins. Foreign genes of interest can be stably integrated into the host yeast genome via homologous or heterologous recombination. These microorganisms are capable of growing to very high biomass (more than 100 g/l) resulting into high product yields per fermentation. Induction of protein production is evoked by methanol, which is a low cost input raw material. Yeasts can exert post-translational modifications, like glycosylation and disulfide bond formation. On the other hand, they posses authentic processing and secretion abilities, which can make the isolation and purification procedure much easier and cheaper [1]. Their advantages were the reason to use *Pichia pastoris* methylotrophic yeast as host for expression of hG-CSF, a hormone-like protein with 174 amino acids, two disulfide bonds and one free cysteine residue. Since it stimulates granulocyte colony formation and affects proliferation, differentiation and activation of hematopoietic cells of the neutrophilic granulocyte lineage, it is widely applicable for treatment of neutropenia in cancer therapy and HIV-associated neutrophil defects [2, 3]. Recombinant hG-CSF has already been expressed in various hosts, such as bacteria, conventional yeast and mammalian cells [3]. Genetically engineered *P. pastoris* is capable of secreting numerous recombinant proteins, especially successfully when employing the signal sequence of native *Saccaromyces cerevisiae* alpha mating factor (α-MF) prepro leader, which targets the secretory pathway. In yeast, the secretory proteins pass through the endoplasmatic reticulum and Golgi apparatus, and are packaged into secretory vesicles, which fuse with the plasma membrane to release the fully processed protein into the medium. The sequences encoding this α-MF signal have been incorporated into the *P. pastoris* pPIC9 commercial expression vector, which has been used in our experiments. Processing of the α-MF prepro leader requires three proteolytic activities: 1) signal peptidase that cleaves the 19 amino acid pre region, 2) a dibasic endopeptidase, the product of KEX 2 gene that cleaves on the carboxy-terminal side of Lys-Arg at the junction between α-MF prepro and the foreign protein and 3) a dipeptidyl aminopeptidase, STE 13 gene product, that removes Glu-Ala spacer residues from the N-terminal of the foreign protein. Efficiency of the secretion and processing of the particular foreign protein using the above pre-type signal sequences varies for different proteins and can not be predicted [1, 4]. We constructed two hybrid hG-CSF protein precursors with two different N-termini where we exposed different sites for proteolytic cleavage in order to study protein processing and yeast secretion efficiency, as well as different purification possibilities.

Materials and methods

Host strains. *P. pastoris*, strain GS 115 methylotrophic yeast was used as a heterologous protein expression system. *Escherichia coli*, TOP 10 F′ and One Shot™ (Invitrogen) strains were used as hosts for expression and cloning vectors.

Cloning vectors. BBG13 (British Biotechnology), plasmid with hG-CSF includes N-terminal Met gene cloned between the *Hind* III and *Eco*RI sites in the polylinker of pUC 18. It was used as a template for PCR amplification of the G-CSF gene. PCR 2.1 vector (Invitrogen) was used for direct cloning of PCR products.

Pichia expression vector. pPIC 9 (Invitrogen) plasmid with strong inducible AOX1 (alcohol oxidase) promoter and with α-MF signal secretion was used to prepare G-CSF plasmid constructs with two different N-terminal sequences.

Cultivation media. LB medium: 1% tryptone, 0.5% yeast extract, 1% NaCl, pH 7.0. LB/ampicillin agar plates: LB medium plus 15 g/l agar-agar, 50 μg/ml medium ampicillin added. LB/ampicillin/X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside): LB-ampicillin plate with 40 μl (40 mg/ml) X-gal on the plate. YPD plates (per litre): 10 g yeast extract, 20 g bacto peptone. MD plates: 1.34% YNB, $4x10^{-5}$% biotin, 1% dextrose. BGY growth medium: 100 mM K-phosphate, 1.34% YNB; $4x10^{-5}$% biotin, 1% glycerol, MM minimal production medium: 1.34% YNB; 4x 10^{-5}% biotin; 0.5% methanol. BMM medium: with 100 mM K-phosphate pH 6,0 buffered MM medium.

Genetic manipulations. Expression plasmid constructs preparation. pPHGCSF plasmid construct with a hexa-histidine tag (His6) and enterokinase cleavage site (EK= Asp4Lys) on the N-terminus of hG-CSF was prepared by amplification of hG-CSF gene from BBG13, using PCR primers with the following sequences:

5′ G CGC TAC GTA CAT CAT CAT CAT CAT CAT GAC GAC GAC GAC AAG ATG ACC CCC CTG GGC CCT 3′;

5′ GGT GCA GAA TTC TTA TCA GGG CTG CGC AAG GT 3′.

The G-CSF gene for the pPDSGCSF plasmid construct was also PCR-amplified from BBG13, with the following primers:

5′ GGT GCA CTC GAG AAA AGA ATG ACC CCC CTG GGC CCT 3′

5′ GGT GCA GAA TTC TTA TCA GGG CTG CGC AAG GT 3 ′.

Ligations of PCR products in pCR2.1 cloning vector and transformations in *E. coli* One Shot™ cells were performed using Original TA Cloning® Kit (Invitrogen). The correct amplified sequences of the inserts were confirmed by DNA sequencing. The genes were excised from the vectors with restriction enzymes *Sna*BI and *Eco*RI (in the first construct) or *Eco* RI and *Xho*I (in the second construct) getting pPHGCSF and pPDSGCSF plasmid constructs. Inserts were cloned into the pPIC9 vector, using the same restriction sites. The fragments were separated on the 0.7% low melting point electrophoretic gel and ligated in the gel at 14-16°C overnight. Transformations of *E. coli* TOP 10 F′ competent cells were done by electroporation. Gene inserts of constructs were confirmed by restriction analysis.

Integration of plasmid constructs in P. pastoris genome and selection of good producing clones. About 10 μg DNA of pPHGCSF and pPDSGCSF plasmid constructs were linearised by *Bgl*II restriction enzyme and integrated in *Pichia* genome, by spheroplasting using Pichia Spheroplast Kit or electroporation (both Invitrogen protocols). Four days old colonies were cultivated for one day on BMGY medium (10ml), than transferred to 50 ml MM medium for the recombinant protein production (30°C, 160 rpm.). After one day of cultivation, cells were separated from the medium and 1 ml samples of supernatants were slot-blotted on the nitrocellulose membrane for immuno detection and screening of good producing clones.

Protein purification and characterisation. After 70-90 hours of shake-flask fermentation on the MM medium (160 rpm and 30°C), the fermentation broth was centrifuged at 10000 rpm for 15 min., and the supernatant was concentrated at +4°C approximately 10 times, using an Amicon stirred cell device and YM-10 membrane. IMAC chromatography was done on Ni-NTA Agarose (Qiagen) with 0.020 M Na-phosphate, 0.2 M NaCl, pH 7.0 loading buffer and linear gradient of 0.020 M-Na acetate, 0.2 M NaCl, pH 3.8 elution buffer. Different detergents (Triton X-100, Tween 20, Tween 80, Sarcosyl) in the concentration range from 0.2 to 1% and urea (2M and 6M) were tested as additives to the buffers in order to increase the fraction of the bound His6-G-CSF. Ion exchange chromatography was done on SP Sepharose fast flow (Pharmacia) with 50 mM Na-acetate, pH 4.8 loading buffer and elution was accomplished by linear salt gradient. Gel filtration was performed on Sephacryl S-200 (Pharmacia) in 0.050 M Na- phosphate, 0.050 M NaCl, pH 7.0 buffer.

Protein concentration was determined according to the Bradford procedure against BSA standard curve. Concentration of the pure G-CSF was determined from the UV absorbance at 280 nm taking into account a molar extinction coefficient of 16400 $M^{-1}cm^{-1}$. SDS-PAGE was carried out on 0.75 mm thick 15% separating gel with a 4% stacking gel according to Laemmli procedure in the Mini-Protean II electrophoresis unit (Bio-Rad). Coomassie brilliant blue or sensitive silver stain was used to visualise the proteins. Immunodetection was performed as slot blot or Western blot after SDS-PAGE and electrophoretic transfer, using polyclonal goat anti-hG-CSF primary antibodies (R&D) and rabbit anti-goat IgG peroxidase conjugate (Sigma) secondary antibodies. Densitometrical evaluation of Coomassie stained SDS-PAGE gels or immunoblots was done on the Imaging Densitometer GS-670 (Bio-Rad) using analytical imaging software Molecular Analyst 1.1.1 (Bio-Rad). Amino-terminal sequence analysis was performed on Applied Biosystems Model 492 Pulsed Field Liquid Sequenator.

Secretion efficiency of G-CSF, expressed as the ratio of extracellular to total synthesized G-CSF, was determined by densitometric evaluation of immunoblots.

Results and discussion

Recombinant G-CSF has wide pharmaceutical applicability, therefore we explored its expression in methylotrophic yeast *P. pastoris*, which is a versatile host organism capable of efficient secretion. Since micro-heterogeneity of the N-terminus is not acceptable in the case of clinical-grade proteins, we tested the processing and secretion efficiency of G-CSF protein in comparison with the protein bearing the histidine affinity tag, which can be exactly cleaved off. Histidine tag proved to be sufficiently resistant against proteolytic attack during fermentation process and our experiments confirmed its successful elimination by enterokinase (data not shown).

Preparing pDSGCSF plasmid construct we made a deletion of Glu Ala repeats and *Sna*BI restriction site after Lys Arg signal cleavage originally present in pPIC9 plasmid. It is known that foreign gene fused to MF prepro leader at the Lys Arg processing site whithout Glu Ala spacer can lead to properly or non-properly processed protein. On the other hand, Glu Ala spacer between Lys Arg site and the N-terminus of the foreign gene can lead to a secreted protein, which is efficiently cleaved after Lys Arg, but it can retain the Glu Ala spacer at the amino terminus [1, 4]. Both recombinant proteins His6-EK-G-CSF and G-CSF were efficiently secreted (Fig 1) and the amino acid analysis confirmed the exact processing of the amino terminals. Secretion efficiency of 95% in the case of His6-EK-G-CSF and 97% for G-CSF obtained in cultivation on the minimal medium are much better than reported for secretion of G-CSF from *Saccharomyces cerevisiae* [2] where the efficiencies of 26 to 47% were achieved. The expression of our recombinant proteins was higher than 10 mg/l, which is relatively good for shake flask fermentation, where biomasses of approximately 10 g/l are usually achieved. Most probably we could get even better results in the fermenter, where growth to high cell densities was reported, reaching > 130 g/l. Additionally, in the fermenter 3-5 times higher transcription levels can be obtained as a consequence of the controlled methanol concentration [1].

In our experiments both recombinant proteins His6-EK-G-CSF and G-CSF represented majority of the *P. pastoris* secreted proteins (Fig 1). Due to a hexa-histidine tag in the

case of the former, Immobilised Metal Ion Affinity Chromatography (IMAC) was used as a single purification step, while cation exchange chromatography was applied for purification of the latter. Unfortunately, in both cases only a small portion of the protein present in the supernatant bound to the respective column, while majority of the protein flowed through and eluted in the huge peak at the beginning of the chromatogram. In IMAC we performed a series of experiments with different detergents (Triton X-100, Tween 20, Tween 80, Sarcosyl, in the concentration range from 0.2 to 1%) and urea (2M and 6M) added to the sample and chromatographic buffers in order to increase the fraction of the bound His6-EK-G-CSF. The best results were achieved by addition of 6M urea, which indicated that very harsh solubilisation or even denaturation was needed to expose the histidine affinity tag. Most probably, majority of the secreted recombinant proteins is not quite correctly folded, which results in the formation of soluble aggregates. As judged from SDS-PAGE with and without the reducing agent and subsequent immunoblot analysis (data not shown) these aggregates are kept together by hydrophobic interactions and to a lesser extent by intermolecular disulfide bonds. Although numerous heterologous proteins often aggregate and form inclusion bodies when expressed intracellularly in recombinant *E. coli*, formation of soluble aggregates in the case of secreted products is not a frequent phenomenon. Something similar has already been mentioned for G-CSF secreted from recombinant *S. cerevisiae* [2, 3].

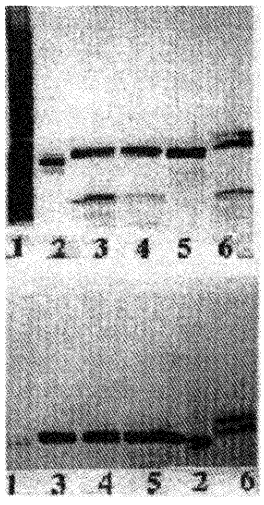

Fig. 1 SDS PAGE, silver stained (top) and immunoblot (bottom) of G-CSF and His6-EK-G-CSF proteins. Lane 1: 2.5 μg intracellular proteins (from the sonicated pellet); Lane 2: 0.3 μg hG-CSF standard (Neupogen syringe, Amgen); Lane 3, 4, 5: extracellular hG-CSF from selected clones (20 μl supernatant); Lane 6: extracellular His6-EK-G-CSF (20 μl supernatant).

By changing cultivation conditions we observed that higher pH of the medium promoted aggregation. Unfortunately, also by addition of Tween 80 to the cultivation medium we were not able to circumvent the formation of aggregates in the *P. pastoris* supernatant. As seen from Table 1, we tried different denaturation and renaturation procedures and achieved a significant increase of the yield of monomeric G-CSF. Over 80% of the monomeric G-CSF were recovered when using denaturation by guanidine-HCl and subsequent separation and renaturation during gel filtration chromatography. Although we obtained almost the same percentage of monomeric protein using classical renaturation during dialysis, renaturation on the gel filtration column is much more effective in the respect of speed and protein yield. To conclude, we succeeded to develop an efficient expression/secretion system for hG-CSF in *P. pastoris* and achieved to isolate the protein in the monomeric, biologically active form.

Table 1 Effectiveness of different denaturation and renaturation procedures, measured as percentage of aggregated versus monomeric GCSF in gel filtration chromatography.

Denaturation	Renaturation	Aggregated Form (%)	Monomeric Form (%)
-	-	95.7	4.3
8 M urea, 2h	Gel filtration	53.5	46.5
4.3 M Gdn-HCl, 2 h	Gel filtration	20.4	79.6
4.3 M Gdn-HCl, 6 h	Gel filtration	12.5	87.5
6.0 M Gdn-HCl,18 h	Gel filtration	13.1	86.9
6.0 M Gdn-HCl, 8 h	Overnight dialysis	18.0	82.0

In each case fractions from peak A (aggregated form) and peak M (monomeric form) in gel filtration chromatography were pooled together and analysed by SDS-PAGE; gels were Coomassie stained and densitometrically evaluated.

Acknowledgements We thank Dr. Menart for suggestions in cloning strategy, Mrs. Jesenko and M. Sc. Vasle Preradov for technical support on protein purification and characterisation, Mrs. Lenarčič for gene sequencing, Mrs. Pezdirc for biological activity measurements and Dr. Križaj for amino-terminal sequence analysis.

References

1. Cereghino JL and Cregg JM (2000) Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiology Reviews 24: 45-66
2. Bae CS, Yang DS, Chang KR, Seong BL, Lee J (1998) 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 57: 600-609
3. Bae CS, Yang DS, Lee J. Park Y-H (1999) Improved process for production of recombinant yeast-derived monomeric human G-CSF. Appl Microbiol Biotechnol 52: 338-344
4. Zsebo KM, Lu H-S, Fieschko JC, Goldstein L, Davis J, Duker K, Suggs SV, Lai PH, Bitter GA (1986) Protein secretion from *Saccharomyces cerevisiae* directed by the prepro-α-factor leader region. J Biol Chem 261: 5858-5865