

# Strategies for the production of difficult-to-express full-length eukaryotic proteins using microbial cell factories: production of human alpha-galactosidase A

Ugutx Unzueta · Felicitas Vázquez · Giulia Accardi · Rosa Mendoza ·  
Verónica Toledo-Rubio · Maria Giuliani · Filomena Sannino · Ermenegilda Parrilli ·  
Ibane Abasolo · Simo Schwartz Jr. · Maria L. Tutino · Antonio Villaverde ·  
José L. Corchero · Neus Ferrer-Miralles

Received: 10 September 2014 / Revised: 12 December 2014 / Accepted: 14 December 2014 / Published online: 24 January 2015  
© Springer-Verlag Berlin Heidelberg 2015

**Abstract** Obtaining high levels of pure proteins remains the main bottleneck of many scientific and biotechnological studies. Among all the available recombinant expression systems, *Escherichia coli* facilitates gene expression by its relative simplicity, inexpensive and fast cultivation, well-known genetics and the large number of tools available for its biotechnological application. However, recombinant expression in *E. coli* is not always a straightforward procedure and major obstacles are encountered when producing many eukaryotic proteins and especially membrane proteins, linked to missing

posttranslational modifications, proteolysis and aggregation. In this context, many conventional and unconventional eukaryotic hosts are under exploration and development, but in some cases linked to complex culture media or processes. In this context, alternative bacterial systems able to overcome some of the limitations posed by *E. coli* keeping the simplicity of prokaryotic manipulation are currently emerging as convenient hosts for protein production. We have comparatively produced a “difficult-to-express” human protein, the lysosomal enzyme alpha-galactosidase A (*hGLA*) in *E. coli* and in the

**Electronic supplementary material** The online version of this article (doi:10.1007/s00253-014-6328-9) contains supplementary material, which is available to authorized users.

U. Unzueta · F. Vázquez · G. Accardi · R. Mendoza ·  
V. Toledo-Rubio · A. Villaverde · J. L. Corchero ·  
N. Ferrer-Miralles (✉)  
Institut de Biotecnologia i de Biomedicina,  
Universitat Autònoma de Barcelona,  
Bellaterra, 08193 Barcelona, Spain  
e-mail: neus.ferrer@uab.cat

U. Unzueta · G. Accardi · A. Villaverde · J. L. Corchero ·  
N. Ferrer-Miralles  
Departament de Genètica i de Microbiologia,  
Universitat Autònoma de Barcelona,  
Bellaterra, 08193 Barcelona, Spain

U. Unzueta · F. Vázquez · R. Mendoza · V. Toledo-Rubio ·  
I. Abasolo · S. Schwartz Jr. · A. Villaverde · J. L. Corchero ·  
N. Ferrer-Miralles  
CIBER de Bioingeniería, Biomateriales y Nanomedicina  
(CIBER-BBN), Barcelona, Spain

M. Giuliani · F. Sannino · E. Parrilli · M. L. Tutino  
Dipartimento di Scienze Biochimiche, Facoltà di Medicina  
e Chirurgia, Università degli studi di Palermo,  
90127 Palermo, Italy

F. Sannino  
Institute of Protein Biochemistry, CNR, Via Pietro Castellino 111,  
I-80131 Naples, Italy

I. Abasolo · S. Schwartz Jr.  
VHIR Vall d’Hebron Inst Recerca, CIBBIM Nanomed,  
08035 Barcelona, Spain

*Present Address:*  
F. Vázquez  
GREENALTECH S.L., 08028 Barcelona, Spain

G. Accardi  
Department of Pathobiology and Medical and Forensic  
Biotechnology, University of Palermo, 90134 Palermo, Italy

V. Toledo-Rubio  
BioIngenium S.L., 08028 Barcelona, Spain

M. Giuliani  
Novartis Vaccines and Diagnostics, via Fiorentina 1, 53100 Siena,  
Italy

psychrophilic bacterium *Pseudoalteromonas haloplanktis* TAC125 cells (*P. haloplanktis* TAC125). While in *E. coli* the production of active hGLA was unreachable due to proteolytic instability and/or protein misfolding, the expression of hGLA gene in *P. haloplanktis* TAC125 allows obtaining active enzyme. These results are discussed in the context of emerging bacterial systems for protein production that represent appealing alternatives to the regular use of *E. coli* and also of more complex eukaryotic systems.

**Keywords** Recombinant protein · Expression systems · *Escherichia coli* · *Pseudoalteromonas haloplanktis* TAC125 · Human alpha-galactosidase A · Fabry's disease

## Introduction

The deep genetic and physiological characterization, short generation time, ease of handling, established fermentation know-how and finally the capacity to accumulate foreign proteins to a high percentage of the total cellular protein content have made *Escherichia coli* the most widely used prokaryotic organism for recombinant protein production. However, there are disadvantages in using *E. coli* as an expression host. For instance, *E. coli* is not capable of producing eukaryotic posttranslational modifications, such as glycosylation, phosphorylation or disulphide bridge formation in the reducing cytoplasm, which can be critical for the production of folded, active proteins. Even so, several strategies can be implemented to obtain specific posttranslational modifications as the use of *E. coli* strains which have been engineered to maintain oxidizing conditions in the cellular milieu or to send the recombinant protein to the periplasmic space to allow cysteine bridging (Inaba 2009; Nozach et al. 2013; Prinz et al. 1997). In addition, a N-linked protein glycosylation system has been recently identified in the human enteropathogen *Campylobacter jejuni* that can be transferred to *E. coli* (Wacker et al. 2006), opening up the possibility of engineering recombinant products needing glycosylation for research and industrial applications.

On the other hand, some proteins, especially large and membrane proteins, simply fail to be produced in *E. coli*, and they occur as proteolysed species or deposited in inclusion bodies (Vallejo and Rinas 2004). Others are subjected to premature termination of translation in the presence of repetitive DNA sequences or rare codons (Daly and Hearn 2005; Sallach et al. 2009) or have a low rate of internal translation initiation (Ferreira et al. 2013; Nakamoto 2009). A large number of studies describe the conversion of proteins accumulated in inclusion bodies into soluble forms. Mainly, these methods can be categorized into three alternative approaches. In the first one, factors influencing the amount of recombinant protein present in the insoluble fraction can be modified through

careful optimization and control of the production conditions, leading to the expression of the recombinant protein in its soluble version. In this context, conventional approaches include gene expression at low temperature, use of promoters with different transcriptional strengths, modifications of growth media and the use of folding modulators (Kolaj et al. 2009). Alternatively, the protein can be extracted from inclusion bodies either under native or denaturing conditions (Martinez-Alonso et al. 2009). Finally, the target protein can be engineered to achieve soluble expression through fusion of solubility-enhancing tags (Sahdev et al. 2008; Torres et al. 2012).

In the recent years, a novel prokaryotic expression system has been developed based on the use of the psychrophilic bacterium *Pseudoalteromonas haloplanktis* TAC125 by driving the expression of genes of interest by both basal or inducible promoters (Duilio et al. 2004a, b). In addition, as was established for *E. coli* expression system, mutation of genes coding for proteases greatly reduces proteolysis of the recombinant protein (Parrilli et al. 2008). The system has demonstrated to be especially useful in improving protein solubility in relation to the widely used *E. coli* expression system and gives higher protein yield for secreted proteins (Cusano et al. 2006; Giuliani et al. 2014; Vigentini et al. 2006). Proteomics analysis points out the trigger factor chaperone as the main factor involved in protein folding during recombinant protein expression under cold temperatures (Piette et al. 2010).

The human  $\alpha$ -galactosidase A (EC 3.2.1.22;  $\alpha$ -Gal A or GLA) is the lysosomal exoglycosidase responsible for the hydrolysis of terminal  $\alpha$ -galactosyl moieties from various glycoconjugates. In the late 1980s, the full-length complementary DNA (cDNA) and the entire genomic sequence encoding mammalian GLA were isolated and characterized (Calhoun et al. 1985; Gotlib et al. 1996). The 1.4-kb full-length cDNA encodes a peptide of 429 residues, which includes a 31-residue amino-terminal signal peptide. This enzyme precursor is a glycopeptide (of approx. 55 kDa) that is processed by cleavage of the signal peptide and by oligosaccharide modifications in the Golgi and lysosomes to form the mature, active, homodimeric enzyme (approx. 100 kDa). After cleavage of the signal peptide, the glycopeptide undergoes modification of its N-linked oligosaccharide moieties in the Golgi apparatus, and then it is transported to the lysosome via the mannose-6-phosphate receptor (M6PR)-mediated pathway. Mutations in the hGLA gene resulting in deficient or absent enzymatic activity are the basis of Fabry's disease, a disorder characterized by progressive glycosphingolipid deposition in vascular lysosomes leading to early demise from renal, cardiac or cerebral vascular disease (Garman and Garboczi 2002).

The subunits of the glycosylated mature enzyme contain four putative N-glycosylation consensus sites at positions

Asn139, Asn192, Asn215 and Asn408 and five disulphide bonds between residues Cys52–Cys94, Cys56–Cys63, Cys142–Cys172, Cys202–Cys223 and Cys378–Cys382 (Garman and Garboczi 2004; Saito et al. 2013). The human enzyme has been purified from a variety of sources, and its physical and kinetic properties have been also characterized (Chen et al. 2000a, b; Corchero et al. 2011, 2012; Yasuda et al. 2004). However, the *hGLA* obtained from prokaryotic expression systems has neither been fully characterized nor purified (Hantzopoulos and Calhoun 1987). Therefore, the objective of this work is the production and purification of active *hGLA* from prokaryotic cell factories.

## Material and methods

### Bacterial strains and plasmids

The *E. coli* strains used in this work were DH5 $\alpha$  for plasmid cloning and maintenance; BL21(DE3), *Rosetta* 2(DE3) and *Rosetta-gami* B(DE3) (Novagen) for recombinant protein expression and S17-1( *$\lambda$ pir*) were used as donor in interspecific conjugation experiments (Tascon et al. 1993). *P. haloplanktis* TAC125 is a Gram-negative bacterium isolated from Antarctic seawater and is deposited and available at the Institut Pasteur Collection (CIP 108707) (Medigue et al.

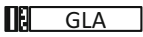

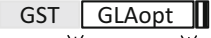

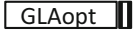
2005). Antarctic bacteria transformation was achieved by intergeneric conjugation as previously reported (Duilio et al. 2004b).

The four *E. coli* expression vectors used in this work (Fig. 1a) are as follows: (a) pReceiver-B01-GLA (product EX-Q0172-B01, OmicsLink ORF Expression Clone), encoding a full-length version of *hGLA* under the control of the T7 promoter; in this vector, 6 $\times$ His-tag is fused to the N-terminus of the *hGLA* open reading frame (ORF) for further detection and purification purposes; (b) pGEX4T2-GLA, encoding the mature form of the *hGLA* with a glutathione S-transferase (GST) fused at its N-terminus; a His-tag and a Tobacco etch Virus protease (TEVp) cleavage site were also added in the designed PCR oligonucleotides of the *hGLA* gene for purification purposes; (c) pGEX4T2-Opt-GLA, synthetic, codon optimized *hGLA* gene (GeneArt, Invitrogen; KF500099) cloned into pGEX4T2 as previously described; and (d) pGEX4T2-GLA-GFP, derived from pGEX4T2-GLA in which the green fluorescent protein (GFP) coding sequence was inserted into the *Eco*RI cloning site. Some other four expression vectors used in *E. coli* strains are described in Suppl. Fig. S1a.

Finally, the psychrophilic gene expression vector pPM13psDs-GLA (e) was constructed as follows: the pPM13psDs shuttle vector, containing the strong constitutive promoter PM13 (Duilio et al. 2004a) correctly located with

**Fig. 1 a.** Expression vectors used in recombinant *hGLA* production assays. Expression vectors used in *E. coli* strains: **a–d** Expression vector used in

*Pseudoalteromonas haloplanktis* is shown in **e**. *GST*, glutathione S-transferase; *GLA*, *hGLA* gene; *GLAopt*, *hGLA* gene codon optimized for *E. coli* in **c** and gene codon optimized for *Pseudoalteromonas haloplanktis* in **e**. *Tobacco Etch Virus* protease cleavage site (TEVp) location is marked by an *arrow tip*. His-tag is marked by a *striped box*. Signal peptide of *hGLA* in **a** is marked by a *stippled box*. **b** Amino acid sequence of *hGLA*. The first 31 amino acids correspond to the peptide signal that is marked in *italics*. The residues coded by rare codons are shown in *bold*. Putative translation termination sequence is *underlined*

a	Recombinant GLA	Cell compartment	Size	Molecular weight
a) pReceiver-B01-GLA		cytoplasm	435 aa	49.6 kDa
b) pGEX4T2-GLA		cytoplasm	640 aa	73.6 kDa
c) pGEX4T2-Opt-GLA		cytoplasm	640 aa	73.6 kDa
d) pGEX4T2-GLA-GFP		cytoplasm	876 aa	100.1 kDa
e) pPM13psDs-GLA		periplasm	405 aa	46.3 kDa

### b

```

1.  MQLRNPELHL  GCALALRFLA  LVSWDIPGAR
31.  ALDNGLARTP  TMGWLHWERF  MCNLDCQEEP
61.  DSCISEKLFM  EMAELMVSEG  WKDAGYEYLC
91.  IDDCWMAPQR  DSEGRLOADP  QRFPHGIRQL
121. ANYVHSGKGLK  LGIYADVGNK  TCAGFPGSFG
151. YYDIDAQTFA  DWGVDLLKFD  GCYCDSLENL
181. ADGYKHMSLA  LNRTGRSIVY  SCEWPLYMWP
211. FQKPNYTEIR  QYCNHWRNFA  DIDDSWSKSIK
241. SILDWTSEFNQ  ERIVDVAGPG  GWNDPDLVI
271. GNFGLSWNQQ  VTQMALWAIM  AAPLFMSNDL
301. RHISPOAKAL  LQDKDVIAIN  QDPLGKQGYQ
331. LRQGDNFEVW  ERPLSGLAWA  VAMINRQEIG
361. GPRSytiAVA  SLGKGVACNP  ACFITOLLPV
391. KRKLGfyEWT  SRLRSHINPT  GTVLLQLENT
421. MQMSLKDLL

```

respect to a psychrophilic Shine-Dalgarno sequence and the 93-bp encoding the signal peptide of *P. haloplanktis* TAC125 DsbA was double digested by *Sall-EcoRI*. The synthetic codon optimized *hGLA* gene (KF500099) was double digested by *Sall-EcoRI* and cloned into pPM13psDs corresponding sites (Fig. 1a).

#### Recombinant human GLA production and cellular fractionation

The production process in recombinant *E. coli* cells was performed by the Protein Production Platform (<http://ibb.uab.es/ibb>) of the Institute of Biotechnology and Biomedicine (IBB-UAB) and the Biomedical Networking Center (CIBER-BBN). In short, shake flask cultures were set at 37 °C and 250 rpm in LB-rich medium, plus 100 µg/mL of ampicillin for plasmid maintenance in all expression *E. coli* vectors. Expression of recombinant *hGLA* gene was induced when the absorbance at 550 nm reached values around 0.5, adding isopropyl β-D-1-thiogalactopyranoside (IPTG) at concentrations ranging from 0.1 to 1 mM. Incubation proceeded then at different temperatures of 16, 20, 25 and 37 °C depending on the strain as indicated. After induction of gene expression, 1.5-mL culture samples were withdrawn at different times and soluble and insoluble fractions were separated. In short, cells were harvested by centrifugation at 5000g (at 4 °C) for 15 min and resuspended in 200 µL of phosphate-buffered saline (PBS) (7.5 mM Na<sub>2</sub>HPO<sub>4</sub>, 110 mM NaCl, 2.5 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.4) and further sonicated. Soluble and insoluble fractions were separated by centrifugation at 15,000g for 15 min at 4 °C. Insoluble fractions were resuspended in PBS (same volumes than those from their respective soluble fractions). Both fractions were stored at –80 °C until further analysis.

*P. haloplanktis* TAC125 (pPM13psDs-GLA) recombinant cells were cultured in aerobic conditions at 4 and 15 °C in TYP broth (16 g/L yeast extract, 16 g/L bacto tryptone, 10 g/L marine salt mix) at pH 7.5, supplemented with 100 µg/mL ampicillin, and the recombinant *hGLA* production and its cellular localization were evaluated in cell samples withdrawn at different growth phases. Cell pellets (corresponding to OD<sub>600 nm</sub>=10) were resuspended in 1 mL of 50 mM Tris–HCl at pH 8.0, 50 mM EDTA and disrupted by ultrasonic treatment consisting of six cycles of 30 s on/1 min off on ice. The mixture was centrifuged for 15 min at 10,000g at 4 °C; the resulting supernatant and pellet were collected as soluble and insoluble fractions, respectively. Total *P. haloplanktis* TAC125 (pPM13psDs-GLA) cell extract was obtained by resuspending 0.5 OD<sub>600 nm</sub> cell pellet into 400 µL of sodium dodecyl sulphate (SDS) loading dye, followed by 30 min treatment at 90 °C.

The periplasmic cell extract was obtained by resuspending bacterial pellets in 1/20 of culture volume of borate buffer

(200 mM Na<sub>2</sub>B<sub>4</sub>O<sub>7</sub>, 130 mM NaCl, 5 mM EDTA, pH 8) and incubating the mixture for 18 h at 4 °C. The suspension was then centrifuged at 8000 rpm for 15 min at 4 °C. The supernatant was stored as periplasmic extract, while the pellet was resuspended in 1/20 of original culture volume of SDS loading dye, followed by 30 min of treatment at 90 °C and stored as cytoplasmic extract.

#### Recombinant protein detection

Total cell extracts, soluble, insoluble and periplasmic cell fractions were analysed by SDS-polyacrylamide gel electrophoresis (PAGE) and Coomassie staining. For Western blot studies, upon SDS-PAGE proteins were blotted onto nitrocellulose membranes, *hGLA* immunoreactive bands were developed using either a rabbit polyclonal anti-GLA serum from Santa Cruz Biotechnology, Inc. (α-gal A H-104: sc-25823) raised against an epitope corresponding to amino acids 316–429 at the C-terminus of the mature form of *hGLA* (Fig. 1b), a rabbit polyclonal anti-GLA serum from Sigma Prestige Antibodies (product HPA000237) raised against amino acids 302–412 of the mature form of *hGLA*, or a monoclonal antibody anti-His-tag from GE Healthcare (product 27-471001), and the corresponding secondary antibodies. GST was detected with mouse monoclonal anti-GST antibody (Santa Cruz Biotechnology, Inc.; SC-138) and GFP with rabbit polyclonal anti-GFP antibodies (Santa Cruz Biotechnology, Inc.; SC-8334).

For comparison purposes, gels were loaded with sample volumes adjusted according to the OD of the culture. Samples for quantitative comparison were run in the same gel and processed as a set. Bands were analysed with the Quantity One analysis software (BioRad).

#### Protein purification

GST-containing *hGLA* recombinant proteins were purified from Rosetta-gami B cell cultures induced at 0.1 mM IPTG for 16 h at 20 °C at 250 rpm. Cell pellets were resuspended in PBS buffer (7.5 mM Na<sub>2</sub>HPO<sub>4</sub>, 110 mM NaCl, 2.5 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.4) in the presence of EDTA-free protease inhibitors (Roche Applied Science), and cells were lysed using a French Press (Thermo FA-078A) at 1100 psi. The soluble fraction was separated by centrifugation at 15,000g for 45 min at 4 °C, filtered through 0.22-µm filters and loaded onto a 1-mL GSTrap HP column (GE Healthcare, 17-5281-05). Bound GST-GLA protein was eluted with five column volumes of elution buffer (50 mM Tris–HCl and 10 mM reduced glutathione, pH 8.0). Positive fractions were collected and dialysed against 0.01 M acetic buffer (pH 4.5), and protein concentration was estimated using the Bradford method.

The pellets of *E. coli* or *P. haloplanktis* TAC125 cells containing His-tagged *hGLA* proteins were resuspended in

buffer A (20 mM Tris–HCl pH 7.5, 500 mM NaCl and 10 mM imidazole) and cells disrupted by sonication in the presence of EDTA-free protease inhibitors (Complete, 11873580001 from Roche). The soluble cell fraction was separated by centrifugation at 15,000g for 15 min at 4 °C. After filtration through 0.22- $\mu$ m filters, recombinant proteins were purified by affinity chromatography in Ni<sup>2+</sup> columns (HiTrap Chelating HP columns, 17-0408-01 from GE Healthcare) in an ÄKTA™ Purifier (GE healthcare) fast protein liquid chromatography system. Positive fractions in elution buffer (20 mM Tris–HCl pH 7.5, 150 mM NaCl and 500 mM Imidazole) were collected and dialyzed against 0.01 M acetic buffer (pH 4.5). Protein concentration was estimated using the Bradford method.

Purified proteins were characterized by N-terminal sequencing by Edman's automated degradation using in an Applied Biosystems Procise 492 protein sequencer, and molecular weight was experimentally determined by mass spectrometry (UltraFlex MALDI-TOF mass spectrometer, Bruker Daltonics, Bremen, Germany). Both analyses were performed at the Proteomics and Bioinformatics Unit of the Scientific Technical Service, SepBioEs, of the Autonomous University of Barcelona.

#### Time course of TEV cleavage reaction of GST-GLA-GFP

Purified GST-GLA-GFP was dialysed against reaction buffer (50 mM Tris–HCl, pH 8.0, 0.5 mM EDTA). Ten units of TEVp (Invitrogen, 12575-015) was added to 20  $\mu$ g of GST-GLA-GFP and incubated at 30 °C; 30- $\mu$ L aliquots were removed at specified incubation times. Reaction was stopped by adding 30  $\mu$ L of stop solution (125 mM Tris–HCl, pH 6.8; 4 % SDS; 1.4 M  $\beta$ -mercaptoethanol; 20 % (v/v) glycerol; 0.01 % bromophenol blue), and samples were stored at –20 °C until gel electrophoresis analysis. Protein samples when performing *hGLA* activity assays were used without buffer addition after protease incubation.

#### Enzyme assay of $\alpha$ -galactosidase A activity

The enzymatic *hGLA* activity was assayed with 4-methylumbelliferyl  $\alpha$ -D-galactoside (4MU- $\alpha$ -Gal, ref. M7633, Sigma Chemical, St. Louis, MO) as a substrate, at a concentration of 2.46 mM in 0.01 M acetic buffer (pH 4.5). A typical 4MU- $\alpha$ -Gal assay was performed with a reaction mixture containing 100  $\mu$ L of substrate and 25  $\mu$ L of enzyme solution. Enzymatic reactions took place in agitation, at 37 °C for 1 h. Reaction was stopped adding 1.25 mL of 0.2 M glycine–NaOH buffer (pH 10.4). The released 4-methylumbelliferone (4-MU) was determined by fluorescence measurement at 365 and 450 nm as excitation and emission wavelengths, respectively. Samples ranging from 5 to 500 ng/mL of 4-MU (ref. M1381, Sigma, St. Louis, MO) in 0.2 M glycine–NaOH buffer (pH 10.4) were used as standard curve.

One unit of enzyme activity was defined as the amount of enzyme releasing 1 nmol of 4-MU per milligram of enzyme and per hour. *hGLA* obtained from human fibroblasts or HEK-293 T cells were used as positive controls (Corchero et al. 2011).

## Results

### Recombinant expression of human *GLA* ORF in *E. coli*

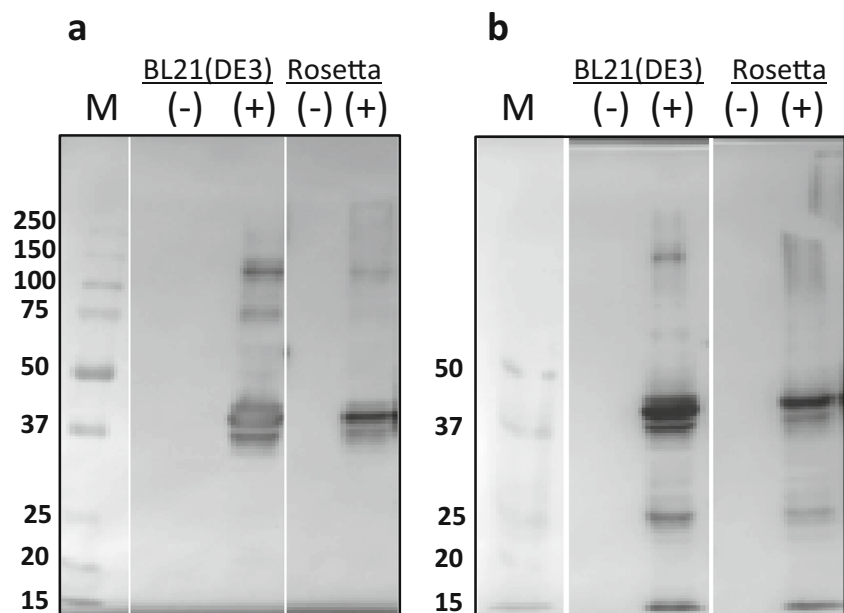
Plasmid pReceiver-B01-GLA, which contains the complete ORF of *hGLA* gene, was transformed in *E. coli* BL21(DE3) and Rosetta(DE3) *E. coli* strains (Fig. 1a). The Rosetta strain is able to compensate low concentrations of certain tRNAs of *E. coli* and to improve recombinant protein translation. It should be noticed that in the *hGLA* ORF, at least four critical arginine codons are present in the DNA sequence (Fig. 1b). Strains BL21(DE3) and Rosetta(DE3), with and without the expression plasmid, were submitted to gene expression induction conditions and analysed by SDS-PAGE followed by Western blot. Nitrocellulose membranes were developed either with a polyclonal anti-GLA antibody (Sigma-Aldrich Co.) raised against an epitope corresponding to amino acids 302–412 of the *hGLA* protein or with the monoclonal anti-His-tag (Fig. 2a, b). Results showed that the polyclonal antibody from Sigma and the anti-His antibody detect a main protein band corresponding to the degradation product of about 40–42 kDa (expected MW 49.6 kDa), both antibodies detecting this band as the main immunoreactive product, while samples from not transformed cells did not show any immunoreactive band. Also, both antibodies detected bands that might be associated to protein aggregates or oligomers, while proteolysis fragments were detected with the anti-His-tag monoclonal antibody but not with the polyclonal antibody, indicating a specific C-terminal cleavage. Finally, equivalent protein bands were detected upon tRNA supplementation. According to these results, only a product of approximately 42 kDa and probably other aggregated and degraded protein products, recognized by the monoclonal anti-His-tag and the anti-GLA antibodies, could be obtained in these strains transformed with the pReceiver-B01-GLA expression vector.

In addition, cell fractionation studies performed in *E. coli* BL21(DE3)/pReceiver-B01-GLA cultures showed that the *hGLA* produced at different induction temperatures was exclusively detected in the insoluble cell fraction (Suppl. Fig. S2).

### Recombinant production of mature form of *hGLA* in *E. coli*

*hGLA* ORF present in pReceiver-B01-GLA codes for the 429 amino acid *hGLA* precursor including 31 amino acid residues corresponding to the signal peptide, which is naturally processed in the ER to a 398-amino acids

**Fig. 2** Crude extracts of induced *E. coli* strains BL21(DE3) and Rosetta(DE3), not transformed (*minus sign*) or transformed with pReceiver-B01-GLA (*plus sign*). Samples were collected 3 h postinduction at 30 °C, and equivalent amounts of protein sample were loaded in each lane. **a** Western blot developed with polyclonal anti-GLA from Sigma. **b** Western blot developed with monoclonal anti-His-tag. Molecular weight marker standards are indicated in kilodaltons (Dual color, BioRad)



mature form (Lemansky et al. 1987) (Fig. 1b). Since eukaryotic signal peptides are not processed in prokaryotic hosts, nucleotides coding exclusively for the mature form of the *hGLA* ORF (32-429; Fig. 1b) were amplified by PCR and transferred to *E. coli* expression vectors to improve protein folding in the heterologous expression system (Suppl. Fig. S1a, expression vectors a and b). In addition, to promote the formation of disulphide bonds in the *hGLA* product, the experiments were performed in either Rosetta-gami *E. coli* strain or the protein was sent to the oxidizing environment present in the periplasm. Expression of the mature form of *hGLA* in *E. coli* under these conditions produced aggregated forms of the recombinant protein and protein bands compatible with degradation products (Suppl. Fig. S3 and Suppl. Fig. S4).

It has been reported that purification of full-length human proteins can be achieved by fusion to soluble proteins as GST in *E. coli* (Braun et al. 2002). This DNA fragment was also inserted into pGEX4T-2, which adds a GST fusion protein at the N-terminal end of the desired protein product (Fig. 1a).

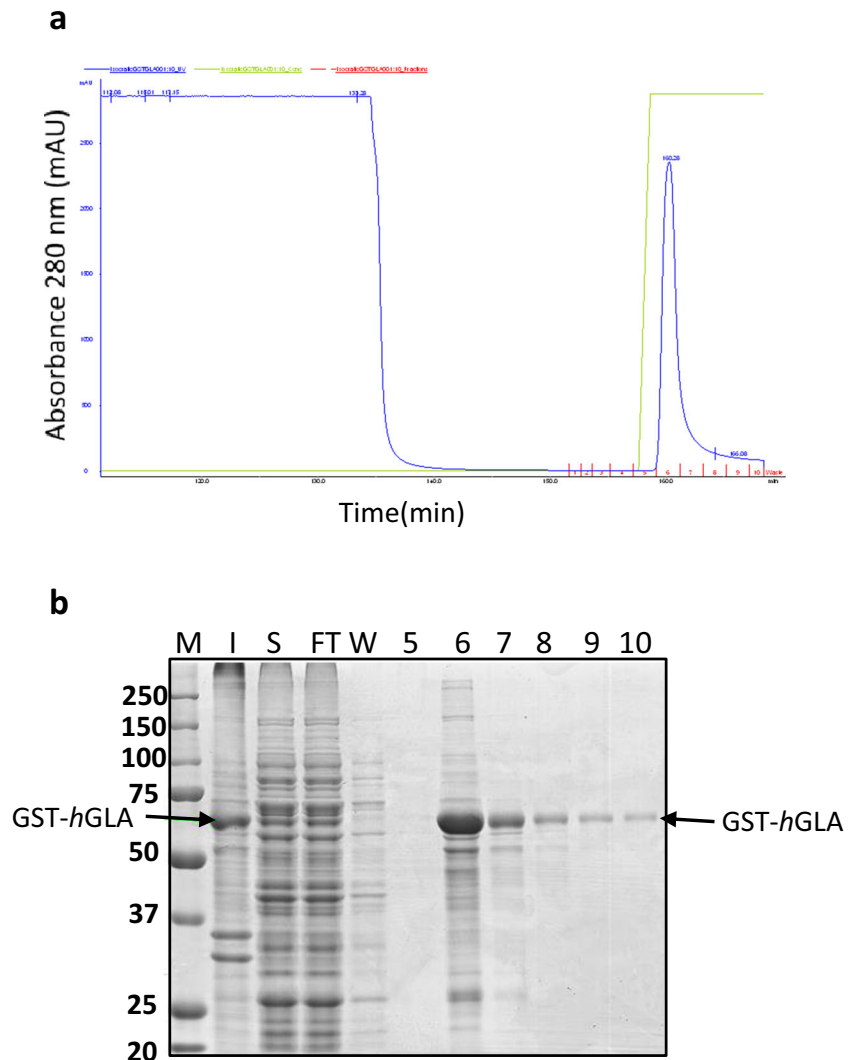
Protein expression experiments were carried out with transformed Rosetta-gami B(DE3) cells with pGEX4T2-GLA with the aim of using the fused GST for solubilization and purification purposes. Several protein induction conditions were assayed (Suppl. Fig. S5). Specific protein bands were detected in all tested conditions in both soluble and insoluble fractions. A predominant band was observed in the insoluble fraction with an apparent weight below 75 kDa that would correspond to the expected GST tagged recombinant *hGLA*, apart from other protein bands compatible with aggregates and proteolysed forms.

#### Purification and characterization of the produced human GLA

Soluble GST-fused protein was purified by affinity chromatography using a GSTrap HP 1 mL column (GE Healthcare). Elution peak contains highly pure protein sample obtained in one single chromatographic step, indicating that the chosen purification strategy was appropriate (Fig. 3). However, molecular weight of the purified protein was again smaller than expected. N-terminal sequencing of the protein band demonstrated that corresponded to the initial amino acids of the recombinant protein, although the first methionine was only partially removed. In *E. coli*, methionine aminopeptidase removes initial methionine with variable efficiency depending on the penultimate amino acid, and it has been demonstrated that the presence of an acidic residue as L present in *hGLA* reduces its activity up to 20 % (Liao et al. 2004; Xiao et al. 2010a). Moreover, mass spectrometry of the affinity purified sample revealed that the exact molecular weight of the product was 68.9 kDa (theoretical weight 73.6 kDa). This size reduction in 4.7 kDa is compatible with the elimination of 29 amino acids at the C-terminus of the protein since as it has been demonstrated, the N-terminal end remains intact. Therefore, in this proteolysed *hGLA* protein, the last amino acid corresponded then to S401 after which an arginine coded by a rare codon has to be added at position 402. In addition, at position 392, another arginine rare codon is present (Fig. 1b).

On the basis of the above-mentioned results, and to avoid the presence of rare codons close to the proteolytic point of the polypeptide chain, site-directed mutagenesis of the mentioned arginine residues in the *hGLA* (R402 and R392 rare codons) were performed using the QuikChange Lightning site-directed mutagenesis kit (Stratagene) with oligonucleotides designed by PrimerX ([www.Bioinformatics.org](http://www.Bioinformatics.org)). The resulting mutant

**Fig. 3 a** Chromatogram of recombinant GST-GLA protein on GStrap FF 1 mL column obtained from soluble cellular fraction of Rosetta-gami B(DE3) induced cell culture. **b** Coomassie blue stained SDS-PAGE of soluble, insoluble cell fractions and protein samples of affinity chromatography. Flow through (FT), wash (W), fractions 5–10 of elution peak. Molecular weight marker standards are indicated in kilodaltons (Dual color, BioRad)



pGEX4T2-GLA (pGEX4T2-GLA<sub>mut1</sub>; Suppl. Fig. S1a) was sequenced and used in expression experiments. Western blot analyses of the resulting protein samples showed that a GST-*h*GLA protein could be obtained, but its molecular weight corresponded to the 69-kDa truncated protein previously obtained (data not shown).

Further analysis of the RNA sequence of the *h*GLA revealed the presence of a putative *E. coli* transcriptional terminator motif surrounding the nucleotide position at which translation seems to finish (Nudler and Gottesman 2002) (Suppl. Fig. S1b). Oligonucleotides were designed to eliminate hairpin loop secondary structures in that region and were used to mutate the pGEX4T2-GLA DNA sequence using the QuikChange Lightning site-directed mutagenesis kit (Stratagene) generating pGEX4T2-GLA<sub>mut2</sub> (Suppl. Fig. S1a). Protein expression experiments showed that under these conditions, the resulting recombinant protein was still

truncated, showing an apparent molecular weight of 69 kDa (data not shown). Protein aggregation was observed in protein samples stored at 4 °C for 2 weeks and confirmed in Superdex 200 10/300 GL (GE Healthcare) gel filtration chromatography since a protein peak was obtained in the void volume of the column. Activity assays of purified GLA protein fused to GST were performed using purified human GLA from a mammalian expression system as positive control (Corchero et al. 2012) as well as cellular extracts of *E. coli* cells transformed with pReceiver-B01-GLA and pET22b-GLA but enzymatic activity was not detected in any sample (Table 1).

A global codon optimisation of *h*GLA gene was attempted next. For this purpose, a synthetic GLA gene was designed and constructed (GeneArt; Invitrogen) which was cloned within the pGEX4T2 vector to produce a pGEX4T2-Opt-GLA expression vector with an optimized codon usage

**Table 1** Enzymatic activity of purified *h*GLA obtained from different cell origins

Sample	Origin	Enzymatic activity ( $\mu\text{mol/h/mg GLA}$ )
Replagal	Human fibroblasts	1945.8 $\pm$ 24.2
GLA-HEK	Human embryonic kidney	1398.6 $\pm$ 47.7
GLA- <i>E. coli</i>	<i>E. coli</i>	ND
GLA- <i>P. haloplanktis</i>	<i>P. haloplanktis</i>	77.4 $\pm$ 0.3

ND not detected

(Fig. 1a, expression vector c). In that construct, the His-tag purification module was fused at the C-terminus to differentiate from the previous GST-*h*GLA constructs in which the His-tag module was located at the N-terminus of *h*GLA. The resulting expression vector was used to transform the Rosetta-gami B(DE3) strain. Protein expression experiments of the resulting strain indicated that once again, a truncated protein was produced and hence, production of the intact *h*GLA does not depend on optimized codon usage (Fig. 4a). In addition, the detection of the GST-GLA recombinant protein with anti-GST antibodies demonstrated the presence of GST in the final protein product (Fig. 4b). However, the same samples were not detected by anti-His antibodies, indicating that truncation of GLA protein occurred at the C-terminus (Fig. 4c). On the other hand, GST-GLA obtained from the expression of the non-optimized gene contained in pGEX4T2-GLA was detected by both anti-GST and anti-His antibodies since in this construction, His-tag peptide was located N-

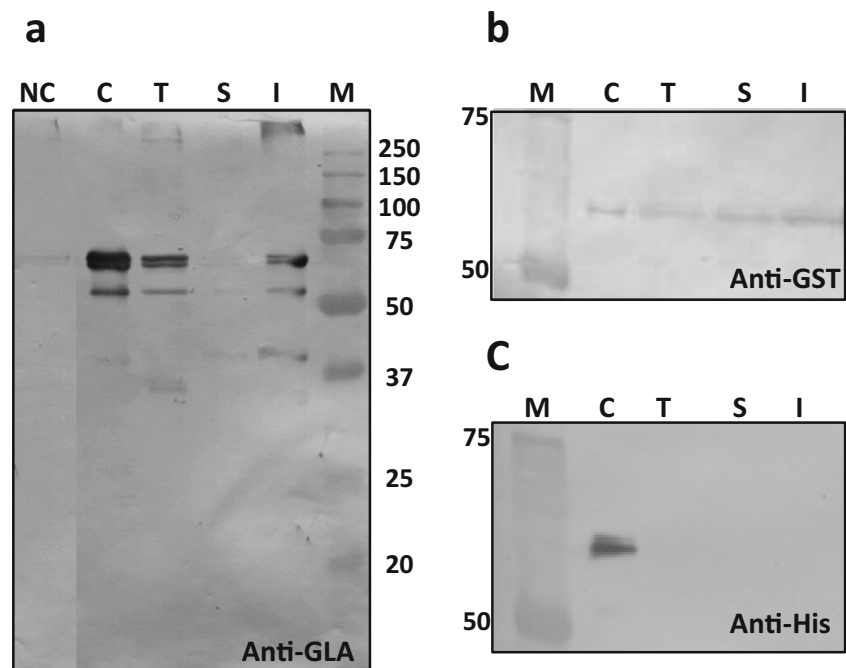
terminally between the GST and GLA coding sequences (Fig. 1a).

Since an incomplete product was obtained regardless of the strategies attempted so far, a protease cleavage analysis was carried out using the ExPaSy PeptideCutter program. However, no specific protease cleavage site was found at the C-terminal end of the truncated GLA product obtained. Nevertheless, to confirm these findings, the next attempted strategy was the fusion of another protein (GFP (Hsieh et al. 2010)) at the C-terminus of the GST-GLA with the aim to protect GLA from a possible proteolysis at this end (Murby et al. 1991). pGEX4T2-GLA-GFP (Fig. 1a, expression vector d) was obtained and transformed into Rosetta-gami B(DE3). Protein expression assays using anti-GLA (Fig. 5a) and specifically anti-GFP Western blot experiments (Fig. 5b) demonstrated that the whole fused protein GST-His-GLA-GFP could be produced by this strain in contrast to previous experiments. Purification and subsequent TEV proteolysis of the fused protein (Fig. 5c) did not succeed in producing an active *h*GLA product despite the strategies attempted (GST and His-tag affinity chromatography), and only background signal was detected when determining the enzymatic activity of the obtained product corresponding to the signal background of the negative control (data not shown).

Recombinant production of mature form of *h*GLA in *P. haloplanktis* TAC125

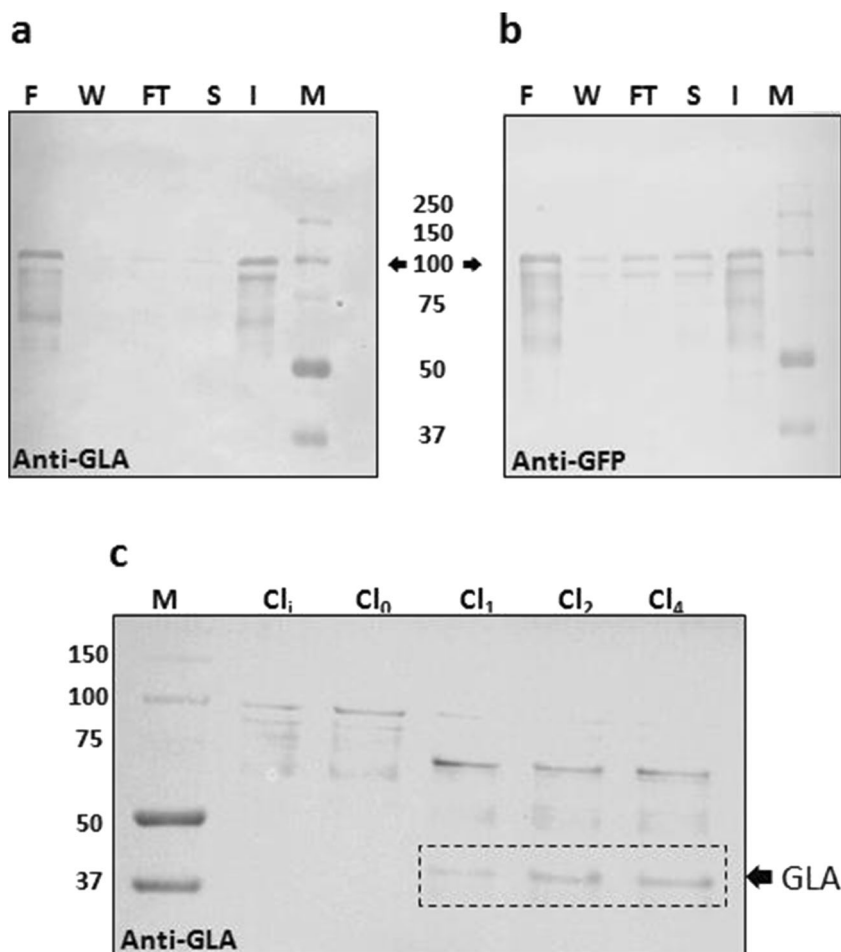
Recombinant *P. haloplanktis* (pPM13psDs-GLA) cells were aerobically grown in complex rich medium at two different

**Fig. 4** Detection of GST-GLA protein in Rosetta-gami B(DE3)/pGEX4T2-Opt-GLA induced cell cultures by Western blot analysis using **a** polyclonal anti-GLA antibody (Sigma), **b** polyclonal anti-GST, **c** monoclonal anti-His. NC, non-induced cell cultures; C, purified GST-GLA protein; T, total cellular fraction; S, soluble fraction; I, insoluble fraction. Molecular weight marker standards are indicated in kilodaltons (Dual color, BioRad)





**Fig. 5** Detection of GST-GLA-GFP protein in Rosetta-gami B(DE3)/pGEX4T2-GLA-GFP induced cell cultures by Western blot analysis. **a** polyclonal anti-GLA antibody (Sigma). **b** Polyclonal anti-GFP antibody. *F*, pooled positive protein fractions obtained in GST affinity chromatography; *W*, wash fraction; *FT*, flow through; *S*, soluble cell fraction; *I*, insoluble cell fraction. **c** Release of *h*GLA protein from purified GST-GLA-GFP recombinant protein by Tobacco Etch Protease cleavage. *Cl<sub>i</sub>*, initial protein sample after protein purification; *Cl<sub>0</sub>*, protein sample at time 0; *Cl<sub>1</sub>*–*Cl<sub>4</sub>*, protein sample at 1–4 h. Molecular weight marker standards are indicated in kilodaltons (Dual color, BioRad)



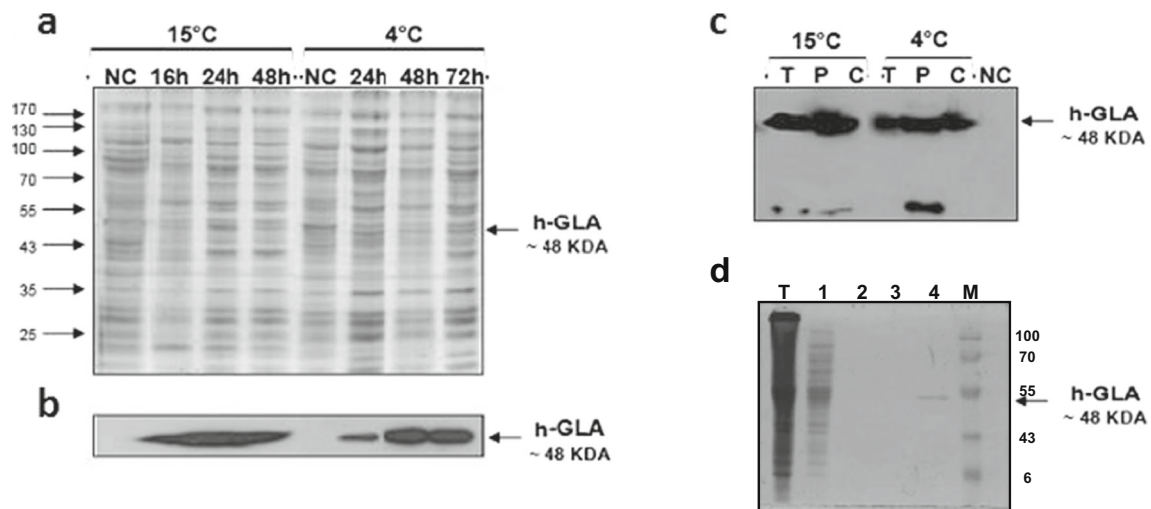
temperatures (4 and 15 °C), and samples were withdrawn at different times of cultivation. The *h*GLA production was evaluated by SDS-PAGE followed by Western blotting analysis (using specific anti-GLA polyclonal antibodies) of total soluble protein extracts, and it revealed the highest protein yield after 24 h of cultivation at 15 °C (Fig. 6a, b). No recombinant GLA protein was detected in insoluble protein extracts (data not shown). A cell fractioning followed by Western blotting analysis demonstrated that recombinant GLA produced at higher temperature is also fully localized in the bacterial periplasm (Fig. 6c), while when produced at 4 °C, a fraction of the protein was still associated to the cytoplasm, likely deriving from a less efficient translocation across bacterial inner membrane. Furthermore, at the lower temperature, the recombinant *h*GLA translocated in the periplasm was also subjected to host encoded proteolytic cleavage, as demonstrated by the appearance of a lower molecular weight specific band on Western blotting analysis shown in Fig. 6c.

Soluble protein extracts from *P. haloplanktis*/pPM13psDs-GLA grown at 15 °C for 24 h were used for His-trap affinity purification (His-trap 1 mL column, GE Healthcare)

according to the protocol reported in the methods section. A linear gradient was used for elution (20–500 mM imidazole in 15 min at 0.5 mL/min flow), thus obtaining a partial purified protein (about 70 % purity) which has been used for the determination of enzymatic activity (Fig. 6d). Although the purified enzyme has the tendency to precipitate, it turned out to be active and a value of  $77.4 \pm 0.3 \mu\text{mol h}^{-1} \text{mg}^{-1}$  prot of activity was recorded (Table 1).

## Discussion

Recombinant protein production in heterologous expression systems represents a major bottleneck in the development of biopharmaceutical, industrial and research applications (Liras 2008). Even though high throughput techniques for cloning and protein production have been already described (Barnard et al. 2010; Nettleship et al. 2010; Savitsky et al. 2010; Xiao et al. 2010b), purification of challenging proteins is still a matter of trial-and-error approaches. As expression system, *E. coli* offers many advantages over eukaryotic systems



**Fig. 6** **a** Coomassie blue stained SDS-PAGE of soluble cell fractions of *P. haloplanktis* transformed with pPM13psDs-GLA at different temperatures. **b** Western blot analysis of the same samples using a polyclonal anti-GLA antibody (Santa Cruz Biotechnology). **c** Detection of recombinant *hGLA* in induced *P. haloplanktis* cell cultures. *T*, total cell

extract; *P*, periplasm; *C*, cytoplasm; *NC*, negative control. **d** Coomassie blue stained SDS-PAGE of *hGLA* purification by affinity chromatography. *T*, total cell extract; *1*, unbound protein; *2* and *3*, washes; *4*, eluted protein

referring to rapidity, simplicity and expenditure (Demain and Vaishnav 2009). However, it has some disadvantages which, in some cases, are difficult to override as the lack of many posttranslational modifications and high tendency to form aggregated species (Panda 2003). In the present study, a human full-length protein, GLA, which has been purified from eukaryotic expression systems (Chen et al. 2000a, b; Corchero et al. 2012; Yasuda et al. 2004) and non-prokaryotic source (Hantzopoulos and Calhoun 1987), has been selected to try available expression methods in *E. coli* and the Antarctica psychrophilic bacterium *P. haloplanktis* TAC125. In 1987, Hantzopoulos and coauthors described the production of an active *hGLA* in *E. coli* (Hantzopoulos and Calhoun 1987). However, such results have not been reproduced or continued since then, and moreover, no purification protocol has been published to date. The interest in developing a production and purification system of human GLA from a prokaryotic source relies on the high cost of the actual enzymatic replacement therapy in Fabry's disease treatment which consists in recombinant protein obtained from mammalian cells having a great economical impact in patient expenditure. We have shown that *hGLA* can be detected in *E. coli* as either aggregated and proteolytically cleaved forms in any experimental condition tested. Aggregation of proteins seems to be driven by the presence of specific amino acid sequences (Espargaro et al. 2012), saturation of the folding machinery (Kolaj et al. 2009) or to the lack of the required posttranslational modifications as glycosylation (Ioannou et al. 1998). However, full-length recombinant *hGLA* has been produced when fused to GST and GFP, although purified protein is not correctly folded lacking biological activity. The absence of

enzymatic activity might be related to either unfolded protein species or deletion of the C-terminal end of the protein which seems to be critical (Miyamura et al. 1996). In addition, GLA has four glycosylation sites added after signal peptide removal in the Golgi apparatus. This sugar moieties (especially glycosylation site located at residue Asn215) play a crucial role in both folding and solubility and consequently in enzymatic activity (Ioannou et al. 1998). However, when producing *hGLA* in the psychrophilic expression system based in the use of the Antarctic bacteria *P. haloplanktis* TAC125, the functional full-length enzyme can be produced and purified. It has been described in comparative studies that recombinant protein production at low temperature has a positive effect in protein yield (Dragosits et al. 2011; Duilio et al. 2004a) as well as in protein conformation (Vigentini et al. 2006). The ability of the *P. haloplanktis* TAC125 to produce active *hGLA* over the proteolytic and aggregation tendency of the protein in *E. coli* might be related to the presence of high peptidyl prolyl *cis-trans* isomerases genes found in *P. haloplanktis* TAC125 genome and more specifically to the upregulation of the main molecular chaperone, trigger factor (Giuliani et al. 2014; Piette et al. 2010). The folding activity in the psychrophilic expression system can counteract the high tendency of the *hGLA* to adopt unsuitable conformations. On the other hand, in the *E. coli* cellular environment, the nascent *hGLA* polypeptides adopt unstable conformational species that are detected by cellular proteases, mainly at the C-terminus of the protein, leading to the formation of inactive aggregation-prone protein species. Even though the full-length protein can be produced in *E. coli*, the limited chaperone activity in this expression system renders inactive protein. Therefore,

microbial host range might be expanded to psychrophilic expression systems when difficult-to-express proteins such as full-length mammalian proteins are to be produced.

**Acknowledgments** This work was supported by ERANET-IB08-007 project from the European Union and its linked national project EUI2008-03610 to AV. We also appreciate the support from EME2007-08 to NFM from Universitat Autònoma de Barcelona, from Antartide 2010 to MLT and EP, from MIUR Azioni Integrate Italia-Spagna 2010 Prot. IT10LECLM9 to MLT, from MINECO (IT2009-0021) to AV and LT, from AGAUR (2009SGR-108) to AV. AV is also supported by The Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN, Spain), an initiative funded by the VI National R&D&i Plan 2008–2011, Iniciativa Ingenio 2010, Consolider Program, CIBER Actions and financed by the Instituto de Salud Carlos III with assistance from the European Regional Development Fund. PS has received predoctoral fellowship from ISCIII, and AV has been distinguished with an ICREA ACADEMIA award (Catalonia, Spain).

**Conflict of interests** The authors declare no conflict of interests.

## References

- Barnard GC, Kull AR, Sharkey NS, Shaikh SS, Rittenhour AM, Burnina I, Jiang Y, Li F, Lynaugh H, Mitchell T, Nett JH, Nysten A, Potgieter TI, Prinz B, Rios SE, Zha D, Sethuraman N, Stadheim TA, Bobrowicz P (2010) High-throughput screening and selection of yeast cell lines expressing monoclonal antibodies. *J Ind Microbiol Biot* 37:961–971
- Braun P, Hu YH, Shen BH, Halleck A, Koundinya M, Harlow E, LaBaer J (2002) Proteome-scale purification of human proteins from bacteria. *Proc Natl Acad Sci U S A* 99:2654–2659
- Calhoun DH, Bishop DF, Bernstein HS, Quinn M, Hantzopoulos P, Desnick RJ (1985) Fabry disease: isolation of a cDNA clone encoding human alpha-galactosidase A. *Proc Natl Acad Sci U S A* 82:7364–7368
- Chen YS, Jin M, Egborge T, Coppola G, Andre J, Calhoun DH (2000a) Expression and characterization of glycosylated and catalytically active recombinant human alpha-galactosidase A produced in *Pichia pastoris*. *Protein Expr Purif* 20:472–484
- Chen YS, Jin M, Goodrich L, Smith G, Coppola G, Calhoun DH (2000b) Purification and characterization of human alpha-galactosidase A expressed in insect cells using a baculovirus vector. *Protein Expr Purif* 20:228–236
- Corchero JL, Mendoza R, Lorenzo J, Rodriguez-Sureda V, Dominguez C, Vazquez E, Ferrer-Miralles N, Villaverde A (2011) Integrated approach to produce a recombinant, His-tagged human alpha-galactosidase A in mammalian cells. *Biotechnol Prog* 27:1206–1217
- Corchero JL, Mendoza R, Ferrer-Miralles N, Montras A, Martinez LM, Villaverde A (2012) Enzymatic characterization of highly stable human alpha-galactosidase A displayed on magnetic particles. *Biochem Eng J* 67:20–27
- Cusano AM, Parrilli E, Marino G, Tutino ML (2006) A novel genetic system for recombinant protein secretion in the Antarctic *Pseudoalteromonas haloplanktis* TAC125. *Microb Cell Fact* 5:40
- Daly R, Hearn MT (2005) Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J Mol Recog* 18:119–138
- Demain AL, Vaishnav P (2009) Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv* 27:297–306
- Dragosits M, Frascotti G, Bernard-Granger L, Vazquez F, Giuliani M, Baumann K, Rodriguez-Carmona E, Tokkanen J, Parrilli E, Wiebe MG, Kunert R, Maurer M, Gasser B, Sauer M, Branduardi P, Pakula T, Saloheimo M, Penttila M, Ferrer P, Tutino ML, Villaverde A, Porro D, Mattanovich D (2011) Influence of growth temperature on the production of antibody Fab fragments in different microbes: a host comparative analysis. *Biotechnol Prog* 27:38–46
- Duilio A, Madonna S, Tutino ML, Pirozzi M, Sannia G, Marino G (2004a) Promoters from a cold-adapted bacterium: definition of a consensus motif and molecular characterization of UP regulative elements. *Extremophiles* 8:125–132
- Duilio A, Tutino ML, Marino G (2004b) Recombinant protein production in Antarctic Gram-negative bacteria. *Methods Mol Biol (Clifton, N J)* 267:225–237
- Espargaro A, Villar-Pique A, Sabate R, Ventura S (2012) Yeast prions form infectious amyloid inclusion bodies in bacteria. *Microb Cell Factories* 11:89
- Ferreira JP, Overton KW, Wang CL (2013) Tuning gene expression with synthetic upstream open reading frames. *Proc Natl Acad Sci U S A* 110:11284–11289
- Garman SC, Garboczi DN (2002) Structural basis of Fabry disease. *Mol Genet Metab* 77:3–11
- Garman SC, Garboczi DN (2004) The molecular defect leading to Fabry disease: structure of human alpha-galactosidase. *J Mol Biol* 337:319–335
- Giuliani M, Parrilli E, Sannino F, Apuzzo GA, Marino G, Tutino ML (2014) Recombinant production of a single-chain antibody fragment in *Pseudoalteromonas haloplanktis* TAC125. *Appl Microbiol Biotechnol* 98:4887–4895
- Gotlib RW, Bishop DF, Wang AM, Zeidner KM, Ioannou YA, Adler DA, Distche CM, Desnick RJ (1996) The entire genomic sequence and cDNA expression of mouse alpha-galactosidase A. *Biochem Mol Med* 57:139–148
- Hantzopoulos PA, Calhoun DH (1987) Expression of the human Alpha-galactosidase-A in *Escherichia coli*-K-12. *Gene* 57:159–169
- Hsieh JM, Besserer GM, Madej M, Bui HQ, Kwon S, Abramson J (2010) Bridging the gap: a GFP-based strategy for overexpression and purification of membrane proteins with intra and extracellular C-termini. *Protein Sci* 19:868–880
- Inaba K (2009) Disulfide Bond formation system in *Escherichia coli*. *J Biochem* 146:591–597
- Ioannou YA, Zeidner KM, Grace ME, Desnick RJ (1998) Human alpha-galactosidase A: glycosylation site 3 is essential for enzyme solubility. *Biochem J* 332:789–797
- Kolaj O, Spada S, Robin S, Wall JG (2009) Use of folding modulators to improve heterologous protein production in *Escherichia coli*. *Microb Cell Factories* 8:9
- Lemansky P, Bishop DF, Desnick RJ, Hasilik A, Von Figura K (1987) Synthesis and processing of alpha-galactosidase A in human fibroblasts. Evidence for Different Mutations in Fabry Disease. *J Biol Chem* 262:2062–2065
- Liao YD, Jeng JC, Wang CF, Wang SC, Chang ST (2004) Removal of N-terminal methionine from recombinant proteins by engineered *E. coli* methionine aminopeptidase. *Protein Sci* 13:1802–1810
- Liras A (2008) Recombinant proteins in therapeutics: haemophilia treatment as an example. *Int Arch Med* 1:4
- Martinez-Alonso M, Gonzalez-Montalban N, Garcia-Fruitos E, Villaverde A (2009) Learning about protein solubility from bacterial inclusion bodies. *Microb Cell Fact* 8:4
- Medigue C, Krin E, Pascal G, Barbe V, Bernsel A, Bertin PN, Cheung F, Cruveiller S, D'Amico S, Duilio A, Fang G, Feller G, Ho C, Mangenot S, Marino G, Nilsson J, Parrilli E, Rocha EPC, Rouy Z, Sekowska A, Tutino ML, Vallenet D, von Heijne G, Danchin A (2005) Coping with cold: the genome of the versatile marine

- Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. *Genome Res* 15:1325–1335
- Miyamura N, Araki E, Matsuda K, Yoshimura R, Furukawa N, Tsuruzoe K, Shirotani T, Kishikawa H, Yamaguchi K, Shichiri M (1996) A carboxy-terminal truncation of human alpha-galactosidase A in a heterozygous female with Fabry disease and modification of the enzymatic activity by the carboxy-terminal domain—increased, reduced, or absent enzyme activity depending on number of amino acid residues deleted. *J Clin Invest* 98:1809–1817
- Murby M, Cedergren L, Nilsson J, Nygren PA, Hammarberg B, Nilsson B, Enfors SO, Uhlen M (1991) Stabilization of recombinant proteins from proteolytic degradation in *Escherichia coli* using a dual affinity fusion strategy. *Biotechnol Appl Biochem* 14:336–346
- Nakamoto T (2009) Evolution and the universality of the mechanism of initiation of protein synthesis. *Gene* 432:1–6
- Nettleship JE, Assenberg R, Diprose JM, Rahman-Huq N, Owens RJ (2010) Recent advances in the production of proteins in insect and mammalian cells for structural biology. *J Struct Biol* 172:55–65
- Nozach H, Fruchart-Gaillard C, Fenaile F, Beau F, Ramos OHP, Douzi B, Saez NJ, Moutiez M, Servent D, Gondry M, Thai R, Cuniasse P, Vincentelli R, Dive V (2013) High throughput screening identifies disulfide isomerase DsbC as a very efficient partner for recombinant expression of small disulfide-rich proteins in *E. coli*. *Microb Cell Fact* 12:37
- Nudler E, Gottesman ME (2002) Transcription termination and anti-termination in *E. coli*. *Genes Cells* 7:755–768
- Panda AK (2003) Bioprocessing of therapeutic proteins from the inclusion bodies of *Escherichia coli*. *Adv Biochem Eng Biotechnol* 85:43–93
- Parrilli E, De VD, Cirulli C, Tutino ML (2008) Development of an improved *Pseudoalteromonas haloplanktis* TAC125 strain for recombinant protein secretion at low temperature. *Microb Cell Fact* 7:2
- Piette F, D'Amico S, Struvay C, Mazzucchelli G, Renaut J, Tutino ML, Danchin A, Leprince P, Feller G (2010) Proteomics of life at low temperatures: trigger factor is the primary chaperone in the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. *Mol Microbiol* 76:120–132
- Prinz WA, Aslund F, Holmgren A, Beckwith J (1997) The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *J Biol Chem* 272:15661–15667
- Sahdev S, Khattar SK, Saini KS (2008) Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol Cell Biochem* 307:249–264
- Saito S, Ohno K, Sakuraba H (2013) Comparative study of structural changes caused by different substitutions at the same residue on alpha-galactosidase A. *PLoS One* 8:e84267
- Sallach RE, Conticello VP, Chaikof EL (2009) Expression of a recombinant elastin-like protein in *Pichia pastoris*. *Biotechnol Prog* 25:1810–1818
- Savitsky P, Bray J, Cooper CD, Marsden BD, Mahajan P, Burgess-Brown NA, Gileadi O (2010) High-throughput production of human proteins for crystallization: the SGC experience. *J Struct Biol* 172:3–13
- Tascon RI, Rodriguez-Ferri EF, Gutierrez-Martin CB, Rodriguez-Barbosa I, Berche P, Vazquez-Boland JA (1993) Transposon mutagenesis in *Actinobacillus pleuropneumoniae* with a Tn10 Derivative. *J Bacteriol* 175:5717–5722
- Torres LL, Ferreras ER, Cantero A, Hidalgo A, Berenguer J (2012) Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb Cell Fact* 11:105
- Vallejo LF, Rinas U (2004) Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb Cell Fact* 3:11
- Vigentini I, Merico A, Tutino ML, Compagno C, Marino G (2006) Optimization of recombinant human nerve growth factor production in the psychrophilic *Pseudoalteromonas haloplanktis*. *J Biotechnol* 127:141–150
- Wacker M, Feldman MF, Callewaert N, Kowarik M, Clarke BR, Pohl NL, Hernandez M, Vines ED, Valvano MA, Whitfield C, Aebi M (2006) Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc Natl Acad Sci U S A* 103:7088–7093
- Xiao Q, Zhang F, Nacev BA, Liu JO, Pei D (2010a) Protein N-terminal processing: substrate specificity of *Escherichia coli* and human methionine aminopeptidases. *Biochem-US* 49:5588–5599
- Xiao R, Anderson S, Aramini J, Belote R, Buchwald WA, Ciccocanti C, Conover K, Everett JK, Hamilton K, Huang YJ, Janjua H, Jiang M, Kornhaber GJ, Lee DY, Locke JY, Ma LC, Maglaqui M, Mao L, Mitra S, Patel D, Rossi P, Sahdev S, Sharma S, Shastry R, Swapna GV, Tong SN, Wang D, Wang H, Zhao L, Montelione GT, Acton TB (2010b) The high-throughput protein sample production platform of the Northeast Structural Genomics Consortium. *J Struct Biol* 172:21–33
- Yasuda K, Chang HH, Wu HL, Ishii S, Fan JQ (2004) Efficient and rapid purification of recombinant human alpha-galactosidase A by affinity column chromatography. *Protein Expr Purif* 37:499–506