Improvement in the production of the human recombinant enzyme N-acetylgalactosamine-6-sulfatase (rhGALNS) in *Escherichia coli* using synthetic biology approaches

Luis H. Reyes¹, Carolina Cardona¹, Luisa Pimentel¹, Alexander Rodríguez-López⁵,¹,² & Carlos J. Alméciga-Díaz¹

Previously, we demonstrated production of an active recombinant human N-acetylgalactosamine-6-sulfatase (rhGALNS) enzyme in *Escherichia coli* as a potential therapeutic alternative for mucopolysaccharidosis IVA. However, most of the rhGALNS produced was present as protein aggregates. Here, several methods were investigated to improve production and activity of rhGALNS. These methods involved the use of physiologically-regulated promoters and alternatives to improve protein folding including global stress responses (osmotic shock), overexpression of native chaperones, and enhancement of cytoplasmic disulfide bond formation. Increase of rhGALNS activity was obtained when a promoter regulated under $\sigma^s$ was implemented. Additionally, improvements were observed when osmotic shock was applied. Noteworthy, overexpression of chaperones did not have any effect on rhGALNS activity, suggesting that the effect of osmotic shock was probably due to a general stress response and not to the action of an individual chaperone. Finally, it was observed that high concentrations of sucrose in conjunction with the physiological-regulated promoter *proUmod* significantly increased the rhGALNS production and activity. Together, these results describe advances in the current knowledge on the production of human recombinant enzymes in a prokaryotic system such as *E. coli*, and could have a significant impact on the development of enzyme replacement therapies for lysosomal storage diseases.

Mucopolysaccharidosis IVA (MPS IVA, Morquio A disease, OMIM 253000) is a rare autosomal-recessive disease characterized by the deficiency of the lysosomal enzyme N-acetylgalactosamine-6-sulfate sulfatase (GALNS, EC 3.1.6.4) with an estimated incidence of 1:200,000 born alive¹,². This enzyme is indispensable for the degradation of glycosaminoglycans (GAGs) keratan sulfate and chondroitin-6-sulfate. The progressive accumulation of these GAGs within the lysosomes of multiple tissues such as ligaments, connective tissues, bone and cartilage, leads to the classical clinical manifestations of the disease including laxity of joints, skeletal dysplasia, hearing loss, corneal clouding, and pulmonary dysfunction, among others³.

Currently, the leading therapeutic option for MPS IVA is the enzyme replacement therapy (ERT) by using a recombinant enzyme produced in Chinese hamster ovaries (CHO) cells (elosulfase alfa)⁴. Although elosulfase alfa is a therapeutic option for MPS IVA patients, current limitations include⁴: (i) a limited effect on skeletal, corneal, and heart valvular tissues (ii) a short half-life of the enzyme and rapid clearance from the circulation (iii) immunological problems, and (iv) a high cost. An improved ERT with a long circulating enzyme and a bone-targeting enzyme have been proposed⁵, and a recombinant GALNS produced in other sources may potentially help to resolve some of the listed issues⁶.

¹Institute for the Study of Inborn Errors of Metabolism, Faculty of Science, Pontificia Universidad Javeriana, Bogotá, Colombia. ²Chemistry Department, Faculty of Science, Pontificia Universidad Javeriana, Bogotá, Colombia. Correspondence and requests for materials should be addressed to L.H.R. (email: lh.reyes@uniandes.edu.co) or C.J.A.-D. (email: cjalme@ja@veriana.edu.co)

Received: 1 December 2016
Accepted: 20 June 2017
Published online: 19 July 2017
The human GALNS complementary DNA is composed by 1569 bp, encoding a 522 amino acids peptide (herein known as precursor peptide). This protein undergoes several posttranslational modifications by the trafficking through the endoplasmic reticulum, Golgi apparatus and lysosome. These posttranslational modifications include the removal of the signal peptide (first 26 amino acids – 3 kDa), the active-site activation by the formylglycine-generating enzyme, the addition of two N-glycosylations, and the proteolytic processing to obtain a ~58 kDa mature enzyme formed by two polypeptides of 40 kDa and 18 kDa.a,b In the case of the recombinant GALNS produced in *Escherichia coli* and *Pichia pastoris*, it was reported a similar processing to that observed for the enzyme produced in mammalian cells.c,d

*E. coli* has been extensively used as the prokaryotic model organism for production of proteins with therapeutic and industrial interest.e-g The key problem in the production of human recombinant proteins in this bacterial host is related with the lack of post-translational modifications, as glycosylations, and poor protein folding, leading to loss of enzyme activity and the formation of insoluble protein aggregates.h

Previously, we demonstrated the production of a recombinant human GALNS enzyme (rhGALNS) in *E. coli* BL21(DE3) as a potential therapeutic alternative for MPS IVA.i–j However, most of the produced protein was present as protein aggregates. Culture conditions were optimized, including inducer concentrations and temperature shifts, which maximized rhGALNS activity, but most of its production was present in the insoluble protein fraction.k Here we explored different approaches to increase the production and activity of rhGALNS in *E. coli*. These approaches included the use of physiologically-regulated promoters to modulate gene expression, induction of osmoprotectants as helpers in protein folding, the overexpression of chaperone proteins, the improvement in the formation of disulfide bonds, and the combination of these different approaches to explore additive effects.

**Results and Discussion**

**Effect of physiologically-regulated promoters on rhGALNS activity.** There are several factors contributing to the formation of inclusion bodies in prokaryotic systems at the transcriptional level, which are generally attributed to a high transcriptional rate.l and transcription leakage due to poor promoter repression.m The staple prokaryotic promoter used in recombinant protein production in *E. coli* is the lac promoter, which has been widely utilized due to its ability to be straightforwardly controlled with isopropyl-β-D-1-thiogalactopyranoside (IPTG) and tightly repressed via *lacz*.*n Other promoters commonly used in recombinant protein production are thermal or chemically induced.o Previous studies demonstrated the production of active rhGALNS in *E. coli* BL21(DE3) using the tac promoter, a synthetic promoter created from the combination of promoters from the *trp* and *lac* operons, which is inducible by IPTG.p However, most of the rhGALNS was recovered from protein aggregates.q,r,s We hypothesized that the expression of rhGALNS is either controlled poorly at the transcriptional level or it is highly overexpressed, which leads to the accumulation of rhGALNS as protein aggregates.

To facilitate the production process (i.e. avoid the use of inductors) and considering a necessary strong regulation over rhGALNS expression, the promoters *osmY* and *proUmod* were selected for gene expression. These two promoters have been described to be tightly controlled under σs and to induce the gene transcription at the late-exponential and at the onset of stationary growth phases.q–s The promoters *osmY* and *proUmod* were cloned to obtain the plasmids pGEXosmY and pGEXproUmod (Table 1 and Supplementary Data). Subsequently, the plasmids were transformed into *E. coli* BL21(DE3) for rhGALNS production as described in the Materials and Methods section. The cells were grown in M9 minimal media, and *E. coli* BL21(DE3)/pGEX-5X-GALNSopt was used as control.

As shown in Fig. 1, we did not observe a statistically significant increment in rhGALNS activity, in comparison to the levels observed using the strain BL21(DE3)/pGEX-5X-GALNSopt, when gene expression was driven by the *osmY* promoter, either intra- or extracellularly. This promoter has been reported as a weak promoter strongly regulated by σs, in correspondence to the low enzyme activities obtained. On the other hand, rhGALNS activity increased significantly when the *proUmod* promoter was used. The enzymatic activities obtained at the intracellular level (Fig. 1C and D) were 0.08 U/(mg total protein) and 0.16 U/mL, corresponding to increments of 695% and 245%, respectively, in comparison to the levels observed using the strain *E. coli* BL21(DE3)/pGEX-5X-GALNSopt. At the extracellular fraction, the enzyme activities, as shown in Fig. 1A and B, were 0.67 U/(mg total protein) and 0.86 U/mL of rhGALNS, equivalent to increments of 751% and 309%, respectively, in comparison to the levels observed using pGEX-5X-GALNSopt.

It is important to highlight that secretion of rhGALNS is a process mediated by the human native GALNS signal peptide. Extracellular secretion of rhGALNS in *E. coli* BL21(DE3), associated to the presence of a native signal peptide, was previously observed.t,v In the study, by using a similar plasmid construction to the one used in this report (i.e. GALNS with a native signal peptide downstream of the glutathione S-transferase [GST]), it was observed that removal of the signal peptide completely abolished secretion of the recombinant enzyme. Those results suggested that GALNS signal peptide is recognized by *E. coli* secretion machinery, even if it is located downstream of the GST peptide. In addition, removal of signal peptide severely affected the activation process of rhGALNS.

In order to evaluate the presence of rhGALNS in different protein fractions, we performed a Western-blot using an anti-GALNS antibody, produced against the N-terminal region of the enzyme.t As expected, the antibody reacted with the high molecular mass rhGALNS polypeptide (~60 kDa) but it did not detect the small rhGALNS subunit (~18 kDa) (Fig. 1E). Although the sequence encoding for the GST tag was present in the recombinant plasmids and in-frame with GALNS sequence (see plasmids maps in the Supplementary Data), the expected 26 kDa size shift corresponding to this tag was not observed. In this sense, these results might suggest that either the fusion protein GST-rhGALNS was not produced or a post-translational processing of the recombinant enzyme might eliminate the GST tag. The absence of the fusion protein was validated using an affinity chromatography (see Supplementary Data). The results showed that GST tag was produced but it was not fused to recombinant enzyme. Although, further experiments are necessary to elucidate the reasons of the production of a
rhGALNS mRNA, and the protein amount of rhGALNS in the different protein fractions. After the first 6 h of cellular and extracellular total soluble protein, the intracellular and extracellular rhGALNS activity, amounts of σ profile of promoters governed by the elevation of the rhGALNS transcripts. The expression profile is an expected gene expression dynamics intracellular soluble protein (Fig. 2). At this point, the expression of rhGALNS increased significantly, supported by the reduction of rhGALNS gene expression under the control of proUmod promoter. This increment in rhGALNS expression under the promoter proUmod was determined as shown in Fig. 2. Different parameters were monitored in 100 ml culture during the first 24 h of cultivation including the intracellular and extracellular total soluble protein, the intracellular and extracellular rhGALNS activity, amounts of rhGALNS mRNA, and the protein amount of rhGALNS in the different protein fractions. After the first 6 h of cultivation the cells decelerate their growth, entering in stationary phase, as evidenced in the steady amounts of protein lacking the GST tag, the Western-blot analysis, in addition to the enzyme activity results, confirm the production of the rhGALNS in E. coli BL21(DE3). Furthermore, the Western-blot analysis showed a lower rhGALNS production in all the protein fractions studied by using the proUmod promoter, opposite to the production under the lac promoter (Fig. 1E). In addition, the ~60 kDa rhGALNS precursor was observed in extra and intracellular fractions, while the precursor and the processed polypeptides were observed in the protein aggregates. These results might suggest an improvement in protein folding, since rhGALNS activity increases using the promoter proUmod even when the corresponding enzyme amounts decrease. A densitometry analysis of the Western-blot for Fig. 1E was performed using the software ImageJ and allowed us to roughly estimate increments in protein folding of approximately 20X and 15X in the intracellular and extracellular fractions, respectively.

The dynamics of rhGALNS expression under the promoter proUmod was determined as shown in Fig. 2. These results demonstrate the importance of controlling gene expression in the production of rhGALNS, not only in terms of promoter strength, but also regarding transcriptional dynamics and regulation. These results also support the versatility of using physiologically-controlled promoters for protein expression in the sense that it will facilitate the controllability of production processes and reduce the process costs since no inducer is required.

### Induction of osmoprotectants.

Two strategies to improve the amount of soluble recombinant proteins were used in this study, which involve increasing the concentration of osmolytes and the overexpression of chaperones. Overproduction of bacterial chaperones, some of which actively drive folding processes whereas others prevent protein aggregation, can be obtained by different extracellular stresses such as heat-shock or osmotic shock.

We implemented osmotic shock as a method for inducing proper protein folding, exposing the bacterial cultures of E. coli BL21(DE3)/pGEX-5X-GALNSopt high concentrations of sucrose. Two concentrations of

| Plasmid | Description | Selection Marker | Source of reference |
|---------|-------------|------------------|---------------------|
| pGEX-5X-GALNSopt | pGEX-5X-3 plasmid encoding the enzyme N-acetylgalactosamine-6-sulfatase codon-optimized for expression in E. coli | AmpR | This study |
| pGEXosmY | Modification of the pGEX-5X-GALNSopt plasmid where the promoter tac was replaced by the osmY promoter with its corresponding RBS | AmpR | This study |
| pGEXproUmod | Modification of the pGEX-5X-GALNSopt plasmid where the promoter tac was replaced by the proUmod promoter with its corresponding RBS | AmpR | This study |
| pACYCDuet™-1 | Bicistronic plasmid used as a control and to clone all the chaperones used in this study | CmR | Novagen, Merck Milipore |
| pDuet::GroS | Plasmids for the overexpression of the genes groS, groL, groD and groE, dnaK, dnaJ and dnak, dnap, dnap and dnap, dhsB, dhsA, dhsB, dhsA and dhsB, and clpB, and rpsL (respectively) driven by the lac promoter. Backbone pACYCDuet™-1 | CmR | This study |

| Source of reference | Plasmid Description | Selection Marker |
|---------------------|---------------------|------------------|
| pACYCDuet™-1 | Bicistronic plasmid used as a control and to clone all the chaperones used in this study | CmR | Novagen, Merck Milipore |

| Selection Marker | Source of reference |
|------------------|---------------------|
| AmpR | This study |
| CmR | This study |

Table 1. List of plasmids used in this work.
sucrose were tested for this purpose, 5% and 10% (w/v). As shown in Fig. 3, we observed a statistically significant increase in rhGALNS production when the two conditions were tested, although the highest response was observed using a concentration of 5% (w/v) sucrose. The lower effect observed with 10% (w/v) sucrose, in comparison with 5% (w/v) sucrose, agrees with Barth, et al. report, who suggested that an excessive osmotic stress in the absence of cytoplasmic compatible solutes to rescue cell growth from inhibitory conditions (e.g. organic osmolytes as betaine, carnitine, trehalose, proline, mannitol, and small peptides), could hinder recombinant protein folding.

The largest increments in rhGALNS activity were obtained at the intracellular level, with enzyme activity levels of rhGALNS of 0.06 U/(mg total protein) and volumetric activities as U/ml (E) Western blotting analysis of E. coli extracts. Here we evaluated the effect of the promoters tac and proUmod on the level expression of rhGALNS soluble protein at optimal (0.5 mM) IPTG concentration (in the case of tac) and 37°C. Both soluble (intracellular [INT] and extracellular [EXT]) and protein aggregates (IB) fractions are shown. Samples from the intracellular fraction were normalized by the total protein amount, solubilized protein aggregates fractions were normalized by biomass, and the extracellular samples were normalized by culture volume. A non-cropped version of this figure can be found in the Supplementary Data.

Figure 1. Production of rhGALNS in E. coli BL21(DE3) using different promoters. (A) Specific activity in the extracellular fraction. (B) Volumetric activity in the extracellular fraction. (C) Specific activity in the intracellular fraction. (D) Volumetric activity in the intracellular fraction. Enzyme activity was assessed 24 hours after induction (no induction was necessary in the case of the promoters studied). A Student’s t-test was performed to assess significance, where (***) corresponds to a p-value < 0.0001. Specific activities are expressed as U rhGALNS/(mg total protein) and volumetric activities as U/ml.
On the other hand, an interesting effect occurred with the protein secreted to the culture medium. As shown in Fig. 3A and B, although a statistically significant increment in specific activity after 24 h of induction was not observed, there was a significant increase in the volumetric activity. The same tendency was evidenced when 10% (w/v) sucrose was used. Elevated amounts of secreted rhGALNS were quantified using an indirect ELISA, as shown in Supplementary Data, where approximately 300% more rhGALNS was found in the medium when the culture was exposed to osmotic stress. In this sense, these results indicate an elevation in the amount of secreted protein but not in its folding, as well as that osmotic stress does not increase the proper folding of proteins at the periplasmic space. This effect can be explained since bacterial cells, under osmotic stress, accumulate small molecules known as osmolytes in the cytoplasm, acting as chemical chaperones.

Overexpression studies were carried out to determine the effect of individual chaperones genes on the enzyme activity of rhGALNS, or on the contrary, if the observed benefit via osmotic stress was due to a global stress response. Several chaperones were selected for these overexpression studies including the chaperonin system GroESL (genes groS and groL); the chaperones DnaK, DnaJ and GrpE capable of repairing heat-inducing protein damage; heat-shock proteins IbpA and IbpB involved in the binding of protein aggregates in E. coli; the complex DsbA-DsbB involved in the formation of disulfide bonds in the periplasmic space and the chaperone ClpB. The role of these chaperones on protein folding is schematized in Fig. 4A. The plasmids for the overexpression of the chaperones proteins are summarized in Table 1 and Supplementary Data. The E. coli strain BL21(DE3) was co-transformed with the plasmids pGEX-5X-GALNSopt and a recombinant pACYCDuet ™-1 plasmid for overexpression of the different chaperone genes (Table 1). The clones were screened based on the double antibiotic selection and tested for rhGALNS production in 100 mL as described in the Materials and Methods section. We did not observe a benefit neither in specific nor in volumetric rhGALNS activity when the chaperones were co-expressed alongside rhGALNS (Fig. 4). Even though chaperones are cataloged as folding modulators in the production of recombinant proteins, their co-production has been shown to reduce the yield and quality of several recombinant proteins produced in E. coli, as observed in the production of horseradish peroxidase, guinea pig liver transglutaminase, fibroblast growth factor, cyclodextrin glycosyltransferase and different antibodies fragments. In addition, overexpression of chaperones can contribute to metabolic burden, thereby leading to growth rate reduction as well as decreased final biomass yields. It remains to be determined what genes are involved in the increment of rhGALNS activity once the cells are exposed to high concentrations of sucrose, but we hypothesize based on the findings described above, that this is the effect of a global stress response and not to the action of individual chaperones.

![Graph](image-url)
Improving formation of disulfide bonds. The disulfide bond is the most common link between amino acids after the peptide bond and around 15% of human produced proteins are predicted to have disulfide bonds. In the instance of the human GALNS, it contains three disulfide bonds per monomer, increasing its stability and activity. The formation of disulfide bonds in all organisms are compartmentalized in the extra-cytoplasmic sections such as the endoplasmic reticulum in eukaryotes, or the periplasmic space of gram-negative bacteria, such as *E. coli*. This phenomenon is due to the presence of enzymes devoted to the reduction of disulfide bonds in the cytoplasm. In *E. coli*, the oxidative environment in the periplasm allows the formation of disulfide bonds, crucial for the activity and stability of several proteins, promoting resistance against proteases and harsh environments. Therefore, proteins requiring disulfide bonds for their folding and stability, such as the rhGALNS, are prone to be misfolded and not active when expressed in the cytoplasm of *E. coli*. We hypothesized that the lower activity of the intracellular rhGALNS, in comparison to its extracellular counterpart, could be associated with inefficient protein folding due to the reducing cytoplasmic environment. For this reason, we tested the *E. coli* strain SHuffle® T7. This strain is an engineered version of *E. coli* BL21, modified to allow the formation of stable disulfide bonded proteins within the cytoplasm. This modification was generated by the deletion of the thioredoxin reductase (trxB) and glutathione reductase (gor) genes, the presence of the mutant peroxidase AhpC* to restore reducing power to the bacteria, and the overexpression of the chaperone DsbC lacking the signal sequence to be co-expressed in the cytoplasm.

Figure 3. Production of rhGALNS in *E. coli* BL21(DE3) with cultures exposed to high osmotic stress. (A) Specific activity in the extracellular fraction. (B) Volumetric activity in the extracellular fraction. (C) Specific activity in the intracellular fraction. (D) Volumetric activity in the intracellular fraction. Enzyme activity was assessed 24 hours after induction. A Student’s t-test was performed to assess significance, where (*) corresponds to a p-value < 0.05 and (****) p-value < 0.0001. Specific activities are expressed as U rhGALNS/(mg total protein) and volumetric activities as U/ml. (E) Western-blot of rhGALNS levels in the intracellular and insoluble protein fractions using high osmotic stress, with production controlled by the promoter tac using optimal (0.5 mM) IPTG concentration. The results showed a significant increase of rhGALNS protein in both, intracellular and solubilized protein aggregates fraction under tac control supplemented with 5% of sucrose. In addition, increased expression was observed in both processed and mature protein.
We transformed the E. coli strain SHuffle® T7 either with the plasmids pGEX-5X-GALNSopt or pGEXproU-mod. As observed in Fig. 5A and B, there were not statistically significant differences in the extracellular specific and volumetric activities of rhGALNS in E. coli strain SHuffle® T7 when the gene expression was controlled by the tac promoter, in comparison to the results observed in E. coli BL21(DE3). When the promoter proU_mod was used, the outcome was similar to the previously described in regard of the specific activity, but we observed a reduced volumetric activity of rhGALNS. This effect can be attributed to the lower biomass reached by the strain E. coli SHuffle® T7 (1.8 g/L) than the one obtained using E. coli BL21(DE3) (3.2 g/L), which could be associated with the numerous genetic modifications in the strain E. coli SHuffle® T7. Nevertheless, by using the proU_mod promoter, a large improvement occurred at the cytoplasm with an increase of the specific activity of 1,283% [0.14 U/(mg total protein)] (Fig. 5C and D). We evaluated the increments of protein folding by using a Western-blot (Fig. 5E). The Western-blot analysis indicates a reduction of the intracellular levels of the precursor rhGALNS using E. coli SHuffle® T7, in comparison with the levels observed by using E. coli BL21(DE3). Through a densitometric analysis of the Western-blot results, we estimated the concentrations of rhGALNS present in the evaluated samples, in order to assess the specific enzyme activity (U rhGALNS/mg rhGALNS), thus determining an approximate increment in enzyme activity of 7X at the intracellular fraction when either plasmids (pGEX-5X-GALNSopt or pGEXproUmod) were used in conjunction to the strain E. coli SHuffle® T7. However, we did not observe important variations at the extracellular fraction. This result is in accordance with the expected effect using E. coli SHuffle® T7, since all the genetic modifications in this strain are designed to improve intracellular protein folding due to an oxidative cytoplasmic conditions, but not at the periplasmic space where the disulfide forming environment is already present. Similar results were reported in the production of human herpesvirus type-6, with production increments up to 5-fold in comparison to the control strain E. coli BL21(DE3)56 and the production of the neurosecretory protein GM, with presence of intracellular soluble protein in contrast to the E. coli BL2157.

Figure 4. Production of rhGALNS in E. coli BL21(DE3) when different chaperones were co-expressed. (A) The biological role of the overexpressed chaperones in protein folding in E. coli. The scheme shows the importance of the different chaperones in proper folding of nascent peptides, recovery from partially/misfolded proteins and the response to stress. (B) Specific activity in the extracellular fraction. (C) Volumetric activity in the extracellular fraction. (D) Specific activity in the intracellular fraction. (E) Volumetric activity in the intracellular fraction. Enzyme activity was assessed 24 hours after induction using optimal (0.5 mM) IPTG concentration. A Student’s t-test was performed to assess significance, where (***) corresponds to a p-value < 0.01. Specific activities are expressed as U rhGALNS/(mg total protein) and volumetric activities as U/ml. As noticed, the overexpression of individual chaperone proteins did not aid in the increase of rhGALNS activity. The names shown in this panels B to E correspond to the overexpressed chaperones. Full names of the plasmids used can be found in Table 1.
Lastly, we tested the combination of the studied approaches to determine possible additive effects. The largest increment of rhGALNS activity in the extracellular fraction occurred with the use of \textit{proU mod} promoter and 5\% (w/v) sucrose, obtaining activities of 1.23 U/(mg total protein) and 1.20 U/ml, which represent an improvement of 1,463\% and 470\%, respectively, in comparison to the results obtained with \textit{E. coli} BL21(DE3)/pGEX-5X-GALNSopt. These substantial improvements also occurred at the intracellular level as seen in Table 2, with activities of 0.16 U/(mg total protein) and 0.20 U/ml, which represent increments of 1,475\% and 335\%, respectively in comparison to BL21(DE3)/pGEX-5X-GALNSopt.

We tested if the use of a different osmotic agent, in a similar osmolarity as implemented with sucrose, leads to the same effect on rhGALNS activities. As seen in Fig. 6, using sodium chloride at 0.05 M, in conjunction with the use of the promoter \textit{proU mod} allows the production of similar enzymatic activities to the observed using sucrose in combination to the same promoter. This result suggests that the osmotic effect on rhGALNS activity is not due to the nature of the osmoreagent. Aspedon et al., using a transcriptomic analysis in \textit{Pseudomonas aeruginosa} under osmotic stress with sodium chloride and sucrose indicated that stress response to osmotic shock is not dependent of a specific compound or its ionic nature, but it is rather a general stress response process 58.

Finally, we observed that the highest intracellular specific activity was obtained by combining all the approaches studied in this work (i.e., 5\% (w/v) of sucrose, the strain \textit{E. coli} SHuffle® T7 and the promoter \textit{proU mod}). This improvement corresponds to 2,095\% of the activity observed under the initial culture conditions [0.22 U/(mg total protein)]. This arrangement permitted statistically significant improvements of the other studied variables, but it was not as high as with other combinations.

**Figure 5.** Production of rhGALNS using the promoters \textit{tac} and \textit{proU mod} with two different \textit{E. coli} strains: BL21(DE3) and SHuffle® T7. (A) Specific activity in the extracellular fraction. (B) Volumetric activity in the extracellular fraction. (C) Specific activity in the intracellular fraction. (D) Volumetric activity in the intracellular fraction. Enzyme activity was assessed 24 hours after induction (no induction was necessary in the case of the promoter \textit{proU mod}). A Student’s t-test was performed to assess significance, where (*) corresponds to a p-value < 0.05, (***p-value < 0.001 and (****) p-value < 0.0001. Specific activities are expressed as U rhGALNS/(mg total protein) and volumetric activities as U/ml. (E) Western-blot of rhGALNS levels in the intracellular and extracellular fractions in the different studied conditions. All samples were normalized by the amount of total protein. Non-cropped versions of these figures can be found in the Supplementary Data.
Figure 6. Production of rhGALNS in *E. coli* using the combinations of the different studied approaches. (A) Specific activity in the extracellular fraction. (B) Volumetric activity in the extracellular fraction. (C) Specific activity in the intracellular fraction. (D) Volumetric activity in the intracellular fraction. Enzyme activity was assessed 24 hours after induction (no induction was necessary in the case of the promoter proUmod). A Student's t-test was performed to assess significance, where (*) corresponds to a p-value < 0.05, (**) p-value < 0.01, (***) p-value < 0.001 and (****) p-value < 0.0001. Specific activities are expressed as U rhGALNS/(mg total protein) and volumetric activities as U/ml.

| Studied strains                  | Extracellular | Intracellular |
|----------------------------------|---------------|---------------|
|                                 | U/mg total protein (U/ml (% of improvement)) | U/ml (% of improvement) | U/mg total protein (U/ml (% of improvement)) | U/ml (% of improvement) |
| **BL21(DE3)/pGEX-5X-GALNSopt**   | 0.08 ± 0.01   | 0.21 ± 0.04   | 0.01 ± 0.00   | 0.05 ± 0.01 |
| **BL21(DE3)/pGEX-5X-GALNSopt 5% suc** | 0.23 ± 0.03 (197%) | 0.67 ± 0.09 (216%) | 0.06 ± 0.02 (507%) | 0.14 ± 0.04 (219%) |
| **BL21(DE3)/pGEX-5X-GALNSopt 10% suc** | 0.20 ± 0.04 (149%) | 0.45 ± 0.00 (116%) | 0.01 ± 0.00 (0%) | 0.10 ± 0.03 (122%) |
| **BL21(DE3)/pGEXproUmod**        | 0.67 ± 0.03 (751%) | 0.86 ± 0.04 (309%) | 0.08 ± 0.00 (695%) | 0.16 ± 0.01 (245%) |
| **BL21(DE3)/pGEXproUmod 5% suc** | 1.23 ± 0.16 (1,463%) | 1.20 ± 0.26 (470%) | 0.16 ± 0.01 (1,475%) | 0.20 ± 0.02 (333%) |
| **BL21(DE3)/pGEXproUmod 10% suc** | 0.62 ± 0.07 (682%) | 0.76 ± 0.09 (259%) | 0.07 ± 0.00 (589%) | 0.10 ± 0.00 (129%) |
| **BL21(DE3)/pGEXproUmod 0.05 N NaCl** | 0.97 ± 0.03 (1,134%) | 0.96 ± 0.03 (355%) | 0.14 ± 0.01 (1,234%) | 0.13 ± 0.01 (191%) |
| **SHuffle/pGEX-5X-GALNSopt**     | 0.20 ± 0.04 (156%) | 0.34 ± 0.07 (63%) | 0.02 ± 0.00 (52%) | 0.05 ± 0.01 (8%) |
| **SHuffle/pGEX-5X-GALNSopt 5% suc** | 0.21 ± 0.04 (171%) | 0.31 ± 0.05 (48%) | 0.02 ± 0.01 (84%) | 0.04 ± 0.02 (22%) |
| **SHuffle/pGEX-5X-GALNSopt 10% suc** | 0.29 ± 0.03 (274%) | 0.23 ± 0.00 (7%) | 0.06 ± 0.01 (527%) | 0.04 ± 0.01 (19%) |
| **SHuffle/pGEXproUmod**          | 0.67 ± 0.01 (748%) | 0.59 ± 0.02 (138%) | 0.14 ± 0.00 (1,283%) | 0.07 ± 0.00 (65%) |
| **SHuffle/pGEXproUmod 5% suc**   | 0.37 ± 0.04 (369%) | 0.31 ± 0.06 (48%) | 0.15 ± 0.02 (1,336%) | 0.07 ± 0.01 (47%) |
| **SHuffle/pGEXproUmod 10% suc**  | 0.63 ± 0.15 (699%) | 0.52 ± 0.12 (147%) | 0.22 ± 0.07 (2,095%) | 0.10 ± 0.01 (118%) |

Table 2. Summary of the enzymatic rhGALNS activities obtained in this work after 24 hours of induction. Values in bold correspond to statistically significant improvements using a t-student test with a p-value < 0.05.
Conclusions

In summary, several approaches were used to increase the production of the human recombinant enzyme GALNS produced in E. coli. These approaches included the control of rhGALNS expression using a promoter regulated under osmotic shock; the induction of osmoprotectans production through osmotic shock by using sucrose or sodium chloride, the improvement in the formation of disulfide bonds in the cytoplasmic space, and the combination of different angles to explore additive effects. The use of the promoter proUmod allowed an improvement in protein folding, with increased integrations of rhGALNS specific activities of 20X and 15X in the intracellular and extracellular fractions, respectively. This improvement could be attributed to the reduction of gene expression, as evidenced in Fig. 1E, permitting the protein folding machinery of the bacteria to cope with the folding of the recombinant protein. The nature of proUmod eliminates the use of inducers, due to the regulation of the promoter by the sigma factor σ32, thus facilitating the protein production bioprocess. We also utilized 5% (w/v) sucrose to expose bacterial cultures to osmotic shock, favoring a proper protein folding. We found important improvements in the intracellular level (Table 2), possibly due to a global stress response. This hypothesis was supported when several chaperones were co-expressed alongside rhGALNS. In this instance, none of the tested chaperones increased the rhGALNS activity, which agrees with the idea that the increase in the rhGALNS activity after osmotic shock induction was probably due to a general stress response and not to the action of a particular chaperone protein. On the other hand, as an approximation to increase proper folding of the recombinant protein by boosting the formation of disulfide bonds, we used the strain E. coli SHuffle® as a host for protein production. We observed important improvements in protein activity in the intracellular fraction when rhGALNS was driven by either the promoter lac or proUmod. However, the low biomass yield hindered the use of this strain in the production of large amounts of rhGALNS. Lastly, we reported that high concentrations of sucrose in conjunction with the physiological regulated promoter proUmod significantly increased the rhGALNS production and activity. Taken together, these results represent valuable information for the production of human lysosomal enzymes in E. coli, and could have a significant impact on the development of enzyme replacement therapies for lysosomal storage diseases.

Materials and Methods

Bacterial strains and plasmids construction. The E. coli strains, BL21(DE3) (B-ompTGil dcm lon hsdSB (rB mB–) λ (DE3)[lac lacUV5-T7 gene1 indi sam7 nin5]) [malB+ K-12 (λs)] and SHuffle® T7 (New England Biolabs, Ipswich, MA, USA) (F- lac pro lacO(Δ-ara-leu)7697 araD197 flm2 lacZ-T7 gene1 Δ (phoA)34) were used in this study.

Previously, we showed that recombinant GALNS produced in E. coli BL21(DE3) had a lower enzyme activity in the crude extract (i.e. production) than that reported using CHO cells9. In an attempt to increase the enzyme production, GALNS cDNA sequence (GenBank accession number NM_000512.4) was optimized by adapting the previously described.10 All the corresponding genes for the chaperones in study were amplified via PCR from genomic DNA of E. coli DH5α by boosting the formation of disulfide bonds, we used the strain E. coli SHuffle® as a host for protein production. We observed important improvements in protein activity in the intracellular fraction when rhGALNS was driven by either the promoter lac or proUmod. However, the low biomass yield hindered the use of this strain in the production of large amounts of rhGALNS. Lastly, we reported that high concentrations of sucrose in conjunction with the physiological regulated promoter proUmod significantly increased the rhGALNS production and activity. Taken together, these results represent valuable information for the production of human lysosomal enzymes in E. coli, and could have a significant impact on the development of enzyme replacement therapies for lysosomal storage diseases.

Conclusions

In summary, several approaches were used to increase the production of the human recombinant enzyme GALNS produced in E. coli. These approaches included the control of rhGALNS expression using a promoter regulated under osmotic shock; the induction of osmoprotectans production through osmotic shock by using sucrose or sodium chloride, the improvement in the formation of disulfide bonds in the cytoplasmic space, and the combination of different angles to explore additive effects. The use of the promoter proUmod permitted an improvement in protein folding, with increased integrations of rhGALNS specific activities of 20X and 15X in the intracellular and extracellular fractions, respectively. This improvement could be attributed to the reduction of gene expression, as evidenced in Fig. 1E, permitting the protein folding machinery of the bacteria to cope with the folding of the recombinant protein. The nature of proUmod eliminates the use of inducers, due to the regulation of the promoter by the sigma factor σ32, thus facilitating the protein production bioprocess. We also utilized 5% (w/v) sucrose to expose bacterial cultures to osmotic shock, favoring a proper protein folding. We found important improvements in the intracellular level (Table 2), possibly due to a global stress response. This hypothesis was supported when several chaperones were co-expressed alongside rhGALNS. In this instance, none of the tested chaperones increased the rhGALNS activity, which agrees with the idea that the increase in the rhGALNS activity after osmotic shock induction was probably due to a general stress response and not to the action of a particular chaperone protein. On the other hand, as an approximation to increase proper folding of the recombinant protein by boosting the formation of disulfide bonds, we used the strain E. coli SHuffle® T7 as a host for protein production. We observed important improvements in protein activity in the intracellular fraction when rhGALNS was driven by either the promoter lac or proUmod. However, the low biomass yield hindered the use of this strain in the production of large amounts of rhGALNS. Lastly, we reported that high concentrations of sucrose in conjunction with the physiological regulated promoter proUmod significantly increased the rhGALNS production and activity. Taken together, these results represent valuable information for the production of human lysosomal enzymes in E. coli, and could have a significant impact on the development of enzyme replacement therapies for lysosomal storage diseases.

Materials and Methods

Bacterial strains and plasmids construction. The E. coli strains, BL21(DE3) (B-ompTGil dcm lon hsdSB (rB mB–) λ (DE3)[lac lacUV5-T7 gene1 indi sam7 nin5]) [malB+ K-12 (λs)] and SHuffle® T7 (New England Biolabs, Ipswich, MA, USA) (F- lac pro lacO(Δ-ara-leu)7697 araD197 flm2 lacZ-T7 gene1 Δ (phoA)34) were used in this study.

Previously, we showed that recombinant GALNS produced in E. coli BL21(DE3) had a lower enzyme activity in the crude extract (i.e. production) than that reported using CHO cells9. In an attempt to increase the enzyme production, GALNS cDNA sequence (GenBank accession number NM_000512.4) was optimized by adapting the previously described.10 All the corresponding genes for the chaperones in study were amplified via PCR from genomic DNA of E. coli DH5α by boosting the formation of disulfide bonds, we used the strain E. coli SHuffle® as a host for protein production. We observed important improvements in protein activity in the intracellular fraction when rhGALNS was driven by either the promoter lac or proUmod. However, the low biomass yield hindered the use of this strain in the production of large amounts of rhGALNS. Lastly, we reported that high concentrations of sucrose in conjunction with the physiological regulated promoter proUmod significantly increased the rhGALNS production and activity. Taken together, these results represent valuable information for the production of human lysosomal enzymes in E. coli, and could have a significant impact on the development of enzyme replacement therapies for lysosomal storage diseases.

Materials and Methods

Bacterial strains and plasmids construction. The E. coli strains, BL21(DE3) (B-ompTGil dcm lon hsdSB (rB mB–) λ (DE3)[lac lacUV5-T7 gene1 indi sam7 nin5]) [malB+ K-12 (λs)] and SHuffle® T7 (New England Biolabs, Ipswich, MA, USA) (F- lac pro lacO(Δ-ara-leu)7697 araD197 flm2 lacZ-T7 gene1 Δ (phoA)34) were used in this study.

Previously, we showed that recombinant GALNS produced in E. coli BL21(DE3) had a lower enzyme activity in the crude extract (i.e. production) than that reported using CHO cells9. In an attempt to increase the enzyme production, GALNS cDNA sequence (GenBank accession number NM_000512.4) was optimized by adapting the previously described.10 All the corresponding genes for the chaperones in study were amplified via PCR from genomic DNA of E. coli DH5α by boosting the formation of disulfide bonds, we used the strain E. coli SHuffle® as a host for protein production. We observed important improvements in protein activity in the intracellular fraction when rhGALNS was driven by either the promoter lac or proUmod. However, the low biomass yield hindered the use of this strain in the production of large amounts of rhGALNS. Lastly, we reported that high concentrations of sucrose in conjunction with the physiological regulated promoter proUmod significantly increased the rhGALNS production and activity. Taken together, these results represent valuable information for the production of human lysosomal enzymes in E. coli, and could have a significant impact on the development of enzyme replacement therapies for lysosomal storage diseases.
48% glycerol, 9% 2-mercaptoethanol, and 0.03% bromophenol blue) and boiled at 95 °C for 10 min. 2 pellets were processed for analysis of the insoluble fraction. The protein aggregates were solubilized by mixing sodium chloride to a concentration of 0.5 mM. All assays were done in triplicate.

To evaluate the effect of osmotic stress on protein production and activity, sucrose was added to the cultures at 5% and 10% (w/v) or (extracellular).

Table 3. Primers used in this study.

| Name       | Sequence                                                                 |
|------------|--------------------------------------------------------------------------|
| pGEX-5X-f  | 5′- CCT ATT TTT ATA GGT TAA TGT CAT GAT AAT AAT GGT TTC TTA GAC GTC-3′    |
| pGEX-5X-r  | 5′- CAT TAA CTT ATA AAA ATA GGC GTA TCA CGA GGC CCT TTC-3′                |
| osmY-f     | 5′- TCT ATA TTT AAC AAA CCA AAT CAA CAA AAG CCG TTA CTA ACA TGT CCC CTA TAC TAG TTT ATT G-3′ |
| osmY-r     | 5′- TTT GGT TGG TTA AAT ATA GAT CAC AAT TTT GAA ACC GCT GCG GAT AGG AAT TAA TGC GGT GTG AAA TAC-3′ |
| groS-f     | 5′- ACT CGA ATT CAT GAA TAT CCG ATT GCA TGA TCA T-3′                      |
| groS-r     | 5′- CGC AAC CTT TTA CGC TTC AAT AAT TGC CAG-3′                            |
| groL-f     | 5′- TAT ACA TAT GAT GGC AGC ATA AGA CGT AA-3′                             |
| groL-r     | 5′- CAG ACT CGA GTT ACA TCA TGC GGC CCA TG-3′                             |
| dsbA-f     | 5′- ATT CCA GCT CAT GGG TAA AAT TGG TAT CGA C-3′                          |
| dsbA-r     | 5′- CAG CAA GCT TTT ATT TTT GTI CTT TGA CTT CAA ATT C-3′                  |
| ibpA-f     | 5′- TAT ACA TAT GAT GGC TAA GCA AGA AGA TTA TTA CGA G-3′                  |
| ibpA-r     | 5′- CAG ACT CGA GTT AGC GAC CGA ACA GAT CAG-3′                            |
| dhaB-f     | 5′- ATC CGA ATT CAT GAA AAA GAT TGT GCT GGC G-3′                          |
| dhaB-r     | 5′- CGG CAA GCT TTT ATT TTT TCT CGG ACA GAT ATT TCA CT-3′                 |
| dhaB-f     | 5′- TAT ACA TAT GAT GGT CCG ATT TTT GAA CCA GCA ATG-3′                    |
| dhaB-r     | 5′- CAG ACT CGA GTT AGC GAC CGA ACA GAT CAG-3′                            |
| bphA-f     | 5′- ATC CGA ATT CAT GCG TAA CTT TTA GAT TAC CCC-3′                        |
| bphA-r     | 5′- CGG CAA GCT TTT AGT TTA TGT CGG TAC GCC GC-3′                         |
| bphB-f     | 5′- TAT ACA TAT GAT GGC TAA CTT CTA TAT CCC-3′                           |
| bphB-r     | 5′- AGA CTC GAG TTA GCT ATT TAA CCG GGG AGC T-3′                          |
| gplF-f     | 5′- ATC CGA ATT CAT GAG TAG TAA AGA ACA GAA AAC G-3′                      |
| gplF-r     | 5′- ATG CGG CGG CTT AAG CTT TGT CTT TCG CTA CAG-3′                        |
| cphA-f     | 5′- TAT ACA TAT GAT GGC TCT GGA TCG TCT TAC-3′                           |
| cphA-r     | 5′- CAG ACT CGA GTT ACT GGA CCG CGA CAA TC-3′                            |
| GALNS-f    | 5′- GTG AAC CGC GCC GTC AAA CC-3′                                        |
| GALNS-r    | 5′- ACC AAT GCA CAT GCA CGC AA-3′                                         |

of the chaperone protein and rhGALNS, the corresponding plasmid was transformed using heat-shock into the strain E. coli BL21(DE3)/pGEX-5X-GALNSopt and screened using double antibiotic selection (AmpR + CmR).

Culture conditions. Overnight cultures were grown in LB liquid medium or on solid LB agar plates supplemented with appropriate antibiotics and incubated at 37 °C. Production of rhGALNS was carried out in 100 mL cultures using M9 minimum medium [2 mM MgSO 4, 0.1 mM CaCl 2, 6.78 g/L Na 2HPO 4, 3 g/L KH 2PO 4, 4.5 g/L NaCl, 1 g/L NH 4Cl, 1 mg/L thiamine and trace elements. 1000X trace elements solution: 13.4 mM ethylenediamine tetraacetic acid (EDTA), 3.1 mM FeCl 3.6H 2O, 0.62 mM ZnCl 2, 76 µM MnCl 2, 162 µM CoCl 2.2H 2O, 8.1 µM NiCl 2, 8.1 µM CuCl 2.2H 2O, 42 µM AlCl 3, 36.3 mg/L H 3BO 3, 9.4 mg/L H 2SiO 3, 8.1 µM MgCl 2.6H 2O, 0.58 mM CaCl 2, 0.58 mM Na 2EDTA, 0.58 mM Mg 2EDTA, 1.4 mg/L MnCl 2, 1.4 mg/L ZnCl 2, 1.4 mg/L H 2BO 3] supplemented with 20 g/L of D-glucose under aerobic conditions, which was also supplemented with appropriated antibiotics. When necessary, gene expression was induced using isopropyl-β-D-1-thiogalactopyranoside (IPTG, Gold Biotechnology, St. Louis, MO, USA) at 0.5 mM after 8 hours of incubation. All cultures were maintained at 180 RPM during cultivation. Cells were cultivated for 24 hours after induction. Biomass was estimated using optical density at 540 nm. To evaluate the effect of osmotic stress on protein production and activity, sucrose was added to the cultures at 5% and 10% (w/v) or (extracellular).

Crude protein assays. Fifty mL cultures were centrifuged and the growth media was saved for further analysis. The pellets were resuspended in 5 mL lysis buffer (25 mM Tris, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 5% (v/v) glycerol, and 1% (v/v) Triton X-100, to a pH 7.2). Samples were sonicated during 1 minute at 4 °C and 25% amplitude (Vibra-Cell, Sonics & Materials Inc., Newtown, CT, USA) and centrifuged at 3000 × g and 4 °C during 20 minutes. The soluble fraction (supernatant) was stored at −20 °C for further analysis. Recovered pellets were processed for analysis of the insoluble fraction. The protein aggregates were solubilized by mixing 2 µL of lysed cells in 30 µL of 6X SDS loading buffer (375 mM Tris-HCl pH 6.8, 6% sodium dodecyl sulfate (SDS), 48% glycerol, 9% 2-mercaptoethanol, and 0.03% bromophenol blue) and boiled at 95 °C for 10 min.

Quantification of GALNS activity. rhGALNS activity was assessed using the fluorescent substrate 4-methylumbelliferyl-β-D-galactopyranoside-6-sulfate (Toronto Chemicals Research, North York, ON, Canada) as described elsewhere. One unit (U) was defined as the amount of rhGALNS catalyzing 1 nmol substrate per hour, and rhGALNS activity was expressed as U/(mg total protein) (total protein determined by Lowry protein assay) or as U/mL. Enzyme activity was assayed in the soluble fraction (intracellular) and growth media (extracellular).
**SDS–PAGE Analysis.** Crude proteins extracts and growth media were analyzed by SDS - polyacrylamide gel electrophoresis (SDS–PAGE) under reducing conditions as described by Laemmli et al. Crude proteins extracts were loaded and processed by SDS–PAGE gels and electroblotted onto nitrocellulose membranes (Amershan Protran™ 0.45 μm, GE Healthcare, Little Chalfont, UK). The membranes were blocked with a blocking buffer containing TTBS [0.3% Tween 20, and 5% BSA (Sigma-Aldrich, St. Louis, MO, USA)] at room temperature for 1 hour. Membranes containing resuspended protein aggregates, extracellular and intracellular proteins were incubated overnight at 4 °C with a specific primary antibody rabbit anti human GALNS (1:1000 in blocking buffer). The specificity of this antibody was previously assessed, and no band was recognized in samples from non-induced and plasmid-free strains. A peroxidase conjugated goat anti-rabbit (Sigma–Aldrich) was added (1:2000 in blocking buffer) for 1 hour at room temperature. The specific bands were visualized using enhanced chemiluminescence (SuperSignal ™ West Pico Chemiluminescent Substrate, Thermo Fisher Scientific®, Walthera, MA, USA).

**Western blotting.** The proteins from different fractions were homogenized in lysis buffer and equivalent amounts of total protein extracts were loaded and processed by SDS–PAGE gels and electroblotted onto nitrocellulose membranes (Amershan Protran™ 0.45 μm, GE Healthcare, Little Chalfont, UK). The membranes were blocked with a blocking buffer containing TTBS [0.3% Tween 20, and 5% BSA (Sigma–Aldrich, St. Louis, MO, USA)] at room temperature for 1 hour. Membranes containing resuspended protein aggregates, extracellular and intracellular proteins were incubated overnight at 4 °C with a specific primary antibody rabbit anti-human GALNS (1:1000 in blocking buffer). The specificity of this antibody was previously assessed, and no band was recognized in samples from non-induced and plasmid-free strains. A peroxidase conjugated goat anti-rabbit (Sigma–Aldrich) was added (1:2000 in blocking buffer) for 1 hour at room temperature. The specific bands were visualized using enhanced chemiluminescence (SuperSignal ™ West Pico Chemiluminescent Substrate, Thermo Fisher Scientific®, Walthera, MA, USA).

**qRT-PCR experiments.** Frozen stocks of the strain BL21(DE3)/pGEXproUmod (0.5 mL) were transferred into 10 mL of fresh LB media supplemented with 100 ng/mL of ampicillin and incubated for 18 h at 37 °C and 180 rpm. Production of rhGALNS was carried out in 100 mL cultures using M9 minimum medium supplemented with 20 g/L of D-glucose under aerobic conditions, inoculated with 10 mL of overnight culture, kept at 37 °C and 180 rpm. Samples (10 mL) were taken 2, 4, 7, 12 and 24 hours after inoculation of all three biological replicates. The samples were pelleted by centrifugation at 3000 x g and 4 °C during 10 minutes. The supernatant was discarded and the pellets were immediately placed at −80 °C until further processing. The samples were processed in a period lower than 12 hours after pelleting. Total RNA was extracted using the ZR Fungal/Bacterial RNA MiniPrep™ Kit (Zymo Research, Irvine, CA, USA), following manufacturer’s indications and immediately stored at −80 °C. The cDNA was generated using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) using 100 ng of total RNA and following the random hexamer primed synthesis manufacturer’s protocol. Two primers (GALNS-f and GALNS-r (Table 3)) were designed for the quantification of rhGALNS transcripts using NCBI Primer Design Tool and they were used for the qPCR experiments. Primer amplification efficiency was assessed using the plasmid pGEXproUmod (purified using the Wizard SV Gel and PCR Clean-up System (Promega, Madison, WI, USA)) as template. The amplification efficiency of these primers was estimated as 94%. To generate a standard curve for absolute rhGALNS copy number quantification, the plasmid pGEXproUmod was serially diluted to obtain different plasmid copies ranging from 300,000 to 30. Triplicates of each data-point were used to ensure reliability of the data. All qPCR experiments were carried out using Luna® Universal qPCR Master Mix (New England Biolabs) following manufacturer’s indications and read in a QuantStudio 3 Real-Time PCR System (Thermo Fisher Scientific). The results as reported as rhGALNS copy number per ng of total RNA.

**Statistical analysis.** Statistical significance was evaluated using a Student’s t-test, employing the software GraphPad Prism version 6 for Windows (GraphPad Software, San Diego, California, USA).

**References**

1. Montano, A. M., Tomatsu, S., Gottesman, G. S., Smith, M. & Orii, T. International Morquio A Registry: clinical manifestation and natural course of Morquio A disease. *J Inherit Metab Dis* 30, 165–174, doi:10.1007/s10545-007-0529-7 (2007).
2. Tomatsu, S. et al. Mucopolysaccharidosis type IVA (Morquio A disease): clinical review and current treatment. *Curr Pharm Biotechnol* 12, 931–945, doi:10.3189/1011-5880.00.00.00.100 (2011).
3. Tomatsu, S. et al. Morquio A syndrome: diagnosis and current and future therapies. *Pediatr Endocrinol Rev* 12(Suppl 1), 141–151 (2014).
4. Hendriksz, C. J. et al. Efficacy and safety of enzyme replacement therapy with BMN 110 (cloustrafase alfa) for Morquio A syndrome (mucopolysaccharidosis IVA): a phase 3 randomised placebo-controlled study. *J Inherit Metab Dis* 37, doi:10.1054/jimd.2004.0650 (2010).
5. Tomatsu, S. et al. Enzyme replacement therapy for treating mucopolysaccharidosis type IVA (Morquio A syndrome): effect and limitations. *Expert Opin on Orphan Drugs* 3, 1279–1290, doi:10.1517/21678787.2015.1086640 (2015).
6. Tomatsu, S. et al. Enhancement of drug delivery: enzyme-replacement therapy for murine Morquio A syndrome. *Mol Ther* 18, 1094–1112, doi:10.1038/mt.2010.32 (2010).
7. Espejo-Mojica, A. et al. Human recombinant lysosomal enzymes produced in microorganisms. *Mol Genet Metab* 116, 13–23, doi:10.1016/j.ymgme.2015.06.001 (2015).
8. Bielicki, J. et al. Expression, purification and characterization of recombinant human N-acetylgalactosamine-6-sulfatase. *Biochem J* 371(Pt 1), 333–339 (1999).
9. Rodríguez, A. et al. Enzyme replacement therapy for Morquio A: an active recombinant N-acetylgalactosaminase-6-sulfatase produced in *Escherichia coli* BL21. *J Ind Microbiol Biotechnol* 37, 1193–1201, doi:10.1007/s10295-010-0766-x (2010).
10. Rodríguez-López, A. et al. Recombinant human N-acetylgalactosaminase-6-sulfatase (GALNS) produced in the methylotrophic yeast *Pichia pastoris*. *Sci Rep* 6, 29329, doi:10.1038/srep29329 http://www.nature.com/articles/srep29329-supplementary-information (2016).
11. Kamionka, M. Engineering of therapeutic proteins production in *Escherichia coli*. *Curr Pharm Biotechnol* 12, 268–274 (2011).
12. Critchley, R. J. et al. Potential therapeutic applications of recombinant, invasive *E. coli*. *Gene Ther* 11, 1224–1233, doi:10.1038/si.gt.3002281 (2004).
13. Baehner, M. N. et al. Production of Biopharmaceuticals in *E. coli*: Current Scenario and Future Perspectives. *J Microbiol Biotechnol* 25, 953–962, doi:10.4144/jmjb.1412.12079 (2015).
14. Banexy, F. & Mujacic, M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol* 22, 1399–1408, doi:10.1038/nb2012 (2004).
15. Mosquera, A. et al. Characterization of a recombinant N-acetylgalactosaminase-6-sulfatase produced in *E. coli* for enzyme replacement therapy of Morquio A disease. *Process Biochem*. 47, 2097–2102 (2012).
16. Hernandez, A. et al. Effect of culture conditions and signal peptide on production of human recombinant N-acetylgalactosamine-6-sulfate sulfatase in Escherichia coli BL21. J Microbiol Biotechnol 23, 689–698 (2013).

17. Carrio, M. & Villaverde, A. Construction and deconstruction of bacterial inclusion bodies. J Biotechnol 96, 3–12 (2002).

18. Muller-Hill, B., Crapo, L. & Gilbert, W. Mutants that make more lac repressor. Proc Natl Acad Sci USA 59, 1259–1264 (1968).

19. Sweet, C. R. Expression of recombinant proteins from lac promoters. Methods Mol Biol 235, 277–288, doi:10.1385/1-59259-409-7:3-277 (2003).

20. Hannig, G. & Makrides, S. C. Strategies for optimizing heterologous protein expression in Escherichia coli. Trends Biotechnol 16, 54–60 (1998).

21. de Boer, H. A., Comstock, L. J. & Vasser, M. The tac promoter: a functional hybrid derived from the trp and lac promoters. Proc Natl Acad Sci USA 80, 21–25 (1983).

22. Wise, A., Brems, R., Ramakrishnan, V. & Villarejo, M. Sequences in the — region of Escherichia coli rpoS-dependent genes promote transcription by E sigma S. J Bacterial 178, 2785–2793 (1996).

23. Kim, S. G., Kweon, D. H., Lee, D. H., Park, Y. C. & Seo, J. H. Coexpression of folding accessory proteins for production of active cyclodextrin glycosyltransferase of Bacillus macerans in recombinant Escherichia coli. J Biotechnol 127, 244–257, doi:10.1016/j.jbiotec.2006.07.004 (2007).

24. Ikura, K. et al. The role of the amino-and carboxyl-terminal domains. J Bacteriol 180, 519–524 (1998).

25. Schroder, H., Langer, T., Hartl, F. U. & Bukau, B. DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. EMBO J 12, 4137–4144 (1993).

26. S. et al. Compatible-solute-supported periplasmic expression of functional recombinant proteins under stress conditions. Appl Environ Microbiol 66, 1572–1579 (2000).

27. Smart, A. N. Physiological and genetic responses of bacteria to osmotic stress. Microbiological Reviews 53, 121–147 (1989).

28. Chen, J., Acton, T. B., Basu, S. K., Montelione, G. T. & Inouye, M. Enhancement of the solubility of proteins overexpressed in Escherichia coli by heat shock. J Mol Microbiol Biotechnol 4, 519–524 (2003).

29. Carrio, M. M. & Villaverde, A. Construction and deconstruction of bacterial inclusion bodies. Biotechniques 20, 176, 100–107 (1994).

30. Kempf, B. U. & Bremer, E. Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. Biochimie 84, 374–381 (1999).

31. Bar, S. et al. Disulfi de bond acquisition through eukaryotic protein evolution. J Protein Chem 15, 867–671 (2006).

32. Weissman, J. S. & Kim, P. S. Efficient catalysis of disulphide bond rearrangements by protein disulphide isomerase. J Biochemistry (Mosc) 77–92 (John Wiley & Sons, Inc., 2014).

33. de Marco, A. In Protein Aggregation in Bacteria 77–92 (John Wiley & Sons, Inc., 2014).

34. de Marco, A. Protein expression and purification of U24 from human herpesvirus type-6 in Escherichia coli. J Microbiol Biotechnol 23, 689–698 (2013).

35. Lobstein, J. et al. Shuffling, a novel Escherichia coli protein expression strain capable of correctly folding disulfi de bonded proteins in its cytoplasm. Microb Cell Fact 11, 56, doi:10.1186/1475-2859-5-56 (2012).

36. Tait, A. R. & Strauss, U. K. Overexpression and puriﬁcation of U24 from human herpesvirus type-6 in E. coli: unconventional use of oxidizing environments with a maltose binding protein-hexahistidine dual tag to enhance membrane protein yield. Microbial Cell Factories 10, 51–51, doi:10.1186/1475-2859-10-51 (2011).
57. Masuda, K., Furumitsu, M., Taniuchi, S., Iwakoshi-Ukena, E. & Ukena, K. Production and characterization of neurosecretory protein GM using *Escherichia coli* and Chinese Hamster Ovary cells. *FEBS Open Bio* 5, 844–851, doi:10.1016/j.fob.2015.10.002 (2015).

58. Aspedon, A., Palmer, K. & Whiteley, M. Microarray Analysis of the Osmotic Stress Response in *Pseudomonas aeruginosa*. *Journal of Bacteriology* 188, 2721–2725, doi:10.1128/JB.188.7.2721-2725.2006 (2006).

59. Espah Borujeni, A., Channarasappa, A. S. & Salis, H. M. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res* 42, 2646–2659, doi:10.1093/nar/gkt1139 (2014).

60. Li, M. Z. & Elledge, S. J. Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. *Nat Methods* 4, 251–256, doi:10.1038/nmeth1010 (2007).

61. Thomas, J. G. & Baneyx, F. Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing Heat-shock proteins. *J Biol Chem* 271, 11141–11147 (1996).

62. van Diggelen, O. P. et al. A fluorimetric enzyme assay for the diagnosis of Morquio disease type A (MPS IV A). *Clin Chim Acta* 187, 131–139 (1990).

63. Dawson, J. M. & Heatlie, P. L. Lowry method of protein quantification: evidence for photosensitivity. *Anal Biochem* 140, 391–393 (1984).

64. Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227, 680–685 (1970).

Acknowledgements
The authors wish to thank Juan David Patiño Salazar for the help in the qRT-PCR experiments and Oscar Alejandro Hidalgo Lanussa with his assistance with the creation of the illustrations. LHR and CC were supported by Pontificia Universidad Javeriana and COLCIENCIAS [Grant No. FP44842-232-2015, ID 6700 (Mejoramiento en la producción de proteínas recombinantes mediante el uso de biología sintética e ingeniería de proteínas) and FP44842-233-2015, ID 6699]. ARL received a doctoral scholarship from Pontificia Universidad Javeriana. CJAD was also supported by Pontificia Universidad Javeriana (Grant ID 3964, 5537, and 7204) and COLCIENCIAS (Grant ID 5174, contract No. 120356933205; and Grant ID 5170, contract No. 120356933427).

Author Contributions
L.H.R. and C.J.A.-D. designed the research; L.H.R., C.C. and L.N.P. performed the experiments; L.H.R., C.C., A.R.-L. and C.J.A.-D. analyzed the data; L.H.R., C.C., A.R.-L. and C.J.A.-D. wrote the paper.

Additional Information
Supplementary information accompanies this paper at doi:10.1038/s41598-017-06367-w

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017