



Cost-effective production of recombinant peptides in Escherichia coli

Gaglione, R., Pane, K., Dell'Olmo, E., Cafaro, V., Pizzo, E., Olivieri, G., ...
Arciello, A.

This is a "Post-Print" accepted manuscript, which has been published in "New
Biotechnology"

This version is distributed under a non-commercial no derivatives Creative Commons



([CC-BY-NC-ND](https://creativecommons.org/licenses/by-nc-nd/4.0/)) user license, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited and not used for commercial purposes. Further, the restriction applies that if you remix, transform, or build upon the material, you may not distribute the modified material.

Please cite this publication as follows:

Gaglione, R., Pane, K., Dell'Olmo, E., Cafaro, V., Pizzo, E., Olivieri, G., ... Arciello, A. (2019). Cost-effective production of recombinant peptides in Escherichia coli. *New Biotechnology*, 51, 39-48. DOI: 10.1016/j.nbt.2019.02.004

You can download the published version at:

<https://doi.org/10.1016/j.nbt.2019.02.004>

Cost-effective production of recombinant peptides in *Escherichia coli*

Rosa Gaglione¹, Katia Pane², Eliana Dell'Olmo¹, Valeria Cafaro³, Elio Pizzo³, Giuseppe Olivieri^{4,*}, Eugenio Notomista³, Angela Arciello^{1,5,*}

¹Department of Chemical Sciences, University of Naples Federico II, 80126 Naples, Italy; ²IRCCS SDN, 80143 Naples, Italy; ³Department of Biology, University of Naples Federico II, 80126 Naples, Italy; ⁴Bioprocess Engineering, AlgaePARC, Wageningen University and Research, PO Box 16, 6700 AA Wageningen, The Netherlands; ⁵Istituto Nazionale di Biostrutture e Biosistemi (INBB), Italy.

***Correspondence:** Dr. Angela Arciello (email: anarciel@unina.it); Dr. Giuseppe Olivieri (email: giuseppe.olivieri@wur.nl).

Abstract

Among bioactive peptides, cationic antimicrobial peptides (AMPs), also referred to as host defence peptides (HDPs), are valuable tools to treat infections, since they are able to kill a wide variety of microbes directly and/or to modulate host immunity. HDPs have great therapeutic potential against antibiotic-resistant bacteria, viruses and even parasites. However, high manufacturing costs have greatly limited their development as drugs, thus highlighting the need to develop novel and competitive production strategies. Here, a cost-effective procedure was established to produce the high amounts of peptides required for basic and clinical research. Firstly, a novel culture medium was designed, which was found to support significantly higher cell densities and recombinant expression levels of peptides under test compared to conventional media. The procedure has been also efficiently scaled up by using a 5 L fermenter, while the costs have been significantly lowered by developing a successful auto-induction strategy, which has been found to support significantly higher yields of target constructs and cell biomass compared to conventional strategies based on expression induction by IPTG. Interestingly, it was estimated that by increasing the production scale from 100 to 1,000 mg/batch, unit costs decreased strongly from 253 to 42 €/mg. These costs appear highly competitive when compared to chemical synthesis strategies. Altogether, the data indicate that the strategy represents an important starting point for the future development of large-scale manufacture of HDPs.

Keywords: host defence peptides; apolipoprotein B; apolipoprotein A-I; auto-inducing expression procedure; cost-effective recombinant production; techno-economic analysis.

Abbreviations: **HDPs**, host defence peptides; **AMPs**, antimicrobial peptides; **TB**, Terrific Broth; **SB**, Super Broth; **RNase**, ribonuclease; **ONC**, onconase; **LB**, Luria-Bertani broth; **NAB**, Notomista-Arciello broth; **iNAB**, auto-inducing Notomista-Arciello broth; **IPTG**, isopropyl- β -D-thiogalactopyranoside; **ApoB**, apolipoprotein B; **ApoA-I**, apolipoprotein A-I; **TF-acetate**, trifluoroacetate.

Introduction

Both basic research and clinical applications require high quality peptides to be readily available in a cost-effective manner. This is the case for antimicrobial peptides (AMPs), also referred to as host defence peptides (HDPs), in view of their immunomodulatory properties. To date, more than 2,500 HDPs have been identified [1]. Since they exhibit a wide spectrum of biological activities, their potential applications range from biomedical therapies to the food industry and agricultural area. In general, isolation of HDPs from natural sources is a labor intensive and time-consuming process, not suitable to obtain peptides in large amounts. Chemical synthesis, although very efficient, is complex and expensive [2], suited to obtain small amounts of peptides, but prohibitive when peptides longer than 30 amino acids or large quantities of peptide are required. Costs are even higher in the case of lanthipeptides, ribosomally synthesized and post-translationally modified peptides having thioether cross-linked amino acids, lanthionines, as a structural element [3]. Indeed, industrial scale synthesis of lanthipeptides appears unprofitable, due to expensive raw materials and overall low yields, which culminate in exorbitant production costs in large scale process [4]. Based on this, chemical synthesis appears as an unsuitable platform for large-scale peptide production.

In this scenario, recombinant DNA technology provides economical means for peptide manufacture. However, although many HDPs have been successfully obtained through recombinant production in *Escherichia coli* [5, 6], several difficulties have been encountered [6-8]. First, the antibacterial nature of the peptides makes them potentially lethal to the producing host. Secondly, their small size and cationic property make them

highly susceptible to proteolytic degradation [7, 9]. A strategy that effectively overcomes both obstacles is peptide fusion to a carrier protein [7, 10], from which the peptide of interest can be released by enzymatic or chemical cleavage at a specific site around the carrier-peptide junction [8]. The fusion design mimics the peptide precursor structure, with the carrier protein playing a similar role to the natural peptide pro-segment, thus protecting the host from the toxic peptide and the peptide from bacterial proteases [11]. Recently, the rational development of a new carrier protein for high yield production of recombinant peptides in *E. coli* has been described [12]. The denatured form of Onconase (ONC), a ribonuclease (RNase) from *Rana pipiens* [13], was found to be a highly suitable partner, since it can be expressed at very high levels in inclusion bodies (about 200–250 mg/L in Terrific Broth), is a small protein (104 aa), and its solubility is pH dependent (the denatured protein is soluble only at pH<4.0), thus allowing the purification of peptides soluble at pH 7.0 by selective precipitation of the carrier. Moreover, the chimeric construct has been designed to contain a His tag sequence, located between the ONC and the peptide moieties, suitable for easy purification, a flexible linker (Gly-Thr-Gly) and a dipeptide (Asp-Pro), which is cleaved in acidic conditions, thus allowing the release of the peptide from the carrier [12]. The optimized procedure allows the purification of about 10-15 mg of pure peptide from 1 L of bacterial culture [12, 14, 15]. However, a cost analysis to evaluate the competitiveness of the production process has not yet been performed. Here, we describe the development of a novel growth medium and the setting up of an auto-induction expression procedure with the ultimate goal of scaling up an efficient and commercially competitive production strategy, as confirmed by techno-economic analyses. The experimental strategy has been successfully applied to GKY20 peptide, a

short cationic AMP derived from the C-terminus of human thrombin [12, 16] and to two novel HDPs identified in an isoform of human Apolipoprotein B (ApoB) [15, 17].

Materials and methods

Bacterial strains, plasmids and bacterial culture media

E. coli BL21(DE3) and *E. coli* BL21(DE3)pLysS strains were from AMS Biotechnology (Abingdon, UK). To produce recombinant constructs ONC-r(P)GKY20 [12], ONC-r(P)ApoB_L [15] and ONC-r(P)ApoB_S [15] in bacterial cells, the cDNAs encoding the antimicrobial peptides were cloned in the pET-22b(+) expression vector (Novagen, Merck Group, Darmstadt, Germany). In the case of His6-ApoA-I, the pET-20b(+) expression vector (Novagen) was used. Bacterial cultures were carried out accordingly to [18]. NZY medium was prepared accordingly to [18]. All the reagents were from Sigma-Aldrich (Milan, Italy), unless otherwise specified.

Shake flask expression

Expression of recombinant constructs was carried out as previously described [12, 14, 15] by using IPTG (isopropyl- β -D-thiogalactopyranoside) at a final concentration of 0.7 mM. Expression of recombinant constructs was carried out by shaking bacterial samples at 180 rpm at 37 °C. In the case of apolipoprotein A-I (ApoA-I) protein, an ApoA-I expressing pET20 plasmid in *E. coli* strain BL21(DE3)pLysS (ThermoFisher Scientific, Waltham, MA) was used, as previously described [19-21].

Bioreactor cultivation

Bioreactor cultivation was performed in a BioFlo 3000 system from New Brunswick Scientific Co. (Connecticut, USA). The instrument is endowed with a 5 L vessel, temperature control by a heating sleeve and an integrated cooling system, monitoring of airflow rate, and monitoring of pH values and dissolved O₂ levels through specific probes (Mettler Toledo S.p.A., Novate Milanese, Milan, Italy). The medium in bioreactor (3 L) was inoculated with *E. coli* BL21(DE3) or with *E. coli*

BL21(DE3)pLysS cells, previously transformed with recombinant vectors, and incubated overnight at 37 °C in 150 mL of antibiotic-supplemented medium ($OD_{600\text{ nm}} = 2-4$). Sterile antifoam A emulsion (Sigma-Aldrich, Milan, Italy) was added manually (0.01% v/v final concentration) under aseptic conditions prior to overnight incubation. To avoid undesirable growth of non-transformed bacterial cells upon ampicillin degradation, antibiotic was regularly added to bacterial cells at a rate of 100 mg/h by using a fermenter peristaltic pump. During the whole fermentation process, the stirring speed was kept constant at 300 rpm, with an air flow rate of 10 cm³/sec, and temperature was kept constant at 37 °C. When bacterial culture in the bioreactor reached an $OD_{600\text{ nm}}$ of 3-5, protein expression was induced by adding 0.7 mM IPTG by using a fermenter peristaltic pump. A time of 30 min after IPTG addition was defined as the start time ($t = 0$ h). In the case of the auto-induction expression procedure, 30 min after bacteria inoculation into the bioreactor was defined as the start time ($t = 0$ h). Bioreactor cultivation was carried out for a total of 24 h. For each experimental condition, bioreactor cultivations were performed at least twice for each expressed protein construct.

Biomass determination

For both shake flask and bioreactor cultivations, $OD_{600\text{ nm}}$ measurements were performed on triplicate samples by using a Photometer UV-2600 (Shimadzu Europa GmbH, Duisburg, F.R. Germany), with samples diluted in complete medium. For cell dry weight determination, samples (25 mL) were harvested, centrifuged for 20 min at 6,000 rpm, and washed 3 times with 0.9% NaCl, then dried overnight at 80 °C, and lyophilized before weighing.

Analyses of expression levels

For both shake flask and bioreactor cultivations, at established time intervals, bacterial aliquots (corresponding to 0.125 OD_{600 nm}) were withdrawn and solubilized in lysis buffer (0.125 M Tris-HCl pH 6.8 containing 2% SDS and 10% glycerol), in order to determine recombinant construct expression levels by 15% SDS-PAGE. Protein band intensities were detected by densitometric analyses using a ChemiDoc Imaging System (Bio-Rad, California, USA). In order to estimate recombinant constructs expression levels, the intensity of protein bands with the expected electrophoretic mobility was referred to a calibration curve obtained by analyzing increasing defined amounts of bovine serum albumin (BSA, Sigma-Aldrich, Milan, Italy).

Techno-economic analysis

The analysis of the costs of peptide production was performed by implementing the process sections in Superpro Designer 9.0 (Intelligen Inc., Thessaloniki, Greece). In the case of each unit operation, mass and energy balances were performed to evaluate chemical demand, electricity consumption and utilities duties. In Supplementary Fig. S1 and S2, the flowsheet of the entire process is reported. As shown, the process consisted of 3 sections: (i) fermentation, (ii) protein extraction and (iii) purification (Supplementary Fig. S1 and S2). Step 1 (fermentation) comprised 3 stages: (i) pre-culture carried out in shake flasks, (ii) fermentation and (iii) sample harvesting by centrifugation. Extraction consisted of four washing-centrifugation steps according to [12]. Purification was performed by expanded bed chromatography followed by diafiltration. Thereafter, peptide was released by acidic hydrolysis of the chimeric construct, *i.e.* onconase-peptide. Onconase was then selectively precipitated and removed by centrifugation, whereas the soluble peptide was dried by lyophilization and stored prior to solubilization in an appropriate buffer. The chosen scenario describes

first the production costs on a small production scale of 100 mg of peptide *per* batch, which is related to a 5 L fermenter volume. The production process has then been scaled up to 10 times by simulating the case of a small-medium enterprise. Italy was chosen as the location to determine the costs of labor, electricity and other utilities, as reported in Table S1.

Capital (CAPEX) and Operating Expenditures (OPEX) were calculated according to the procedure reported in Table S2. The overall Lang factor of the process was about 8, which is in line with indications for high value product from microbial processes [22]. Costs of chemicals were taken from bulk quotations kindly provided by Sigma-Aldrich (Milan, Italy), PanReac AppliChem (Maryland Heights, MO, USA) and GE Healthcare Europe GmbH (Milan, Italy), as reported in Table S3. A detailed batch schedule was implemented by reporting operations associated with each unit, according to the procedure reported in Supplementary Figures S1 and S2. Major equipment depreciation was evaluated by assuming a 10-year lifetime and an interest rate of 2-3%. To take into account maintenance, 330 days of production *per* year were considered. The process was initially fixed at a laboratory scale by considering a working fermenter volume of 5 L, and then scaled up to 50 L for future scenarios. Fermenter cost was scaled according to power law exponents suggested by Kalk and Langlykke [22]. Costs of other equipment were scaled up according to the model provided by SuperPro Designer. Wastewater treatment was fixed to 0.4 € m⁻³ [23]. Since the entire process was carried out under batchwise conditions, the labour demand was fixed to one direct labour hour *per* hour for all the procedures described in each step.

Results

Bacterial growth medium design

A novel semi-defined rich medium was developed, designated Notomista-Arciello broth (NAB), with the aim of obtaining high cell densities and significantly higher expression levels of recombinant peptides and proteins. The composition of NAB is reported in **Table 1**. As will be noted, the medium is composed of inexpensive, readily available components, and contains defined amounts of each of them. In particular, yeast extract has been replaced by known growth factors, in order to obtain a chemically defined medium rather than a complex crude extract whose composition varies not only depending on the manufacturer but also from batch to batch of the same manufacturer [24]. In NAB, yeast extract has been primarily replaced by tryptone, which has a defined amino acidic composition. As different batches of tryptone can contain different amounts of lactose, 4 g/L glucose has been added to NAB (**Table 1**), to avoid undesirable basal induction of recombinant protein expression. NAB also contains a mixture of salts providing optimal amounts of required metals (e.g. Mg, Ca, Zn, Fe, Mn, Cu, etc.), usually acting as cofactors for essential enzymatic reactions. Glycerol (1.2%) was added as the main carbon source. Ammonium citrate was added both as a source of inorganic nitrogen and as an additional buffer system. Furthermore, since citrate acts as a chelating agent, its presence is fundamental to avoid precipitation of transition metal cations and also promotes iron uptake [24]. Finally, betaine was added as one of the best osmolytes for *E. coli* [25], that is able to synthesize it from choline, a component likely present in yeast extract but not in tryptone.

To verify the efficiency of NAB for *E. coli* culture and recombinant protein production, a comparison was first performed between TB (Terrific Broth) and NAB in shake flask cultures. For this, *E. coli* BL21(DE3) strain was used as a host to produce recombinant bioactive peptides. Since the peptides GK Y20, ApoB887-923, and ApoB887-911, upon cleavage of the chimeric construct in acidic conditions, are released with an additional proline residue at the N-terminus, they are termed r(P)ApoB_L, r(P)ApoB_S, and r(P)GKY20. Similarly, fusion proteins with onconase (ONC) are termed ONC-r(P)ApoB_L, ONC-r(P)ApoB_S, and ONC-r(P)GKY20, respectively. To compare TB and NAB, bacterial cells, transformed with pET recombinant plasmids, were grown in 50 mL of medium containing 100 µg/mL ampicillin, and then used to inoculate 1 L of ampicillin containing medium. The culture was incubated at 37 °C up to an OD_{600 nm} of 1-3.5. Expression of recombinant constructs was induced by addition of 0.7 mM IPTG. Cells were harvested after overnight induction by centrifugation and lysed in order to analyze recombinant constructs expression by 15% SDS-PAGE (**Figure 1 a-c**). With NAB, higher expression levels were reached for ONC-r(P)ApoB_L and ONC-r(P)ApoB_S recombinant constructs (**Figure 1 a, b, and Table 2**). In particular, expression levels were found to be 1.5 and 1.6 times higher in NAB than in TB for ONC-r(P)ApoB_L and ONC-r(P)ApoB_S recombinant constructs, respectively (**Table 2**). For the ONC-r(P)GKY20 construct, in contrast, the effects of NAB on expression levels were found to be slight. Moreover, in all the cases, we observed a strong decrease of basal expression in NAB compared to TB (**Figure 1 a-c**). Expression levels were determined on the basis of densitometric analyses of protein bands on SDS-PAGE gels (Materials and Methods). It can also be noted that faster growth kinetics were observed in NAB compared to TB. In fact, by monitoring biomass increase on the basis of OD_{600 nm}

values, it was observed that in the early stationary phase, the point selected for IPTG addition, higher cell densities were achieved in NAB compared to TB. In **Table 2**, results of representative experiments are reported. Time-course experiments were also performed by monitoring bacterial growth on the basis of absorbance values at 600 nm and protein construct expression yield by SDS-PAGE analyses, followed by evaluation of intensity values of protein bands of interest. Both bacterial growth and protein expression yield were monitored at regular time intervals for 24 h. Results for ONC-r(P)ApoB_L and ONC-r(P)ApoB_S constructs in TB and NAB are reported in Supplementary Figures 3 and 4. It can be seen that, over 24 h cultivation, biomass and protein construct expression yields progressively increased in both media. Moreover, final expression yields of protein constructs appeared significantly higher in NAB than in TB. These findings suggest that NAB is able to support higher biomass and protein product expression yields than conventional TB, at least in the case of the linear peptides under test.

Set up of an auto-induction protocol

To further improve and to significantly lower the costs of the production procedure, an auto-induction protocol was set up. An auto-induction medium was first developed, characterized by a well-balanced combination of different carbon sources and other essential nutrients, thus allowing cultures to grow to high cell densities and to support the expression of the target chimeric proteins at high levels. Indeed, NAB was conceived to possess the mineral salt formulation required for a successful auto-induction procedure [27]. Hence, to establish an efficient auto-induction protocol, NAB was supplemented with 0.4% glycerol (54 mM) and 1.92 g/L lactose (5.6 mM) [28].

Glucose was added at a concentration of 0.5 g/L (2.8 mM) [28] to promote rapid cell growth in the early stages of the culture, while also preventing uptake and metabolism of inducing sugar (lactose). In this way, lactose was consumed only after glucose exhaustion, with a consequent automatic induction of recombinant protein expression under the control of the T7 promoter. Indeed, *E. coli* has specific preferences with respect to alternative carbon sources and can switch between them in a specific order (diauxic shift) [29]. Unlike experimental procedures based on the use of IPTG, induction of protein expression is automatic and thus bacterial cell growth does not need to be monitored. Furthermore, to achieve sustained bacterial growth during the induction phase, glycerol is provided in auto-inducing media together with lactose [28, 30]. NAB containing the above reported doses of glucose, glycerol and lactose has been termed auto-inducing NAB (iNAB).

To set up the auto-induction procedure, bacteria were transformed with pET recombinant plasmids, and grown in 50 mL of standard NAB containing 100 µg/mL ampicillin, at 37 °C overnight. Bacterial cultures were then used to inoculate 1 L of ampicillin containing iNAB. The culture was then incubated at 37 °C for 24 h, in order to allow automatic induction of chimeric protein expression. Following incubation, cells were harvested by centrifugation and lysed in order to analyze recombinant construct expression by 15% SDS-PAGE (**Figure 1 a-c**). Significantly higher expression levels were observed for all the constructs under test (**Figure 1 a-c**, and **Table 2**), which were found to be 1.5-2 times higher than those observed in NAB under IPTG induction (**Table 2**). However, it should be noted that saturation densities ($OD_{600\text{ nm}}$ values at the end of the experiment) were significantly lower than those observed in the “classical”

protocol (NAB with the addition of IPTG). This is in agreement with previous findings indicating that, in shake flask cultivations, lactose auto-induction generally occurs in the active growth phase, with a possible consequent inactivation of bacterial cell growth [28, 31].

Scale up of the production procedure

In order to scale up the production procedure, a fermentation process was set up using a 5 L fermenter. As before, the production procedure was optimized by using classical NAB and IPTG induction, and then the recombinant construct expression yields were compared with those obtained in iNAB, in order to set up a competitive and cost-effective process. Transformed bacterial cells were inoculated into 150 mL of standard NAB containing 100 µg/mL of ampicillin and 4 g/L glucose, and incubated at 37 °C overnight. Bacterial cultures were then diluted into 3 L (dilution 1:20 v/v) of NAB into a 5 L batch reactor. Since in stirred bioreactors a higher gas–liquid oxygen transfer is generally achieved than in shake flasks [31], oxygen was not expected to be a limiting factor. Oxygen saturation was maintained above 60% throughout the cultivation. Throughout the fermentation process, the stirring speed was kept constant at 300 rpm, with an air flow rate of 10 cm³/sec. To overcome drawbacks associated with ampicillin degradation, the antibiotic was added to bacterial cells at a rate of 100 mg/h using a fermenter peristaltic pump. In fact, a fundamental problem associated with the use of ampicillin resistance as a selection mechanism is the rapid degradation of the antibiotic by the extracellular enzyme β-lactamase [32, 33]. For the BL21(DE3) host strain containing the pET-type vectors, the half-life of 100 µg/mL ampicillin is reported to be 30 min for mid-log phase cells grown in a fermenter [34]. By regularly adding

ampicillin, it was possible to prevent plasmid loss, with consequent improvement of expression yields. The fermentation process was carried out at 37 °C up to OD_{600 nm} of about 4. At this stage, expression of recombinant constructs was induced by addition of 0.7 mM IPTG. Fermentation was carried out for 24 h. At the end of the fermentation process, cells were collected by using a peristaltic pump, harvested by centrifugation, and lysed in order to analyze recombinant constructs expression by 15% SDS-PAGE (**Figures 2-4**). In all cases, expression of recombinant chimeric constructs began at about 2-4 h after IPTG addition, reaching a maximum at 8 h (**Figures 2-4 a, c, e**), and expression levels were found to be significantly higher than those observed in shake flask cultures carried out in NAB upon IPTG induction (**Figure 1** and **Table 2**). Instead, saturation densities (OD_{600 nm} values at the end of fermentation process) were similar to those obtained in shake flasks cultures (Table 2). Biomass was monitored on the basis of cell dry weight and represented as a curve (**Figures 2-4 e**), which was similar to that representing the yield (g/L) of recombinant chimeric constructs (Fig. 2-4 e).

Once the validity of the experimental strategy on a larger scale was verified, it was applied to the auto-induction procedure, in order to develop a cost-effective production process suitable for industrial purposes. The same experimental procedure described above was performed with the exception that *E. coli* BL21(DE3) cells, once transformed with recombinant pET-22b(+) expression vectors, were incubated overnight at 37 °C in 150 mL of NAB supplemented with ampicillin (0.1 mg/mL), but lacking glucose, glycerol and lactose. Following incubation, cell cultures were diluted 1:20 (v/v) in 3 L of iNAB into a 5 L batch reactor. Experimental conditions identical to those described above were applied, with the exception of IPTG addition, since

recombinant expression of chimeric constructs was automatically induced by lactose present in the medium. At defined time intervals, bacterial aliquots (corresponding to 0.125 OD_{600 nm}) were withdrawn and lysed to determine recombinant constructs expression levels by 15% SDS-PAGE (**Figures 2-4 b, d, f**). Recombinant expression of chimeric constructs was found to begin at 2-4 h, indicating that, at this time, glucose probably decreased and lactose began to be metabolized, with a consequent inducing effect. In these experimental conditions, expression levels were found to reach a maximum at 24 h. In all cases, recombinant construct expression levels were significantly higher than those obtained either under the same experimental conditions using classical NAB (**Figure 2-4 and Table 2**) or in shake flask cultures by using iNAB (**Table 2**). Similar results were also obtained by analyzing saturation densities (OD_{600 nm} values at the end of fermentation process), which were higher than those obtained under the same experimental conditions using classical NAB under IPTG induction for all the constructs, with the exception of ONC-r(P)GKY20. In the last case, biomass growth in batch reactor was found to be similar in both experimental conditions (**Table 2**).

It can be highlighted that the set up procedure is of general validity and may be applied to any recombinant soluble protein expressed in *E. coli*. To demonstrate this, the auto-induction protocol was applied to the production of recombinant Apolipoprotein A-I (ApoA-I). Soluble ApoA-I expression levels were significantly higher in iNAB than in classical NZY medium upon IPTG induction in shake flasks (Supplementary Fig. S5). In the case of batch reactor cultivations, ApoA-I expression levels were highest after 8 h

growth and decreased slightly after 24 h (Supplementary Fig. S6), in agreement with data showing ApoA-I instability in bacterial expression systems [35].

Estimation of HDP production costs from laboratory to larger scale

The procedure to isolate pure recombinant HDPs has been described previously [12, 14, 15]. Briefly, upon recovery from inclusion bodies, the chimeric protein was purified by nickel affinity chromatography (**Figure 5**). The peptide was then released from the carrier onconase by hydrolysis in acidic conditions. Since the carrier is insoluble at neutral or alkaline pH, the peptide is isolated from insoluble components by repeated cycles of centrifugation, and finally lyophilized. By applying this procedure to fermentation process carried out in iNAB, peptide purity was 99% with a yield of about 40-50 mg/L. A final gel-filtration step was added, in order to remove salts used along the purification process and that tend to attach to the peptides [36]. This step is responsible for the loss of a significant amount of peptide (~50%), with a consequent final yield of about 20-25 mg of pure peptide *per* L of bacterial culture.

Figure 6 reports the breakdown of peptide unit cost (€/mg) as a function of the production scale expressed as mg/batch, in the case of the fermentation process performed in iNAB. It can be seen that, at a small production scale (5L fermenter), the cost is about 250 €/mg and is mainly determined by the labour cost and by the capital investment due to main equipment costs and additional fixed costs required to maintain equipment. When the production scale is increased up to 200, 500 and 1,000 mg/batch, the unit cost decreases sharply to 136, 65 and 42 €/mg, respectively. In the last case (1,000 mg/batch), labour and chemical costs become the only relevant factors, which is

typical of a production scale [24]. In order to identify possible process bottlenecks, costs have also been analyzed by considering the contribution of each step in the process (fermentation, extraction and purification) (**Figure 7**). More than the half of the costs derive from the final purification step and this contribution increases with increasing production scale. The labour costs, in contrast, decrease by increasing production scale. However, at 1 g/batch scale, labour still accounts for 1/3rd of the whole unit production cost, and the purification process still accounts for more than the half. To further reduce the unit cost, it would be advisable to shorten the schedule and the total time of the procedures associated with the purification step or to automate them. Indeed, a reduction of 50% of the whole labour task in the purification step could lead to a further decrease of the unit cost to 35 €/mg. The demand for chemicals, required to solubilise inclusion bodies and to perform expanded bed chromatography, strongly affects the overall large scale process costs. Indeed, chemicals demand is responsible for up to 1/3rd of the unit cost at a production scale of 1 g/batch. In particular, the requirement for guanidine accounts for more than 80% of the chemical costs and, consequently, is responsible for more than 25% of the whole peptide unit cost. Hence, alternative methods to solubilize inclusion bodies should be tested in the future. It can also be highlighted that determined unit costs appear really competitive when considering the cost of chemical synthesis (**Table 3**). Indeed, those reported in **Table 3** do not take into account trifluoroacetate (TF-acetate) exchange by another counter-ion, a step often required for synthetic cationic peptides, which are mainly obtained as TF-acetate salts [35]. Since TF-acetate strongly affects *in vitro* and *in vivo* studies [37-39], its removal is often necessary, with a consequent increase of chemical synthesis costs of about € 200.

This provides a further key element attesting the competitiveness of the production procedure developed here.

Discussion

The past three decades have seen a rise in regulatory approval and sales of protein- and peptide-based therapeutics. To explore the pharmaceutical and therapeutic potential of HDPs, a cost-effective and scalable method to produce active and effective peptides is still required. Here, we set up a cost-effective procedure to produce high amounts of HDPs by developing a novel growth medium, NAB, comprising cheap, readily available components and defined amounts of each one. Among culture media, Luria-Bertani (LB) broth is the most commonly used for culturing *E. coli*, being rich in nutrients and with an optimal osmolarity for *E. coli* growth at early log phase. However, it is generally responsible for a cessation of bacterial cell growth at relatively low densities, mainly due to the presence of scarce amounts of carbohydrates and divalent cations [41]. Not surprisingly, when the amount of peptone or yeast extract is increased, higher cell densities are reached [28]. Similarly, divalent cation supplementation results in higher cell growth [28]. Glucose supplementation is of limited help because acid generation by glucose metabolism overwhelms the limited buffer capacity of LB [42, 43]. If culture acidification poses a problem, the media can be buffered with phosphate salts. Indeed, this is the case of 2xYT, TB (Terrific Broth) and SB (Super Broth) media, which have been shown to be superior to LB to reach higher cell densities [28, 44].

NAB, when applied to the production of the recombinant host defence peptides under test, was found to support significantly higher cell densities and recombinant expression levels than conventional culture media. The production procedure was also efficiently scaled up and procedure costs significantly lowered by developing a successful auto-induction strategy. By this means, recombinant constructs expression was found to be

mediated by a common and cheap sugar, such as lactose, instead of IPTG. Indeed, even if IPTG is generally required at submillimolar concentrations, its use might result in high cost accumulation for industrial purposes. IPTG substitution was also found to have further advantages, since this compound is not innocuous for bacterial cells [45, 46]. Moreover, the auto-induction procedure was found to support significantly higher yields of target constructs than conventional strategies based on expression induction mediated by IPTG. When the procedure was scaled up by using a 5 L fermenter, expression levels of recombinant constructs (about 2-3 g/L) were found to be significantly higher than those obtained in shake flask cultures (about 1-2 g/L). A further main advantage of this procedure is the possibility of isolating the peptide of interest from its carrier protein by acidic conditions because of the presence of a dipeptide Asp-Pro between the peptide and the onconase moieties. This allows avoidance of the use of specific proteases [47], which would have made the production process too expensive and not suitable for industrial purposes. Moreover, the set up procedure is of general validity for application to recombinant soluble protein expressed in *E. coli*, as demonstrated by applying the auto-induction protocol to the production of recombinant soluble ApoA-I.

Since the protocol to purify HDPs was previously optimized and found to provide high yields of pure peptides through a simple purification procedure (Fig. 5) [12, 14, 15], it is tempting to speculate that this optimized procedure might be successfully scaled up at an industrial scale. Thus, we estimated that by increasing the production scale to 200, 500 and 1,000 mg/batch, unit costs fall to 136, 65 and 42 €/mg, respectively. These values appear highly competitive when considering chemical synthesis costs of the

same peptides (Table 3). Indeed, they also appear competitive against the costs of a different linear helical peptide containing 37 amino acid residues, *i.e.* LL-37 human multifunctional cathelicidin-type host defence peptide endowed with antibacterial, antiviral, and immunomodulatory activities. Indeed, the cost of synthetic human LL-37 peptide is about 400 €/mg, whereas the cost of recombinant protein precursor is about 1,400 \$/mg. Altogether, these findings indicate that the herein set up production strategy appears competitive.

Acknowledgements

This study was partially supported by the Italian Cystic Fibrosis Foundation (grant number 20/2014). We are deeply indebted to volunteers who devote many efforts in fundraising, in particular delegations from Palermo and Ragusa Vittoria Catania 2.

Conflict of interest

The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version of the manuscript.

References

1. Hancock REW, Haney EF, Gill EE. The immunology of host defence peptides: beyond antimicrobial activity. *Nature Reviews Immunology* 2016;16:321–34.
2. Andersson L, Blomberg L, Flegel M, Lepsa L, Nilsson B, Verlander M. Large-scale synthesis of peptides. *Biopolymers* 2000;55:227–50.
3. Alvarez-Sieiro P, Montalbán-López M, Mu D, Kuipers OP. Bacteriocins of lactic acid bacteria: extending the family. *Appl Microbiol Biotechnol.* 2016;100:2939-2951.
4. Ongey EL, Neubauer P. Lanthipeptides: chemical synthesis versus in vivo biosynthesis as tools for pharmaceutical production. *Microb Cell Fact* 2016; 15:97.
5. Ingham AB, Moore RJ. Recombinant production of antimicrobial peptides in heterologous microbial systems. *Biotechnol Appl Biochem* 2007;47:1–9.
6. Li Y, Chen Z. RAPD: a database of recombinantly-produced antimicrobial Peptides. *FEMS Microbiol Lett.* 2008;289:126–29.
7. Li Y. Carrier proteins for fusion expression of antimicrobial peptides in *Escherichia coli*. *Biotechnol Appl Biochem.* 2009;54:1–9.
8. Li Y. Recombinant production of antimicrobial peptides in *Escherichia coli*: a review. *Protein Expr Purif.* 2011;80:260–67.
9. Di Gaetano S, Guglielmi F, Arciello A, Mangione P, Monti M, Pagnozzi D, et al. Recombinant amyloidogenic domain of ApoA-I: analysis of its fibrillogenic potential. *Biochem Biophys Res Commun.* 2006;351:223-28.
10. Guglielmi F, Monti DM, Arciello A, Torrassa S, Cozzolino F, Pucci P, et al. Enzymatically active fibrils generated by the self-assembly of the ApoA-I

- fibrillogenic domain functionalized with a catalytic moiety. *Biomaterials* 2009;30:829-35.
11. Vassilevski AA, Kozlov SA, Grishin EV. Antimicrobial peptide precursor structures suggest effective production strategies. *Recent Pat Inflamm Allergy Drug Discov.* 2008;2:58–63.
 12. Pane K, Durante L, Pizzo E, Varcamonti M, Zanfardino A, Sgambati V, et al. Rational Design of a Carrier Protein for the Production of Recombinant Toxic Peptides in *Escherichia coli*. *PLoS One* 2016;11:e0146552.
 13. Notomista E, Cafaro V, Fusiello R, Bracale A, D'Alessio G, Di Donato A. Effective expression and purification of recombinant onconase, an antitumor protein. *FEBS Lett.* 1999;463:211–15.
 14. Pane K, Sgambati V, Zanfardino A, Smaldone G, Cafaro V, Angrisano T, et al. A new cryptic cationic antimicrobial peptide from human apolipoprotein E with antibacterial activity and immunomodulatory effects on human cells. *FEBS J.* 2016;283:2115-31.
 15. Gaglione R, Dell'Olmo E, Bosso A, Chino M, Pane K, Ascione F, et al. Novel human bioactive peptides identified in apolipoprotein B: evaluation of their therapeutic potential. *Biochem Pharmacol.* 2017;130:34-50.
 16. Papareddy P, Rydengård V, Pasupuleti M, Walse B, Mörgelin M, Chalupka A, et al. Proteolysis of human thrombin generates novel host defense peptides. *PLoS Pathog.* 2010;6:e1000857.
 17. Pane K, Durante L, Crescenzi O, Cafaro V, Pizzo E, Varcamonti M, et al. Antimicrobial potency of cationic antimicrobial peptides can be predicted from

- their amino acid composition: Application to the detection of "cryptic" antimicrobial peptides. *J Theor Biol.* 2017;419:254-65.
18. Green MR, Hughes H, Sambrook J, MacCallum P. *Molecular Cloning: A Laboratory Manual*. 4th edn. Cold Spring Harbor Laboratory Press, New York; 2012.1890 p.
 19. Petrlova J, Duong T, Cochran MC, Axelsson A, Mörgelin M, Roberts LM, et al. The fibrillogenic L178H variant of apolipoprotein A-I forms helical fibrils. *J Lipid Res.* 2012;53:390–8.
 20. Del Giudice R, Arciello A, Itri F, Merlino A, Monti M, Buonanno M, et al. Protein conformational perturbations in hereditary amyloidosis: Differential impact of single point mutations in ApoAI amyloidogenic variants. *Biochim Biophys Acta* 2016;1860:434-44.
 21. Gaglione R, Smaldone G, Di Girolamo R, Piccoli R, Pedone E, Arciello A. Cell milieu significantly affects the fate of AApoAI amyloidogenic variants: predestination or serendipity? *Biochim Biophys Acta* 2017; 1862:377-384.
 22. Kalk JP, Langlykke AF. Cost estimation for biotechnology projects. In: A.L. Demain and N.A. Solomon (eds.) *Manual of Industrial Microbiology and Biotechnology*, Washington, DC: American Society for Microbiology; 1986. p. 363-385.
 23. Heinzle E, Biber A, Cooney C. *Development of Sustainable Bioprocesses*, John Wiley & Sons, West Sussex, England; 2006. 316 p.
 24. Huang CJ, Lin H, Yang X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol.* 2012;39:383-99.

25. Banerjee S, Paul S, Nguyen LT, Chu BC, Vogel HJ. FecB, a periplasmic ferric-citrate transporter from *E. coli*, can bind different forms of ferric-citrate as well as a wide variety of metal-free and metal-loaded tricarboxylic acids. *Metallomics* 2016;8:125-33.
26. Sleator RD, Hill C. Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiol Rev.* 2002;26:49-71.
27. Krause M, Ukkonen K, Haataja T, Ruottinen M, Glumoff T, Neubauer A, et al. A novel fed-batch based cultivation method provides high cell-density and improves yield of soluble recombinant proteins in shaken cultures. *Microb Cell Fact.* 2010;9:11.
28. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif.* 2005;41:207-34.
29. Mostovenko E, Deelder AM, Palmblad M. Protein expression dynamics during *Escherichia coli* glucose-lactose diauxie. *BMC Microbiol.* 2011;11:126.
30. Blommel PG, Becker KJ, Duvnjak P, Fox BG. Enhanced bacterial protein expression during auto-induction obtained by alteration of lac repressor dosage and medium composition. *Biotechnol Prog.* 2007;23:585-98.
31. Mayer S, Junne S, Ukkonen K, Glazyrina J, Glauche F, Neubauer P, et al. Lactose autoinduction with enzymatic glucose release: characterization of the cultivation system in bioreactor. *Protein Expr Purif.* 2014;94:67-72.
32. Neubauer P, Hofmann K, Holst O, Mattiasson B, Kruschke P. Maximizing the expression of a recombinant gene in *Escherichia coli* by manipulation of induction time using lactose as inducer. *Appl Microbiol Biotechnol.* 1992;36:739-44.

33. Jung G, Denèfle P, Becquart J, Mayaux JF. High-cell density fermentation studies of recombinant *Escherichia coli* strains expressing human interleukin-1 beta. *Ann Inst Pasteur Microbiol.* 1988;139:129-46.
34. Hoffman BJ, Broadwater JA, Johnson P, Harper J, Fox BG, Kenealy WR. Lactose fed-batch overexpression of recombinant metalloproteins in *Escherichia coli* BL21 (DE3): process control yielding high levels of metal-incorporated, soluble protein. *Protein Expr Purif.* 1995;6:646-54.
35. Ryan RO, Forte TM, Oda MN. Optimized bacterial expression of human apolipoprotein A-I. *Protein Expr Purif.* 2003;27:98-103.
36. Bommarius B, Jenssen H, Elliott M, Kindrachuk J, Pasupuleti M, Gieren H, et al. Cost-effective expression and purification of antimicrobial and host defense peptides in *Escherichia coli*. *Peptides* 2010;31:1957-65.
37. Roux S, Zekri E, Rousseau B, Paternostre M, Cintrat JC, Nicolas F. Elimination and exchange of trifluoroacetate counter-ion from cationic peptides: a critical evaluation of different approaches. *J Pept Sci.* 2008;14:354–9.
38. Pini A, Lozzi L, Bernini A, Brunetti J, Falciani C, Scali S, et al. Efficacy and toxicity of the antimicrobial peptide M33 produced with different counter-ions. *Amino Acids* 2012;43:467–73.
39. Andrushchenko VV, Vogel HJ, Prenner EJ. Optimization of the hydrochloric acid concentration used for trifluoroacetate removal from synthetic peptides. *J Pept Sci.* 2007;13:37–43.
40. BCCResearch. Global Markets and Manufacturing Technologies for Protein Drugs 2016 (cited 2016 May). Available from:
http://www.researchandmarkets.com/research/4hf2x2/global_markets.

41. Sezonov G, Joseleau-Petit D, D'Ari R. *Escherichia coli* physiology in Luria-Bertani broth. *J Bacteriol.* 2007;189:8746-9.
42. Weuster-Botz D, Altenbach-Rehm J, Arnold M. Parallel substrate feeding and pH-control in shaking-flasks. *Biochem Eng J.* 2001;7:163-70.
43. Scheidle M, Dittrich B, Klinger J, Ikeda H, Klee D, Büchs J. Controlling pH in shake flasks using polymer-based controlled-release discs with pre-determined release kinetics. *BMC Biotechnol.* 2011;11:25.
44. Madurawe RD, Chase TE, Tsao EI, Bentley WE. A recombinant lipoprotein antigen against Lyme disease expressed in *E. coli*: fermentor operating strategies for improved yield. *Biotechnol Prog.* 2000;16:571-76.
45. Andrews KJ, Hegeman GD. Selective disadvantage of non-functional protein synthesis in *Escherichia coli*. *J Mol Evol.* 1976;8:317-28.
46. Malakar P, Venkatesh KV. Effect of substrate and IPTG concentrations on the burden to growth of *Escherichia coli* on glycerol due to the expression of Lac proteins. *Appl Microbiol Biotechnol.* 2012;93:2543-49.
47. Zhou QF, Luo XG, Ye L, Xi T. High-level production of a novel antimicrobial peptide perinerin in *Escherichia coli* by fusion expression. *Curr Microbiol.* 2007;54:366-70.

Figure legends

Fig. 1. Analysis by SDS-PAGE of total proteins extracted from non-induced (N.I., lane 2) and induced bacterial cells (I., lane 3) transformed with pET recombinant plasmids encoding ONC-r(P)ApoB_L (a), ONC-r(P)ApoB_S (b), or ONC-r(P)GKY20 (c) constructs in shake flasks. Densitometric analyses of protein bands with a molecular weight corresponding to that of ONC-r(P)ApoB_L (17.4 kDa), ONC-r(P)ApoB_S (16.1 kDa), or ONC-r(P)GKY20 (14.4 kDa) are reported as histograms in a, b, and c, respectively. In each panel, from left to the right, analysis of protein expression was performed in TB medium with IPTG induction, in NAB with IPTG induction and in iNAB, respectively. For the left gel of panel c, protein standards (lane 1), non-induced cells in NAB (lane 2), induced cells in NAB with IPTG (lanes 3 and 4), non-induced cells in TB (lane 5), induced cells in TB with IPTG (lanes 6 and 7). Molecular weights of expressed protein bands have been determined on the basis of calibration curves, prepared for each gel by plotting electrophoretic mobility (mm) of visible standard proteins as a function of known molecular weights.

Fig. 2. Analysis by SDS-PAGE of total proteins extracted from bacterial cells transformed with pET recombinant plasmid encoding ONC-r(P)ApoB_L construct (a, b). Analyses are representative of fermentation processes carried out either in classical NAB (panels a, c, and e) or in iNAB (panels b, d, and f). Samples were collected at different time intervals along the fermentation process in batch reactor. In (a), protein standards (lane 1), non-induced cells (lane 2), and cells after 2 h induction (lane 3), 4 h induction (lane 4), 6 h induction (lane 5), 8 h induction (lane 6), and 24 h induction (lane 7). In (b), protein standards (lane 1), cells at time 0 (lane 2), and after 2 h growth (lane 3), 4 h growth (lane 4), 6 h growth (lane 5), 8 h growth (lane 6), and 24 h

growth (lane 7). Molecular weight of expressed protein band has been determined on the basis of calibration curves, prepared for each gel by plotting electrophoretic mobility (mm) of visible standard proteins as a function of known molecular weights.

Densitometric analyses of protein bands with a molecular weight corresponding to that of ONC-r(P)ApoB_L construct (17.4 kDa) are reported as histograms (c, d). Chimeric construct expression levels (g/L) and biomass (cell dry weight, g/L) are reported as a function of time (h) in e and f.

Fig. 3. Analysis by SDS-PAGE of total proteins extracted from bacterial cells transformed with pET recombinant plasmid encoding ONC-r(P)ApoB_S construct (a, b). Analyses are representative of fermentation processes carried out either in classical NAB (panels a, c, and e) or in iNAB (panels b, d, and f). Samples were collected at different time intervals along the fermentation process in batch reactor. In a, protein standards (lane 1), non-induced cells (lane 2), after 2 h induction (lane 3), 4 h induction (lane 4), 6 h induction (lane 5), 8 h induction (lane 6) and 24 h induction (lane 7). In b, protein standards (lane 1), cells at time 0 (lane 2), after 2 h growth (lane 3), 4 h growth (lane 4), 6 h growth (lane 5), 8 h growth (lane 6), 9 h growth (lane 7) and 24 h growth (lane 8). Molecular weight of expressed protein band has been determined on the basis of calibration curves, prepared for each gel by plotting electrophoretic mobility (mm) of visible standard proteins as a function of known molecular weights. Densitometric analyses of protein bands with a molecular weight corresponding to that of ONC-r(P)ApoB_S construct (16.1 kDa) are reported as histograms (c, d). Chimeric construct expression levels (g/L) and biomass (cell dry weight, g/L) are reported as a function of time (h) in e and f.

Fig. 4. Analysis by SDS-PAGE of total proteins extracted from bacterial cells transformed with pET recombinant plasmid encoding ONC-r(P)GKY20 construct (a, b). Analyses are representative of fermentation processes carried out either in classical NA medium (panels a, c, and e) or in iNAB (panels b, d, and f). Samples were collected at different time intervals along the fermentation process in batch reactor. In a, protein standards (lane 1), non-induced cells (lane 2), after 2 h induction (lane 3), 4 h induction (lane 4), 6 h induction (lane 5), 8 h induction (lane 6), and 24 h induction (lane 7). In b, protein standards (lane 1), cells at time 0 (lane 2), after 2 h growth (lane 3), 4 h growth (lane 4), 6 h growth (lane 5), 8 h growth (lane 6) and 24 h growth (lane 7). Molecular weight of expressed protein band has been determined on the basis of calibration curves, prepared for each gel by plotting electrophoretic mobility (mm) of visible standard proteins as a function of known molecular weights. Densitometric analyses of protein bands with a molecular weight corresponding to that of ONC-r(P)GKY20 construct (14.4 kDa) are reported as histograms (c, d). Chimeric construct expression levels (g/L) and biomass (cell dry weight, g/L) are reported as a function of time (h) in e and f.

Fig. 5. Block flow diagram representing the production and purification process of recombinant HDPs.

Fig. 6. Breakdown of peptide unit production cost as a function of production scale.

Fig. 7. Impact of production process steps (fermentation, extraction and purification) on peptide unit production costs.

Table legends

Table 1. NAB composition.

Table 2. Expression levels of recombinant constructs and $OD_{600\text{ nm}}$ values of bacterial cultures before and after expression induction in TB and NAB by using IPTG as inducer, or at the end of the expression protocol by using iNAB. Results of representative experiments are reported.

Table 3. Quotations for peptide synthesis obtained from three companies. Company identities not disclosed.

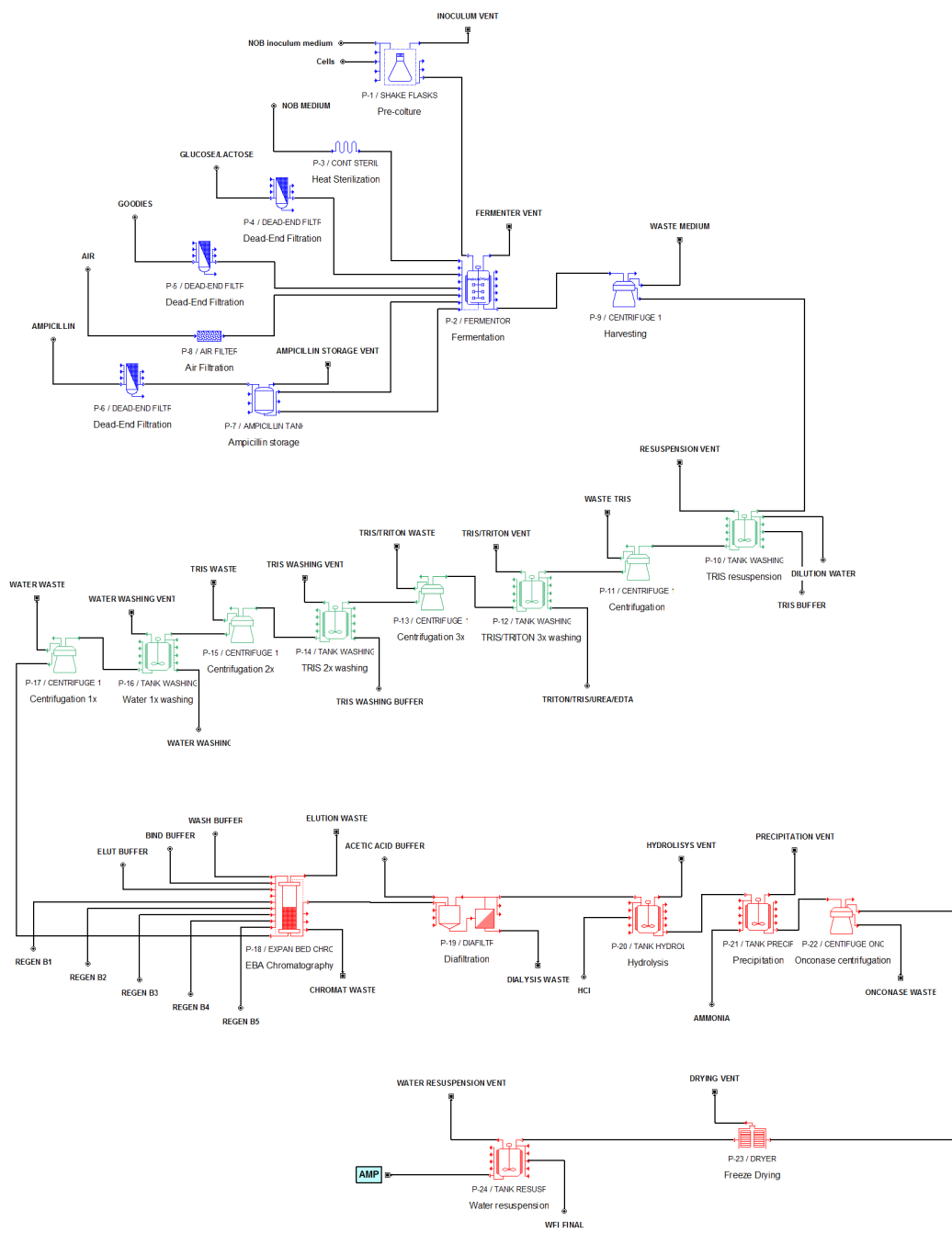
SUPPLEMENTARY DATA

Cost-effective production of recombinant peptides in *Escherichia coli*

Rosa Gaglione¹, Katia Pane², Eliana Dell'Olmo¹, Valeria Cafaro³, Elio Pizzo³, Giuseppe Olivieri^{4,*}, Eugenio Notomista³, Angela Arciello^{1,5,*}

¹Department of Chemical Sciences, University of Naples Federico II, 80126 Naples, Italy; ²IRCCS SDN, 80143 Naples, Italy; ³Department of Biology, University of Naples Federico II, 80126 Naples, Italy; ⁴Bioprocess Engineering, AlgaePARC, Wageningen University and Research, PO Box 16, 6700 AA Wageningen, The Netherlands; ⁵Istituto Nazionale di Biostrutture e Biosistemi (INBB), Italy.

***Correspondence:** Dr. Angela Arciello (email: anarciel@unina.it); Dr. Giuseppe Olivieri (email: giuseppe.olivieri@wur.nl).



Fermentation

Extraction

Purification

Figure S1. Flowsheet of the process illustrating the main sections and batch operations within each step of production.

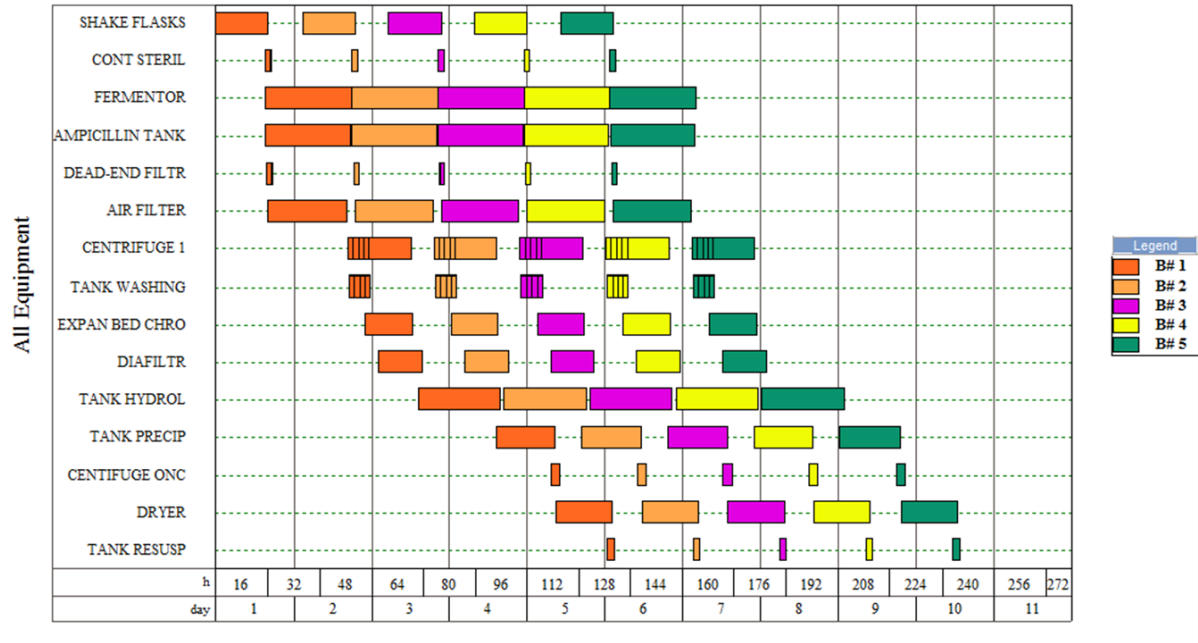


Figure S2. Detailed schedule of the batch operations within each step of the production process.

Table S1. Unit cost of electricity, labour and utilities referred to the Italian scenario of industrial production.

	Unit cost
Operator labor	26.3 € hr ⁻¹ *
Electricity	0.16 € kWh ⁻¹ **
Freon	0.18 € MT ⁻¹ ***
Chilled water	0.49 € MT ⁻¹ ***
Hot water	0.062 € MT ⁻¹ ***
Steam	14.8 € MT ⁻¹ ***

*source: Bureau of Labor Statistics - www.bls.gov; lumped cost considering employer contribution, supervision, operating supplies and overheads;

**source: Eurostat - http://ec.europa.eu/eurostat/statistics-explained/index.php/Energy_price_statistics;

***source: SuperPro databank; data are linearly scaled on the basis of electricity cost in USA and in Italy.

Table S2. CAPEX and OPEX estimation.

		<i>Cultivation</i>	
TOTAL CAPITAL INVESTMENT	Direct Cost (DC)	Major equipment cost	MEC
		Installation costs	50% MEC
		Instrumentation and control	50% MEC
		Piping	30% MEC
		Insulation	3% MEC
		Electrical	20% MEC
		Buildings	100% MEC
		Land improvements	15% MEC
		Service facilities	70% MEC
	Indirect Cost (IC)	Construction expenses	25% DC
Engineering and supervision		35% DC	
Other Cost (OC)	Contractor's fee	5% (DC+IC)	
	Contingency (Major equipment)	10% (DC+IC)	
Direct fixed capital cost (DFC)		DC+IC+OC	
Working capital		OPEX first month of operation	
Start-up and validation		5% DFC	
Total capital investment		DFC + working capital + Start-up and validation	
CAPEX	Depreciation	(DC+IC+OC)/10 years	
	Interest	8% of depreciation	
	Property tax	2% of depreciation + interest	
	Insurance	1% of depreciation + interest	
	Purchase tax	5% of depreciation + interest	
OPEX	Energy	Calculated from MEC consumption	
	Labor	Salaries + Employer's contribution (40%) + Operating Supplies (10%) + Supervision (20%) + Administration(60%)	
	Raw materials	Calculated from mass balances	
	Utilities	Calculated from MEC consumption and mass balances	
	Laboratory/Quality control/Quality assurance	15% Labor cost	
	Wastewater treatment	Calculated from mass balances	
	Consumables	Calculated from MEC design	
	Others	Maintenance	6% DFC
		Operating supplies	0.4% (Electricity + Raw materials + Utilities)
		Contingencies	15% (Raw materials + Utilities)
Overheads		55% (Labor + Maintenance)	

Table S3. Unit cost of chemicals.

Chemical	Cost*
Tryptone	Euro 94.35 (1 kg)
Betaine	Euro 29.80 (50 gr)
Ammonia	Euro 31.3 (5 l)
Citric acid	Euro 26.10 (1 kg)
Potassium dihydrogen phosphate	Euro 28.70 (500 gr)
Potassium monohydrogen Phosphate	Euro 83.34 (500 gr)
C-goodies:	
- Magnesium sulfate	Euro 76.59 (500 gr)
- Iron(II) sulfate heptahydrate	Euro 18.40 (250 gr)
- Magnesium oxide	Euro 66.60 (250 gr)
- Calcium carbonate	Euro 55.50 (500 gr)
- Zinc sulfate heptahydrate	Euro 44.40 (100 gr)
- Manganese sulfate monohydrate	Euro 44.82 (100 gr)
- Copper sulfate	Euro 45.60 (100 gr)
- Cobalt(II) sulfate hydrate	Euro 52.83 (100 gr)
- Boric acid	Euro 56.60 (500 gr)
- Nickel(II) chloride	Euro 64.44 (100 gr)
- Sodium molybdate	Euro 43.02 (100 gr)
Ampicillin	Euro 42.21 (5 gr)
Glucose	Euro 67.32 (1kg)
Lactose	Euro 35.01 (500 gr)
Glycerol	Euro 51.4 (2,5 l)
Trizma base	Euro 109 (500 gr)
Hydrochloric acid	Euro 20.20 (1 l)
Ethylenediaminetetraacetic acid	Euro 39.78 (500 gr)
Triton X-100	Euro 109 (500 gr)
Urea	Euro 27.60 (1 kg)
IMAC Sepharose [®] 6 Fast Flow	Euro 299 (25 ml)
Acetic acid	Euro 20.07 (1 l)
Guanidine hydrochloride	Euro 123.40 (1 kg)
Sodium acetate	Euro 52.11 (250 gr)
Sodium phosphate monobasic	Euro 38 (1 kg)
Disodium hydrogen phosphate	Euro 109 (500 gr)
Sodium chloride	Euro 84.15 (1 kg)
Imidazole	Euro 55.98 (100 gr)
Nickel(II) sulfate	Euro 52.29 (10 gr)

*Source: Sigma-Aldrich, PanReac AppliChem and GE Healthcare Europe GmbH

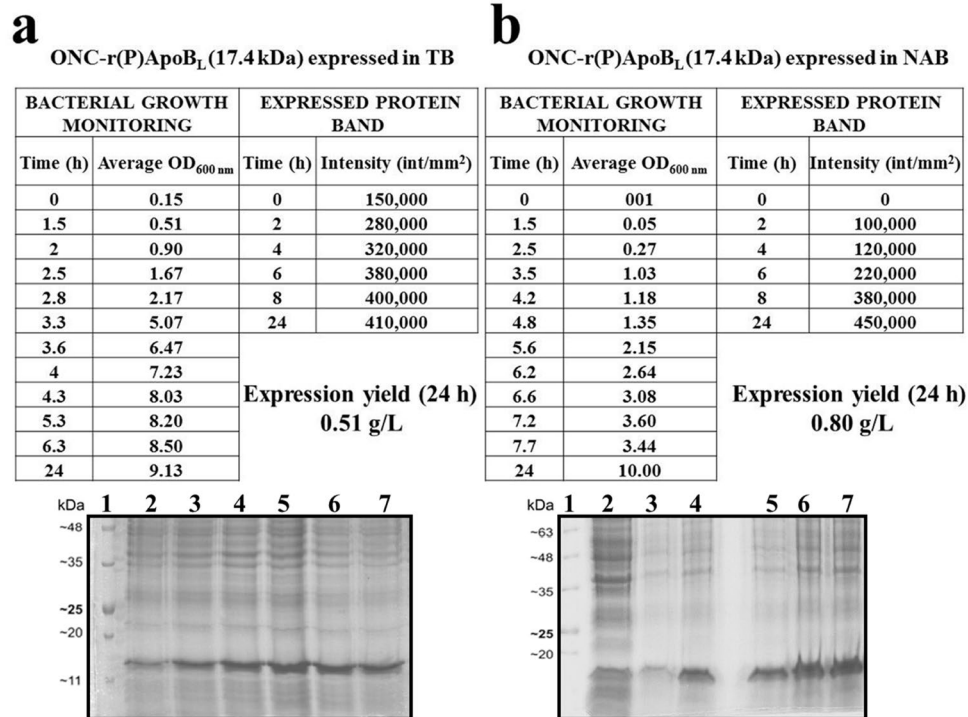


Figure S3. Analysis of ONC-r(P)ApoB_L construct expression levels in classical TB (a) and in NA medium (b). Analyses are representative of shake flask cultivations. Bacterial growth was monitored by determining optical density values at 600 nm at regular time intervals. Values for each time point are reported in the tables. Total proteins were extracted from bacterial cells collected at different time intervals, and analysed by SDS-PAGE. In the case of samples analysed in TB, protein standards (lane 1), not induced cells (lane 2), cells upon 2 h induction (lane 3), cells upon 4 h induction (lane 4), cells upon 6 h induction (lane 5), cells upon 8 h induction (lane 6), cells upon 24 h induction (lane 7). In the case of samples analyzed in NAB, protein standards (lane 1), not induced cells (lane 2), cells upon 2 h induction (lane 3), cells upon 4 h induction (lane 4), cells upon 6 h induction (lane 5), cells upon 8 h induction (lane 6), cells upon 24 h induction (lane 7). Molecular weight of expressed protein band has been determined on the basis of calibration curves, prepared for each gel by plotting electrophoretic mobility (mm) of visible standard proteins as a function of known molecular weights. Densitometric analyses of protein bands with a molecular weight corresponding to that of ONC-r(P)ApoB_L construct (17.4 kDa) are reported in the tables.

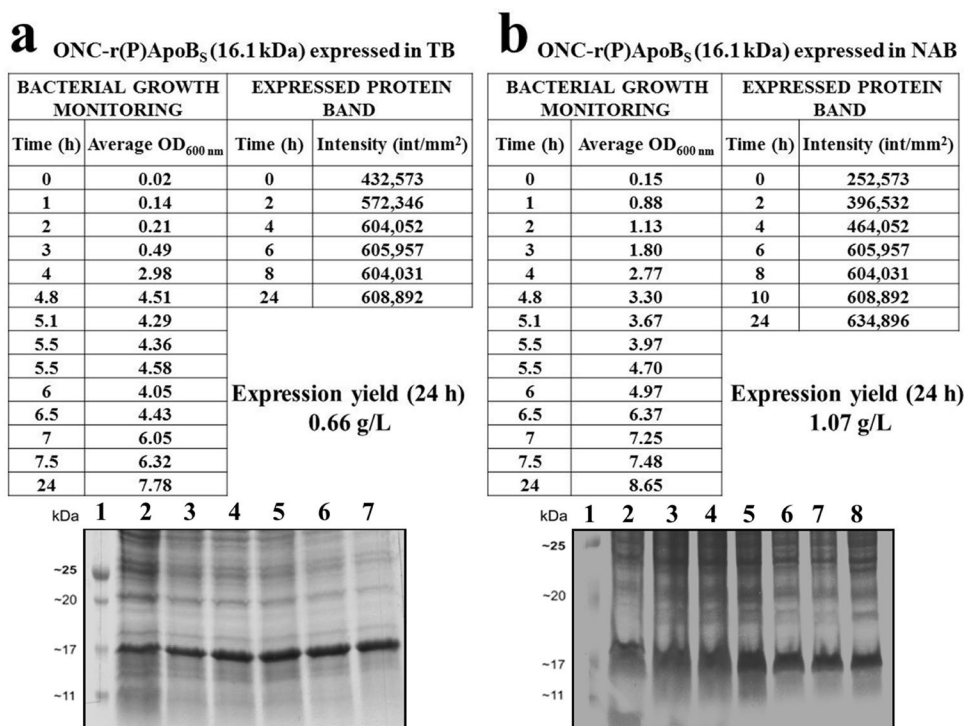


Figure S4. Analysis of ONC-r(P)ApoB_s construct expression levels in classical TB (a) and in NA medium (b). Analyses are representative of shake flask cultivations. Bacterial growth was monitored by determining optical density values at 600 nm at regular time intervals. Values for each time point are reported in the tables. Total proteins were extracted from bacterial cells collected at different time intervals, and analysed by SDS-PAGE. In the case of samples analysed in TB, protein standards (lane 1), not induced cells (lane 2), cells upon 2 h induction (lane 3), cells upon 4 h induction (lane 4), cells upon 6 h induction (lane 5), cells upon 8 h induction (lane 6), cells upon 24 h induction (lane 7). In the case of samples analyzed in NAB, protein standards (lane 1), not induced cells (lane 2), cells upon 2 h induction (lane 3), cells upon 4 h induction (lane 4), cells upon 6 h induction (lane 5), cells upon 8 h induction (lane 6), cells upon 10 h induction (lane 7), cells upon 24 h induction (lane 8). Molecular weight of expressed protein band has been determined on the basis of calibration curves, prepared for each gel by plotting electrophoretic mobility (mm) of visible standard proteins as a function of known molecular weights. Densitometric analyses of protein bands with a molecular weight corresponding to that of ONC-r(P)ApoB_s construct (16.1 kDa) are reported in the tables.

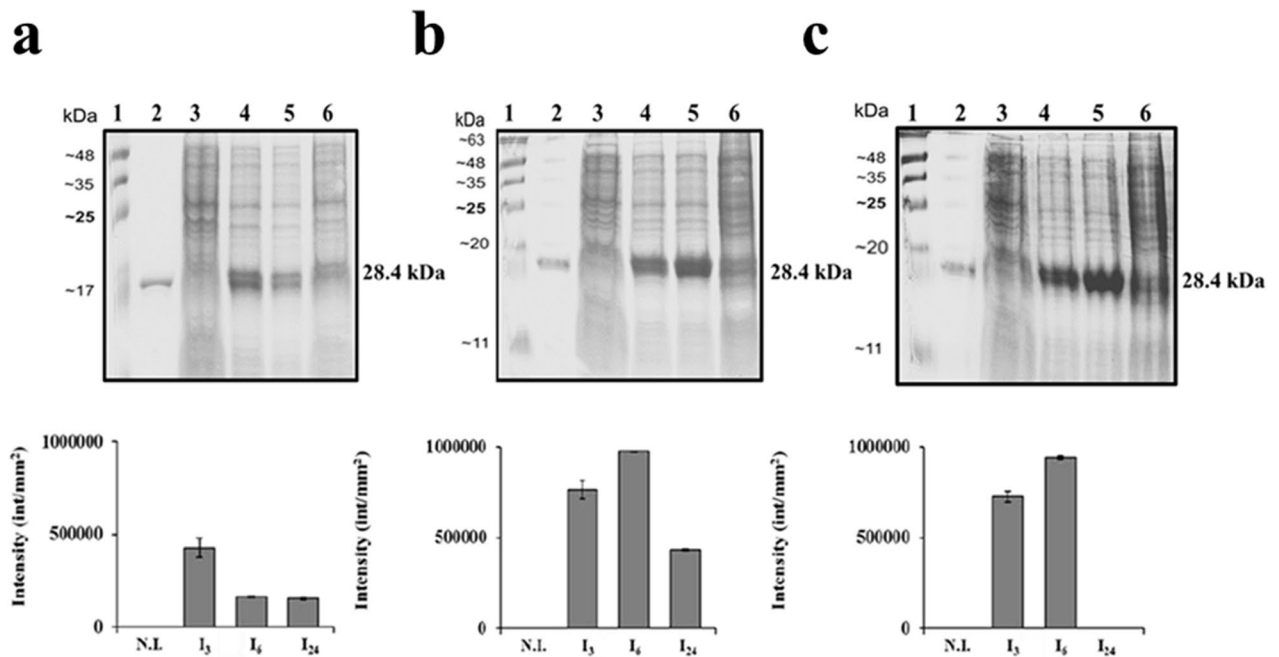


Figure S5. Analysis by SDS-PAGE of total proteins extracted from not induced (N.I.) and induced (I) bacterial cells transformed with pET recombinant plasmid encoding ApoA-I and cultured in shake flasks. Analysis of protein expression was performed in NZY medium upon IPTG induction (a), in NA medium upon IPTG induction (b), and in auto-inducing NA medium (c). In a and b, protein standards (lane 1), pure recombinant apolipoprotein A-I (2 μ g, lane 2), not induced cells (lane 3), cells upon 3 h induction (lane 4), cells upon 6 h induction (lane 5), cells upon 24 h induction (lane 6). In c, protein standards (lane 1), pure recombinant apolipoprotein A-I (2 μ g, lane 2), not induced cells (lane 3), cells upon 3 h growth (lane 4), cells upon 6 h growth (lane 5), cells upon 24 h growth (lane 6). Molecular weight of expressed protein band has been determined on the basis of calibration curves, prepared for each gel by plotting electrophoretic mobility (mm) of visible standard proteins as a function of known molecular weights. Densitometric analyses of protein bands with a molecular weight corresponding to that of ApoA-I (28.4 kDa) are reported as histograms.

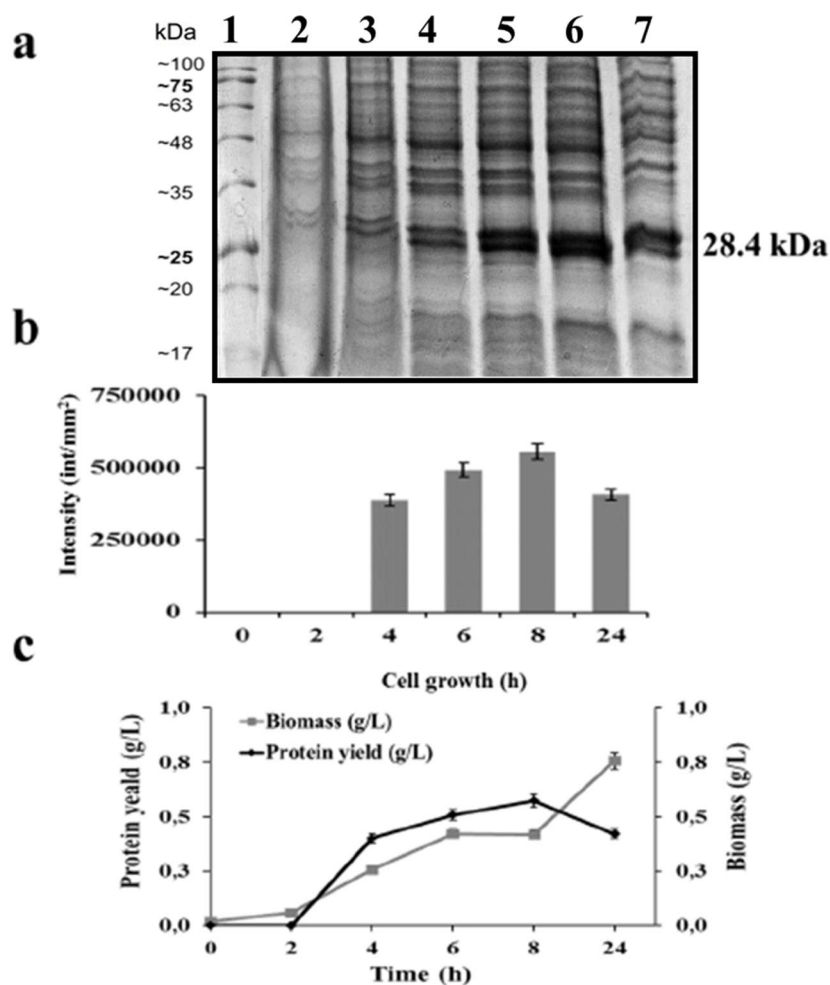


Figure S6. Analysis by SDS-PAGE of total proteins extracted from bacterial cells transformed with pET recombinant plasmid encoding ApoA-I. Analyses are representative of fermentation processes carried out in auto-inducing NA medium. Samples were collected at different time intervals along the fermentation process in batch reactor. In a, protein standards (lane 1), not induced cells (lane 2), cells upon 2 h growth (lane 3), cells upon 4 h growth (lane 4), cells upon 6 h growth (lane 5), cells upon 8 h growth (lane 6), cells upon 24 h growth (lane 7). Molecular weight of expressed protein band has been determined on the basis of calibration curves, prepared for each gel by plotting electrophoretic mobility (mm) of visible standard proteins as a function of known molecular weights. Densitometric analyses of protein bands with a molecular weight corresponding to that of ApoA-I (28.4 kDa) are reported as histograms (b). ApoA-I expression levels (g/L) and biomass (cell dry weight, g/L) are reported in c.