scientific reports

Check for updates

OPEN

# Taguchi L9 optimization of *Bs*PdaC-CD production in *Escherichia coli*

Marc Carnicer[1,2]✉, Pablo Haasnoot[1,2], Mònica Texidó[1] & Antoni Planas[1,2,3]✉

**The production of sequence-defined partially acetylated chitooligosaccharides (paCOS) is a promising strategy for the valorization of chitin-rich waste into high-value bioactive compounds. The catalytic domain of a deacetylase from *Bacillus subtilis* (*Bs*PdaC-CD), known to act on N-acetylglucosamine (GlcNAc) residues of chitooligosaccharides, represents a promising biocatalyst for paCOS production. In this work, a fed-batch process to produce the recombinant *Bs*PdaC-CD in *Escherichia coli* BL21 was developed and optimized using a Taguchi L9 ($3^4$) design of experiments. Across nine independent fed-batch processes, *Bs*PdaC-CD titers and space-time yield (STY) varied widely, ranging from 167 to 1197 mg *Bs*PdaC-CD/L and 2.2 to 22.5 mg *Bs*PdaC-CD/(L*h), respectively. Temperature and feed rate were identified as the most influential factors affecting these outcomes. In both cases, the main variability was observed at low feeding rate and temperature, 5 mL/L/h and 22 °C respectively, which showed lower *Bs*PdaC-CD productions. Subsequent experiments at 10 mL/L/h and 30 °C confirmed the robustness of the optimized conditions, consistently yielding 1.2 g/L of *Bs*PdaC-CD within 50 h. These results highlight the utility of Taguchi-based experimental design for rapid identification of critical parameters in recombinant protein bioprocess optimization.**

**Keywords** Chitin, Chitooligosaccharides (COS), Taguchi, Design of experiments, Bioprocess optimization, Recombinant protein production

**Abbreviations**

| | |
|---|---|
| COS | Chitooligosaccharides |
| paCOS | Partially acetylated chitooligosaccharides |
| DP | Degree of polymerization |
| FA | Fraction of acetylation |
| PA | Pattern of acetylation |
| GlcNAc | N-acetylglucosamine |
| GlcN | Glucosamine |
| CDA | Chitin deacetylase |
| DoE | Design of experience |
| *Bs*PdaC-CD | Catalytic domain of a deacetylase from *Bacillus subtilis* |
| STY | Space-time yield |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| WCW | Wet cell weight |
| DCW | Dry cell weight |

Chitin is the second most abundant organic compound in nature just after cellulose. It is a water-insoluble amino-polysaccharide composed of β-1,4-linked N-acetylglucosamine units that is commonly found in insect cuticles, fungal cell walls and crustacean shells[1]. Among these, crustacean shells are a particularly relevant source, as the sea-food industry generates approximately 106 tons of waste annually, most of which is discarded or processed into low-value products such as animal feed or fertilizer[2]. A more sustainable and valuable use of this shell waste involves its conversion into high-value compounds such as purified chitin and its derivatives. However, one of the main limitations of chitin is its lack of solubility in water. To overcome this limitation, chitin is frequently converted into more soluble derivatives including chitosan, chitooligosaccharides (COS) and partially acetylated chitooligosaccharides (paCOS), which have diverse industrial and biomedical application[3].

Chitin depolymerization by chemical or enzymatic processes render chitin-oligosaccharide[1]. The resulting short-chain oligomers, typically with a degree of polymerization (DP) of less than 20 and an average molecular

[1]Laboratory of Biochemistry, Institut Químic de Sarrià, University Ramon Llull, 08017 Barcelona, Spain. [2]Bioprocess Pilot Plant, Institut Químic de Sarrià, University Ramon Llull, 08017 Barcelona, Spain. [3]Chemistry Section, Royal Academy of Sciences and Arts of Barcelona, Barcelona, Spain. ✉email: marc.carnicer@iqs.url.edu; antoni.planas@iqs.url.edu

weight below 3.9 kDa, are known as COS. Deacetylation of chitin leads to chitosans with variable amounts of glucosamine units (GlcN). In contrast to chitin, chitosans are soluble in slightly acidic aqueous solutions due to their high percentage of GlcN, expressed as degree of deacetylation (DDA) or degree of acetylation (DA). Their polycationic nature, due to partial protonation of the free amino groups at physiological pH, enables them to interact with polyanionic biomolecules such as membrane phospholipids, DNA and proteins.

Depolymerization of chitosans or deacetylation of COS gives partially acetylated chitooligosaccharides (paCOS)[1]. These molecules are defined by three structural parameters: the degree of polymerization (DP), the degree of acetylation (DA) and the pattern of acetylation (PA), the latter describes the specific sequence of GlcNAc and GlcN residues along the oligomer chain[2]. Despite the wide range of potential applications described for COS and paCOS[3], many of these biological activities have shown significant inconsistencies strongly linked to the structural features of the oligomers (DP, DA and PA). Although DP and DA can be controlled to generate well-defined second-generation COS[2], the role of PA has recently gained considerable attention. For example, specific PA motifs in chitosan oligomers have been shown to play a key role in priming immune responses in rice plants without triggering full defense activation[4].

As mentioned, traditional production of paCOS relies on chemical depolymerization and deacetylation of chitin, which typically generates complex mixtures of oligomers with heterogeneous DP and DA. Although these mixtures can be separated based on these two parameters, PA cannot be selectively controlled or resolved by this approach[5]. Moreover, chemical methods tends to be harsh and complex processes, often involving the use of noxious chemicals and generating environmentally harmful waste[3].

As an alternative, enzymatic synthesis has emerged as a more promising method due to their higher specificity and milder reaction conditions. Chitinases and chitosanases enable controlled depolymerization and improved control of DP compared with chemical methods, although the resulting product distributions still depend strongly on enzyme specificity and substrate composition. To overcome these limitations, our research group has contributed to develop a bottom-up approach in which COS are first produced by engineered *E. coli* strains expressing rhyzobial chitin synthases to produce COS with defined DP from DP4 to DP7 depending on the enzyme and engineered active site mutants[6–8]. Next, enzymatic modification using chitin deacetylases (CDAs) enables the generation of sequence-defined paCOS with controlled acetylation patterns. In recent years, several CDAs with distinct regioselectivities have been identified and characterized expanding the available toolbox for paCOS synthesis[9,10].

Our research group has expressed and characterized a number of CDAs[9,11–13] and proposed the "subsite capping model" to rationalize substrate specificity[10]. More recently, some peptidoglycan deacetylases have also shown to be able to deacetylate COS, thus expanding the portfolio of enzymes potentially leading to new patterns of deacetylation. Particularly, we have identified and characterized a remarkable peptidoglycan deacetylase from *Bacillus subtilis*, the *Bs*PdaC. While its natural substrate is in peptidoglycan, acting on N-acetyl muramic acid (MurNAc) residues, this enzyme is also able to deacetylate COS following a distributive mechanism that preferentially targets central GlcNAc residues and ultimately leaves the reducing-end unit acetylated[12].

The aim of this study was to develop and optimize the recombinant production of the catalytic domain of *Bs*PdaC (*Bs*PdaC-CD) in *Escherichia coli* BL21 using a fed-batch bioprocess. Process optimization was performed using a Taguchi L9 ($3^4$) design of experiments to systematically evaluate key cultivation parameters and identify conditions enabling robust and efficient enzyme production[14].

## Methods

Analytical grade reagents were supplied by ITW Reagents if not stated otherwise.

### Bacterial strains and plasmids

*E. coli* BL21 (DE3) STAR (InVitrogen™ Thermoscientific; Massachusetts, USA) was utilized in this study. *E. coli* BL21 (DE3) STAR competent cells were prepared following established protocols. The plasmid pET22b_BsPdaC-CD, which harbors the isolated catalytic domain of the deacetylase *Bs*PdaC from *Bacillus subtitlis (Bs*PdaC-CD) and a C-terminal Strep-tag for protein purification, was obtained from previous research studies in the group[12]. Competent *E. coli* cells were transformed with pET22b_BsPdaC-CD using the conventional heat-shock protocol described elsewhere.

### Media and shake flask cultures

The basal M9 medium used in all cultures consisted of 13.2 g/L of $K_2HPO_4$; 2.65 g/L of $KH_2PO_4$; 2.04 g/L of NaCl; 4.1 g/L of $(NH_4)_2SO_4$, 0.5 g/L of $MgSO_4{\cdot}7H_2O$, 0.043 g/L of $FeCl_3{\cdot}6H_2O$; 0.05 g/L of citric acid and 1.5 mL/L of a Trace Salt solution. Trace salt solution composition was 0.82 g/L $CaCl_2{\cdot}H_2O$; 0.02 g/L of $AlCl_3{\cdot}6H_2O$; 0.44 g/L of $ZnSO_4{\cdot}7H_2O$; 0.08 g/L of $CoCl_2{\cdot}6H_2O$; 1.1 g/L of $CuSO_4{\cdot}5H_2O$; 0.005 g/L of $H_3BO_3$; 0.71 g/L of $MnCl_2{\cdot}4H2O$; 0.005 g/L of $NiCl_2{\cdot}6H_2O$ and 0.01 g/L of $Na_2MoO_4$; pH adjusted to 2 with HCl.

To complete the media used in the inoculums and in the batch phases of Bioreactor processes, the basal M9 medium was supplemented by 8 g/L or 20 g/L of glucose (Vadequimica, Sentmenat, Spain) respectively plus 1 g/L of yeast extract. Process 1 and 2 of the preliminary cultures, see results and discussion section, were developed without the yeast extract addition. Antifoam 204 (Merck, New Jersey, USA) was added initially in the bioreactor tank with the batch medium at a ratio of 0.0001 L/L. Carbenicillin was consistently added to maintain the selection pressure over transformed BL21 (DE3) pET22b_BsPdaC-CD; 0.1 g/L of carbenicillin.

The feeding solution used in the fed-batch processes were prepared by sterilizing a concentrated glucose solution of 400 g/L. Post-sterilization, the feed solutions were supplement with 10 mL/L of 50 g/L of $MgSO_4{\cdot}7H_2O$ stock solution, 5 mL/L of Trace salt solution and yeast extract to match a concentration of 1, 5–10 g/L depending on the process.

Shake flask cultures were performed preparing 50 mL of the previously described media in 250 mL baffled shake flasks. Glucose concentration was fixed at 8 g/L and antifoam was not added at this stage. Flasks were regularly cultivated at 30 °C and 250 rpms until $OD_{600}$ reached a value of 2 ± 0.2, approximately 6–8 h.

## Fed-batch cultures

The main fed-batch cultures were performed in two 3 L Bioreactor systems (BIONET F1, Murcia, Spain). The initial batch volume was fixed at 2.2 L. The batch phase was performed at 30 °C, pH at 7 maintained with 25% ammonia solution, 5% antifoam 204 solution used for foam control and $pO_2$ was maintained over 25% with stirring control ranged between 350 and 900 rpm. Overpressure was applied to specified cultures to 0.2 Ba to avoid oxygen limitation.

Fed-batch started automatically when $pO_2$ peak (> 60%) was recorded after 10 h of culture indicating the complete depletion of the glucose present in the Batch medium. The feeding rate was constant at 5, 7.5 and 10 mL/L/h as specified in each culture. These rates were selected to obtain an initial growth rate in the fed-batch phase of 0.10, 0.15 and 0.20 $h^{-1}$, respectively, assuming an average yield of 0.40 g of Biomass per g of glucose assimilated[15]. As expected, these initial growth rate decreased during the fed-batch phase due to the constant feeding rate. For all preliminary processes, the feeding rate was fixed at 5 mL/L/h. The $pO_2$, the pH and the foam controls were maintained from batch to fed-batch phase in all cultures. The temperature was kept at 30 °C until 5 h before induction time, where temperature was readjusted to 22–25 °C if required.

Induction was performed with a 10 mM IPTG solution to obtain a final concentration of 0.05 mM. The volume of IPTG required was determined considering the added ammonia, feed and antifoam as well as the extracted volumes during sampling for each process. Samples after 6 and 8 h post-induction were consistently taken for volumetric production determination.

## *Bs*PdaC-CD purification and quantification

To reduce variation in the purification process of *Bs*PdaC-CD produced intracellular, every sample analyzed always contained the same amount of biomass. Specifically, after $OD_{600}$ determination, the equivalent sample volume to collect a biomass corresponding to 30 mL with an $OD_{600}$ of 20 was aliquoted, centrifuged at 4000 g for 10 min and the pellet stored at -20 °C.

Cell lysis consisted of resuspending the frozen biomass with 30 mL of lysis buffer (50 mM $K_2HPO_4$, 300 mM NaCl, pH 8.5 with Phenylmethanesulfonyl fluoride (PMSF) 1 mM) and sonicate the solution for 8 min cycle (5' ON, 10' OFF) in an ice-ethanol bath (VCX 130, Vibra Cell, Newtown, USA). The disrupted cell suspensions were centrifuged at 10 000 g for 30 min at 4 °C. Pellets were stored if necessary for insoluble protein visualization through SDS-PAGE gels. Supernatants (SN) were filtered throw a 0.45 µm PVDF syringe filter (Teknokroma, Sant Cugat del Vallès, Spain) and the obtained volumes registered.

Volumes of 10 or 15 ml of filtered SN were purified using StrepTrapTM XT column in ÄKTA start chromatography system (Cytiva, Cornellà de Llobregat, Spain). After loading the samples, a PBS solution (50 mM $K_2HPO_4$; 300 mM NaCl; pH: 8.5) was used to clean non-bound proteins until absorbance reaches base level. Biotin solution (49 mM in PBS) was then introduced to elute bound proteins which were recovered in 15 ml chilled Falcons until absorbance reaches base level.

Purified *Bs*PdaC-CD, approximately 2 ml, was dialyzed using Amicon Ultra-15 10KDa (Millipore; Massachusetts, USA). All samples were dialyzed at least 100-fold to reduce biotin concentration using PBS. The exact volumes recovered after dialysis were registered. Protein concentration was quantified using the BCA assay (Pierce™ BCA Protein Assay Kits, Thermo Fisher, Alcobendas, Spain).

## *Bs*PdaC-CD activity assay

The activity of the purified *Bs*PdaC-CD was assessed using a Fluorescamine-based assay. Firstly, 160 µl of 2.5 mM of tetraacetyl-chitotetraose (A4) substrate were added into 96-well microplate (Thermo Scientific Sterilin, Alcobendas, Spain) and heated to 37 °C for 1 min. Secondly, up to 40 µl of the purified enzyme was mixed with the seeded A4 to obtain a final enzyme concentration between 2 and 6 µM. Plate was incubated at 37 °C and 20 µl samples were withdrawn after 5, 10, 15 and 20 min of incubation and mixed with 90 µl of STOP solution (1-propanol: $H_2O$; 50:50 v/v) in 96-well fluorescence plate (Thermoscientific NunclonTM Delta Surface, Alcobendas, Spain).

Glucosamine standard curve was prepared by mixing 20 µl of specific glucosamine concentrations, 0.5, 0.4, 0.3, 0.2, 0.1, 0 mM glucosamine, with 90 µl of STOP solution in the same fluorescence microplate. 20 µl of the fluorescamine solution (2 mg/ml in DMF; Biosynth, Staad, Switzerland) was added to samples and standards. The fluorescence microplate was incubated in the dark at room temperature for 10 min with shaking. Lastly, 150 µl of Labelling STOP solution (DMF: H2O; 50:50 v/v) was added to all samples to terminate the reaction. Fluorescence was then measured using fluorescence microplate reader (BioTek FLx800, Santa Clara, USA) with an excitation: emission range of 340/360:460/480 nm.

## Analytics

Culture growth was monitored by measuring light absorbance at 600 nm, $OD_{600}$, in a JENWAY 6305 spectrophotometer (Illinois, USA). To measure the Wet cell weight (WCW) and Dry cell weight (DCW), 1 ml cell suspension was transferred to 1 ml pre-weighed tubes and centrifuged at 16 600 g for 5 min. Supernatant was completely removed and stored at -20 °C to later analyze glucose and acetic acid. Tubes with wet pellets were weighed again to calculate WCW. Then, wet pellets were dried at 80 °C in a drying stove (19 L Conterm, Selecta, Barcelona, Spain) until stable weight, approximately between 48 and 72 h. WCW and DCW were performed in triplicates for each sampling time. DCW variability among replicates was consistently below 5%, as shown in Table 5.

Glucose and acetic acid concentrations were analyzed using the supernatant fractions from WCW and DCW samples. The quantification of the analytes was performed using the Y15 system (Biosystems, Barcelona, Spain).

### Design of experiments – L9 ($3^4$) Taguchi matrix

An L9 Taguchi matrix was used during the design of experience (DoE). The L9 matrix is a 9×4 table that allows the evaluation of four independent factors each of them at three levels. In this study, Temperature at induction, yeast extract concentration in the feeding solution and feed rate were the evaluated factors. The fourth factor was used to assess the error of the system. DoE was aimed to maximize the production of *Bs*PdaC-CD, so response values were adjusted according to the purpose higher the better using a commonly used conversion to S/N (Eq. 1):

$$\frac{S}{N} = -10 * \log\left(\frac{\sum \frac{1}{Y^2}}{n}\right) \quad (1)$$

Where *S/N* is the Signal-to-noise ratio, *n* is the number of measurements performed in each run and *y* the response variable. All statistical analysis and graphics were performed using Microsoft Excel and R-project[16].

## Results and discussion

The production of sequence-defined partially acetylated chitooligosaccharides (paCOS) is a promising strategy for the valorisation of chitin-rich waste into high-value bioactive compounds. This study aimed the production optimization of the catalytic domain of a deacetylase from *Bacillus subtilis* (*Bs*PdaC-CD), known to act not only on N-acetylmuramic acid (MurNAc) residues of peptidoglycan but also on GlcNAc of COS, a strong candidate to produce paCOS.

### Preliminary cultures

To identify the key variables influencing *Bs*PdaC-CD production, a series of bioreactor cultures were conducted. Initially, isopropyl β-D-1-thiogalactopyranoside (IPTG) was selected as the inducer, as previous data demonstrated greater efficiency in promoting intracellular accumulation of *Bs*PdaC-CD compared to lactose (data not shown). At the shake flask scale, IPTG is typically used at a concentration of 1 mM[17,18]. However, several studies at the bioreactor scale have reported successful recombinant protein expression using substantially lower IPTG concentrations, such as 0.1 mM and 0.05 mM[18], with comparable yields. Nevertheless, this outcome, while desirable, may not necessarily apply to *Bs*PdaC-CD so needed to be assessed. For these preliminary processes, the production was initially evaluated at 6 h post-induction.

In Process 1, the IPTG concentration was set at 0.5 mM to ensure full induction yielding 254 mg *Bs*PdaC-CD/L (Table 1). In Process 2, IPTG concentration was reduced to 0.1 mM reaching resemblant volumetric production compared to Process 1, 283 mg BsPdaC-CD/L culture, thereby, reflecting similar behaviour compared to other *E. coli* production processes[19].

Many other studies use enriched media to bust the production of different types of recombinant proteins[20]. In Process 3 and 4, the effect of supplementing the medium with 1 g/L yeast extract during the preculture, batch and feeding phases was evaluated, together with a further reduction of the IPTG concentration to 0.05 mM. In Process 3, the volumetric production increased 4-fold compared to Process 2, an effect directly linked to the addition of yeast extract. Such increase revealed the relevance of yeast extract to produce *Bs*PdaC-CD like other recombinant proteins. Moreover, reducing the IPTG concentration to 0.05 mM (Process 4) resulted in production levels comparable to those obtained in Process 3, demonstrating that full induction can be achieved with one-twentieth of the IPTG concentration typically used in shake-flask cultures. IPTG concentration of 0.05 mM was consistently applied to the following cultures since there was no evidence that suggested that induction of protein expression was not complete. Reducing IPTG concentration even further, although possible, might represent a minimal cost-related benefit and was preferred to ensure an efficient induction in all the following processes.

In Process 5 and 6, the YE concentration was maintained at 1 g/L during the preculture and batch phases while its concentration in the feeding solution was increased to 10 g/L and 30 g/L, respectively, to burst further the *Bs*PdaC-CD production. In contrast to Processes 3 and 4, volumetric production in these processes returned

| | IPTG (mM) | YE_b (g/L) | YE_fb (g/L) | DCW (g/L) | mg *Bs*PdaC-CD/L |
|---|---|---|---|---|---|
| Process 1 | 0.5 | – | – | 25.2 | 254 |
| Process 2 | 0.1 | – | – | 29.1 | 283 |
| Process 3 | 0.1 | 1 | 1 | 24.6 | 1146 |
| Process 4 | 0.05 | 1 | 1 | 23.7 | 1261 |
| Process 5 | 0.05 | 1 | 10 | 34.3 | 281 |
| Process 6 | 0.05 | 1 | 30 | 32.2 | 418 |

**Table 1**. *Bs*PdaC-CD volumetric production and Biomass production for Processes 1 to 6. YE_b: yeast extract in batch media; YE_fb: yeast extract in feeding solution; DCW: dry cell weight; mg *Bs*PdaC-CD/L: amount of purified enzyme per liter of culture.

to levels comparable to those observed in yeast-free cultures (281 and 418 mg BsPdaC-CD/L, respectively; Table 1).

Regarding the carbon source, glucose was selected as it performed well as commonly used is *E. coli* cultivations. Glycerol could have been considered as alternative carbon source, but it would have required evaluation of basal (leaky) expression prior to induction, which we preferred to avoid due to the potential metabolic burden imposed by premature recombinant protein synthesis.

In terms of biomass, the six preliminary processes yielded DCW ranging from 25 to 34 g/L. Although values of 30 to 40 g/L were expected, variation on *Bs*PdaC-CD production could not be attributed to differences in biomass production but to the specific condition applied in each process.

## L9 Taguchi matrix. Design

To increase the efficacy for obtaining high levels of *Bs*PdaC-CD by testing different factors at the same time we decided to use the Taguchi's method[14]. Specifically, we used an L9 ($3^4$) orthogonal array to design the experiments and evaluate the effects of 3 factor at 3 levels: yeast extract concentration in the feeding solution, feeding rate and temperature at induction. The extra factor available in the matrix was used to assess the variability of the processes (commonly named error). Table 2 shows the Factors and the levels used in the Taguchi experiment design.

*Factor 1 – Yeast extract feeding*

In the preliminary processes, the addition of 1 g/L yeast extract to both the batch media and the feed solution increased *Bs*PdaC production. However, when the yeast extract concentration in the feed was increased to 10 and 30 g/L, production declined, showing the opposite trend. To further clarify the effect of this compound, the amount of yeast extract in the batch phase was kept constant at 1 g/L, while its concentration in the feed was included as a factor within a narrower range, from 1 to 10 g/L.

*Factor 2 – Feeding rate*

An important factor in fed-batch processes is the feeding rate, which can influence not only *Bs*PdaC-CD production but also the overall duration of the process. In our study, we opted to evaluate only constant feed rates rather than linear or exponential feeding profiles to maintain a constant oxygen uptake rate in the bioreactors and avoid potential process limitations caused by oxygen scarcity, which can arise due to the increasing oxygen demand associated with linear or exponential feeding strategies.

During the batch phase of preliminary processes, *E. coli* achieved near-maximum specific growth rates under the tested conditions ranging from 0.6 to 0.7 $h^{-1}$. In the Taguchi experimental design, fixed feeding rates of 5, 7.5, and 10 mL/L/h were selected to achieve initial specific growth rates of 0.10, 0.15, and 0.20 $h^{-1}$, respectively, during the fed-batch phase. These calculations were based on an assumed average yield of 0.40 g of biomass per g of glucose consumed[15]. These specific growth rates were considerably lower than the maximum growth rates of the strain, thereby minimizing the risk of glucose accumulation. With these settings the specific growth rate gradually declines as biomass concentration increases without a proportional increase in the feeding rate.

To compare the production levels between different feeding rate, the duration of the induction phase was kept constant regardless of the applied feed rate, thereby avoiding potential protein degradation due to excessively long induction or incomplete induction resulting from insufficient induction time. This fact implied that we would allow more glucose assimilation and growth in those processes with higher feeding rates values (See Table 3). The induction was set to be triggered at an $OD_{600}$ of 80 based on simple mechanistic fed-batch models. However, the actual $OD_{600}$ at induction varied between 75 and 100 (see Table 3).

*Factor 3 – Temperature at induction*

Temperature remained constant at 30 °C in all preliminary processes. This temperature is adequate for *E. coli* growth despite not being the optimum. A common methodology for recombinant protein productions is to perform an initial growth phase at 37 °C followed by an induction at 25 °C to allow proper folding and minimize inclusion bodies[21]. In Process 3, inclusion bodies were clearly observed after cell disruption (Supplementary Information, Figure S1). To mimic the common methodology of shake flask studies, we included temperature at induction time as third factor in the DoE. The selected induction temperatures were 22 °C, 25 °C and 30 °C. To prevent potential glucose accumulation in the fed-batch phase, since lower temperatures reduce the growth rate, the temperature adjustment was implemented 5 h before induction with the system stabilizing to the lower temperature within 20 min (see Fig. 1).

| | | Level 1 | Level 2 | Level 3 |
|---|---|---|---|---|
| Factor 1 | Yeast extract feeding (g/L) | 1 | 5 | 10 |
| Factor 2 | Feeding rate (mL/L/h) | 5 | 7.5 | 10 |
| Factor 3 | Induction temp (°C) | 22 | 25 | 30 |

**Table 2**. Factors (parameters) and levels in Taguchi experimental design.

| | Factor: variability | Factor: yeast extract (g/L) | feeding rate (mL/L/h) | Factor: temperature (°C) |
|---|---|---|---|---|
| Run 1 | 1 | 1 (1) | 1 (5) | 1 (22) |
| Run 2 | 1 | 2 (5) | 2 (7.5) | 2 (25) |
| Run 3 | 1 | 3 (10) | 3 (10) | 3 (30) |
| Run 4 | 2 | 1 (1) | 2 (7.5) | 3 (30) |
| Run 5 | 2 | 2 (5) | 3 (10) | 1 (22) |
| Run 6 | 2 | 3 (10) | 1 (5) | 2 (25) |
| Run 7 | 3 | 1 (1) | 3 (10) | 2 (25) |
| Run 8 | 3 | 2 (5) | 1 (5) | 3 (30) |
| Run 9 | 3 | 3 (10) | 2 (10) | 1 (22) |

**Table 3**. Factors and levels in Taguchi experimental design. In brackets there are the translation of each level (1 to 3) to the value of each factor.

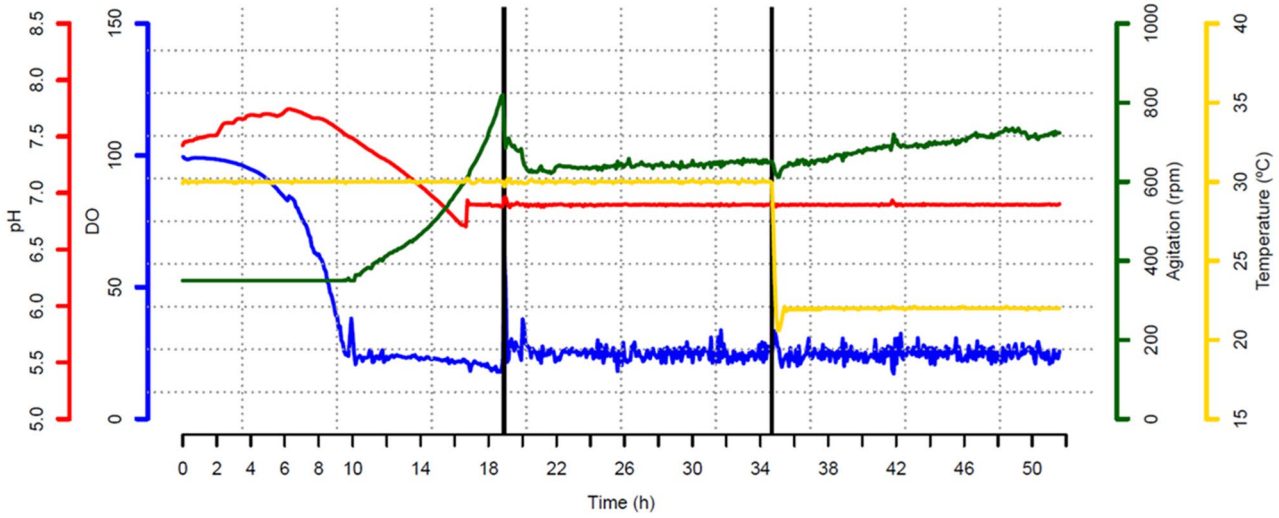



**Fig. 1**. Principal process parameters recorded from Run 5. Vertical black lines represent the time the process shift automatically from batch to fed-batch phase (first line) and from fed-batch phase to induction phase at lower temperature. DO: Dissolved Oxygen.

*Factor 4 – Variability*

The last factor available in the design was used to assess the variability between the different runs proposed. With the defined factors and levels of Tables 2 and 3 shows the layout of the L9 ($3^4$) Taguchi orthogonal array design with the 9 independent runs to be performed.

## L9 Taguchi matrix. Raw data analysis

Before assessing the *Bs*PdaC-CD production obtained in each of the 9 independent runs, Fig. 1 exemplifies the profile obtained in each of the runs in terms of dissolved oxygen (DO), pH, agitation and temperature. The full recorded parameters as well as the only offline measurements of $OD_{600}$, glucose, acetate, WCW and DCW for each run are available in Supplementary Information.

The graph in Fig. 1 is divided into batch, fed-batch and induction phases (vertical black lines). The batch phase of Run 5 lasted 18.5 h and it can be observed how the $pO_2$ control maintained a value of 25% by increasing the agitation from 350 up to 820 rpms. The culture started to acidify the pH and base addition started at 16 h to adjust the pH to 7. At 18.5 h, a $pO_2$ peak of 65% was recorded and automatically triggered the fed-batch phase initiation. In this Run 5, as well as in all runs with 10 mL/L/h of feeding rate, the $pO_2$ control just with agitation was not enough to maintain the setpoint value and 0.2 bars of overpressure was applied to the system to increase the oxygen solubility and consequently the oxygen transfer rate of the system. In this phase, the agitation was rather stable in the range of 650 to 720 rpms, indicating constant oxygen demands as expected for a constant feed rate. Moreover, in any of the runs performed neither accumulation of glucose nor acetate was detected.

At 34 h, 5 h before the induction, the temperature setpoint was changed to 22 °C and it decreased in less than 20 min. At this temperature, the maximum growth rate was significantly lowered but not enough to accumulate glucose as desired. At 39 h the system was induced without any impact on the main parameters recorded.

The initial Runs of the L9 ($3^4$) designed, Run 3 (30 °C), 5 (22 °C) and 8 (30 °C), were assessed at 6, 8 and 10 h post-induction time to determine the optimal induction time (Table 4). As can be clearly seen, 8 h was the best in terms of volumetric production in all cases, observing a partial degradation of *Bs*PdaC-CD at longer times

| Process | Volumetric production (mg *Bs*PdaC-CD /L) | | |
|---|---|---|---|
| | Post-induction 6 h | Post-induction 8 h | Post-induction 10 h |
| Run 3 | 618 [563–673] | 947 [926–968] | 613 [593–633] |
| Run 5 | 87 [80–94] | 215 [193–237] | 199 [163–234] |
| Run 8 | 217 [211–222] | 345 [320–371] | 299 [295–303] |

**Table 4**. *Bs*PdaC-CD volumetric production comparison of 6 h, 8 h and 10 h post-induction time points. Values in brackets represent the min – max values of 2 technical replicates of the *Bs*PdaC-CD purification and quantification methodology.

| | R1 | R2 | R3 | R4 | R5 | R6 | R7 | R8 | R9 |
|---|---|---|---|---|---|---|---|---|---|
| **Yeast feed (g/L)** | 1 | 5 | 10 | 1 | 5 | 10 | 1 | 5 | 10 |
| **Feed rate (mL/L/h)** | 5 | 7.5 | 10 | 7.5 | 10 | 5 | 10 | 5 | 7.5 |
| **Temperature (°C)** | 22 | 25 | 30 | 30 | 22 | 25 | 25 | 30 | 22 |
| Batch time (h) | 18.5 | 14 | 17 | 14 | 18.5 | 16.4 | 16 | 22.5 | 17.5 |
| Volume added feed (L) | 0.654 | 0.653 | 0.727 | 0.631 | 0.680 | 0.632 | 0.745 | 0.588 | 0.608 |
| Total time (h) | 77 | 53.1 | 47.3 | 51.5 | 47.3 | 73.4 | 49.2 | 77 | 53.4 |
| $OD_{600}$ at induction | 77 | 86 | 105 | 84 | 90 | 97 | 88 | 73 | 76 |
| Final $OD_{600}$ | 90 | 104 | 107 | 106 | 126 | 77 | 117 | 87 | 91 |
| Final DCW (g/L) | 34.4 ± 0.9 | 39.9 ± 0.9 | 39 ± 1 | 38 ± 1 | 42 ± 1 | 30.3 ± 0.6 | 39.9 ± 0.9 | 30.3 ± 0.5 | 35.1 ± 0.6 |
| $Y_{x/s}$ (g biomass/g substrate) | 0.356 | 0.391 | 0.468 | 0.392 | 0.424 | 0.371 | 0.381 | 0.317 | 0.376 |
| Volumetric production (mg *Bs*PdaC-CD /L) | 167 [151–183] | 1197 [1169–1226] | 947 [926–968] | 507 [495–521] | 215 [193–237] | 345 [322–367] | 945 [925–965] | 345 [320–371] | 776 [758–796] |
| $Y_{p/x}$ (mg *Bs*PdaC-CD/g biomass) | 4.9 [4.4–5.3] | 30.0 [29.3–30.7] | 24.3 [23.7–24.8] | 13.3 [13.0–13.7] | 5.1 [4.6–5.6] | 11.4 [10.6–12.1] | 23.7 [23.2–24.2] | 11.4 [10.6–12.2] | 22.1 [21.6–22.7] |
| STY (mg *Bs*PdaC-CD/(L*h)) | 2.2 [2.0–2.4] | 22.5 [22.0–23.1] | 20.0 [19.6–20.5] | 9.8 [9.6–10.1] | 4.5 [4.1–5.0] | 4.7 [4.4–5.0] | 19.2 [18.8–19.6] | 4.5 [4.2–4.8] | 14.5 [14.2–14.9] |

**Table 5**. Summary of the volumetric productions, the STYs and yields after 8 h induction with 0.05 mM IPTG. Values in brackets represent the min – max values of 2 technical replicates of the *Bs*PdaC-CD purification and quantification methodology.

independently of the temperature at induction. Due to these results, the main comparison between the Runs was performed at 8 h. Table 5 shows a summary of the volumetric productions, the STYs and yields extracted from all 9 runs at 8 h post-induction. Data at 6 h post-induction were anyway analyzed and, as expected from the three first Runs, resulted in lower values and can be seen in Table S1. Data at 10 h was no longer analyzed because of the observed product degradation.

One of the main concerns of the L9 ($3^4$) matrix design was to obtain similar biomass production in the range of high cell density cultures, 40 g/L, across all runs to not bias the overall analysis (Final DCW; Table 5). The mean biomass production considering all runs was 36.5 ± 1.5 with a minimum value of 30 in Run 8 and a maximum value of 42 in Run 5. This variation was acceptable and did not stem from differences in terms of carbon source addition (Volume added feed (L); Table 5) but from biomass yield ($Y_{x/s}$ (g biomass/g substrate); Table 5). This fact indicates that the main factors chosen in the DoE design not only affect the *Bs*PdaC-CD production but also the biomass/carbon source yield, although it will not be analyzed further in this study.

Regarding the *Bs*PdaC-CD production, this can be evaluated using either the Volumetric production (mg *Bs*PdaC-CD /L) or space-time yield (STY; mg BsPdaC-CD/(L·h)). Choosing one of these outcomes in the optimization would depend on the techno-economic analysis of the process[22] but both of them are analyzed in the ANOVA section (below) for comparison. Strictly, we should also include the product produced per substrate consumed, $Y_{P/S}$, as a factor to evaluate in the *Bs*PdaC-CD production. Nevertheless, this factor is not highly important for middle-high value products, such as the *Bs*PdaC-CD, where the substrate, glucose in our case, is not a major cost.

As shown in Table 5, there is a large variation in *Bs*PdaC-CD volumetric production, ranged from 167 to 1197 mg BsPdaC-CD /L, as well as in *Bs*PdaC-CD STY, ranged from 2.2 to 22.5 mg *Bs*PdaC-CD/(L·h). The impact of each factor on both response variables is detailed in the next section.

## L9 Taguchi matrix. ANOVA

To evaluate the importance of each factor in the variation of response variables, volumetric production or STY, the initial raw values were transformed to S/N ratios using the higher the better conversion and used to perform an ANOVA. Figure 2 shows the main effect plots for the S/N ratio using volumetric production and STY respectively. Clearly, the feeding rate and the temperature had more impact on the response variables than yeast

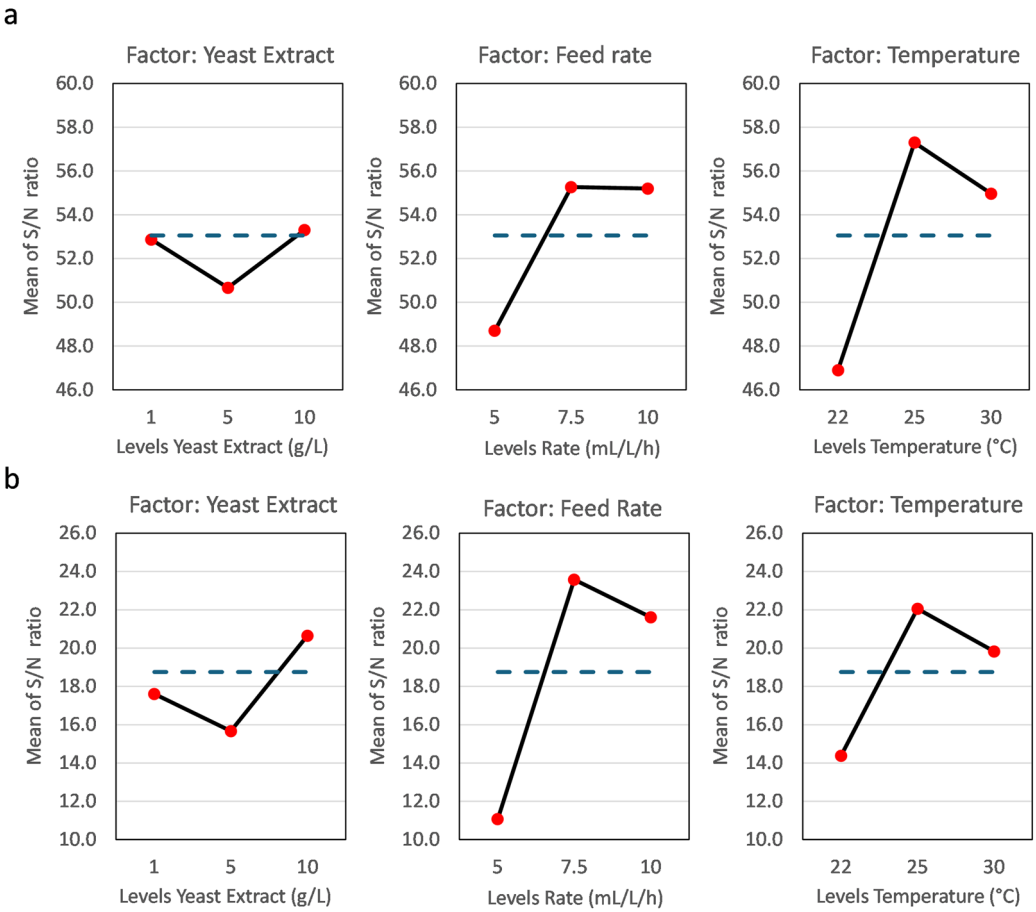



**Fig. 2**. Main effect plots for S/N ratio using volumetric production (**a**) or STY (**b**) as response variable. The blue discontinued lines represent the global S/N mean.

| | SS | DF | MS | F | *p*-value |
|---|---|---|---|---|---|
| Yeast extract (g/L) | 0.28 | 2 | 0.14 (0.1%) | 0.009 | 0.990 |
| Feed rate (mL/L/h) | 86.94 | 2 | 43.47 (29.4%) | 2.8430 | 0.260 |
| Temperature (°C) | 178.14 | 2 | 89.07 (60.2%) | 5.825 | 0.146 |
| Variability (error) | 30.58 | 2 | 15.29 (10.3) | | |
| Total | 295.94 | 8 | | | |

**Table 6**. ANOVA for S/N ratio using volumetric production as response variable. In brackets there are the MS variations in % scaled up to overall MS variation. SS: Sum of Squares; DF: Degrees of Freedom; MS: Mean Square.

extract. In both cases, the main variability was observed in level 1 of feeding rate and temperature, 5 mL/L/h and 22 °C respectively, which showed lower *Bs*PdaC-CD productions.

When performing the ANOVA of the S/N ratios, we determined that, independently of the response variable, only 10% of the Mean Square (MS) variation could be associated with undetermined factors (variability, Tables 6 and 7). Considering the high complexity of a fed-batch process as well as the *Bs*PdaC-CD analysis performed, we consider this value a success, indicating highly robust methodology.

Nevertheless, ANOVA also showed that none of the factors obtained a p-value lower than 0.05. Variation of the yeast extract concentration in the feeding solution had almost a null contribution in both responses, thereby, for cost reasons, 1 g/L was selected as the best level.

When evaluating volumetric production, temperature showed a p-value of 0.146 confirming the effect with a confidence level of 85%. Analyzing the main effect plots, 22 °C could be discarded as the optimal level while retaining 25 °C and 30 °C as favorable conditions. This observation contrasts with the commonly applied strategy in shake-flask cultures, which has been successfully used in several previous projects within our research group and in other studies[21]. To deeper understand *Bs*PdaC-CD production behavior, Fig. 3 shows how *Bs*PdaC-CD splits into soluble forms and inclusion bodies depending on the temperature used at induction. The results

| | SS | DF | MS | F | *p*-value |
|---|---|---|---|---|---|
| Yeast extract(g/L) | 20.92 | 2 | 10.46 (4.9%) | 0.480 | 0.676 |
| Feed rate (mL/L/h) | 271.55 | 2 | 135.77 (63.2%) | 6.232 | 0.138 |
| Temperature (°C) | 93.77 | 2 | 46.88 (21.8%) | 2.152 | 0.317 |
| Variability (error) | 43.57 | 2 | 21.78 (10.1%) | | |
| Total | 429.81 | 8 | | | |

**Table 7**. ANOVA for S/N ratio using STY as response variable. In brackets there are the MS variations in % scaled up to overall MS variation. SS: Sum of Squares; DF: Degrees of Freedom; MS: Mean Square.

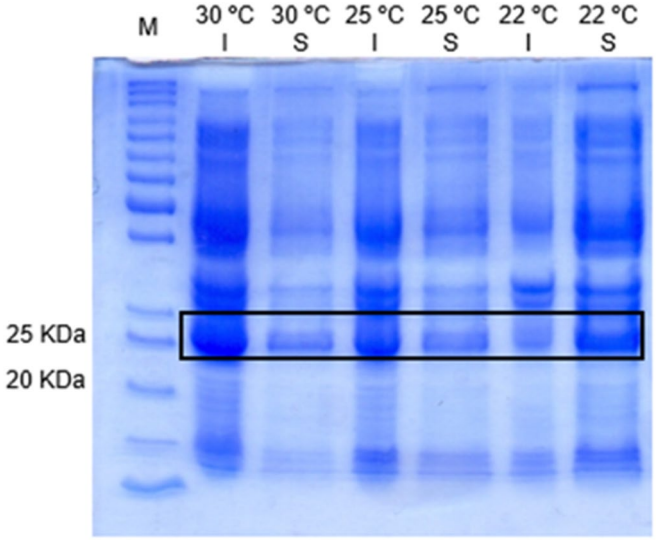



**Fig. 3**. SDS-PAGE gel comparing the ratio of soluble (S) and insoluble (I) fractions at different temperatures. Fractions at 30 °C (Run 4), 25 °C (Run 7), and 22 °C (Run 9).

| | Total time (h) | Final DCW (g/L) | Volumetric production (mg *Bs*PdaC-CD /L) | STY (mg *Bs*PdaC-CD/(L*h)) |
|---|---|---|---|---|
| Verification 1 | 50.4 | 36.7 ± 1.3 | 1261 [1245–1278] | 25.0 [24.7–25.3] |
| Verification 2 | 51.0 | 38.1 ± 1.3 | 1370 [1129–1612] | 26.9 [22.1–31.6] |

**Table 8**. *Bs*PdaC-CD volumetric production and STY of the two validation runs. Values in brackets represent the min – max values of 2 technical replicates of the *Bs*PdaC-CD purification and quantification methodology.

demonstrated that indeed inclusion bodies were reduced at 22 °C compared to 25–30 °C, but it did not correlate with higher *Bs*PdaC-CD production. Therefore, testing lower temperatures was not considered.

When evaluating STY, the feeding rate was the factor that pooled the higher variability as could be expected due to the direct link between this factor and the global duration of the processes. The p-value of the feed rate indicated 86% of confidence on STY allowed to discard of the 5 mL/L/h level as optimal level and keep 7.5 and 10 mL/L/h (See Fig. 2).

## L9 Taguchi matrix. Verification

To further consolidate the primary findings of the L9 ($3^4$) Taguchi optimization process, two verification runs were conducted using 1 g/L of yeast extract in the feeding solution, a feeding rate of 10 mL/L/h and a constant temperature of 30 °C (Table 8). The feed rate of 10 mL/L/h and temperature of 30 °C were preferred over 7.5 mL/L/h and 25 °C respectively for simplicity, as this approach eliminates the need for temperature changes during the process and shortens the overall process duration. As shown in Table 8, we could reproducibly produce more than 1.2 g/L of *Bs*PdaC-CD in approximately 50 h which not only confirmed DoE results but also improved the best condition obtained previously (Run 2, Table 5). Moreover, the BsPdaC-CD specific activity was in the expected range of 3 to 5 $min^{-1}$ [12].

Overall, the development of DoE allowed us to optimize and verify the robustness of the proposed methodology to produce *Bs*PdaC-CD in a fed-batch process in *E. coli* BL21.

## Conclusions

This study successfully developed and optimized a fed-batch process for *Bs*PdaC-CD production in *E. coli* BL21, applying a systematic Taguchi L9 ($3^4$) design of experiments approach. Production variability among runs highlighted the complex interactions between process conditions, cell metabolism and production capacities. Analysis of volumetric production and space-time yield revealed temperature and feed rate as the most significant factors influencing *Bs*PdaC-CD production. As expected, lower induction temperatures reduced inclusion body formation but did not improve overall production, differing from trends seen in shake-flask experiments. Additionally, the study highlights the value of the DoE approach for identifying key factors driving process improvement. Finally, follow-up runs at optimal conditions defined by the L9 ($3^4$) design confirmed the reproducibility and robustness of the process, achieving consistent *Bs*PdaC-CD production of 1.2 g/L suitable for scale-up strategies and further application on paCOS production.

## Data availability

All the data relevant to the conclusion are presented in the main manuscript or in the materials presented in Supplementary Information.

Received: 12 March 2026; Accepted: 20 May 2026
Published online: 30 May 2026

## References

1. Martínez, J. P., Falomir, M. P. & Gozalbo, D. Chitin: A Structural Biopolysaccharide with Multiple Applications. *eLS*, 1–10 ( Wiley, 2014).
2. Santos, V. P. et al. Seafood waste as attractive source of chitin and chitosan production and their applications. *Int. J. Mol. Sci.* **21**, 4290 (2020).
3. Liaqat, F. & Eltem, R. Chitooligosaccharides and their biological activities: A comprehensive review. *Carbohydr. Polym.* **184**, 243–259 (2018).
4. Basa, S. et al. The pattern of acetylation defines the priming activity of chitosan tetramers. *J. Am. Chem. Soc.* **142**, 1975–1986 (2020).
5. Cord-Landwehr, S. et al. Patterns matter part 2: Chitosan oligomers with defined patterns of acetylation. *React. Funct. Polym.* **151**, 104577 (2020).
6. de Mey, M., Guidi, C., Biarnes, X. & Planas, A. Method to Synthesize chitin oligosaccharides. Patent: PCT/EP2022/082232; WO 2023/089009; US20250011824A1. doi:WO2023089009. (2021).
7. Guidi, C., Biarnés, X., Planas, A. & De Mey, M. Expanding the chitin oligosaccharide portfolio by engineering NodC chitin synthases in *Escherichia coli*. *Curr. Opin. Biotechnol.* **8**, 100255 (2024).
8. Snoeck, S. et al. Understanding bottlenecks in the microbial producion of partially acetylated chitooligosaccharides. *bioRxiv Febr.* **24** https://doi.org/10.1101/2025.02.24.639843 (2025).
9. Pascual, S. & Planas, A. Carbohydrate de-N-acetylases acting on structural polysaccharides and glycoconjugates. *Curr. Opin. Chem. Biol.* **61**, 9–18 (2021).
10. Aragunde, H., Biarnés, X. & Planas, A. Substrate recognition and specificity of chitin deacetylases and related family 4 carbohydrate esterases. *Int. J. Mol. Sci.* **19**, 1–30 (2018).
11. Andrés, E. et al. Structural basis of chitin oligosaccharide deacetylation. *Angew Chem. Int. Ed. Engl.* **53**, 6882–6887 (2014).
12. Grifoll-Romero, L. et al. Structure-function relationships underlying the dual N-acetylmuramic and N-acetylglucosamine specificities of the bacterial peptidoglycan deacetylase PdaC. *J. Biol. Chem.* **294**, 19066–19080 (2019).
13. Grifoll-Romero, L., Pascual, S., Aragunde, H., Biarnés, X. & Planas, A. Chitin deacetylases: Structures, specificities, and biotech applications. *Polym. (Basel).* **352** https://doi.org/10.3390/polym10040352 (2018).
14. Roy, R. K. *A Primer Onn the Taguchi Methods (2nd Edition)*. https://doi.org/10.2307/2583510 (Society of Manufacturing Engineers, 2010).
15. Kopp, J. et al. Repetitive fed-batch: A promising process mode for biomanufacturing with E. coli. *Front. Bioeng. Biotechnol.* **8**, 1–13 (2020).
16. R Development Core Team. R: a language and environment for statistical computing. http://www.R-project.org (2013).
17. Fong, B. A. & Wood, D. W. Expression and purification of ELP-intein-tagged target proteins in high cell density *E. coli* fermentation. *Microb. Cell. Fact.* **9**, 77 (2010).
18. Larentis, A. L. et al. Evaluation of pre-induction temperature, cell growth at induction and IPTG concentration on the expression of a leptospiral protein in *E. coli* using shaking flasks and microbioreactor. *BMC Res. Notes*. **7**, 671 (2014).
19. Pasini, M., Fernández-Castané, A., Caminal, G., Overton, T. W. & Ferrer, P. Process intensification at the expression system level for the production of 1-phosphate aldolase in antibiotic-free *E. coli* fed-batch cultures. *J Ind. Microbiol. Biotechnol* **49**, (2022).
20. Li, X., Robbins, J. W. & Taylor, K. B. The production of recombinant beta-galactosidase in *Escherichia coli* in yeast extract enriched medium. *J. Ind. Microbiol.* **5**, 85–93 (1990).
21. Rosano, G. L. & Ceccarelli, E. a. Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front. Microbiol.* **5**, 1–17 (2014).
22. Konzock, O. & Nielsen, J. TRYing to evaluate production costs in microbial biotechnology. *Trends Biotechnol.* **42**, 1339–1347 (2024).

## Acknowledgements

The authors thank Paula Morer for her assistance in bioreactors settings in the Bioprocess Pilot Plant. This work was supported by following grants: GLYCOENGIN (PID2022-138252OB-I00) and paCOSVALOR (PDC2022-133580-I00) from MICINN, Spain and the CROP-SAFE European Project (ref: 101209410) from HORIZON-JU-CBE-2024-RIA-03.

## Author contributions

P.H. performed all cultivations, analysis and interpretation of results and first draft of the manuscript; M.T. enzyme purification protocols and activity assays; M.C. conceptualization, data calculation, analysis of results and manuscript preparation. A.P. conceptualization, reviewed the manuscript, funding acquisition and supervision. All the authors read and approved the final manuscript.

## Declarations

### Competing interests
The authors declare no competing interests.

### Additional information
**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-54839-9.

**Correspondence** and requests for materials should be addressed to M.C. or A.P.

**Reprints and permissions information** is available at www.nature.com/reprints.

**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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026