

RESEARCH

Open Access



# Enhanced recombinant protein capture, purity and yield from crude bacterial cell extracts by *N*-Lauroylsarcosine-assisted affinity chromatography

Jose Vicente Carratalá<sup>1,2,3†</sup>, Jan Atienza-Garriga<sup>1,2,3†</sup>, Hèctor López-Laguna<sup>1,2,3</sup>, Esther Vázquez<sup>1,2,3</sup>, Antonio Villaverde<sup>1,2,3\*</sup>, Julieta M. Sánchez<sup>1,2,3,4\*</sup> and Neus Ferrer-Miralles<sup>1,2,3\*</sup>

## Abstract

**Background** Recombinant proteins cover a wide range of biomedical, biotechnological, and industrial needs. Although there are diverse available protocols for their purification from cell extracts or from culture media, many proteins of interest such as those containing cationic domains are difficult to purify, a fact that results in low yields of the final functional product. Unfortunately, this issue prevents the further development and industrial or clinical application of these otherwise interesting products.

**Results** Aiming at improving the purification of such difficult proteins, a novel procedure has been developed based on supplementing crude cell extracts with non-denaturing concentrations of the anionic detergent *N*-Lauroylsarcosine. The incorporation of this simple step in the downstream pipeline results in a substantial improvement of the protein capture by affinity chromatography, an increase of protein purity and an enhancement of the overall process yield, being the detergent not detectable in the final product.

**Conclusion** By taking this approach, which represents a smart repurposing of *N*-Lauroylsarcosine applied to protein downstream, the biological activity of the protein is not affected. Being technologically simple, the *N*-Lauroylsarcosine-assisted protein purification might represent a critical improvement in recombinant protein production with wide applicability, thus smothering the incorporation of promising proteins into the protein market.

**Keywords** Recombinant proteins, Difficult-to-purify proteins, *E. coli*, H6 tag, Detergent, Repurposing

<sup>†</sup>Jose Vicente Carratalá and Jan Atienza-Garriga contributed equally to this work

\*Correspondence:  
Antonio Villaverde  
Antoni.villaverde@uab.cat  
Julieta M. Sánchez  
Julieta.sanchez@uab.cat  
Neus Ferrer-Miralles  
Neus.Ferrer@uab.cat

<sup>1</sup> Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Barcelona, Spain

<sup>2</sup> CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Instituto de Salud Carlos III, 08193 Cerdanyola del Vallès, Barcelona, Spain

<sup>3</sup> Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Barcelona, Spain

<sup>4</sup> Instituto de Investigaciones Biológicas y Tecnológicas (IIBYT) (CONICET-Universidad Nacional de Córdoba), ICTA, FCFEYN, UNC., Av. Velez Sarsfield 1611, X 5016GCA Córdoba, Argentina



## Background

The large-scale production of recombinant proteins has enabled their exploitation in a wide range of sectors such as biomedicine and biotechnology, for diagnostics, therapy and vaccination [1–5], as well as molecular tools in genetic engineering or catalysts in the biotech industry [3, 6]. The production of recombinant proteins, especially at large scale, suffers from important bottlenecks that minimize the yield of functional, usable products [7]. Among them, protein aggregation commonly occurs irrespective of the type of cell factory used for biofabrication [8, 9]. In bacteria, in which the recombinant proteins are in general not secreted to the media, aggregation of the recombinant protein results in large cytoplasmic structures called inclusion bodies (IBs) [10]. Soluble aggregates, probably intermediates in IB formation, also abound [11–13], since aggregation of recombinant proteins is a complex event that involves a wide spectrum of conformational conformers [14–17], ranging from properly folded versions to misfolded, amyloid forms [16]. Globally, the initial step in recombinant protein purification is the separation, after cell lysis, of the soluble cell fraction from the insoluble cell fraction. In this scenario, the populations of recombinant protein that form IBs are retained in the insoluble cell fraction and therefore discarded from the protein purification process, that only involves soluble forms. For aggregation-prone proteins, this fact represents an immediate loss of an important portion of the product that is straightforward excluded from the purification pipeline.

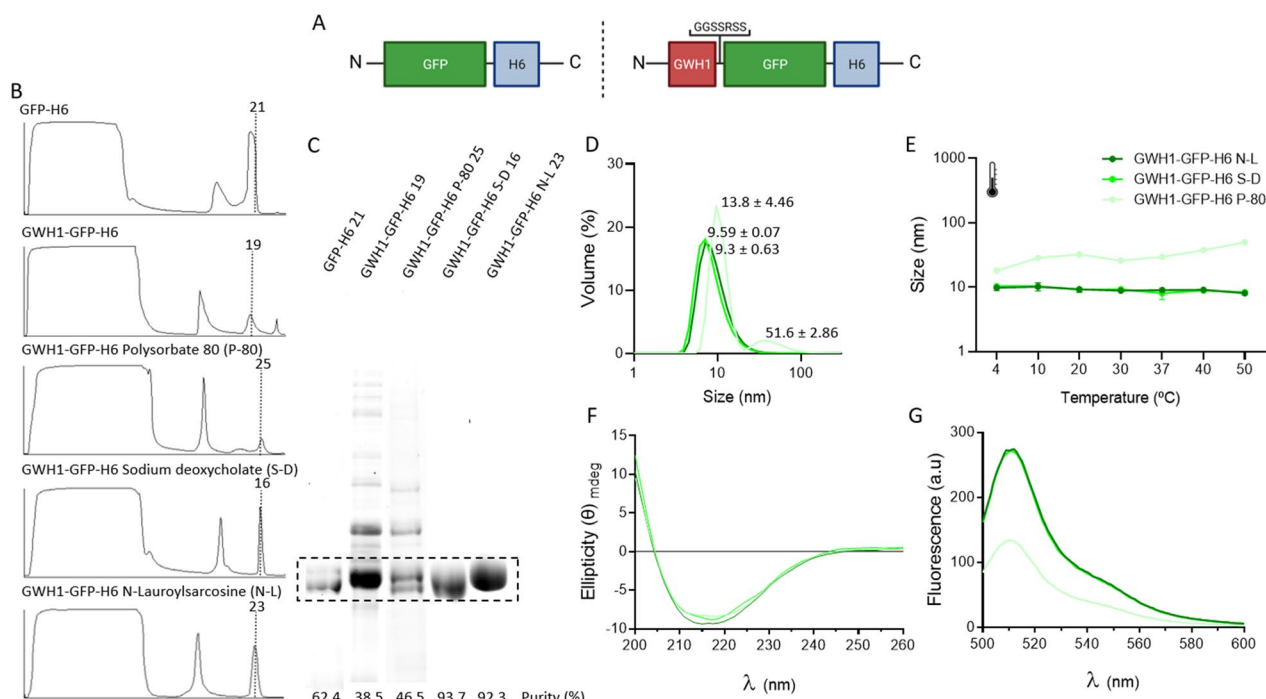
However, proteins can also be recovered from IBs once these have been separated from the cell debris. In this regard, several procedures have been described for the recovery of IB protein, based on either denaturing or non-denaturing conditions. In the first, more conventional approach, the IB protein is completely denatured, solubilized, and subjected to a refolding process aiming at recovering the product in the native conformation and with full functionality [18, 19]. This methodology shows variable and highly product-dependent success rates. For non-denaturing protein extraction, which is still emerging, standard protocols have been established that allow the immediate release of IB proteins with full biological activity. This is achieved by using mild detergents like *N*-Lauroylsarcosine (N-L, also known as sarkosyl) as solubilizing agents [20]. The success of this approach is based on the important amounts of recombinant proteins with native or native-like conformations contained in the IBs [17].

Regardless of the generic optimization of protein recovery from large IB aggregates, many industrially or clinically interesting proteins fall into the category of difficult-to-purify proteins, including those that have

solvent-exposed hydrophobic or cationic domains [21, 22]. Although not deeply analyzed, failure of these proteins to be purified with sufficient efficiency (e.g., by affinity chromatography) may be due to their tendency to aggregate, even as soluble versions. Consequently, there is a steric sheltering of the purification tags or the undesired acquisition of a sticky character that makes the polypeptide to interspecifically interact with other proteins. Despite the success in the application of N-L for IB protein recovery [20], this detergent had not been tested as a tool for improving the purification of soluble, difficult-to-purify protein species. In this study and by using several model proteins, we have tested the capability of N-L to favor the chromatographic purification of complex recombinant proteins from *Escherichia coli* cell extracts, keeping their biological activity and avoiding traces of the detergent in the final product. According to the presented data, we propose the repurposing of N-L from an agent for the solubilization of protein aggregates to an additive for the purification of soluble protein species and the incorporation of this detergent as a highly valuable tool for improved purification protocols to be applied to the soluble cell fraction.

## Results

GWH1-GFP-H6 is a modular protein (Fig. 1A, top) that contains GWH1, a cationic antimicrobial peptide (AMP) of clinical interest for the treatment of infectious diseases and in oncology [23]. The presence of the hexahistidine (H6) at the carboxy terminus of the construct allows the one-step purification of the protein by Ni<sup>2+</sup>-based affinity chromatography and, in addition, its self-assembling as regular homo-oligomers of around 10 nm. The oligomeric disposition of AMPs in multivalent nanoparticles is highly desirable, as such format increases the local concentration and effectiveness of the drug [24, 25]. The presence of GWH1 at the N-terminus of this construct is a clear obstacle in the downstream of the protein, as the purity of the resulting product is much lower than that of its counterpart GFP-H6 (Figure B, C). However, when different detergents are added to the cell extracts, after mechanical cell lysis and before chromatography, the purity of the products is enhanced from around 38% to more than 90% (Fig. 1C). While polysorbate 80 (P-80) had only a moderate effect (46% purity), sodium deoxycholate (S-D) and N-L were highly effective reagents. In addition, P-80 also rendered a product that appeared to be less proteolytically stable than alternative versions, since a double protein band was observed in the gels (Fig. 1C). When checking the oligomeric architecture of the proteins, S-D and N-L-managed products showed a nanometric size compatible with previous observations, in a monodisperse peak (Fig. 1D). On the contrary, when



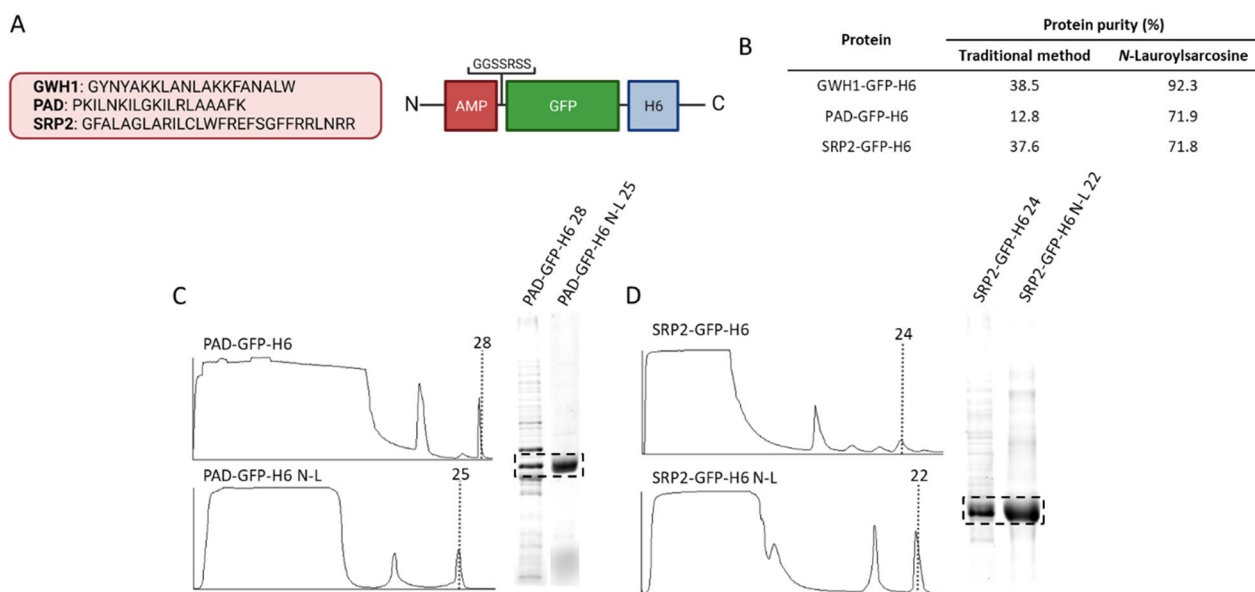
**Fig. 1** Purification of GFP-H6 and GWH1-GFP-H6. **A** Modular architecture of GFP-H6 and GWH1-GFP-H6. Box sizes are only approximate. **B** Chromatographic profile for affinity chromatography of GFP-H6 and GWH1-GFP-H6 constructs present in the *E. coli* soluble cell fraction in the absence or presence of different detergents. Vertical dashed lines and numbers indicate the fraction selected for further analysis. **C** SDS-PAGE coupled to TGX stain-free gel Technology (Bio-Rad) of the protein fractions indicated through vertical lines in the panel **B**. **D** DLS size determination of GWH1-GFP-H6 purified in the presence of alternative detergents. Numbers indicate peak sizes. **E** Temperature-dependent DLS size data of GWH1-GFP-H6 purified in the presence of alternative detergents. **F** Circular Dichroism (CD) spectra of GWH1-GFP-H6 protein samples purified in the presence of alternative detergents. **G** GFP fluorescence spectrum of GWH1-GFP-H6 purified in the presence of alternative detergents

P-80 was involved, a tendency to aggregate was suspected due to the high polydispersion index showed by the protein material. Such tendency to aggregate was confirmed by increasing the temperature of the protein sample, followed by its DLS analysis (Fig. 1E). Then, a temperature-dependent increase in the size of the materials was only observed when P-80 had been involved in their recovery. The higher stability of GWH1-GFP-H6 obtained with S-D and N-L was confirmed from the structural point of view. The  $\beta$  structure of the protein revealed by circular dichroism was highly conserved in the three detergent strategies (Fig. 1F). However, the sample treated with P-80 exhibited a lower intensity value in the whole GFP emission fluorescence spectrum when compared to the samples treated with S-D and N-L (Fig. 1G). This result demonstrated a particular tertiary structure of GWH1-GFP-H6 obtained with P-80 that was in agreement with the high size of protein oligomers and with the observed instability (Fig. 1D, E).

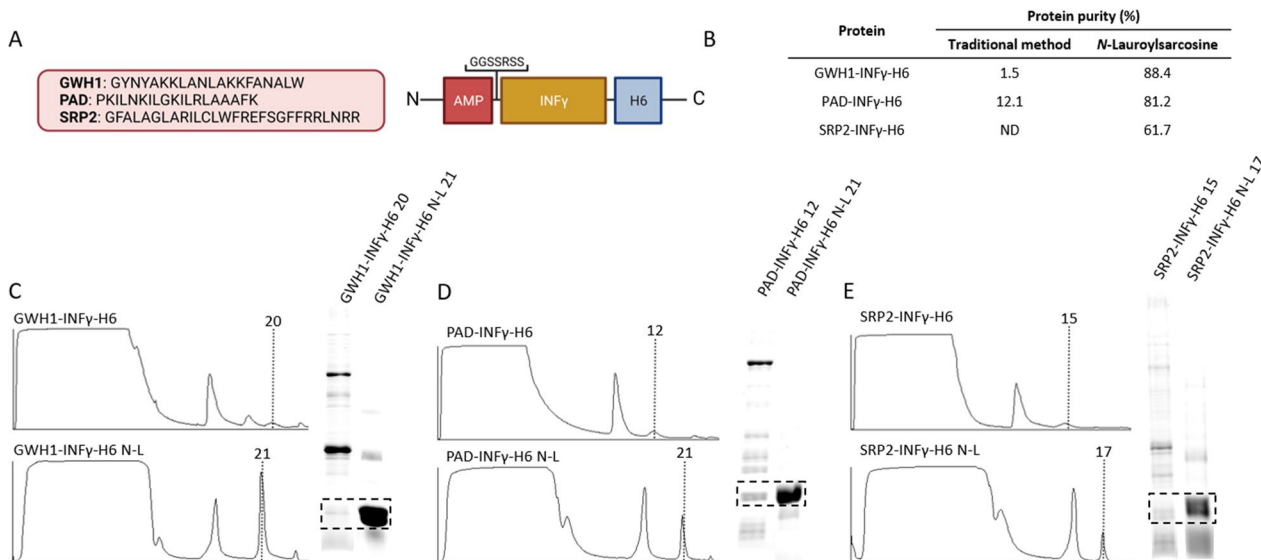
To test if the observed beneficial effect observed on protein purification might be protein-specific, two additional modular proteins based on GFP (Fig. 2A) were produced in *E. coli* and purified as described above, using

N-L as an additive in the cell extracts. These proteins contained different AMPs at the N-terminus. In general, AMPs show a cationic character expected to impair purification. Therefore, being these constructs models of difficult-to-purify proteins, they were well suited for the proposed analysis. As observed (Fig. 2B), the purity of all these proteins dramatically improved when using the detergent-assisted chromatographic method for at least two-fold. Data are summarized in Fig. 2C. Again, to further discard specific links between the improved process and a particular protein domain, all these modular proteins were produced as new modular versions by substituting the scaffold GFP by a central scaffold based on the murine interferon gamma (INF $\gamma$ ) (Fig. 3A). The purification process of all these proteins was largely improved (Fig. 3B), and the final purity was dramatically enhanced, reaching almost 60-fold in the case of GWH1-INF $\gamma$ -H6 (Fig. 3C).

Of course, the presence of a detergent in the cell extracts, even at low concentrations, might be a risk in a production process regarding the potential loss of conformation and functionalities of the target protein, which despite an improved yield could show a



**Fig. 2** Comparative purification of cationic modular proteins based on GFP. **A** General architecture of the fusion proteins used here. Box sizes were only indicative. **B** Purity data summary. **C** and **D** Chromatographic profiles for affinity chromatography of PAD-GFP-H6 (**C**) and SRP2-GFP-H6 (**D**) from the *E. coli* soluble cell fraction in absence or presence of N-L in the cell extracts. Vertical dashed lines and numbers indicate the fraction selected for further analysis. The plots are sided by SDS-PAGE coupled to TGX stain free gel Technology (Bio-Rad) of protein fractions indicated in vertical lines

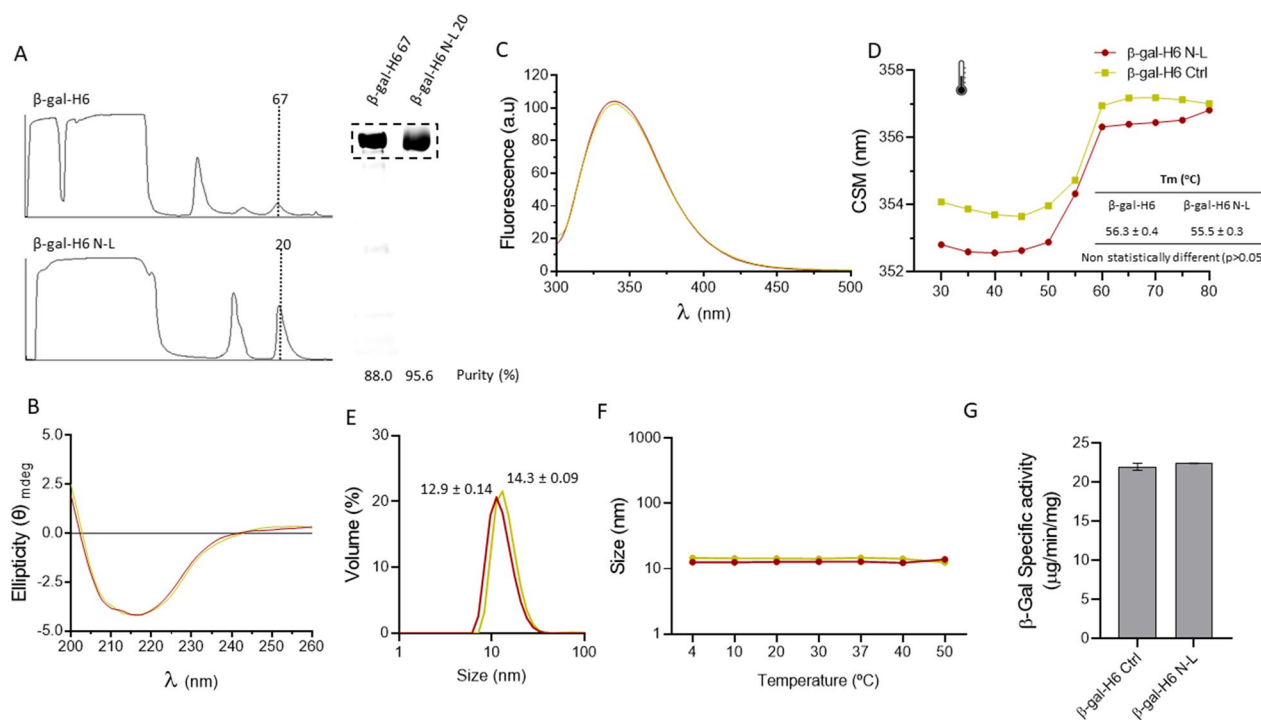


**Fig. 3** Comparative purification of cationic modular proteins based on INF $\gamma$ . **A** General architecture of the fusion proteins used here. Box sizes were only indicative. **B** Purity data summary. **C–E** Chromatographic profiles for affinity chromatography of GWH1-INF $\gamma$ -H6 (**C**), PAD-INF $\gamma$ -H6 (**D**) and SRP2-INF $\gamma$ -H6 from the *E. coli* soluble cell fraction in absence or presence of N-L in the cell extracts. Vertical dashed lines and numbers indicate the particular protein fraction selected for further analysis. The plots are sided by SDS-PAGE coupled to TGX stain free gel Technology (Bio-Rad) of protein fractions indicated by vertical lines

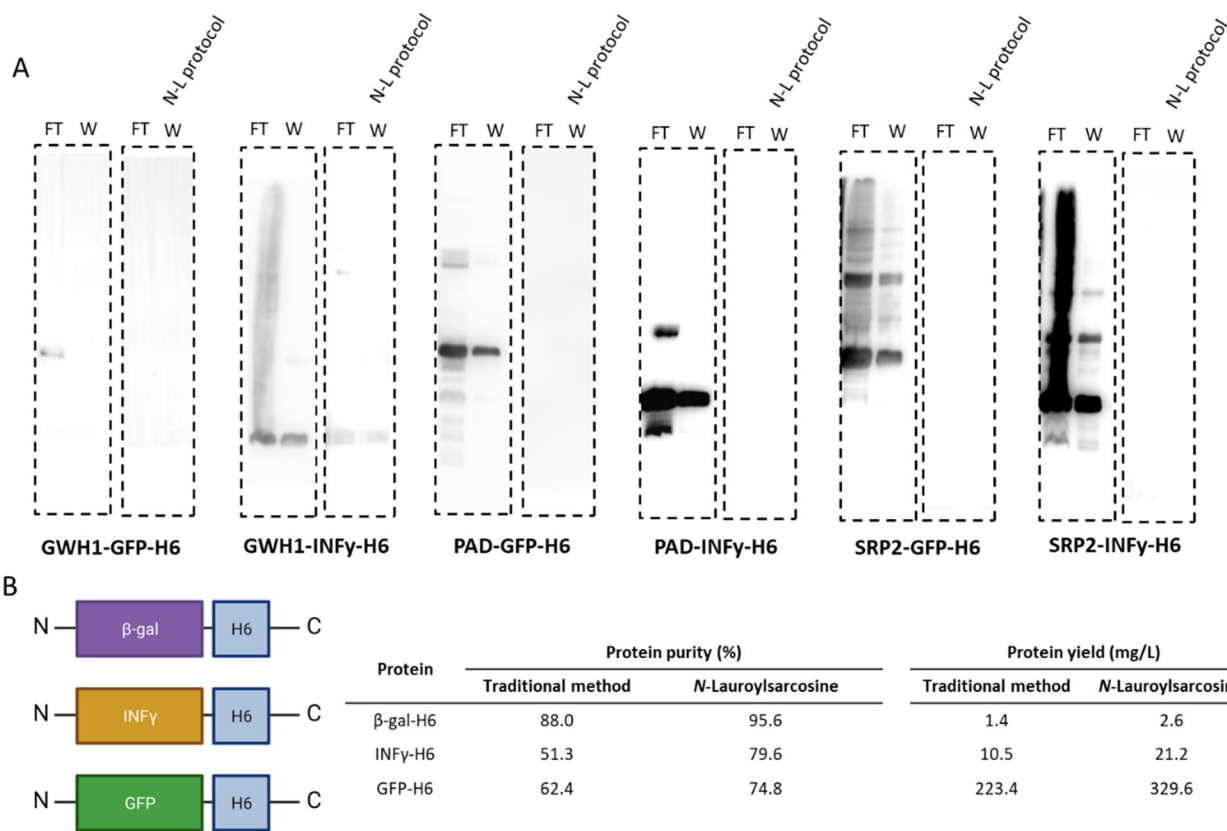
decreased biological activity. To check and to eventually discard this possibility, a conventional, his-tagged *E. coli*  $\beta$ -galactosidase was produced in recombinant form, and its enzymatic activity used as reporter to evaluate a potential impairment of the conformational quality. Being a tetrameric enzyme with four active sites at the protein–protein interfaces, this enzyme is sensitive to denaturing agents and any enzymatic loss should be indicative of a deleterious effect of the method. As in the case of previous tested proteins, N-L enhanced the protein purity after affinity chromatography (Fig. 4A). When testing structural parameters we noted that both protein samples, resulting from conventional purification and from detergent-assisted purification, showed indistinguishable properties. An overlap of the circular dichroism (Fig. 4B) and tryptophan fluorescence (Fig. 4C) spectra, and the DLS profiles (Fig. 4E) of both detergent-treated and untreated  $\beta$ -galactosidase samples confirmed that the secondary (Fig. 4B) and tertiary structures (Fig. 4C and E) of the enzyme were preserved under the tested procedure. Moreover, the thermal stability of the protein until 50 °C was unaffected (Fig. 4D and F) and there was not statistical difference between the unfolding

temperature of both protein samples determined from the CSM vs temperature curve (Fig. 4D). Importantly, and according to the absence of detectable conformational modifications, the specific activity of the enzyme remained unchanged when comparing both purification protocols (Fig. 4G).

The improved purity of all the tested proteins (Figs. 1, 2, 3, 4) found upon detergent-assisted chromatographic purification might be due, as presumed, to the hindrance of heterogeneous protein interactions but also to a higher solvent-exposure of H6, expected to result in a better binding of his-tagged polypeptides to  $\text{Ni}^{2+}$  in the columns. To test this hypothesis, the amount of recombinant protein in the flow-through and wash eluates was measured for the six complex constructs produced in the study, using either the conventional purification protocol or the detergent-assisted purification method. As observed (Fig. 5A), all the tested proteins were retained by the columns with higher efficacies when the detergent was present in the extracts than when it was absent. Such high affinity can account, by itself, the enhanced purity levels observed in the final product. Of course, since high retention in the affinity columns should not only result in



**Fig. 4** Purification of a recombinant *E. coli*  $\beta$ -galactosidase. **A** Comparative chromatographic profile of mAU signal at 280 nm for affinity chromatography of purification of  $\beta$ -gal-H6 with and without N-L. The corresponding SDS-PAGE coupled to TGX stain free gel Technology (Bio-Rad) of selected elution fractions is also shown. **B** Circular Dichroism (CD) spectra of  $\beta$ -gal-H6 protein samples purified with the N-L method. **C** Tryptophan fluorescence spectrum of  $\beta$ -gal-H6 protein samples purified with N-L. **D** Center of Spectral Mass (CSM) of tryptophan fluorescence spectrum of  $\beta$ -gal-H6 samples purified with N-L, versus temperature. Inset: T<sub>m</sub> values calculated from a sigmoidal model of CSM vs temperature. **E** DLS measurements of particle size distribution (by volume) of  $\beta$ -gal-H6 samples. **F** Thermal profile of  $\beta$ -gal-H6 protein size. **G** Comparative enzymatic activity of  $\beta$ -gal-H6 samples. Differences were not statistically different ( $p > 0.05$ )



**Fig. 5** Binding of H6-tagged recombinant proteins to  $\text{Ni}^{2+}$  columns. **A** Western Blot (WB) of protein samples purified by the traditional protocol or assisted by N-L. For each protein, the binding efficiency of the proteins to the Ni-column is shown through the flow through (FT) and the wash (W) sample analyses. **B** General architecture of the fusion proteins used here to determine final yield. Box sizes were only indicative. **C** Summary of yield and purity data

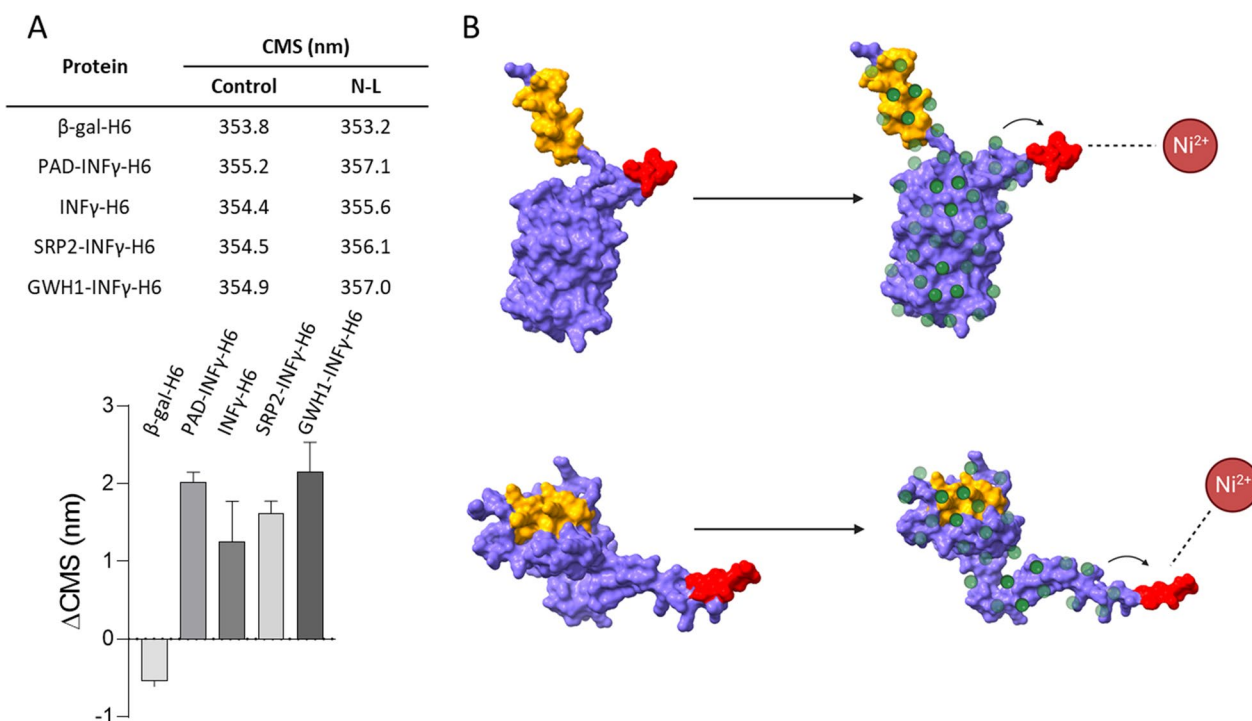
higher purity but in higher protein yields, this parameter was tested in the tagged proteins for which yield could be precisely estimated when using the conventional purification approach (that means, when the protocol rendered sufficient protein yields for their quantification) (Fig. 5B). As expected, for the three tagged proteins yield, in addition to purity, was systematically improved, (Fig. 5C).

The hypothesis that the presence of detergent leads to an increased solvent exposure of H6 in recombinant proteins would be further supported if conformational changes could be detected in the protein upon incubation with N-L at the working concentrations. This possibility was tested with the set of INF $\gamma$ -based proteins and the recombinant  $\beta$ -galactosidase. In all cases, the sets of data demonstrated that in the presence of the detergent, a negligible or a slightly change in the CMS values calculated from the Trp emission of the proteins occurred (Fig. 6A). As shown, the CMS values moved to different or higher values in most of the samples ( $\Delta\text{CMS} > 0$ ) in which N-L was present. This fact reflects moderate rearrangements of the protein structure mediated by N-L, that might be compatible with a better solvent-exposure

of H6 in this situation than when the protein was stored in absence of the surfactant, as modelled in Fig. 6B. Importantly, N-L was not detected in the final protein samples upon the dialysis steps (Appendix A), what was of course relevant regarding the potential industry-oriented uses of the proposed protocol.

## Discussion

The final protein recovery yield in downstream recombinant protein production is a critical parameter for defining the clinical or industrial applicability of promising proteins. Importantly, many of those valuable products are discarded from industrial production and marketing because their low yield or unaffordable efforts required for purification. Apart from the seminal protein fractioning into soluble and insoluble versions, it is progressively accepted from multiple independent observations that a recombinant protein produced in bacteria occurs in a continuum of conformations, and that soluble aggregates are common in the cytoplasm or recombinant bacteria [14–16, 26]. Taking this idea as a conceptual basis, the conformational spectrum of the soluble protein versions



**Fig. 6** **A** Center of spectral mass (CMS) of the Trp emission (Appendix B) in the absence (control) or in the presence of N-L. These data were applied to calculate the  $\Delta$ CMS = CMS<sub>+N-L</sub> - CMS<sub>control</sub> and all the  $\Delta$ CMS values were statistically different from zero ( $p < 0.05$ ). **B** Speculative model of the mechanism of action of N-L, whose presence could place the his-tag in a more convenient way for its binding to the Ni<sup>2+</sup> in the column. GWH1-GFP-H6 is shown at the top and GWH1-INF $\gamma$ -H6 at the bottom. Colour codes are as it follows: yellow (antimicrobial peptide), red (his-tag) purple (scaffold protein) and green (N-L molecules)

might have a differential impact on the purification efficiencies of such conformers, and at least partially, the presence of these forms might also account for the occurrence of difficult-to-purify proteins even if mostly present in the soluble cell fraction.

In this context, N-L is an anionic surfactant with multiple applications in biotechnology and biomedicine, including transdermal drug delivery [27–29], fabrication of biomaterials and nanomaterials [30–32], structural protein analysis [33, 34], separation of co-purified tag-less proteins from protein complexes [35] and mild protein solubilization/extraction [36–38], among many others. In fact, it was the first non-denaturing agent used in the solubilization of recombinant proteins from IBs [39]. This early application in the handling of IBs fully supported the concept, developed much later [40], that functional proteins could be recovered from insoluble protein aggregates by mild detergents [20], in contrast to conventional and less technologically friendly denaturing/refolding-based protocols [18].

In this line, N-L is among the most currently used agents to favor non-denaturing protein extraction from IBs, but its applicability as a generic assistant in the purification of soluble proteins had not been explored. Based

on the hypothesis that at least a fraction of difficult-to-purify protein populations might have conformational defects, the use of a mild detergent might improve their downstream. This would be especially relevant regarding tag-based affinity chromatography, in which the affinity tag in the recombinant protein should be properly exposed to the solvent for a proper binding to the immobilized ligand [41–43]. We have proved here a dramatic improvement in the column retention, final purity and recovery yield of several modular proteins (Figs. 1, 2, 3, 4, 5), in which the cationic character of a fused N-terminus domain clearly impairs its downstream (compare GFP-H6 and GWH1-GFP-H6 in Fig. 1B and C). Although a compensation of charges between the cationic domain and the anionic detergent and a consequent improvement of solubility cannot be discarded, the positive effect that the surfactant shows on the purification of a conventional  $\beta$ -galactosidase (Fig. 4) indicates that such beneficial impact has a transversal character irrespective of the protein surface charge.

Therefore, a model has been proposed to account for the positive impact that N-L shows over H6-tagged protein absorption in Ni<sup>2+</sup>-based chromatography in which the detergent slightly and reversibly unfolds the protein

(Fig. 6A), making H6 more solvent-exposed (Fig. 6B). This is well supported by the moderate changes in the model protein conformation reported here (Fig. 6A) as induced by N-L. Interestingly, such modifications do not affect the domain organization and probably involve local protein zones. It must be noted that the enzymatic activity of  $\beta$ -gal-H6, purified in absence or in presence of N-L, is indistinguishable (Fig. 4), and that the active sites of this enzyme are located in the monomer–monomer interfaces [44], making this enzyme sensitive to global conformational impacts. Also, in the line of potential obstacles for the use of N-L as a regular purification tool, the presence of detergent traces in the final product might pose severe health concerns. Indeed the presence of N-L and related surfactants in intravenous, oral, transdermal and colorectal drug formulations as protein stabilizer inhibits critical enzymes and raises several toxicological issues [28, 45–47]. Regarding the handling of N-L-treated samples, the working concentration used here (0.2%) is far below the limit of 1–30%, which corresponds to Eye Irritation 2, H319, and below the limit of  $>30\%$ , which corresponds to Skin Irritation 2, H315. Furthermore, N-L is not classified as carcinogenic, mutagenic, or toxic for reproduction, and there are no evidence of chronic toxicity, according to the 2nd ATP of Regulation (EC) No 1272/2008 (CLP) and Directive 67/548/EEC (see <https://echa.europa.eu/web/guest>). Finally, N-L is not currently restricted by REACH. Regarding the final product, fine analytical methods have not detected detergent in the final samples obtained by the proposed procedure, with an analytical detection limit of 1 ppm and a global detection limit of 10 ppm for a ten-fold diluted original sample. This is achieved by conventional dialysis procedures that might be further refined if convenient. Therefore, altogether, the data presented here suggest the repurposing of N-L from a solubilizer of IB proteins to its regular use in the purification of difficult-to-purify soluble proteins by affinity methods, as a valuable supporting agent able to improve the final yield and purity of the protein up to more than 50-fold.

## Conclusions

N-Lauroylsarcosine is a detergent commonly used for the mild solubilization of recombinant proteins from bacterial inclusion bodies. Here we have demonstrated that the repurposing of this surfactant as a generic assistant in the affinity purification of His-tagged soluble proteins, upon its addition to the crude cell extracts, dramatically improves the binding of H6-tagged proteins to  $\text{Ni}^{2+}$ . Consequently, both the yield and purity of the eluted material increases in more than one order of magnitude, resulting in improved processes and in more industrially appealing products. This approach is effective for

proteins with or without cationic domains and it does not alter their biological activities. Also, upon simple dialysis, the surfactant is removed in the final sample below the detection limits of very fine analytical methods and far from any toxicologically relevant range. The structural analyses of N-Lauroylsarcosine-exposed soluble protein indicate light conformational adjustments that might be compatible with a subtle conformational relaxation and a higher solvent-exposure of the purification tags.

## Methods

### Strains and genes

Three different *Escherichia coli* strains were employed as hosts for the recombinant production of the different proteins constructs, namely BL21 (DE3), BL21 $\lambda$  Codon plus and Origami B (DE3). The modular proteins used here as models (Fig. 1A) followed the same architectonic principle. From N-terminus to C-terminus; an antimicrobial peptide (AMP), a peptidic linker (GGSSRSS), a scaffold protein (either GFP or INF $\gamma$ ) and a H6-Tag. In GFP-H6, no AMP was present at the N-terminus. All DNA sequences were synthesized by GeneArt (Waltham, MA, USA) and codon-optimized for *Escherichia coli*. The DNA segments encoding INF $\gamma$ -H6 (mouse interferon gamma), GFP-H6, GWH1-GFP-H6, GWH1-INF $\gamma$ -H6, PaDBS1R1-GFP-H6 (abbrev. PAD-GFP-H6) [48], PAD-INF $\gamma$ -H6, SRP2-GFP-H6 [49] and SRP2-INF $\gamma$ -H6 were cloned into pET22b (Amp<sup>R</sup>). On the other hand, a  $\beta$ -galactosidase-H6 ( $\beta$ -gal-H6) encoding gene was cloned into pET26b (Kan<sup>R</sup>). The different vectors containing the synthetic genes were transformed by heat shock (42 °C for 45 s) in chemically competent *Escherichia coli* BL21 (DE3) cells for the production INF $\gamma$ -H6, GWH1-GFP-H6, GWH1-INF $\gamma$ -H6, PAD-GFP-H6, PAD-INF $\gamma$ -H6, SRP2-GFP-H6 and SRP2-INF $\gamma$ -H6, *Escherichia coli* BL21 $\lambda$  Codon plus for the production of  $\beta$ -gal-H6 or *Escherichia coli* Origami B (DE3) for the production of GFP-H6. Protein sequences are indicated in the [Appendix C](#).

### Gene expression and recombinant protein production

All *E. coli* cultures (1–2 L) were grown at 37 °C and 250 rpm in LB broth with ampicillin at 100  $\mu\text{g}/\text{ml}$  or kanamycin at 34  $\mu\text{g}/\text{ml}$  (only for  $\beta$ -gal-H6 production). *E. coli* Origami B (DE3) cultures were grown in LB with ampicillin at 100  $\mu\text{g}/\text{ml}$ , kanamycin at 34  $\mu\text{g}/\text{ml}$  and tetracycline at 12.5  $\mu\text{g}/\text{ml}$ . Once an optical density ( $\text{OD}_{600}$ ) of around 0.5 was reached, isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) was added at 0.1 mmol/L and the expression temperature set at 20 °C for and overnight culture. Cells were harvested by centrifugation (5000g for 15 min at 4 °C) and stored at  $-80$  °C.



### Protein purification protocols

For protein purification, the cell pellet was resuspended in wash buffer (Buffer A=40 mM Tris HCl (pH 8) / 500 mM NaCl) with the protease inhibitor complete EDTA-free (Roche). Bacterial cell lysis was performed by high-pressure homogenization using the Avestin Emulsiflex C5 (ATA scientific). After the disruption, the cellular lysate was divided into aliquots of approximately 25 ml. To each one of these aliquots a specific volume of a solution containing 40 mM Tris HCl (pH 8) and 500 mM NaCl, 2% N-L was added so that the final concentration of N-L was 0.2%. The sample containing the detergent was incubated at room temperature with little agitation during 15–18 h. The next day, the soluble and insoluble fractions were separated by centrifugation (45 min, 15,000g at 4 °C) and the soluble fraction was filtrated firstly through a 0.45 µm filter and then through a 0.22 µm filter (Millex<sup>®</sup>-GP, Millipore Express<sup>®</sup> PES Membrane Filter Unit). Recombinant proteins were purified in an Äkta Pure FPLC system (GE Healthcare) by immobilized metal affinity chromatography (IMAC). After selective binding to a His-Trap HP 5 ml column (GE Healthcare), proteins were eluted by a linear gradient of elution buffer (40 mM Tris HCl (pH 8), 500 mM NaCl, 500 mM Imidazol and 0.2% N-L). The eluted fractions were analyzed through Mini-PROTEAN TGX Stain-Free Gels and Western blot. The selected protein fractions were dialyzed against sodium bicarbonate salt buffer (166 mmol/L NaCO<sub>3</sub>H and 333 mmol/L NaCl, pH 8.0) or bicarbonate buffer (166 mmol/L NaCO<sub>3</sub>H, pH 8.0) and the final protein concentration was determined by NanoDrop One Microvolume UV–Vis Spectrophotometer (Thermo Scientific). The regular purification protocol does not include the presence of the detergent N-L and the overnight incubation at room temperature.

### Dynamic light scattering

The volume size distribution (expressed in nm) of all protein candidates was determined by Dynamic Lights Scattering (DLS) at 633 nm and both standard (25 °C) and increasing temperatures (from 4 to 50 °C) in a Zetasizer Advance Pro (Malvern Panalytical) using a ZEN2112 3 mm quartz cuvette. Samples were measured at least in triplicate, gaussian curves represented, and data expressed as mean ± standard error respectively.

### Fluorescence measurements

Fluorescence spectra were obtained with a Cary Eclipse spectrofluorometer (Agilent Technologies) with a quartz cell of 2 mm path length.

### Trp fluorescence

In proteins other than GFP we determined the tryptophan (Trp) fluorescence. The excitation and emission slit were set at 5 nm. The excitation wavelength ( $\lambda_{ex}$ ) was 295 nm. Fluorescence emission spectra were acquired within a range from 300 to 500 nm. To evaluate the effect of N-L on the Trp emission, 0.2% of N-L was added to each protein sample. To analyse the thermal stability of proteins each spectrum was acquired in an increasing temperature range (from 25 to 80°C) and the temperature of unfolding ( $T_m$ ) was calculated from the center of spectral mass (CSM) vs temperature curve as described elsewhere [50]. For GFP versions, the excitation slit was set at 2.5 nm and the emission slit at 5 nm. The excitation wavelength ( $\lambda_{ex}$ ) was 488 nm. The fluorescence emission spectra were acquired within a range from 500 to 600 nm.

### Circular dichroism

Data were collected in a Jasco J-715 spectropolarimeter (JASCO, Oklahoma City, OK, USA) with a thermostatic device by a Peltier system spectropolarimeter, using a 0.2 mm path length quartz cell. Each spectrum was an average of seven scans. The protein concentration was 0.15 mg/ml in each protein buffer. CD spectra were collected from 260–200 nm. Each final spectrum was obtained from three replicas. Finally, we applied a negative exponential equation for the data smoothing.

### Determination of $\beta$ -galactosidase enzymatic activity

$\beta$ -Galactosidase activity was measured in PBS 1X by monitoring the colorimetric signal at 420 nm produced by the degradation of an artificial substrate, o-nitrophenyl- $\beta$ -D-galactopyranoside (ONPG) [51].

### Detection of N-L in protein samples by mass spectrometry

N-L in dialyzed protein samples was determined using mass spectrometry by the Chemical Analysis Service (SAQ), a technological unit at the Universitat Autònoma de Barcelona (UAB). For that, an HPLC system was used equipped with a DAD detector from Agilent Technologies and a micrOTOF-Q (time-of-flight) mass spectrometer from Bruker Daltonics. Electrospray ionization in negative polarity was used to register the surfactant. 10 µL of samples were injected using water:acetonitrile (ACN) (1:1) at 0.1 ml/min to the mass spectrometer. The parameters were adjusted for sample injection from the liquid chromatograph using Flow Injection Analysis (FIA) and detection in the mass spectrometer, without using any chromatographic column, and taking a 20 ppm N-L standard to obtain maximum instrumental sensitivity. The mass spectrometer was focused on low masses ( $m/z < 1000$ ), given the molecular

weight of the surfactant ( $m/z = 270.2$ ), and the registrations were performed in negative polarity. To ensure that there was no carryover between protein samples, the injection was changed to positive polarity, and the protein was then injected. Data analysis indicated that there was no accumulation of protein in the instrument. Protein samples were diluted 1:10 in water to reduce the carbonate concentration and to make it compatible with the operational conditions of the instrument. Control, N-L-enriched samples were prepared using the same procedure but adding N-L in the vial to 1 ppm. This method allows then the detection of 1 ppm of surfactant per vial, that for ten-fold diluted samples makes a detection limit of 10 ppm per protein sample.

### Protein design and Three-Dimensional (3D) structure prediction

The 3D structures of the stable folded state of GWH1-GFP-H6 and GWH1-INF $\gamma$ -H6 were predicted in silico using the AlphaFold2 [52] algorithm integrated in ColabFold [53] and

using the default settings after introducing each primary FASTA sequence as query, respectively. Molecular graphics and analyses were performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with the support from NIH P41-GM103311 [54].

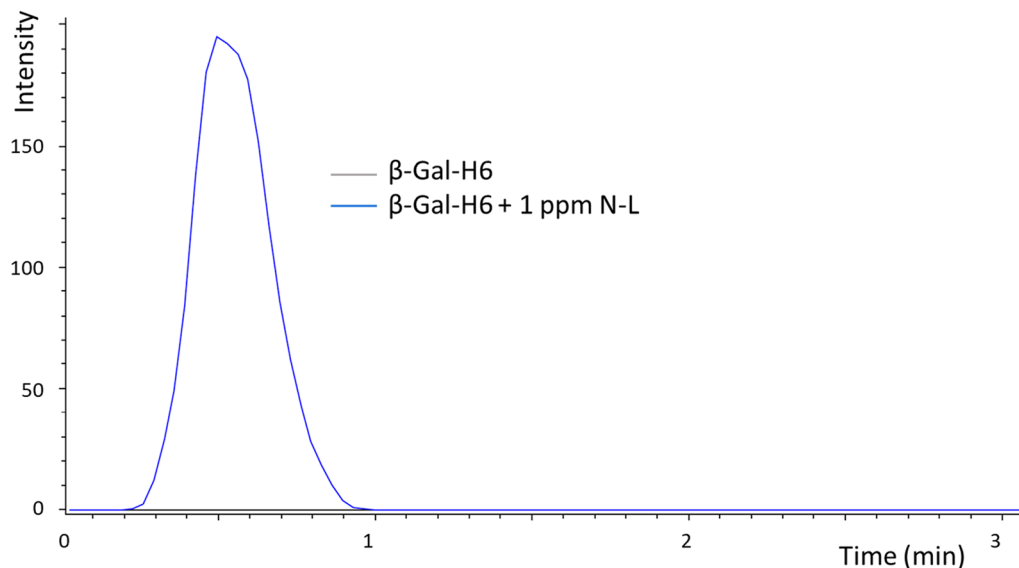
### Statistical analysis

All statistical analyses were conducted in GraphPad Prism version 8.0.0 for Windows (GraphPad Software, San Diego, CA, USA) with at least two independent replicates unless otherwise indicated. T tests to assess the differences in  $T_m$  values or in the specific enzymatic activity values were applied assuming unequal variances. All quantitative values were expressed as mean  $\pm$  standard error of the mean. The significance of the statistical difference was included in each experiment. A nonlinear regression analysis was developed to determine the melting temperature ( $T_m$ ) of  $\beta$ -gal-H6 in both conditions from a sigmoidal model.

## Appendices

### Appendix A

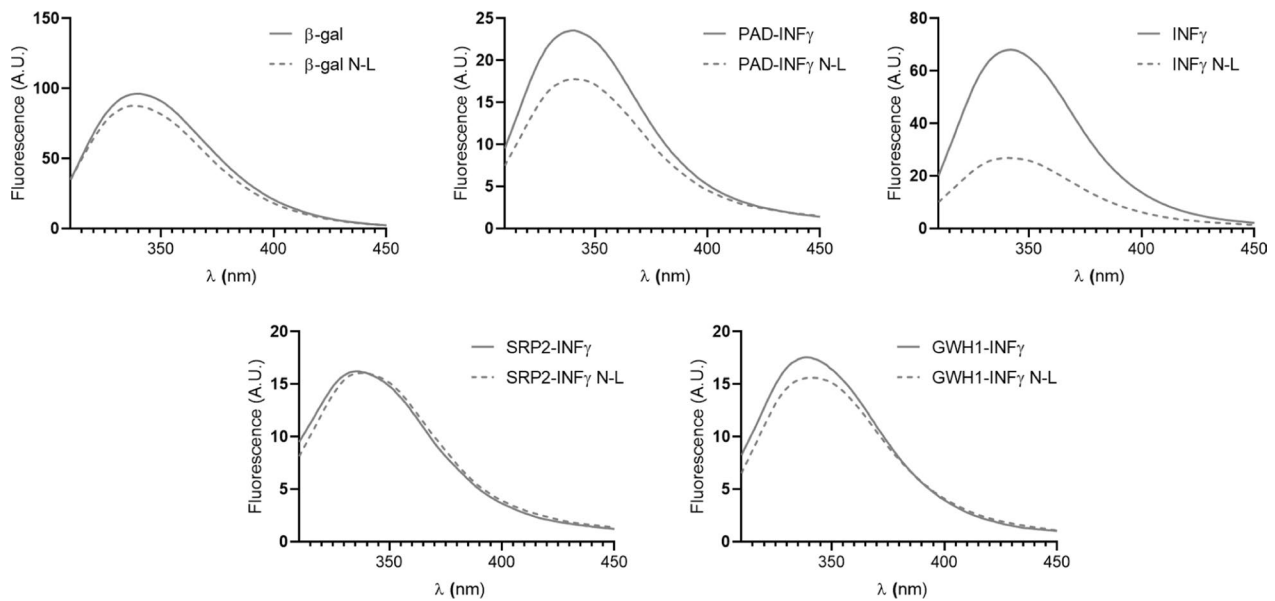
Analysis of N-L in dialyzed protein samples.



Superimposed Extracted Ion Chromatograms (EICs) at  $m/z=270.2$  of a purified  $\beta$ -Gal-H6 sample (grey) and an equivalent sample in presence of added 1 ppm N-L (blue).

## Appendix B

Tryptophan emission spectra of non GFP proteins (0.2 mg/ml) in the presence or absence of 0.2% N-L.



## Appendix C

Amino acid sequences of the proteins used in the present study.

- eGFP-H6 MSKGEELFTGVVPILVELDGDVNG-HKFSVSGEGEGDATYGKLTCLKFICTTGKLPVP-WPTLVTTLTLYGVQCFSRYPDHMKRHDFFK-SAMPEGYVQERTISFKDDGNYKTRAEVKFEFEG-DTLVNRIELKGIDFKEDGNILGHKLEYNYNSH-NVYITADKQKNGIKANFKIRHNIEDGSVQLAD-HYQQNTPIGDGPVLLPDNHYLSTQSALS KDP-NEKRDHMLVLEFVTAAGITHGMDELYKHHH-HHH
- $\text{INF}\gamma$ -H6 MAQQQFFREIENLKEYFNASSPDVAK-GGPLFSEILKNWKDESDKKIISQIVSFYFKLFEN-LKDNQVIQRSMDIKQDMFQKFLNGSSEKLED-FKKLIQIPVDDLQIQRKAINELIKVMNDLSPKSN-LRKRKRSQNLFRGRRRASTKHHHHHHH
- GWH1-GFP-H6 MGYNYAKKLANLAKKFANALWGGSSRSSKGEELFTGVVPILVELDGD-

VNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLTLYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMLVLEFVTAAGITHGMDELYHHH-HHHH

- GWH1- $\text{INF}\gamma$ -H6 MGYNYAKKLANLAKKFANALWGGSSRSSHGTVIESLESLNNYFNSSGIDVEEKSLFLDIWRNWQKDGMMKILQSQIISFYLR-LFEVLKDNQAISNNISVIESHLITTFFSNSKAK-DAFMSIAKFEVNNPQVQRQAFNELIRVVQLL-PESSLRKRKRSRCHHHHHH
- PaDBS1R1-GFP-H6 MPKILNKILGKILRLAAAFKGGSSRSSKGEELFTGVVPILVELDGDVNG-HKFSVSGEGEGDATYGKLTCLKFICTTGKLPVP-WPTLVTTLTLYGVQCFSRYPDHMKRHDFFK-SAMPEGYVQERTISFKDDGNYKTRAEVKFEFEG-DTLVNRIELKGIDFKEDGNILGHKLEYNYNSH-NVYITADKQKNGIKANFKIRHNIEDGSVQLAD-HYQQNTPIGDGPVLLPDNHYLSTQSALS KDP-NEKRDHMLVLEFVTAAGITHGMDELYHHH-HHH
- PaDBS1R1- $\text{INF}\gamma$ -H6 MPKILNKILGKILRLAAAFKGGSSRSSHGTVIESLESLNNYFNSSGIDVEEKSLFLDIWRNWQKDGMMKILQSQIISFYLR-LFEVLKDNQAISNNISVIESHLITTFFSNSKAKDAFM-

SIKFEVNNPQVQRQAFNELIRVVHQLLPES-  
LRKRKRSRCHHHHHH

- SRP2-GFP-H6 MGFALAGLARILCLWFREFSGF-FRRLNRRFAMRRRGSSRSSKGEELFTGVVPIL-VELDGDVNGHKFSVSGEGDATYGLKTLK-FICTTGKLPVPWPTLVTTLTYGVQCFSRYPD-HMKRHDFFKSAMPEGYVQERTISFKDDGNYK-TRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK-LEYNYNSHNVYITADKQKNGIKANFKIRH-NIEDGSVQLADHYQQNTPIGDGPVLLPDNHYL-STQSALS KDPNEKRDMVLLFVTAAGITHGM-DELYHHHHHHH
- SRP2-INF $\gamma$ -H6 MGFALAGLARILCLWFREFSGF-FRRLNRRFAMRRRGSSRSHGTVIESLESNNY-FNSSGIDVEEKSLFLDIWRNWQKDGDMKILQS-QIISFYLRLEFVLDKNQAINNISVIESHLITTF-SNSKAKDAFMSIAKFEVNNPQVQRQAFNELIR-VVHQLLPESLRKRKRSRCHHHHHH
- $\beta$ -gal-H6 GSHMLEDPVVLQRRD WENPGVTQLN-RLAAHPFASWRNSEEARTDRPSQQLRSLNGE-WRFAWFPAPAVPESWLECDLPEADTVVPSN-WQMHGYDAPYITNVYTPITVNPFPVPTENPT-GCYSLTFNVDES WLQEGQTRIFDGVNSAF-HLWCNGRWVGYGQDSRLPSEFDLSAFLRA-GENRLAVMVLRWSDGSYLEDQDMWRMSGI-FRDVSL LHKPTTQISDFHVATRFNDDFSRAV-LEAEVQMC GELRDYLRVTVSLWQGETQVAS-GTAPFGGEIIDERGGYADRVTLRLNVENPKL-WSAEIPNLYRAVELHTADGTLIEAEACD-VGFREVRIENGLLLLNGKPLLIRGVNRHE-HHPLHGQVMDEQTMVQDILLMKQN-NFNAVRC SHYPNHPLWYTLCDRYGLYV-VDEANIETHGMVPMNRLTDDPRWLPAM-SERVTRMVQRDRNHPSVIIWSLGNESGH-GANHDALYRWIKSVDPSRPVQYEGGGA-DTTATDIICPMYARVDEDQFPAPVPKWSIK-KWLSLPGETRPLILCEYAHAAAGNSLGGFAKY-WQAFRQYPRLQGGFVWDVVDQSLIKYDENG-NPWSAYGGDFGDTPNDRQFCMNGLVFADRT-PHPALTEAKHQQQFFQFRLSGQTIEVTSEY-LFRHSDNELLHWMVALD GKPLASGEVPLD-VAPQGKQLIELPELPQPESAGQLWLTVRV-VQPNATAWSEAGHISAWQQWRLAENLSVTL-PAASHAIPHLLTSEMDFCIELGNKRWQFNRS-GFLSQMWIGDKKQLLTPLRDQFTRAPLDN-DIGVSEATRIPNAWVERWKAAGHYQAEAL-LQCTADTLADAVLITTAHAWQHKGKTLFISRK-TYRIDGSGQMAITVDVEVASDTPHPARIGLNC-QLAQAERVNWLGLGPQENYDRLTAACFDR-WDLPLSDMYTPYVFPSENGLRGCTRELNYG-PHQWRGDFQFNISRYSQQLMETSHRHLL-HAEEGTWLNIDGFHMGIGGDDSWSPSVS-AEFQLSAGRYHYQLVWCQKLGHHHHHHH

#### Acknowledgements

Protein production was partially performed by the ICTS "NANBIOSIS", more specifically by the Protein Production Platform of CIBER in Bioengineering, Biomaterials & Nanomedicine (CIBER-BBN)/ IBB, at the UAB (<http://www.nanbiosis.es/portfolio/u1-protein-production-platform-ppp/>). The authors also acknowledge the UAB scientific-technical services LLEB and SAQ (<https://www.uab.cat/web/research/scientific-technical-services/scientific-and-technical-services--1345667278676.html>) for the physicochemical characterization of the recombinant proteins.

#### Author contributions

JVC, JAG, HLL and JMS performed the experimental, analyzed the data and prepared the figures. EV, AV and NFM funded the study and designed and supervised the experimental. JVC, JAG and NFM designed the whole study and its concept. AV prepared the draft of the manuscript, that was revised and edited by all the authors. All authors read and approved the final manuscript.

#### Funding

The authors appreciate the financial support received for the development of multimeric recombinant drugs, from AEI (PID2020-116174RB-I00 to A.V.; PID2019-105416RB-I00/AEI/10.13039/501100011033 and PDC2022-133858-I00 to E.V.; PID2019-107298RB-C22 granted to N.F.-M) and AGAUR (SGR202100092 to A.V.). A.V. received an ICREA ACADEMIA award. JVC is supported with a Margarita Salas grant for the training of young doctoral graduates (722713). J.A.-G. is supported by a predoctoral fellowship from the Ministerio de Universidades (FPU20/02260). HLL is supported by a predoctoral fellowship from AGAUR (2019 FI\_B 00352). JMS is supported with a María Zambrano postdoctoral researcher contract (677904) from Ministerio de Universidades and European Union ("Financed by European Union-Next GenerationEU").

#### Availability of data and materials

The dataset supporting the conclusions of this article is available at <https://doi.org/10.34810/data616>

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

Received: 2 February 2023 Accepted: 4 April 2023

Published online: 25 April 2023

#### References

1. Langer E, Rader R. Biopharmaceutical manufacturing: historical and future trends in titers, yields, and efficiency in commercial-scale bioprocessing. *BioProcessing J.* 2015;13:47–54
2. Sanchez-Garcia L, Martín L, Mangués R, Ferrer-Miralles N, Vázquez E, Villaverde A. Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microb Cell Fact.* 2016;15:33.
3. Puetz J, Wurm FM. Recombinant proteins for industrial versus pharmaceutical purposes: a review of process and pricing. *Processes.* 2019;7:476.
4. Pavlou AK, Reichert JM. Recombinant protein therapeutics—success rates, market trends and values to 2010. *Nat Biotechnol.* 2004;22:1513–9.
5. de Pinho Favaro MT, Atienza-Garriga J, Martínez-Torró C, Parladé E, Vázquez E, Corchero JL, et al. Recombinant vaccines in 2022: a perspective from the cell factory. *Microb Cell Fact.* 2022;21:203.
6. Deckers M, Deforce D, Fraiture MA, Roosens NHC. Genetically modified micro-organisms for industrial food enzyme production: an overview. *Foods.* 2020;9:326.

7. Wingfield PT. Overview of the purification of recombinant proteins. *Curr Protoc Protein Sci.* 2015. <https://doi.org/10.1002/0471140864.ps0601s80>.
8. Corchero JL, Gasser B, Resina D, Smith W, Parrilli E, Vázquez F, et al. Unconventional microbial systems for the cost-efficient production of high-quality protein therapeutics. *Biotechnol Adv.* 2013;31:140–53.
9. Kőszagová R, Nahálka J. Inclusion bodies in biotechnology. *J Microbiol Biotechnol Food Sci.* 2020;9:1191–6.
10. Rinas U, García-Fruitós E, Corchero JL, Vázquez E, Seras-Franzoso J, Villaverde A. Bacterial inclusion bodies: discovering their better half. *Trends Biochem Sci.* 2017;42:726–37.
11. Haacke A, Fendrich G, Ramage P, Geiser M. Chaperone over-expression in *Escherichia coli*: apparent increased yields of soluble recombinant protein kinases are due mainly to soluble aggregates. *Protein Expr Purif.* 2009;64:185–93.
12. van der Henst C, Charlier C, Deghelt M, Wouters J, Matroule JY, Letesson JJ, et al. Overproduced *Brucella abortus* PdhS-mCherry forms soluble aggregates in *Escherichia coli*, partially associating with mobile foci of IbpA-YFP. *BMC Microbiol.* 2010;10:248.
13. Fatima U, Singh B, Subramanian K, Guptasarma P. Insufficient (sub-native) helix content in soluble/solid aggregates of recombinant and engineered forms of IL-2 throws light on how aggregated IL-2 is biologically active. *Protein Journal.* 2012;31.
14. Martínez-Alonso M, González-Montalbán N, García-Fruitós E, Villaverde A. The functional quality of soluble recombinant polypeptides produced in *Escherichia coli* is defined by a wide conformational spectrum. *Appl Environ Microbiol.* 2008;74:7431–3.
15. Vera A, González-Montalbán N, Arís A, Villaverde A. The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnol Bioeng.* 2007;96:1101.
16. Schrödel A, de Marco A. Characterization of the aggregates formed during recombinant protein expression in bacteria. *BMC Biochem.* 2005;6:10.
17. González-Montalbán N, García-Fruitós E, Villaverde A. Recombinant protein solubility—does more mean better? *Nat Biotechnol.* 2007;25:718–20.
18. Singh SM, Panda AK. Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng.* 2005;99:303.
19. Vallejo LF, Rinas U. Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb Cell Fact.* 2004;3:11.
20. Singhvi P, Saneja A, Srichandan S, Panda AK. Bacterial inclusion bodies: a treasure trove of bioactive proteins. *Trends Biotechnol.* 2020;38:474–86.
21. Zhang C-Y, Zhao S-Q, Zhang S-L, Luo L-H, Liu D-C, Ding W-H, et al. Database study on the expression and purification of membrane proteins. *Protein Pept Lett.* 2021;28:972–82.
22. Antil M, Gouin SG, Gupta V. Truncation of C-terminal intrinsically disordered region of mycobacterial Rv1915 facilitates production of “difficult-to-purify” recombinant drug target. *Front Bioeng Biotechnol.* 2020;8:522.
23. Serna N, Céspedes MV, Sánchez-García L, Unzueta U, Sala R, Sánchez-Chardi A, et al. Peptide-based nanostructured materials with intrinsic proapoptotic activities in CXCR4<sup>+</sup> solid tumors. *Adv Funct Mater.* 2017;27:1700919.
24. Carratalá JV, Serna N, Villaverde A, Vázquez E, Ferrer-Miralles N. Nanostructured antimicrobial peptides: the last push towards clinics. *Biotechnol Adv.* 2020;44:107603.
25. Carratalá JV, Brouillette E, Serna N, Sánchez-Chardi A, Sánchez JM, Villaverde A, et al. In vivo bactericidal efficacy of gwh1 antimicrobial peptide displayed on protein nanoparticles, a potential alternative to antibiotics. *Pharmaceutics.* 2020;12:1217.
26. González-Montalbán N, García-Fruitós E, Villaverde A. Recombinant protein solubility—does more mean better? *Nat Biotech.* 2007;25:718–20.
27. Kim YC, Park JH, Ludovice PJ, Prausnitz MR. Synergistic enhancement of skin permeability by N-lauroylsarcosine and ethanol. *Int J Pharm.* 2008;352:129–38.
28. Lee H, Park J, Kim YC. Enhanced transdermal delivery with less irritation by magainin pore-forming peptide with a N-lauroylsarcosine and sorbitan monolaurate mixture. *Drug Deliv Transl Res.* 2018;8:54–63.
29. Kováčik A, Kopečná M, Vávrová K. Permeation enhancers in transdermal drug delivery: benefits and limitations. *Expert Opin Drug Deliv.* 2020;17:145–55.
30. Wang H, Wang JG, Zhou HJ, Liu YP, Sun PC, Chen TH. Facile fabrication of noble metal nanoparticles encapsulated in hollow silica with radially oriented mesopores: Multiple roles of the N-lauroylsarcosine sodium surfactant. *Chem Commun.* 2011;47:7680–2.
31. Titkov AI, Bulina NV, Ulihin AS, Shundrina IK, Karpova EV, Gerasimov EY, et al. N-Lauroylsarcosine capped silver nanoparticle based inks for flexible electronics. *J Mater Sci Mater Electron.* 2017;28:2029.
32. Kurosaki T, Kishikawa R, Matsumoto M, Kodama Y, Hamamoto T, To H, et al. Pulmonary gene delivery of hybrid vector, lipopolyplex containing N-lauroylsarcosine, via the systemic route. *J Controlled Release.* 2009;136:213.
33. Thibeault J, Patrick J, Martin A, Ortiz-Perez B, Hill S, Zhang S, et al. Sarkosyl: a milder detergent than SDS for identifying proteins with moderately high hyperstability using gel electrophoresis. *Anal Biochem.* 2019;571:21.
34. Mukherjee S, Dubois C, Perez K, Nisbet RM, Li QX, Varghese S, et al. Quantitative proteomics of detergent insoluble tangles derived from Alzheimer's disease brains. *Alzheimers Dement.* 2021;17:e055040.
35. Ukwatthage TO, Goodwin OY, Songok AC, Tafaro AM, Shen L, Macnaughtan MA. Purification of Tag-Free *Chlamydia trachomatis* Scc4 for structural studies using sarkosyl-assisted on-column complex dissociation. *Biochemistry.* 2019;58:4284–92.
36. Balaban CL, Suárez CA, Boncompain CA, Peressutti-Bacci N, Ceccarelli EA, Morbidoni HR. Evaluation of factors influencing expression and extraction of recombinant bacteriophage endolysins in *Escherichia coli*. *Microb Cell Fact.* 2022;21:40.
37. Gifre-Renom L, Baltà-Foix R, Arís A, García-Fruitós E. Nondenaturing solubilization of inclusion bodies from lactic acid bacteria. *Methods Mol Biol.* 2022;2406:389–400.
38. Bader JA, Shoemaker CA, Klesius PH. Immune response induced by N-lauroylsarcosine extracted outer-membrane proteins of an isolate of *Edwardsiella ictaluri* in channel catfish. *Fish Shellfish Immunol.* 2004;16:415–28.
39. Frankel S, Sohn R, Leinwand L. The use of sarkosyl in generating soluble protein after bacterial expression. *Proc Natl Acad Sci U S A.* 1991;88:1192.
40. Peternel Š, Grdadolnik J, Gaberc-Porekar V, Komel R. Engineering inclusion bodies for non denaturing extraction of functional proteins. *Microb Cell Fact.* 2008;7:34.
41. Kimple ME, Brill AL, Pasker RL. Overview of affinity tags for protein purification. *Curr Protoc Protein Sci.* 2013. <https://doi.org/10.1002/0471140864.ps0909s7>.
42. Mahmoodi S, Pourhassan-Moghaddam M, Wood DW, Majdi H, Zarghami N. Current affinity approaches for purification of recombinant proteins. *Cogent Biol.* 2019;5:1665406.
43. Arnau J, Lauritzen C, Petersen GE, Pedersen J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expr Purif.* 2006;48:1–13.
44. Matthews BW. The structure of *E. coli*  $\beta$ -galactosidase. *C R Biol.* 2005;328:549–56.
45. Bocci V, Naldini A, Corradeschi F, Lencioni E. Colorectal administration of human interferon- $\alpha$ . *Int J Pharm.* 1985;24:109–14.
46. Whitehead K, Karr N, Mitragotri S. Discovery of synergistic permeation enhancers for oral drug delivery. *J Controlled Release.* 2008;128:128–33.
47. Sattler J, Hesterberg R, Schmidt U, Crombach M, Lorenz W. Inhibition of intestinal diamine oxidase by detergents: a problem for drug formulations with water insoluble agents applied by the intravenous route? *Agents Actions.* 1987;20:270–3.
48. Irazazabal LN, Porto WF, Fensterseifer ICM, Alves ESF, Matos CO, Menezes ACS, et al. Fast and potent bactericidal membrane lytic activity of PaDBS1R1, a novel cationic antimicrobial peptide. *Biochim Biophys Acta Biomembr.* 2019;1861:178–90.
49. Lützelberger M, Groß T, Käufer NF. Srp2, an SR protein family member of fission yeast: In vivo characterization of its modular domains. *Nucleic Acids Res.* 1999;27:2618–26.
50. Carratalá JV, Cano-Garrido O, Sánchez J, Membrado C, Pérez E, Conchillo-Solé O, et al. Aggregation-prone peptides modulate activity of bovine interferon gamma released from naturally occurring protein nanoparticles. *N Biotechnol.* 2020;57:11–9.
51. Sánchez JM, López-Laguna H, Álamo P, Serna N, Sánchez-Chardi A, Nolan V, et al. Artificial Inclusion Bodies for Clinical Development. *Advanced Science.* 2020;7:1902420.

52. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583–9.
53. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods*. 2022;19:679–82.
54. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25:1605–12.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

