Aggregation-prone peptides modulate activity of bovine interferon gamma released from naturally occurring protein nanoparticles

José Vicente Carratalá[a,b], Olivia Cano-Garrido[a,b,c,1], Julieta Sánchez[a,2], Cristina Membrado[a,b], Eudald Pérez[a,b], Oscar Conchillo-Solé[a], Xavier Daura[a,d], Alejandro Sánchez-Chardi[e], Antonio Villaverde[a,b,c], Anna Arís[f], Elena Garcia-Fruitós[f], Neus Ferrer-Miralles[a,b,c]#

[a]Institute for Biotechnology and Biomedicine, Autonomous University of Barcelona, Bellaterra, Barcelona, Spain

[b]Department of Genetics and Microbiology, Autonomous University of Barcelona, Bellaterra, Barcelona, Spain

[c]Bioengineering, Biomaterials and Nanomedicine Networking Biomedical Research Centre (CIBER-BBN), Bellaterra, Barcelona, Spain

[d]Catalan Institution for Research and Advanced Studies, Barcelona, Spain

[e]Microscopy Service, Autonomous University of Barcelona, Bellaterra, Barcelona, Spain.

[f]Department of Ruminant Production, Institute of Agrifood Research and Technology (IRTA), Caldes de Montbui, Barcelona, Spain

[1]Present address: Nanoligent SL. Eureka Building. Campus of the Autonomous University of Barcelona. Bellaterra, Barcelona, Spain

[2]Permanent address: National University of Cordoba, Faculty of Exact, Physical and Natural Sciences, ICTA and Chemistry Department, CONICET Institute of Biological and Technological Research (IIByT) Cordoba Argentina

#Correspondence: Neus Ferrer-Miralles, neus.ferrer@uab.cat.

Abstract

Efficient protocols for the production of recombinant proteins are indispensable for the development of the biopharmaceutical sector. Accumulation of recombinant proteins in naturally-occurring protein aggregates is detrimental to biopharmaceutical development. In recent years, the view of protein aggregates has changed with the recognition that they are a valuable source of functional recombinant proteins. In this study, bovine interferon-gamma (rBoIFN-γ) was engineered to enhance the formation of protein aggregates, also known as protein nanoparticles (NPs), by the addition of aggregation-prone peptides (APPs) in the generally recognized as safe (GRAS) bacterial *Lactococcus lactis* expression system. The L6K2, HALRU and CYOB peptides were selected to assess their intrinsic aggregation capability to nucleate protein aggregation. These APPs enhanced the tendency of the resulting protein to aggregate at the expense of total protein yield. However, fine physico-chemical characterization of the resulting intracellular protein NPs, the protein released

from them and the protein purified from the soluble cell fraction indicated that the compactability of protein conformations was directly related to the biological activity of variants of IFN-γ, used here as a model protein with therapeutic potential. APPs enhanced the aggregation tendency of fused rBoIFN-γ while increasing compactability of protein species. Biological activity of rBoIFN-γ was favored in more compacted conformations. Naturally-occurring protein aggregates can be produced in GRAS microorganisms as protein depots of releasable active protein. The addition of APPs to enhance the aggregation tendency has a positive impact in overall compactability and functionality of resulting protein conformers.

Abbreviations

NPs, nanoparticles; APPs, aggregation-prone peptides; GRAS, Generally Recognized as Safe; rBoIFN-γ, recombinant bovine IFN-γ; IBs, inclusion bodies; HSA, hot spot area; NHSA, normalized hot spot area; a^4vAHS, average aggregation-propensity hot spot

Keywords

Interferon-gamma, protein nanoparticles, protein aggregation, *Lactococcus lactis*, Generally Recognized as Safe, conformational compactability

Introduction

The efficient production and purification of recombinant proteins in a wide range of expression hosts has driven the launch of a large number of biopharmaceutical products. One of the most-studied and most-used gene expression systems for biopharmaceutical products is *Escherichia coli* [1,2]. Prokaryotic endotoxin-free expression systems are being explored to avoid the presence pro-inflammatory contamination by lipopolysaccharide (LPS) components of the outer leaflet of the outer membrane of *E. coli*, including *E. coli* LPS mutant strains [3,4] and Generally Recognized As Safe (GRAS) microorganisms, such as *Lactococcus lactis* [5-7].

During recombinant gene expression, the stress imposed on the protein quality control machinery leads, in most cases, to the accumulation of the recombinant protein in aggregates that form intracellular nanoparticles (NPs), known as inclusion bodies (IBs) [8-10]. These are dynamic and complex nanostructures with a variable content of recombinant protein [11-13]. The trapped protein was formerly thought to be biologically inactive due to aberrant protein conformations or inactive partially folded species incompatible with biological activity. The recombinant protein can often be recovered, with low efficiency, from the insoluble cell fraction by *in vitro* denaturing/refolding processes [14]. However, this view of naturally occurring protein aggregates has changed radically since the detection of biologically active protein embedded in these aggregates [15-17]. The classic view of protein aggregates as mere inactive folding intermediates has been transformed into one of heterogeneous porous multimeric structures stabilized by a scaffold of cross beta-sheet structures containing conformers of the recombinant protein in which a spectrum of species containing quasi-native conformations are incorporated [9]. It has been reported that

biologically active protein species can be extracted from IBs, indicating the biologically active nature of proteins forming these aggregates [18]. Hence, IBs are envisionaged as non-toxic, biocompatible and mechanically stable materials from which biologically active molecules of the recombinant protein can be released under mild solubilization and physiological conditions [13,16,18-21].

Interest in the possibility of controlling the aggregation of recombinant proteins in these types of nanostructures is increasing, and several aggregation-prone peptides (APPs) have been identified for fusion with recombinant proteins to enhance the aggregation process in the producing cell [22]. In this study, interferon (IFN)-γ was selected as a model protein in order to study the effect of the addition of APPs in naturally occurring protein aggregates due to interest in this activity in biomedicine and its potential use in animal health. IFN-γ is the sole type II IFN. IFN-γ secretion by natural killer (NK) cells and antigen-presenting cells enhances the innate immune response, while T-lymphocytes are involved in the secretion of IFN-γ in the adaptive immune response [23,24]. The activity of IFN-γ depends on its interaction, as a dimer, with the IFN-γ receptor (IFNGR). Approved recombinant human IFN-γ can be obtained from the *E. coli* expression system, but novel protein formulations need to be developed in GRAS expression systems due to safety concerns. In most reported studies of the expression and purification of IFN-γ, the recombinant protein is recovered from the purified IBs through extensive denaturation-refolding processes [25-28].

In this work, the mature form of bovine IFN-γ (rBoIFN-γ) protein (UniProtKB P07353, residues 24 to 166) was produced in GRAS lactic acid bacteria (*L. lactis*) in the form of protein NPs.

The ability of APPs fused to rBoIFN-γ to enhance the aggregation propensity of the recombinant cytokine was analyzed and the link assessed between the biological activities contained in protein NPs of IFN-γ variants and their physicochemical characteristics. It was found that the activity of the IBs is related to the specific biological activity of the recombinant protein they contain, whereas the proportion of released protein is not the main factor. The data presented illustrate the potential of endotoxin-free protein NPs as active biomaterials to formulate, at the nanoscale level, releasable proteins of biomedical interest.

Materials and methods

Bacterial strains and plasmids

E. coli MC4100 (StrepR) [29] was used for cloning genes for protein production in *L. lactis*. *E. coli* DH5α was used for cloning genes in *E. coli*. *L. lactis cremoris* NZ9000 (Boca Scientific, MA, USA), and *ClearColi®* BL21(DE3) (Lucigen, WI, USA) were used in experiments for each expression system. Gene sequences were codon optimized for the *L. lactis* expression host

(Geneart, MA; USA, Suppl. Figure S1). For *L. lactis* expression vectors, IFN-γ of bovine origin (*Bos Taurus*; NM_ 174086.1 in Suppl. Figure S1) was cloned into the CmR pNZ8148 plasmid (MoBiTech, Goettingen, Germany) as described in Supplementary Materials and Methods and [6]. In addition, fusions of rBoIFN-γ with APPs were constructed (rBoIFN-γ_L6K2, rBoIFN-γ_HALRU and rBoIFN-γ_CYOB; **Figure 1**). L6K2 is a surfactant-like peptide with aggregating properties [30]. HALRU and CYOB are aggregating-prone peptides from Cytochrome bo3 ubiquinol oxidase subunit 1 from *E. coli* (UniProtKB P0ABI8), HALRU: Aragonite protein AP7 (UniProtKB Q9BP37) selected with AGGRESCAN [31] (see **Table 1**). For Clearcoli®, the *L. lactis* codon-optimized bovine IFN-γ gene was cloned into pETDuet-1 (Novagen, WI, USA) (Suppl. Figure S2 and Supplementary Materials and Methods). The recombinant proteins were produced as the mature form of the IFN-γ (from Gln24 to Thr166; NP_776511.1) (Figure S2). All genes were C-terminally fused to a His-tag for detection and quantification by western blot analysis and a linker with a predicted random coil conformation was positioned between the IFN-γ and APP as previously described [30].

Selection of APPs

APPs were selected by scanning the Disprot v6.02 database [32] with AGGRESCAN software [31]. The selection was based on the assumption that APPs in solvent-exposed regions were the best candidates for the purposes of this study. Two unstructured regions were selected from two different proteins namely CYOB: Cytochrome bo3 ubiquinol oxidase subunit 1 from *E. coli* (UniProtKB P0ABI8) and HALRU: Aragonite protein AP7 (UniProtKB

Q9BP37). CYOB was selected as the peptide displaying the highest hot spot area (HSA). HALRU showed a high normalized hot spot area (NHSA) and average aggregation-propensity hot spot (a4vAHS) while maintaining a significantly high HSA value relative to the other identified peptides. L6K2 was selected based on previous experimental results [30] after analysis with AGGRESCAN showed that this peptide had a high normalized HSA (NHSA) and high average aggregation-propensity hot spot (a^4vAHS) despite having shorter sequence (**Table 1** and **Figure 1**a).

Production and purification of rBoIFN-γ protein from the soluble cell fraction

Cultures of *ClearColi®* BL21 (DE3) cells transformed with the plasmid pETDuet-rBoIFN-γ (Supplementary Materials and Methods) were incubated in a shake flask at 37 °C and 250 rpm in LB medium supplemented with 100 µg/ml ampicillin. Protein expression was induced by adding 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The cultures were then incubated at 20 °C and 250 rpm overnight for protein production. Cells were collected by centrifugation (15 min, 6,000 x *g*, 4 °C), and soluble rBoIFN-γ protein was purified as described in Supplementary Materials and Methods. Protein expression of *L. lactis* cells transformed with plasmid containing the rBoIFN-γ gene was induced and purified as described in Supplementary Materials and Methods and [6]. The control protein rBoIFN-γ_Std, produced in *E. coli* was obtained from R&D Systems (2300-BG-025, R&D Systems, MN, USA).

Production and purification of rBoIFN-γ protein nanoparticles.

L. lactis cells transformed with expression plasmids (pNZ8148-rBoIFN-γ, pNZ8148-rBoIFN-γ_L6K2, pNZ8148-rBoIFN-γ_HALRU and pNZ8148-rBoIFN-γ_CYOB) were grown as above. NP production was induced by adding 12.5 ng/ml nisin (Sigma-Aldrich, MO, USA) to *L. lactis* cultures. After induction, the cultures were grown for 5 h.

The protein NPs were purified using the protocol described previously (Supplementary Materials and Methods and [6]).

Quantitative protein analysis

Recombinant proteins were quantified by denaturing SDS-PAGE as described previously (Supplementary Materials and Methods and [33]). In addition, the yields of purified proteins in each of the formats are shown in Table S1.

Ultrastructural characterization

To characterize the morphometry of the NPs, microdrops of protein aggregate suspensions were deposited for 2 min on silicon wafers (Ted Pella Inc.), air-dried and observed in a near-native state under a field emission scanning electron microscope (FESEM) Zeiss Merlin (Zeiss, Obercochen, Germany) operating at 1 kV. Micrographs were acquired with a high-resolution in-lens secondary electron (SE) detector. Images were taken at magnifications ranging from 20,000x to 80,000x.

Z potential analysis

Z potential (ZP) characterization of each kind of protein NP was carried out by Dynamic Light Scattering (DLS) (Zetasizer Nano ZS, Malvern Instruments Ltd, Malvern, UK). To prevent the electrodes from burning, the samples were prepared in deionized (MilliQ) water. Each sample was analyzed in triplicate.

Determination of rBoIFN-γ biological activity in bovine cells

The different rBoIFN-γ formulations described here were analyzed by a modified kynurenine bioassay (Supplementary Materials and Methods and [34]). The antiproliferative activity of IFN-γ in this assay is related to the induction of the expression of the indoleamine 2,3-dioxygenase 1 (*IDO1*) gene, which is the first and rate-limiting enzyme in tryptophan catabolism. IDO1 catalyzes oxidative cleavage of tryptophan to N-formylkynurenine. Following a hydrolysis step, the latter is transformed into L-kynurenine by Ehrlich's reagent, giving a yellow-colored compound absorbing at 490 nm [35]. The absorbance vs IFN-γ concentration (nmol/L) curves were adjusted to Eq. 1 [20]. Abs*490* is the absorbance at 490 nm, which represents an indirect measurement of IFN-γ binding to the receptor, Abs*max* is the maximal binding of IFN-γ to the receptor, and K_D is the equilibrium dissociation constant. A low value of K_D indicates high IFN-γ affinity to the receptor.

$$\mathrm{Abs}490 = \frac{\mathrm{Abs}max \times \mathrm{IFN}\gamma}{[\mathrm{IFN}\gamma]+\mathrm{K_D}} \qquad (1)$$

Assay of protein solubilization from protein nanoparticles

The rBoIFN-γ protein NPs (rBoIFN-γ_L, rBoIFN-γ_L6K2, rBoIFN-γ_CYOB and rBoIFN-γ_HALRU) were solubilized in PBS. In all cases, the concentration was adjusted to 1 µmol/L. After manual agitation, each sample was incubated at 37 °C for 96 h to reproduce the conditions used during the biological activity analysis. Protein concentration was quantified and the biological activity determined at a single concentration (0.72 nmol/L) as described in previous section.

Interferon size determination

The volume size distribution of IFN-γ was determined by DLS. A 60-µl aliquot (stored at -80 °C) was thawed, and the volume size distribution of each protein format was immediately determined at 633 nm (Zetasizer Nano ZS, Malvern Instruments Ltd, Malvern, UK).

Analysis of protein conformation by intrinsic tryptophan fluorescence

Fluorescence spectra were recorded on a Cary Eclipse spectrofluorometer (Agilent Technologies, CA, USA). A quartz cell with a 10-mm path length and a thermostatic holder was used. The excitation and emission slits were set at 5 nm. The excitation wavelength (λ_{ex}) was set at 295 nm. Emission spectra were acquired within a range from 310 to 550 nm. The protein concentration was 14 µmol/L in PBS DEFINE. *T*o evaluate conformational differences between the proteins, the center of spectral mass (CSM), was applied, the weighted average of the fluorescence spectrum peak. The CSM was calculated for each of

the fluorescence emission spectra [36] according to Eq.2, where I_i is the fluorescence intensity measured at wavelength λ_i.

$$\lambda = \frac{\sum \lambda_i . I_i}{\sum I_i} \qquad (2)$$

CSM values were analyzed at room temperature and under thermal heating at 5 °C/min rate.

Statistical analysis

Prior to the use of parametric tests, normality and homogeneity of variances were tested using the Shapiro-Wilk test for all quantitative data or the Levene test for raw or transformed data. Divergences between groups were tested with one-way ANOVA, and pairwise comparisons were made with Student's t tests. The results were expressed as the arithmetic mean for non-transformed data ± the standard error of the mean ($\bar{x} \pm$ SEM), except otherwise stated.

The least squares method was applied to fit functions through a regression analysis to determine the K_D values according to Eq. 1. Significance was accepted at $p < 0.05$, and Bonferroni correction was applied for sequential comparisons. All statistical analyses were performed with SPSS v. 18 for Windows.

Results and Discussion

Production of rBoIFN-γ in *L. lactis*

In *L. lactis*, most of the rBoIFN-γ protein was detected in the soluble cell fraction in the absence of any APP (**Figure 1**b, upper panel). This observation is in agreement with previous results for the expression of the natural DNA sequence of the bovine IFN-γ gene in *E. coli* [37]. The presence of the APPs in the recombinant protein caused a noticeable shift of the final products toward the insoluble cell fraction, as expected (**Figure 1**b, lower panel). The purity of the protein aggregates ranged between 50-60 % in all constructs (Suppl. Figure S3). The APP resulting in the highest aggregation tendency was the L6K2 peptide. In addition, the presence of an APP tag also had a negative effect on the total recombinant protein produced in the cell (**Figure 1**a, upper panel). The best APP in terms of aggregation propensity and protein yield in the insoluble cell fraction, corresponded to the IFN-γ L6K2 formulation. The performance of this surfactant-like peptide exceeded the predicted aggregation-prone capabilities of CYOB and HALRU peptides (**Table 1**).

Nanoarchitectonic characterization of protein nanoparticles

The morphometry of purified protein NPs of the rBoIFN-γ variants was examined by FESEM (**Figure 2**a). The images revealed the presence of multimeric complexes comprising discrete NPs in addition to isolated protein NPs (inset **Figure 2**a). The NPs were similar to rBoIFN-γ protein NPs obtained previously in this expression system [6]. ZP measurements showed that all of the NPs presented negatively charged surfaces with negative values ranging from -38 to -28 mV (**Figure 2**b), indicating the stability of the NP suspension. The higher values of ZP obtained for the IFN-γ variants provide information about particle

stability, as NPs displaying higher ZP values (>+30 mV or <-30 mV) exhibit increased stability due to greater electrostatic repulsion between particles [38].

Biological activity of soluble IFN-γ and NPs of IFN-γ

The activity of IFN-γ is usually determined by an antiviral assay [39]. However, alternative assays have been developed to simplify the procedure. One approach to evaluate IFN-γ activity mediated by IFN-γ receptor binding is the detection of L-kynurenine. The activity of IFN-γ is highly species-specific and, a specific assay for the bovine IFN-γ was developed and validated in this study [34]. For validation, the activity of three soluble rBoIFN-γ proteins was tested (**Figure 3**a). rBoIFN-γ_Std exhibited the lowest dissociation constant (K_D) among the proteins purified from the soluble cell fraction (**Figure 3**a) with a similar value determined for the human IFN-γ [40]. The difference in this parameter with in-house IFN-γ produced in *Clearcoli* (rBoIFN-γ_*E. coli*) may be related to the absence of C-terminal variants in this sample, the effect of the fused His-tag at the C-terminus, or other variables [41]. The protein obtained from the *L. lactis* expression system displayed less activity, which may be due to differences in the folding efficiency during the production process among prokaryotic expression systems [42,43]. Once the activity assay was validated, the biological activity contained in the IFN-γ protein NPs produced in *L. lactis* was determined. The results showed that all cells were able to elicit responses to the presence of the protein NPs, and the IFN-γ_L6K2 formulation displayed the highest initial rate and kynurenine production (**Figure 3**b). The addition of HALRU and CYOB APP to IFN-γ had a moderate effect on the cell response.

288 It is of interest to know why the sample corresponding to protein NPs of IFN-γ_L6K2 had the
289 highest activity and initial rate, even compared with commercial IFN-γ. Consistent with this
290 observation, a previous analysis of the activity of recombinant β-galactosidase produced in
291 *E. coli* in the form of protein NPs revealed higher specific activity than the corresponding
292 soluble version of the protein [15]. However, protein NPs obtained from *E. coli* have not
293 been characterized in detail. The activity displayed by *E. coli* IBs has been attributed to the
294 release of a spectrum of conformers of the recombinant protein, which leaves a scaffold
295 that is resistant to proteolysis and has an extensive cross-beta-pleated sheet conformation
296 [44,45]. For protein NPs of rBoIFN-γ produced in *L. lactis*, 30-40 % of the material is resistant
297 to proteolysis, indicating that the protein NPs obtained in this expression system follow
298 similar principles to the *E. coli* system [6]. Thus, the activities displayed by the protein NPs
299 are probably due to the partial release of the IFN-γ that forms part of the macromolecular
300 complex [46]. To evaluate better the ability of the protein NPs to release protein, they were
301 incubated in PBS for 96 h to emulate the protein release conditions established during the
302 biological activity assay (see the experimental design used to obtain the different protein
303 samples in **Figure 3**c). Release of 52.67 %, 5.30 %, 0.42 % and 0.46 % was observed for IFN-
304 γ, IFN-γ_L6K2, IFN-γ_HALRU and IFN-γ_CYOB NPs, respectively. In order to analyze the
305 specific activity of the proteins released from the protein NPs, an activity assay was
306 performed and the results compared with proteins obtained directly from the soluble cell
307 fraction (**Figure 3**d). The results showed that the maximal specific activity corresponded to
308 the IFN-γ_L6K2 protein released from NPs. In addition, the comparison of the specific
309 activity of the rBoIFN-γ protein produced in *L. lactis* and purified from the soluble cell

fraction with that of the corresponding protein released from NPs suggested that the released protein elicited better conformational performance (compare the second and last bars in **Figure 3**d).

The addition of APPs to the rBoIFN-γ protein improved the aggregation profile of the produced protein (**Figure 1**b). However, the presence of this type of peptide had a negative effect on the overall production of the protein and, in the case of HALRU and CYOB, a major impact on biological activity (**Figure 3**). From this, AGGRESCAN software is able to predict the propensity of the resulting APP-containing recombinant IFN-γ to aggregate and is a reliable tool for analyzing solubility performance in the design of recombinant genes [31].

Physicochemical characterization of soluble IFN-γ and nanoparticles of IFN-γ

The precise physicochemical analysis of recombinant proteins is important for safety concerns [47,48]. To further analyze the protein in different formats, DLS measurements were performed (**Figure 4**a-**4**d). The rBoIFN-γ_Std exhibited a peak with a maximum at 7.6 nm, similar to the peak at 6.13 nm for the IFN-γ produced in *L. lactis*. This configuration (6-8 nm) might correspond to the dimeric form of the cytokine. However, the IFN-γ obtained from *E. coli* showed a tendency towards a larger size. Therefore, the specific activity of the different rBoIFN-γ formats is not simply linked to the dimeric configuration, which is the functional conformation when binding to the cell receptor, and some other variables might be involved. When analyzing the size of the purified NPs, a peak above 1,000 nm was detected, exceeding the upper sensitivity limit of the equipment (**Figure 4**b). The NPs were

clustered in higher-order complexes from monomeric versions of 200 nm (**Figure 2**a). All samples exhibited the same profile. After solubilization of the protein embedded in the NPs, the size of the remaining material remained above 1,000 nm since the scaffold of the NPs retained the overall structure after the protein was released (**Figure 4**d). The released protein showed a narrow dispersion ranging from the dimeric size of the protein identified in the samples obtained from the commercial IFN-γ or the soluble version purified from *L. lactis* detected in the upper panel of Figure 4a (**Figure 4**c). In addition, the polydispersity index (PI) of these samples was higher than that of the soluble IFN-γ versions. The PI corresponds to an estimate of the width of the distribution, and higher values of PI are consistent with the data showing a pool of conformers in the folding of recombinant proteins when the proteins are produced in the cell [49]. In contrast, in the protein versions purified from the soluble cell fraction, the downstream processing based on affinity chromatography selects only a narrow collection of conformers (only those that are able to bind to the Ni^{2+} in the resin). This indicates that the protein obtained during solubilization assays from protein NPs is more representative of the diversity in conformations of a single protein that are produced in the expression system.

To further analyze the link between the physicochemical properties and the specific activity of the proteins, the fluorescence emission of Trp was recorded. Each fluorescence emission spectrum was transformed into a CSM value. This parameter is related to the relative exposure of the Trp to the protein environment. The maximum red-shift in the CSM of the Trp spectrum is compatible with large solvent accessibility [50-52], whereas the blue shift in the CSM corresponds to a Trp hidden in a more hydrophobic milieu [53]. The mature

353 form of BoIFN-γ has a unique Trp at position 36, which is partially buried in the 3D structure
354 of the protein (PDB 1D9C) [54] and is not involved in either monomer or in cytokine-
355 receptor interactions, as shown in the 3D structure of the human tetrameric complex of the
356 cytokine dimer with the receptor (PDB 1FG9) [55]. A remarkable aspect of the intrinsic
357 fluorescence analysis is that all the rBoIFN-γ variants within the NPs or after solubilization
358 from the protein NPs exhibited lower CSM values than the samples obtained from the
359 soluble fraction (**Table 2**). These results suggest that the protein forming part of the NPs
360 and the protein solubilized from the aggregates have a more compact conformation than
361 the soluble version. The most active IFN-γ soluble version corresponded to the commercial
362 IFN-γ, which had the lowest CMS due to its highly compacted structure. The proteins
363 obtained from the soluble fraction of *E. coli* and *L. lactis* exhibited higher CMS values than
364 the commercial protein. These differences might be related to the distinct sizes detected
365 (**Figure 4**a). The rBoIFN-γ_*E. coli* was approximately three times larger than the same
366 protein produced in *L lactis*, indicating that the Trp residue was located in a more polar
367 environment compared with the *L lactis* form (**Table 2**). For the protein originating from the
368 particulate form, a blue shift was observed compared with the soluble versions, and the
369 CSM increased as it was resolubilized (lines 4 and 6 of **Table 2)**. The CSM value of the
370 solubilized rBoIFN-γ_*L. lactis* protein sample did not reach that of the soluble counterpart
371 (lines 3 and 4 of **Table 2**). When the APPs were incorporated in the engineered protein
372 constructs, the solubilized proteins showed a decrease in the CSM values compared with
373 the protein NPs samples (lines 5 and 7). This behavior suggests a possible self-arrangement
374 of the tag within the protein that could replace water molecules and increase the

hydrophobicity of the Trp environment. The CYOB construct (line 6 of **Table 2**) required a specific analysis as this tag contributes five additional Trp residues to the whole protein structure. In this case, the solubilized protein spectrum exhibited a modest red shift (higher CMS value) compared with the particulate form, indicating that the solubilization process exposed some of the Trp residues to a hydrophilic environment. The CSM values of the CYOB and HALRU protein NPs remained unaltered after solubilization (**Table 2**, lines 6 and 7). These data are in accordance with the higher stability of the particulate forms, which exhibited low levels of protein release.

The NP form of IFN-γ also favored the specific activity (insets, **Figure 3**, *L. lactis* and non-tagged rBoIFN). This phenomenon is not only due to more active conformation of the protein (**Figure 4** and **Table 2**, line 3 vs line 4) [56] but also to the heterogeneous distribution of the protein and the ability of the protein NPs to increase the effective concentration of protein in the proximity of the receptor. Moreover, the formulation containing L6K2 was the most efficient, even compared with the commercial protein. Solubilization clearly conferred the most active and altered conformation of the protein without the tag. Although a low percentage of protein was released from the NPs containing L6K2, at least in PBS, this released protein seems to be sufficient to surpass the activity of the released protein without a tag (**Figure 3**). Furthermore, the CSM thermal profile of the released proteins demonstrated that L6k2 tag not only confers the highly compact structure (low CSM value, **Table 2**) but also contribute to a unique and complete thermal unfolding profile (Figure S4).

Another interesting aspect is the effect of the size of the tag on the structure-function of the protein. The incorporation of a tag larger than 17 amino acids beyond the linker (**Figure 1**a) could generate steric problems preventing the interaction of tagged IFN-γ with the receptor. As shown in **Figure 1**, L6K2 is only 8 amino acids, compared with 17 amino acids for HALRU and 38 amino acids for CYOB. The short size of the L6K2 tag might reduce the difficulty of the interaction between L6K2-IFN-γ and the receptor compared with the longer IFN-γ tags since the C-terminal end of the protein, where the APPs are fused, is located in close proximity to the receptor in the 3D structure in PDB 1FG9.

In the recombinant protein production platform, the general consensus for improving protein yield is to improve the solubility of the protein. However, solubility and conformational quality are not necessarily coincident parameters [57]. The functionalities of the protein obtained from the soluble cell fraction or the protein NPs of rBoIFN-γ_*L. lactis* in the present work supported these previous findings, as the protein obtained from the soluble cell fraction was less active than that recovered from the protein NPs. The compactabilities of the conformations of these proteins were in agreement with their dissimilar biological activity. Therefore, results obtained in this study may indicate that the compactability of protein conformations is a significant parameter related to stability and function [58,59].

Conclusions

In this study, it was demonstrated that the addition of aggregation-prone peptides (APPs) promoted the production of naturally occurring protein nanoparticles (NPs) of interferon gamma (IFN-γ) in the generally recognized as safe (GRAS) *Lactococcus lactis* expression system. The fine physico-chemical characterization of the resulting proteins revealed that conformational compactability was directly related to the biological performance of the recombinant IFN-γ.

Acknowledgments

This work was supported by grants from INIA, MINECO, Spain to N.F.M. and E.G.F. (RTA2015-00064-C02-01 and RTA2015-00064-C02-02). The authors acknowledge financial support granted to A.V. from AGAUR (2017 SGR-229) and from Bioengineering, Biomaterials and Nanomedicine Networking Biomedical Research Centre (CIBER-BBN), financed by the Carlos III Health Institute, Spain, with assistance from the European Regional Development. We are also indebted to the CERCA Programme (Generalitat de Catalunya) and European Social Fund for supporting our research. J.V.C. received a pre-doctoral fellowship from UAB, O.C.G. received a PhD fellowship from MECD (FPU), and E.G.F. received a post-doctoral fellowship from INIA (DOC-INIA). AV has been distinguished with an ICREA ACADEMIA Award. The authors also acknowledge ICTS ‘‘NANBIOSIS”, more specifically the Protein Production Platform of CIBER- BBN/IBB, at the UAB sePBioEs scientific-technical service (http://www.nanbiosis.es/unit/u1-protein-production-platform-ppp/) and the UAB scientific-technical services LLEB, SM and SCAC

(https://www.uab.cat/web/research/scientific-technical-services/all-scientific-technical-services--1345667278676.html). The authors would like to thank Milena Tileva for her helpful advice on technical issues related to the experimental adjustment of the IFN-γ activity bioassay. Special thanks to Sandra Párraga-Ferrer for the design of **Figure 3**c.

Author contributions

E. Garcia-Fruitós and N. Ferrer-Miralles designed and supervised the experiments. J.V. Carratalà, O. Cano-Garrido, J. Sánchez, C. Membrado, E. Pérez, O. Conchillo-Solé and A. Sánchez-Chardi performed the experiments. J. V. Carratalà, O. Cano-Garrido, J. Sánchez and N. Ferrer-Miralles analyzed the data. All authors edited the manuscript. N. Ferrer-Miralles wrote the paper.

References

[1] Baeshen MN, Al-Hejin AM, Bora RS, Ahmed MM, Ramadan HA, Saini KS et al. Production of Biopharmaceuticals in *E. coli*: Current Scenario and Future Perspectives. J Microbiol Biotechnol 2015;25:953-62. https://doi.org/10.4014/jmb.1412.12079

[2] Sanchez-Garcia L, Martin L, Mangues R, Ferrer-Miralles N, Vazquez E, Villaverde A. Recombinant pharmaceuticals from microbial cells: a 2015 update. Microb Cell Fact 2016;15:33. https://doi.org/10.1186/s12934-016-0437-3

[3] Rueda F, Cespedes MV, Sanchez-Chardi A, Seras-Franzoso J, Pesarrodona M, Ferrer-Miralles N et al. Structural and functional features of self-assembling protein nanoparticles produced in endotoxin-free *Escherichia coli*. Microb Cell Fact 2016;15:59. 10.1186/s12934-016-0457-z [doi]

[4] Mamat U, Wilke K, Bramhill D, Schromm AB, Lindner B, Kohl TA et al. Detoxifying *Escherichia coli* for endotoxin-free production of recombinant proteins. Microb Cell Fact 2015;14:57. 10.1186/s12934-015-0241-5 [doi]

[5] Garcia-Fruitos E. Lactic Acid Bacteria: a promising alternative for recombinant protein production. Microb Cell Fact 2012;11:157. 10.1186/1475-2859-11-157 [doi]

[6] Cano-Garrido O, Sanchez-Chardi A, Pares S, Giro I, Tatkiewicz WI, Ferrer-Miralles N et al. Functional protein-based nanomaterial produced in microorganisms recognized as safe: A new platform for biotechnology. Acta Biomater 2016;43:230-9. https://doi.org/10.1016/j.actbio.2016.07.038

[7] Ferrer-Miralles N, Villaverde A. Bacterial cell factories for recombinant protein production; expanding the catalogue. Microb Cell Fact 2013;12:113. https://doi.org/10.1186/1475-2859-12-113

[8] Kopito RR. Aggresomes, inclusion bodies and protein aggregation. Trends Cell Biol 2000;10:524-30.

[9] Rinas U, Garcia-Fruitos E, Corchero JL, Vazquez E, Seras-Franzoso J, Villaverde A. Bacterial Inclusion Bodies: Discovering Their Better Half. Trends Biochem Sci 2017;42:726-37. https://doi.org/10.1016/j.tibs.2017.01.005

[10] Rodriguez-Carmona E, Mendoza R, Ruiz-Canovas E, Ferrer-Miralles N, Abasolo I, Schwartz S Jr et al. A novel bio-functional material based on mammalian cell aggresomes. Appl Microbiol Biotechnol 2015;99:7079-88. https://doi.org/10.1007/s00253-015-6684-0

[11] Garcia-Fruitos E. Inclusion bodies: a new concept. Microb Cell Fact 2010;9:80. https://doi.org/10.1186/1475-2859-9-80

[12] Jurgen B, Breitenstein A, Urlacher V, Buttner K, Lin H, Hecker M et al. Quality control of inclusion bodies in *Escherichia coli*. Microb Cell Fact 2010;9:41. https://doi.org/10.1186/1475-2859-9-41

[13] Singh A, Upadhyay V, Upadhyay AK, Singh SM, Panda AK. Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. Microb Cell Fact 2015;14:41. https://doi.org/10.1186/s12934-015-0222-8

[14] Yamaguchi H, Miyazaki M. Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. Biomolecules 2014;4:235-51. https://doi.org/10.3390/biom4010235

[15] Garcia-Fruitos E, Gonzalez-Montalban N, Morell M, Vera A, Ferraz RM, Aris A et al. Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. Microb Cell Fact 2005;4:27. https://doi.org/10.1186/1475-2859-4-27

[16] Jevsevar S, Gaberc-Porekar V, Fonda I, Podobnik B, Grdadolnik J, Menart V. Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. Biotechnol Prog 2005;21:632-9. https://doi.org/10.1021/bp0497839

[17] Nahalka J, Nidetzky B. Fusion to a pull-down domain: a novel approach of producing *Trigonopsis variabilis*D-amino acid oxidase as insoluble enzyme aggregates. Biotechnol Bioeng 2007;97:454-61. https://doi.org/10.1002/bit.21244

[18] Gifre-Renom L, Cano-Garrido O, Fabregas F, Roca-Pinilla R, Seras-Franzoso J, Ferrer-Miralles N et al. A new approach to obtain pure and active proteins from *Lactococcus lactis* protein aggregates. Sci Rep 2018;8:13917. 10.1038/s41598-018-32213-8 [doi];10.1038/s41598-018-32213-8 [pii]

[19] Cespedes MV, Fernandez Y, Unzueta U, Mendoza R, Seras-Franzoso J, Sanchez-Chardi A et al. Bacterial mimetics of endocrine secretory granules as immobilized in vivo depots for functional protein drugs. Sci Rep 2016;6:35765. https://doi.org/10.1038/srep35765

[20] de Jong LA, Uges DR, Franke JP, Bischoff R. Receptor-ligand binding assays: technologies and applications. J Chromatogr B Analyt Technol Biomed Life Sci 2005;829:1-25. 10.1016/j.jchromb.2005.10.002 [doi]

[21] Peternel S, Komel R. Active protein aggregates produced in *Escherichia coli*. Int J Mol Sci 2011;12:8275-87. https://doi.org/10.3390/ijms12118275

[22] Krauss U, Jager VD, Diener M, Pohl M, Jaeger KE. Catalytically-active inclusion bodies-Carrier-free protein immobilizates for application in biotechnology and biomedicine. J Biotechnol 2017;258:136-47. https://doi.org/10.1016/j.jbiotec.2017.04.033

[23] Le PC, Genin P, Baines MG, Hiscott J. Interferon activation and innate immunity. Rev Immunogenet 2000;2:374-86. ncbi.nlm.nih.gov/pubmed/11256746

[24] Schoenborn JR, Wilson CB. Regulation of interferon-gamma during innate and adaptive immune responses. Adv Immunol 2007;96:41-101. https://doi.org/10.1016/S0065-2776(07)96002-2

[25] Khalilzadeh R, Shojaosadati SA, Maghsoudi N, Mohammadian-Mosaabadi J, Mohammadi MR, Bahrami A et al. Process development for production of recombinant human interferon-gamma expressed in *Escherichia coli*. J Ind Microbiol Biotechnol 2004;31:63-9. https://doi.org/10.1007/s10295-004-0117-x

[26] Perez L, Vega J, Chuay C, Menendez A, Ubieta R, Montero M et al. Production and characterization of human gamma interferon from *Escherichia coli*. Appl Microbiol Biotechnol 1990;33:429-34.

[27] Vaiphei ST, Pandey G, Mukherjee KJ. Kinetic studies of recombinant human interferon-gamma expression in continuous cultures of *E. coli*. J Ind Microbiol Biotechnol 2009;36:1453-8. https://doi.org/10.1007/s10295-009-0632-x

[28] Kumar M, Singh M, Singh SB. Optimization of conditions for expression of recombinant interferon-gamma in *E.coli*. Mol Biol Rep 2014;41:6537-43. https://doi.org/10.1007/s11033-014-3537-3

[29] Thomas JG, Baneyx F. Roles of the *Escherichia coli* small heat shock proteins IbpA and IbpB in thermal stress management: comparison with ClpA, ClpB, and HtpG In vivo. J Bacteriol 1998;180:5165-72.

[30] Zhou B, Xing L, Wu W, Zhang XE, Lin Z. Small surfactant-like peptides can drive soluble proteins into active aggregates. Microb Cell Fact 2012;11:10. https://doi.org/10.1186/1475-2859-11-10

[31] Conchillo-Sole O, de Groot NS, Aviles FX, Vendrell J, Daura X, Ventura S. AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. BMC Bioinformatics 2007;8:65. https://doi.org/10.1186/1471-2105-8-65

[32] Piovesan D, Tabaro F, Micetic I, Necci M, Quaglia F, Oldfield CJ et al. DisProt 7.0: a major update of the database of disordered proteins. Nucleic Acids Res 2017;45:D219-D227. https://doi.org/10.1093/nar/gkw1056

[33] Cano-Garrido O, Rueda FL, Sanchez-Garcia L, Ruiz-Avila L, Bosser R, Villaverde A et al. Expanding the recombinant protein quality in *Lactococcus lactis*. Microb Cell Fact 2014;13:167. https://doi.org/10.1186/s12934-014-0167-3

[34] Spekker K, Czesla M, Ince V, Heseler K, Schmidt SK, Schares G et al. Indoleamine 2,3-dioxygenase is involved in defense against *Neospora caninum* in human and bovine cells. Infect Immun 2009;77:4496-501. https://doi.org/10.1128/IAI.00310-09

[35] Boyanova M, Tsanev R, Ivanov I. A modified kynurenine bioassay for quantitative determination of human interferon-gamma. Anal Biochem 2002;308:178-81.

[36] Lakowicz JR, Kusba J, Wiczk W, Gryczynski I, Szmacinski H, Johnson ML. Resolution of the conformational distribution and dynamics of a flexible molecule using frequency-domain fluorometry. Biophys Chem 1991;39:79-84.

[37] Li GY, Xiao ZZ, Lu HP, Li YY, Zhou XH, Tan X et al. A simple method for recombinant protein purification using self-assembling peptide-tagged tobacco etch virus protease. Protein Expr Purif 2016;128:86-92. https://doi.org/10.1016/j.pep.2016.08.013

[38] Bhattacharjee S. DLS and zeta potential - What they are and what they are not? J Control Release 2016;235:337-51. https://doi.org/10.1016/j.jconrel.2016.06.017

[39] Meager A. Biological assays for interferons. J Immunol Methods 2002;261:21-36.

[40] Bomsztyk K, Stanton TH, Smith LL, Rachie NA, Dower SK. Properties of interleukin-1 and interferon-gamma receptors in B lymphoid cell line. J Biol Chem 1989;264:6052-7.

[41] Hess AK, Saffert P, Liebeton K, Ignatova Z. Optimization of translation profiles enhances protein expression and solubility. PLoS One 2015;10:e0127039. https://doi.org/10.1371/journal.pone.0127039

[42] Boumaiza M, Colarusso A, Parrilli E, Garcia-Fruitos E, Casillo A, Aris A et al. Getting value from the waste: recombinant production of a sweet protein by *Lactococcus lactis* grown on cheese whey. Microb Cell Fact 2018;17:126. https://doi.org/10.1186/s12934-018-0974-z

[43] Marini G, Luchese MD, Argondizzo AP, de Goes AC, Galler R, Alves TL et al. Experimental design approach in recombinant protein expression: determining medium composition and induction conditions for expression of pneumolysin from *Streptococcus pneumoniae* in *Escherichia coli* and preliminary purification process. BMC Biotechnol 2014;14:1. https://doi.org/10.1186/1472-6750-14-1

[44] Garcia-Fruitos E, Aris A, Villaverde A. Localization of functional polypeptides in bacterial inclusion bodies. Appl Environ Microbiol 2007;73:289-94. https://doi.org/10.1128/AEM.01952-06

[45] Upadhyay AK, Murmu A, Singh A, Panda AK. Kinetics of inclusion body formation and its correlation with the characteristics of protein aggregates in *Escherichia coli*. PLoS One 2012;7:e33951. 10.1371/journal.pone.0033951 [doi]

[46] Unzueta U, Cespedes MV, Sala R, Alamo P, Sanchez-Chardi A, Pesarrodona M et al. Release of targeted protein nanoparticles from functional bacterial amyloids: A death star-like approach. J Control Release 2018;279:29-39. https://doi.org/10.1016/j.jconrel.2018.04.004

[47] Hunig T. The storm has cleared: lessons from the CD28 superagonist TGN1412 trial. Nat Rev Immunol 2012;12:317-8. https://doi.org/10.1038/nri3192

[48] Suntharalingam G, Perry MR, Ward S, Brett SJ, Castello-Cortes A, Brunner MD et al. Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412. N Engl J Med 2006;355:1018-28. https://doi.org/10.1056/NEJMoa063842

[49] Yu Z, Reid JC, Yang YP. Utilizing dynamic light scattering as a process analytical technology for protein folding and aggregation monitoring in vaccine manufacturing. J Pharm Sci 2013;102:4284-90. 10.1002/jps.23746 [doi];S0022-3549(15)30802-9 [pii]

[50] Li TM, Hook JW, III, Drickamer HG, Weber G. Plurality of pressure-denatured forms in chymotrypsinogen and lysozyme. Biochemistry 1976;15:5571-80.

[51] Mohana-Borges R, Silva JL, Ruiz-Sanz J, de Prat-Gay G. Folding of a pressure-denatured model protein. Proc Natl Acad Sci U S A 1999;96:7888-93.

[52] Ruan K, Weber G. Hysteresis and conformational drift of pressure-dissociated glyceraldehydephosphate dehydrogenase. Biochemistry 1989;28:2144-53.

[53] Sanchez JM, Nolan V, Perillo MA. beta-galactosidase at the membrane-water interface: a case of an active enzyme with non-native conformation. Colloids Surf B Biointerfaces 2013;108:1-7. https://doi.org/10.1016/j.colsurfb.2013.02.019

[54] Randal M, Kossiakoff AA. The 2.0 A structure of bovine interferon-gamma; assessment of the structural differences between species. Acta Crystallogr D Biol Crystallogr 2000;56:14-24.

[55] Thiel DJ, le Du MH, Walter RL, D'Arcy A, Chene C, Fountoulakis M et al. Observation of an unexpected third receptor molecule in the crystal structure of human interferon-gamma receptor complex. Structure 2000;8:927-36.

[56] Flores SS, Nolan V, Perillo MA, Sanchez JM. Superactive beta-galactosidase inclusion bodies. Colloids Surf B Biointerfaces 2019;173:769-75. https://doi.org/10.1016/j.colsurfb.2018.10.049

[57] Martinez-Alonso M, Garcia-Fruitos E, Villaverde A. Yield, solubility and conformational quality of soluble proteins are not simultaneously favored in recombinant *Escherichia coli*. Biotechnol Bioeng 2008;101:1353-8. https://doi.org/10.1002/bit.21996

[58] Ghobadi S, Ashrafi-Kooshk MR, Mahdiuni H, Khodarahmi R. Enhancement of intrinsic fluorescence of human carbonic anhydrase II upon topiramate binding: Some evidence for drug-induced molecular contraction of the protein. Int J Biol Macromol 2018;108:240-9. https://doi.org/10.1016/j.ijbiomac.2017.12.011

[59] Wang L, Dong Q, Zhu Q, Tang N, Jia S, Xi C et al. Conformational Characteristics of Rice Hexokinase OsHXK7 as a Moonlighting Protein involved in Sugar Signalling and Metabolism. Protein J 2017;36:249-56. https://doi.org/10.1007/s10930-017-9718-x

FIGURE AND TABLE LEGENDS (DO NOT SUBMIT AS A SEPARATE FILE)

FIGURE 1 IFN-γ constructs produced in *E. coli* and *L. lactis* (a) The general organization of IFN-γ constructs is configured (IFN-γ)-Linker-APP-H6. The amino acid sequences of the linker

632 and the APPS are shown below the schematic representation of the protein designs. (b)
633 Quantification of the production of IFN-γ in IB-like nanoparticles in *L. lactis* (top) and
634 solubility (bottom) of IFN-γ in *L. lactis*. Significant results are shown as * $p \leq 0.05$ and **
635 $p \leq 0.005$.

636

637 **FIG 2** (a) Ultrastructural characterization by FESEM of protein aggregates and purified
638 protein nanoparticles of rBoIFN_γ, rBoIFN- γ_CYOB, rBoIFN- γ_HARLU and rBoIFN- γ_L6K2.
639 Scale bars correspond to 200 nm. (b) ZP of purified protein nanoparticles.

640

641 **FIG 3** Kynurenine levels measured by absorbance at 490 nm after treatment of EBTr cells
642 for 96 h with increasing amounts of rBoIFN-γ from different origins. (a) Soluble rBoIFN-γ
643 produced in the indicated expression system. (b) Protein nanoparticles of rBoIFN-γ
644 produced in *L. lactis*. The K_D values are indicated in the plot. (c) Schematic representation
645 of the protein samples used in the activity assays: soluble protein obtained from the soluble
646 cell fraction, protein NPs purified from the insoluble cell fraction, soluble protein obtained
647 from the protein NPs, and the NP core after a resolubilization procedure. (d) Comparison of
648 the activity between rBoIFN-γ protein obtained from solubilization of protein NPs and
649 purified rBoIFN-γ from the soluble cell fraction as indicated at 0.72 nmol/L. Different letters
650 depict differences between proteins ($p < 0.001$) except rBoIFN-γ from protein NPs and
651 rBoIFN-γ_E ($p = 0.024$).

652

653 **FIG 4** Recombinant IFN-γ sizes in different supramolecular arrangements (purified soluble
654 IFN-γ and INF-γ IBs). (a) Soluble rBoIFN-γ from different origins: commercial rBoIFN-γ_Std,
655 in-house rBoIFN-γ from *E. coli* and *L. lactis*. (b) rBoIFN-γ IBs produced in *L. lactis*. (c) PBS
656 solubilized rBoIFN-γ from IBs after interferon release. (d) Scaffold of rBoIFN-γ IBs incubated
657 for 96 hours at 37 °C. The mean size and polydispersity index are indicated in brackets. The
658 average size data of the soluble proteins were analyzed by one-way ANOVA (^{t} corresponds
659 to $P < 0.07$).

660 **Table 1** Selection of APPs from predictions of "hot spots (HS)" of aggregation in
661 polypeptides by AGGRESCAN [31]. CYOB: Cytochrome bo_3 ubiquinol oxidase subunit 1
662 from *E. coli*, HALRU: Aragonite protein AP7. NA: Not applicable. HS: hot spot. HSA: hot
663 spot area. NHSA: normalized HSA. a_4vAHS: average aggregation-propensity in each HS.

664 **Table 2** Center of spectral mass (CSM) of IFN-γ protein preparations in soluble formats
665 or in protein NPs analyzed before and after the resolubilization protocol.

666

667