Perspectives on protein biopolymers: miniaturized flow field-flow fractionation-assisted characterization of a single-cysteine mutated phaseolin expressed in transplastomic tobacco plants

Marassi V.; De Marchis F.; Roda B.; Bellucci M.; Capecchi A.; Reschiglian P.; Pompa A.; Zattoni A.

Journal of Chromatography A

Volume 1637, 25 January 2021, 461806

Perspectives on protein biopolymers: miniaturized flow field-flow fractionation-assisted characterization of a single-cysteine mutated phaseolin expressed in transplastomic tobacco plants

Valentina Marassi[1,4], Francesca De Marchis[2], Barbara Roda[1,4], Michele Bellucci[2], Alice Capecchi[2], Pierluigi Reschiglian[1,4], Andrea Pompa[3], Andrea Zattoni[1,4]*

1. Department of Chemistry G. Ciamician, University of Bologna, Bologna, Italy
2. Institute of Biosciences and Bioresources—Research Division of Perugia, National Research Council of Italy, via della Madonna Alta 130, 06128, Perugia (PG), Italy
3. Department of Biomolecular Sciences, University of Urbino "Carlo Bo", via Donato Bramante 28, 61029 Urbino (PU), Italy
4. byFlow srl, via dell'Arcoveggio 74, 40128 Bologna (BO), Italy
** corresponding author, andrea.zattoni@unibo.it*

HIGHLIGHTS

- Transplastomic phaseolin (P*) expressed in tobacco thylakoids could be exploited in the production of biopolymers
- Plant extracts were submitted to hollow fiber flow field flow fractionation coupled to multi-angle light scattering (HF5-MALS) with two separation methods
- Characterization of native phaseolin and P* showed differences in aggregation state and molar mass
- Conformation studies confirm the concatenation of P* into elongated forms compatible with successful polymerization

KEYWORDS

hollow fiber flow field flow fractionation
multi-angle light scattering
protein biolpolymers
transplastomic phaseolin

ABSTRACT

The development of plant-based protein polymers to employ in biofilm production represents the promising intersection between material science and sustainability, and allows to obtain biodegradable materials that also possess excellent physicochemical properties. A possible candidate for protein biopolymer production is phaseolin, a storage protein highly abundant in P Vulgaris beans. We previously showed that transformed tobacco chloroplasts could be employed to express a mutated phaseolin carrying a signal peptide (directing it into the thylakoids) also enriched of a cysteine residue added to its C-terminal region. This modification allows for the formation of inter-chain disulfide bonds, as we previously demonstrated, and should promote polymerization.

To verify the effect of the peptide modification and to quantify polymer formation, we employed hollow-fiber flow field-flow fractionation coupled to UV and multi-angle laser scattering detection (HF5-UV-MALS): HF5 allows for the selective size-based separation of phaseolin species, whereas MALS calculates molar mass and conformation state of each population. With the use of two different HF5 separation methods we first observed the native state of P.Vulgaris phaseolin, mainly assembled into trimers, and compared it to mutated phaseolin (P*) which instead resulted highly aggregated. Then we further characterized P* using a second separation method, discriminating between two and distinct high-molecular weight (HMW) species, one averaging 0.8 x 10^6 Da and the second reaching the tens of million Da. Insight on the conformation of these HMW species was offered from their conformation plots, which confirmed the positive impact of the Cys modification on polymerization.

1. INTRODUCTION

Proteins are organic heteropolymers serving as nutrients and structural compounds in living organisms with also important roles in cell function and regulation. They generally exist in nature either in the form of insoluble fibrous proteins or water-soluble globular proteins. Their broad spectrum of functional and structural properties due to polar and nonpolar amino acids make them promising raw materials for the production of bioplastics used for packaging materials. Biodegradable polymers (biopolymers) derived from renewable natural resources to produce bioplastics is thus a research topic of growing interest [1]. At present, a number of protein-based films are produced for food packaging [2]. Indeed, they show excellent optical properties (gloss and transparency), are good fat barriers, at low and intermediate humidity feature an excellent oxygen and organic vapor barrier and have fair mechanical properties [3].

The plant polypeptide phaseolin accounts for 50% of the total seed protein of the leguminous species common bean, *Phaseolus vulgaris* L.[4], which is one of the most commonly produced and consumed food legume worldwide, especially in Asia, South America and Africa [5]. This globulin protein belongs to the vicilin (or 7S) family of seed storage proteins and is a homotrimeric glycoprotein. Phaseolin monomer has a weak internal duplication constituted by two regions with a similar secondary structure, each region made up of a β-barrel domain followed by a α-helical cluster comprising of three helices. The trimer is composed of three similar polypeptides linked mainly by hydrophobic interactions between two α-helical clusters [6, 7]. Phaseolin is encoded by a small gene family and consists of two highly homologous classes of α- and β-polypeptides of molecular weight around 45 – 46 kDa, after that their N-terminal, 24-aminoacid signal peptide is cleaved inside the endoplasmic reticulum [8]. In view of its excellent functional properties [9], phaseolin can be applied in food formulations [10], or fused to proteins of biotechnological interest [11], but its potential can also be explored for the production of protein biopolymers.

One way to obtain biopolymers from plant proteins is to exploit the protein synthesis mechanisms of the plant itself. Transformation of the chloroplast genome (plastome) has been largely used for the production of heterologous proteins [12] [13], including protein biopolymers [14] [15]. Therefore, in this study, we transformed tobacco plastome with a β-phaseolin gene in which a cysteine residue has been added to the C-terminal region. This single insertion is able to allow the formation of phaseolin inter-chain disulfide bonds because the cysteine residue is not hidden in the phaseolin trimers but remains exposed, thus enabling the formation of disulfide bonds among trimers [16]. The β-phaseolin gene has been chosen because the crystal forms used to investigate the phaseolin three-dimensional structure showed only β-polypeptides [6]. Moreover, we previously demonstrated that β-phaseolin signal peptide targets a recombinant phaseolin protein to the thylakoid membrane when it is expressed

in tobacco chloroplasts. Indeed, thylakoid localization increases the possibility to produce phaseolin polymers because, in the oxidative environment of this chloroplast sub-compartment, recombinant phaseolin trimers can form higher order systems through disulfide bonds [17]. Moreover, thylakoid-targeted phaseolin can avoid an autoregolatory mechanism which inhibits its mRNA translation [18].
To verify the potential of the thylakoid compartment for phaseolin polymer production it is necessary to assess whether the extracted protein matter consists of phaseolin polymers, oligomers, or both: this implies that in the case of a heterogeneous extract, MW/size-based separation of protein components is necessary. Separation should also be coupled to characterization techniques able to weigh the species found, to provide information on the aggregation/polymerization state.
Flow field-flow fractionation (F4) is one of the most performing size-separation techniques: together with size-exclusion chromatography (SEC), it is the benchmark technique for protein and antibody size characterization and is suggested by FDA as a technique that complements SEC in the validation of protein products [19]. F4 is a flow-assisted technique ideally suited to size-separate dispersed analytes over a broad size range [20]: in F4 retention time is inversely proportional to the hydrodynamic diffusion coefficient of the analyte and, consequently, it is directly proportional to its hydrodynamic size. The analyses can be performed in physiological conditions, allowing for the separation and detection of delicate species in native conditions [21, 22], [23] [24]. The most used F4 technique is Asymmetrical F4 (AF4), where the sample is separated in a flat capillary channel with a porous wall generating the hydrodynamic field.
Flow-based separation science applied to food and agricultural chemistry proved its efficacy on a variety of analytes such as polysaccharides from grains [25] and fungi [26], beer proteins [27], wine [28] [29] and even intact plant ribosomes [30].
In the first stages of biomaterial development it is desirable to be able to work with a low amount of starting material, to facilitate production optimization prior to larger-scale development. This is the case with phaseolin polymerization into chloroplast compartments where separation and characterization efficacy should meet with a reduced-volume working environment
The microvolume variant of F4, HF5, merges the advantages of such a technique with a lower sample volume and demonstrated to be a successful alternative to SEC [31]: down-scaling of the separation channel is advantageous for application in the bio-pharmaceutical field and to give insight during protocol development, since the lower injection volume makes it possible to work samples such as small amounts of plant extracts. Lastly, disposable usage of the separating channel eliminates sample carry-over or sample contamination issues [32] [33] [34] [35]. HF5 has proved its potential in the characterization of complex protein samples, monitoring of protein drug aggregation, and even monitoring of time-dependent formation of amyloid fibrils [24] [36] [37, 38], but had not breached

into the field of plant/food chemistry up to now. When online coupled to uncorrelated detection methods such as MALS and spectrophotometric detectors, HF5 provides high-resolution size distribution analysis. This platform is effective in the investigation of High MW protein species, since it reduces the occurrence of artifacts, and can highlight the stability of oligomers and higher-order aggregates. In the present work, to investigate the effective formation of phaseolin polymeric species, phaseolin standard obtained from beans and transplastomic phaseolin extracted from tobacco leaves were characterized and compared. To quantify and characterize the phaseolin polymers, we employed hollow-fiber flow field-flow fractionation (HF5) coupled to UV absorption and multi-angle light scattering (MALS) detection.

2. MATERIALS AND METHODS

2.1 Gene constructs and plant transformation.

Beta phaseolin cDNA gene was amplified from plasmid pDHA.T343F [39] with NdeI-P (5'-tccatcggacatatgatgagagca-3') and NotI-P*(5'-ccccctccggatcgcggccgctagtacacaaat**gca**ccctttcttccct-3') oligonucleotides to introduce a *Nde*I and a *Not*I restriction site (underlined) at the 5' and 3' ends of the gene, respectively. Moreover, primer NotI-P*is designed to insert a cysteine residue (in bold) at the C-terminal of the protein to allow inter-chain disulphide bridge formation. The PCR product was digested with *Nde*I/*Not*I and cloned into pCR2.1-5'UTR[40], to obtain the pCR2.1- 5'UTR-P* intermediate plasmid, in which phaseolin is under the control of the plastidial *psbA* promoter/5'UTR. The psbA/5'UTR-P* cassette was excised from pCR2.1-5'UTR-P* by *EcoR*V/*Not*I digestion and subcloned into pLD-CTV tobacco plastid vector [41], generating pLD-CTV-P*. Homoplasmic transplastomic plants were obtained by particle bombardment into tobacco leaves of gold microprojectiles coated with pLD-CtV-P* as described [42]. Transplastomic plants were grown at 25°C in 16 h of light in axenic conditions before being transferred to the greenhouse for seed production.

2.2 Phaseolin analysis.

Phaseolin purification: Phaseolin (P) purification was performed from *P. vulgaris* L. seeds (variety small red) according to Suzuki and colleagues [43] with minor modifications. All reagents used were of analytical grade and were obtained from Merck KGaA, Darmstadt, Germany. All solutions were prepared with deionised water having a conductivity of <1 µS/cm. After acidic extraction, the protein was further purified by velocity centrifugation on Sucrose gradient. One ml of phaseolin extract was loaded on a 13 mL linear 5% to 25% (w/v) Sucrose gradient made in 150 mM NaCl, 1 mM EDTA, 0.1% Triton X-100, 50 mM Tris-Cl, pH 7.5. After centrifugation at 110,000 *g* overnight at 4° C, the fractions containing phaseolin, corresponding to those where proteins of about 150 kDa are localized, were collected and merged. To verify purity of the phaseolin extract, an aliquot of this pool was analyzed by SDS-PAGE followed by Western blotting with anti-phaseolin antibody (1:10,000) as described [44]. The merged fractions containing phaseolin were dialyzed for 3 days in distilled H_2O, lyophilized and finally resuspended at a concentration of 3 mg/ml in PBS.

Single cystein-mutated phaseolin (P) analysis and exctraction from transplastomic plants:* Total leaf proteins were extracted from 0.3 g of tissues from transplastomic plants, grounded in liquid nitrogen, homogenized in 0.8 mL of extraction buffer (200 mM NaCl, 1 mM EDTA, 0.2% Triton X-100, 100

mM Tris–Cl, pH 7.8) supplemented with 'Complete' protease inhibitor cocktail (Merck) and analysed by Western blotting as described above.

For P* purification, one hundred and fifty grams of transformed tobacco young leaves were collected after being let in the dark overnight to prevent starch accumulation. They were grinded in a blender with 1.5 L of grinding buffer (GB: 0.33 M sorbitol, 0.2 M Hepes, 0.1 M EDTA, 0.2 M $MgGl_2$, 1 mM 2-Mercaptoethanol, 4.4 mM isoascorbic acid, pH 6.8) four times at low speed for 10 sec each time. The homogenate was filtered through a layer of Miracloth paper and four layers of gauze before centrifugation at 1,700 *g* for 2' at 4°C. The supernatant was discarded and the pellet, containing both broken and intact chloroplasts, was solubilized in the extraction buffer (EB: 0.2 M NaCl, 2% Triton X-100, 1 mM EDTA, 0.1 M Tris-Cl, pH 7.8, supplemented with Complete protease inhibitor cocktail) to dissolve chloroplasts membranes. To remove residual starch an additional centrifugation at 14,000 *g* for 5' at 4°C was performed. Three hundred µL of the resulting supernatant were loaded on a 13 mL velocity Suc gradient 5% - 25% (w/w) made in 150 mM NaCl, 1 mM EDTA, 0.1% Triton X-100, 50 mM Tris-Cl, pH 7.5. After centrifugation at 110,000 *g*, for 24 h at 4°C, 18 fractions of about 0.8 mL and the sample pelleted at the bottom of the tube were collected. An equal aliquot of each fraction and the pellet was analyzed by SDS-PAGE and Western blot with anti-phaseolin antiserum as described above. According to the Western blot results, the bottom of the tube was recovered with 1 mL of distilled H_2O, dialyzed, lyophilized, resuspended in isotonic, pH 7.4 PBS at 0.3 mg/mL and analysed by HF5-MALS. Considering that the initial concentration of P* in the leaves is 0.5% of the total soluble protein, we calculated the P* extraction efficiency which is around 53%.

2.3 HF5-UV-MALS.

HF5 analyses were performed using an Agilent 1200 HPLC system (Agilent Technologies, Santa Clara, CA, USA) consisting in a degasser, an isocratic pump, with an Agilent 1100 DAD UV/Vis spectrophotometer combined with an Eclipse® DUALTEC separation system (Wyatt Technology Europe, Dernbach, Germany). The hollow fiber was a polyethersulfone (PES) fiber, type FUS 0181 available from Microdyn-Nadir (Wiesbaden, Germany) with the following characteristics: 0.8 mm ID, 1.3 mm OD, and 10 kDa Mw cut-off, corresponding to an average pore diameter of 5 nm. The HF5 channels used for the experimental section were standard cartridges containing a 17 cm long fiber, commercially available. The scheme of the HF5 cartridge, its assembly and the modes of operation of the Eclipse® DUALTEC system have already been described elsewhere[33]. The ChemStation version B.04.02 (Agilent Technologies) data system for Agilent instrumentation was used to set and control the instrumentation and method parameters. The software package Wyatt Eclipse @ ChemStation version 3.5.02 (Wyatt Technology Europe) was used to set and control the

flow rate values. A 3-angle multi-angle light scattering detector model miniDAWN TREOS (Wyatt Technology Corporation, Santa Barbara, CA, USA) operating at a wavelength of 658 nm, was used to measure the radius and molar mass of particles in suspension, with the software ASTRA® version 6.1.7 (Wyatt Technology Corporation). Analyses were carried out in isotonic, pH 7.4 PBS as mobile phase. An HF5 conventional method is composed of four steps: focus, focus–injection, elution and elution–injection. During the focus, the mobile phase is split into two different streams entering from inlet and outlet; during focus–injection, the flow settings remain unvaried while the sample is introduced into the channel through the inlet and focused in a narrow band. Then, in the elution step, the flow of mobile phase enters the channel inlet and it splits in a radial component which exits the fiber's pores (cross-flow), and a longitudinal component (channel flow, Vc) that reaches the detectors. Lastly, during elution–injection, no cross-flow is applied and any remaining sample inside the channel is released; also, the flow is redirected in the injection line as well to clean it before the next injection. Longitudinal flow is indicated as Vc, while cross/focus flow as Vx. The two different separation methods developed in this work are summarized in Table 1.

Due to the parabolic flow profile of the carrier flow, smaller particles experience higher flow rates (on the average) than larger ones. In this normal fractionation mode, the particle retention is a function of its apparent diffusion coefficient. Hence, the particle retention volume can be related to its diffusion coefficient, and consequently, to its hydrodynamic diameter (Dh) or radius (Rh) using the Stoke's equation. The conversion of the retention volume to hydrodynamic radius/molar mass can be accomplished by calculating channel parameters and applying F4 theory: in our work, the software ISIS (Wyatt Technology Europe), allowed us to calculate the expected retention times of proteins eluting under an input method, accounting for membrane swelling, ionic strength and instrumental parameters. FlFFF theory rigors are described elsewhere[45]. Multi-angle light scattering (MALS) was used to calculate the molar mass of eluted proteins after normalization with standard bovine serum albumin monomer. MALS allows for the absolute determination of particle root mean square radius of gyration (Rg) by measuring the net intensity of light scattered by such particles at a range of fixed angles, and by consequence, knowing the dn/dc and absorptivity values of the analysed species, the molar mass value of the eluting species[46]. The correlation between Radius of gyration and molar mass distributions provides information on the particle conformation in suspension through the calculation of a scaling exponent ν, or ν value. The ν value is the slope in a double logarithmic logMW – logRg plot, and is theoretically defined for spheres as $\nu = 0.33$, random-coil $\nu = 0.5–0.6$, and rod-like structures $\nu \sim 1$ [47] [48].

3. RESULTS AND DISCUSSION

3.1 Preparation of *P. Vulgaris* phaseolin and recombinant tobacco phaseolin

Lyophilized phaseolin (P) obtained from *P. vulgaris* was resuspended in PBS at a final concentration of 3 mg/mL, whereas the purified mutant phaseolin (P*) had a concentration of 0.3 mg/mL in PBS. P localization in the bean seed cells depends on the presence of two sorting signals: a N-terminal 24 amino acid signal peptide (SP), and a C-terminal four-amino acid propetide (AFVY). The first allows the post-translational insertion of P polypeptides from the cytoplasm into the endoplasmic reticulum (ER), while the second directs P polypeptides from the ER to vacuoles. Since genetically modified tobacco plants undergo insertion of the P* recombinant gene in the plastome, P* is synthesized inside the chloroplast stroma and the SP directs this mutated phaseolin, with a cysteine residue added to the C-terminal, to the thylakoid compartment (Figure 1) [16] [17].

3.2 Gel electrophoresis

Purified bean P, as expected [8] [49], is detected by SDS-PAGE followed by Coomassie staining as several differentially glycosylated monomers of α- and β-polypeptides of molecular weight around 45 – 50 kDa (Figure 2A). Western blot analysis with the anti-phaseolin antiserum confirms the specificity of these polypeptides (Figure 2A). Conversely, the modified P* sample purified from chloroplasts of transplastomic tobacco leaves, when analyzed by Western blot, shows a pattern formed by different phaseolin polypeptides. Apart from the expected unglycosylated 45-kDa polypeptide (the enzymes for glycosylation are not present inside the chloroplast), additional P* forms, not detected in the P sample, are a dimer, with an apparent molecular mass of around 83 kDa, and higher molecular mass forms (Figure 2B). The 83-kDa dimers derive from disulfide bonds between P* monomers (Pompa et al. 2010), while the P* aggregates around 100 – 180 kDa are unexpected. At physiological pH, the most stable form of bean P is thought to be trimeric, with 46 kDa polypeptide being the protomer [50], thus it is likely that these superstructures are denatured in SDS-PAGE followed by Western blotting (Figure 2B). This is only in part true for P*, whose polypeptide species around 100 – 180 kDa are prevalent, suggesting that P* mainly consists in stable, heavier species (Figure 2B).

SDS denaturation does not provide a truthful characterization of the samples content. However, the presence of P* monomers and oligomers confirms the presence of P* polypeptides localized in the chloroplast thylakoids in different forms [17]. To carry out a more in-depth evaluation of the P * high molecular mass species, chloroplasts from leaves of transplastomic tobacco plants have been isolated and then solubilized with a buffer containing 2% Triton X-100. The resulting homogenate was loaded on a velocity Sucrose gradient and the various fractions obtained were subjected to Western blot

analysis (Fig 3a). Almost all the P* protein was recovered in the bottom of the tube, confirming that P* is localized in chloroplasts as polypeptides of molecular mass higher than 700 kDa. On the contrary, the amount of trimeric P* is extremely low, as indicated by the very weak signals corresponding to the components of the fractions 7 – 9 that migrated as a trimer, around 150 kDa (Fig 3b).

3.3 HF5-UV-MALS Analysis

To effectively estimate the size of phaseolin oligo/polymeric systems, we employed hollow fiber flow field-flow fractionation coupled to UV and MALS detection.

To characterize phaseolin extracts, we developed a first separation method (Method 1) which was able to resolve phaseolin aggregated states present in *P. vulgaris*. Then, the tobacco-derived protein extract was submitted to the same method, but the protein species detected had a high-enough size to be fully retained. Therefore, we switched to a second, softer HF5 separation method (Method 2) to improve characterization, and compare these high-MW species to "standard" phaseolin in the same conditions.

3.3.1 Characterization of standard phaseolin extracted from *P. vulgaris*

Liophylized phaseolin was resuspended in PBS at 3mg/mL and 10 µL were analysed at t0 with method 1. In Figure 4 the fractogram obtained is overlaid with the mass values calculated at peak maxima.

Phaseolin was mainly present as its stable trimeric form (143 kDa), eluted at 16.5 minutes. The mass calculated at 19 min (275 kDa) corresponds instead to phaseolin 6-mer. The third shoulder is phaseolin 9-mer. By calculating the retention time expected for protein species in the developed method (see Materials and Methods), we could also attribute the peak at 14 min to the dimer. The relative abundance (expressed as % of the integration of the absorption signal at 280 nm) of each species and their aggregation state attribution are shown in Table 2.

Even though at physiological pH phaseolin is mostly present at trimer, other aggregation states are visible and separated, showing that phaseolin favours the formation of stable species with various grades of assembly.

3.3.2 Characterisation of phaseolin putative polymers extracted from transplastomic tobacco plants

The pellet of a velocity sucrose gradient (Fig 3), utilized to purify the modified phaseolin (P*) extracted from transplastomic tobacco plants, was recovered, dialyzed and lyophilized. P * was

resuspended in PBS and subjected to Method 1. Given the lower concentration obtained (0.3 mg/mL from Nanodrop measurements), the injection volume was increased to 50 µL.

In the fractogram obtained (Figure 5), there is no evidence of phaseolin trimer and oligomers, apart from a small band between 10 and 14 min. Instead, one single peak corresponding to high molecular weight species is visible starting from 24 min and eluted in proximity and after the field release (at 25.5 min). Its molar mass distribution ranged from 10^6 to 10^7 Da showing that P* is differently arranged than P: the presence of disulfide bond-prone residues seem in fact not only to promote higher order accretion, but limits the ability to form lower order oligomers such as the expected trimer.

3.3.3 Characterisation of the different polymer populations

To better elucidate the actual size of P* species a Method 2 was used to analyze P and P*. In this case, the cross flow gradient is lowered from 0.8 ml/min to 0.3 ml/min (starting rate) to ease the separation and elution of the hydrodynamically bigger species uncovered with Method 1.

With this second method, the oligomeric species of standard phaseolin cannot be resolved anymore, and P is eluted as a single peak at 10 min (Figure 6, panel a). The corresponding calculated molar mass (263 kDa) confirms that the peak is indeed a mixture of trimer and heavier species.

On the other hand, P* shows the presence of trace of oligomers (Figure 6b) with a peak at 8 minutes (in common with P), but is mainly composed of two heavier species which resulted now distinct.

The first was eluted with a maximum at 14 minutes, and a molar mass averaging at 0.8 x 10^6 Da, while the second is even heavier and reaches 10^7 Da. However, it is not clear how P* monomers are organized in these two distinct high-molecular weight (HMW) polymeric species.

One indication on the conformation state of mutated phaseolin HMW-forms can be offered with the calculation of their shape factor. When considering an eluting band, the corresponding molar mass and RMS radius (from light scattering characterization) can be plotted in the logarithmic form to obtain a slope (v value) related to conformation. Generally, a value of 0.33 corresponds to a solid sphere, while higher values indicate a less compact or elongated structure like a random coil (0.5) or a rod (1) [51], [52], [53].

In the case of standard phaseolin, the oligomeric forms are compact, as the 0.26 v value suggests (Table 3).

This is in line with the aggregation model of phaseolin, which envisions a triangular assembly [54]. Instead, different values emerge for P* higher order species, which rather denote the formation of random coils (Band 2) and more extended shapes (Band 3). This supports the idea of the successful polymerization of P* caused by its mutation, which prevents the conventional 3-unit aggregation and favours the concatenation of more units by means of disulfide bonds.

4. Conclusions

The use of protein biopolymers for film production is very appealing, given their excellent properties and their potential in replacing fossil-fuel plastics. Moreover, plant-based proteins can represent a sustainable source and are of great interest. [55]

In this work, we employed a combination of hollow-fiber flow field flow fractionation methods combined to UV and MALS detection, to identify and characterize polymeric forms of P*, a mutated protein obtained from the thylakoid compartment of transplastomic tobacco plants.

Protein extracts derived from *P. vulgaris* showed the presence of native phaseolin, prevalently as trimer, while there was no evidence of polymeric forms, as expected. When extracts from tobacco were instead analyzed, we found evidence of higher molar mass species, and traces of oligomeric phaseolin. Moreover, the high molecular weight (HMW) species, characterized using a second separation method, were two and distinct, one averaging 0.8 x 10^6 Da and the second reaching the tens of million Da. Insight of the conformation of these HMW species was offered from their v values, which corresponded to those of coils and longer chain structures, confirming the positive impact of the Cys modification on polymerization.

The HF5-UV-MALS system was successful in evaluating the different aggregation status of native phaseolin and monitoring the formation of two distinct polymeric species in modified P*. The approach will be useful in assisting the further optimization steps of this process. The formation of phaseolin polymers in chloroplasts, not present in *P. vulgaris*, could be very interesting for industrial purposes. Chloroplasts from cultivated transplastomic tobacco plants could be employed as a reactor to produce a biopolymer derived from an edible protein, with a possible application in the production of biodegradable films.

References

1. Pathak, S., C. Sneha, and B.B. Mathew, *Bioplastics: Its Timeline Based Scenario & Challenges.* Journal of Polymer and Biopolymer Physics Chemistry, 2014. **2**(4): p. 84-90.
2. Calva-Estrada, S.J., M. Jiménez-Fernández, and E. Lugo-Cervantes, *Protein-Based Films: Advances in the Development of Biomaterials Applicable to Food Packaging.* Food Engineering Reviews, 2019. **11**(2): p. 78-92.
3. Bhawani, S.A., et al., *Proteins as Agricultural Polymers for Packaging Production*, in *Bionanocomposites for Packaging Applications*, M. Jawaid and S.K. Swain, Editors. 2018, Springer International Publishing: Cham. p. 243-267.
4. Sun, S.M., et al., *Protein Synthesis and Accumulation in Bean Cotyledons during Growth.* Plant Physiology, 1978. **61**(6): p. 918-923.
5. Nedumaran, S., et al. *Grain Legumes Production, Consumption and Trade Trends in Developing Countries-An Assessment & Synthesis*. 2019.
6. Lawrence, M.C., et al., *Structure of Phaseolin at 2·2 Å Resolution: Implications for a Common Vicilin/Legumin Structure and the Genetic Engineering of Seed Storage Proteins.* Journal of Molecular Biology, 1994. **238**(5): p. 748-776.
7. Lawrence, M.C., et al., *The three-dimensional structure of the seed storage protein phaseolin at 3 A resolution.* The EMBO journal, 1990. **9**(1): p. 9-15.
8. Slightom, J.L., et al., *Nucleotide sequences from phaseolin cDNA clones: the major storage proteins from Phaseolus vulgaris are encoded by two unique gene families.* Nucleic acids research, 1985. **13**(18): p. 6483-6498.
9. Yin, S.-W., et al., *Functional and conformational properties of phaseolin (Phaseolus vulgris L.) and kidney bean protein isolate: A comparative study.* Journal of the Science of Food and Agriculture, 2010. **90**(4): p. 599-607.
10. Yin, S.-W., et al., *Surface charge and conformational properties of phaseolin, the major globulin in red kidney bean (Phaseolus vulgaris L): effect of pH.* International Journal of Food Science & Technology, 2011. **46**(8): p. 1628-1635.
11. de Virgilio, M., et al., *The human immunodeficiency virus antigen Nef forms protein bodies in leaves of transgenic tobacco when fused to zeolin.* Journal of experimental botany, 2008. **59**(10): p. 2815-2829.
12. De Marchis, F., A. Pompa, and M. Bellucci, *Plastid Proteostasis and Heterologous Protein Accumulation in Transplastomic Plants.* Plant Physiology, 2012. **160**(2): p. 571-581.
13. Fuentes, P., T. Armarego-Marriott, and R. Bock, *Plastid transformation and its application in metabolic engineering.* Current Opinion in Biotechnology, 2018. **49**: p. 10-15.
14. Guda, C., S.B. Lee, and H. Daniell, *Stable expression of a biodegradable protein-based polymer in tobacco chloroplasts.* Plant Cell Reports, 2000. **19**(3): p. 257-262.
15. Chebolu, S. and H. Daniell, *Chloroplast-derived vaccine antigens and biopharmaceuticals: expression, folding, assembly and functionality.* Current topics in microbiology and immunology, 2009. **332**: p. 33-54.
16. Pompa, A., et al., *An engineered C-terminal disulfide bond can partially replace the phaseolin vacuolar sorting signal.* The Plant Journal, 2010. **61**(5): p. 782-791.
17. De Marchis, F., et al., *A plant secretory signal peptide targets plastome-encoded recombinant proteins to the thylakoid membrane.* Plant Molecular Biology, 2011. **76**(3): p. 427-441.
18. De Marchis, F., M. Bellucci, and A. Pompa, *Phaseolin expression in tobacco chloroplast reveals an autoregulatory mechanism in heterologous protein translation.* Plant Biotechnology Journal, 2016. **14**(2): p. 603-614.
19. *FDA Guidance for Industry: Immunogenicity-Related Considerations for Low Molecular Weight Heparin*
20. Giddings, J., *Field-flow fractionation: analysis of macromolecular, colloidal, and particulate materials.* Science, 1993. **260**(5113): p. 1456-1465.

21. Zattoni, A., et al., *Asymmetrical flow field-flow fractionation with multi-angle light scattering detection for the analysis of structured nanoparticles.* Journal of Chromatography A, 2009. **1216**(52): p. 9106-9112.
22. Tanase, M., et al., *Hydrodynamic size-based separation and characterization of protein aggregates from total cell lysates.* Nature protocols, 2015. **10**(1): p. 134-148.
23. Reschiglian, P., et al., *Hollow-fiber flow field-flow fractionation with multi-angle laser scattering detection for aggregation studies of therapeutic proteins.* Analytical and Bioanalytical Chemistry, 2014. **406**(6): p. 1619-1627.
24. Marassi, V., et al., *A new approach for the separation, characterization and testing of potential prionoid protein aggregates through hollow-fiber flow field-flow fractionation and multi-angle light scattering.* Analytica Chimica Acta, 2019. **1087**: p. 121-130.
25. Fuentes, C., et al., *Characterization of molecular properties of wheat starch from three different types of breads using asymmetric flow field-flow fractionation (AF4).* Food Chemistry, 2019. **298**: p. 125090.
26. Wu, D.-T., et al., *Molecular characterization of branched polysaccharides from Tremella fuciformis by asymmetrical flow field-flow fractionation and size exclusion chromatography.* Journal of Separation Science, 2017. **40**(21): p. 4272-4280.
27. Lu, Y., B. Bergenståhl, and L. Nilsson, *Interfacial properties and interaction between beer wort protein fractions and iso-humulone.* Food Hydrocolloids, 2020. **103**: p. 105648.
28. Marassi, V., et al., *Characterization of red wine native colloids by asymmetrical flow field-flow fractionation with online multidetection.* Food Hydrocolloids, 2021. **110**: p. 106204.
29. Coelho, C., et al., *Asymmetrical flow field-flow fractionation of white wine chromophoric colloidal matter.* Analytical and Bioanalytical Chemistry, 2017. **409**(10): p. 2757-2766.
30. Pitkänen, L., P. Tuomainen, and K. Eskelin, *Analysis of plant ribosomes with asymmetric flow field-flow fractionation.* Anal Bioanal Chem, 2014. **406**(6): p. 1629-37.
31. Marassi, V., et al., *Hollow fiber flow field-flow fractionation and size-exclusion chromatography with multi-angle light scattering detection: A complementary approach in biopharmaceutical industry.* Journal of Chromatography A, 2014. **1372**: p. 196-203.
32. Andrea, Z., et al., *Hollow-Fiber Flow Field-Flow Fractionation.* Current Analytical Chemistry, 2007. **3**(4): p. 310-323.
33. Reschiglian, P., et al., *On-Line Hollow-Fiber Flow Field-Flow Fractionation-Electrospray Ionization/Time-of-Flight Mass Spectrometry of Intact Proteins.* Analytical Chemistry, 2005. **77**(1): p. 47-56.
34. Zattoni, A., et al., *Hollow-fiber flow field-flow fractionation of whole blood serum.* Journal of Chromatography A, 2008. **1183**(1): p. 135-142.
35. Johann, C., et al., *A novel approach to improve operation and performance in flow field-flow fractionation.* Journal of Chromatography A, 2011. **1218**(27): p. 4126-4131.
36. Rambaldi, D.C., et al., *In vitro amyloid Aβ1-42 peptide aggregation monitoring by asymmetrical flow field-flow fractionation with multi-angle light scattering detection.* Analytical and Bioanalytical Chemistry, 2009. **394**(8): p. 2145-2149.
37. Reschiglian, P., et al., *Hollow-Fiber Flow Field-Flow Fractionation: A Pipeline to Scale Down Separation and Enhance Detection of Proteins and Cells*, in *Field-Flow Fractionation in Biopolymer Analysis*, S.K.R. Williams and K.D. Caldwell, Editors. 2012, Springer Vienna: Vienna. p. 37-55.
38. Zattoni, A., et al., *Hollow-fiber flow field-flow fractionation: a novel pre-MS method for proteomics.* Journal of Biotechnology, 2010. **150**: p. 207-208.
39. Pedrazzini, E., et al., *Protein quality control along the route to the plant vacuole.* The Plant cell, 1997. **9**(10): p. 1869-1880.
40. Watson, J., et al., *Expression of Bacillus anthracis protective antigen in transgenic chloroplasts of tobacco, a non-food/feed crop.* Vaccine, 2004. **22**(31): p. 4374-4384.
41. Dhingra, A., A.R. Portis, and H. Daniell, *Enhanced translation of a chloroplast-expressed RbcS gene restores small subunit levels and photosynthesis in nuclear RbcS antisense plants.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(16): p. 6315-6320.

42. Bellucci, M., et al., *A mutant Synechococcus gene encoding glutamate 1-semialdehyde aminotransferase confers gabaculine resistance when expressed in tobacco plastids.* Plant Cell Reports, 2015. **34**(12): p. 2127-2136.
43. Suzuki, E., et al., *Crystallization of phaseolin from Phaseolus vulgaris.* Journal of Biological Chemistry, 1983. **258**(4): p. 2634-6.
44. Bellucci, M., et al., *Accumulation of maize γ-zein and γ-zein: KDEL to high levels in tobacco leaves and differential increase of BiP synthesis in transformants.* Theoretical and Applied Genetics, 2000. **101**(5): p. 796-804.
45. Beckett, R., Z. Jue, and J.C. Giddings, *Determination of molecular weight distributions of fulvic and humic acids using flow field-flow fractionation.* Environmental Science & Technology, 1987. **21**(3): p. 289-295.
46. Marassi, V., et al., *Hollow-fiber flow field-flow fractionation and multi-angle light scattering as a new analytical solution for quality control in pharmaceutical nanotechnology.* Microchemical Journal, 2018. **136**: p. 149-156.
47. Striegel, A.M., *Stepan Podzimek: Light scattering, size exclusion chromatography and asymmetric flow field flow fractionation. Powerful tools for the characterization of polymers, proteins and nanoparticles.* Analytical and Bioanalytical Chemistry, 2012. **402**(5): p. 1857-1858.
48. Masuelli, M. and D. Renard, *Advances in Physicochemical Properties of Biopolymers (Part 1)*. 2017: Bentham Science Publishers.
49. Bollini, R., A. Vitale, and M.J. Chrispeels, *In vivo and in vitro processing of seed reserve protein in the endoplasmic reticulum: evidence for two glycosylation steps.* The Journal of cell biology, 1983. **96**(4): p. 999-1007.
50. Blagrove, R.J., et al., *Physicochemical and structural studies of phaseolin from French bean seed.* Plant Foods for Human Nutrition, 1983. **33**(2): p. 227-229.
51. Smilgies, D.-M. and E. Folta-Stogniew, *Molecular weight-gyration radius relation of globular proteins: a comparison of light scattering, small-angle X-ray scattering and structure-based data.* Journal of applied crystallography, 2015. **48**(Pt 5): p. 1604-1606.
52. Corporation, W.T.; Available from: https://www.wyatt.com/solutions/properties/conformation-of-macromolecules-and-nanoparticles.html.
53. Feroz Jameel, S.H., *Formulation and Process Development Strategies for Manufacturing Biopharmaceuticals*.
54. Foresti, O., et al., *A phaseolin domain involved directly in trimer assembly is a determinant for binding by the chaperone BiP.* Plant Cell, 2003. **15**(10): p. 2464-75.
55. Chen, H., et al., *Application of Protein-Based Films and Coatings for Food Packaging: A Review.* Polymers (Basel), 2019. **11**(12).

Figure captions

Figure 1. *Localization of Modified Phaseolin (P*) localization in transplastomic chloroplast compartments (thylakoids) of tobacco leaves.*

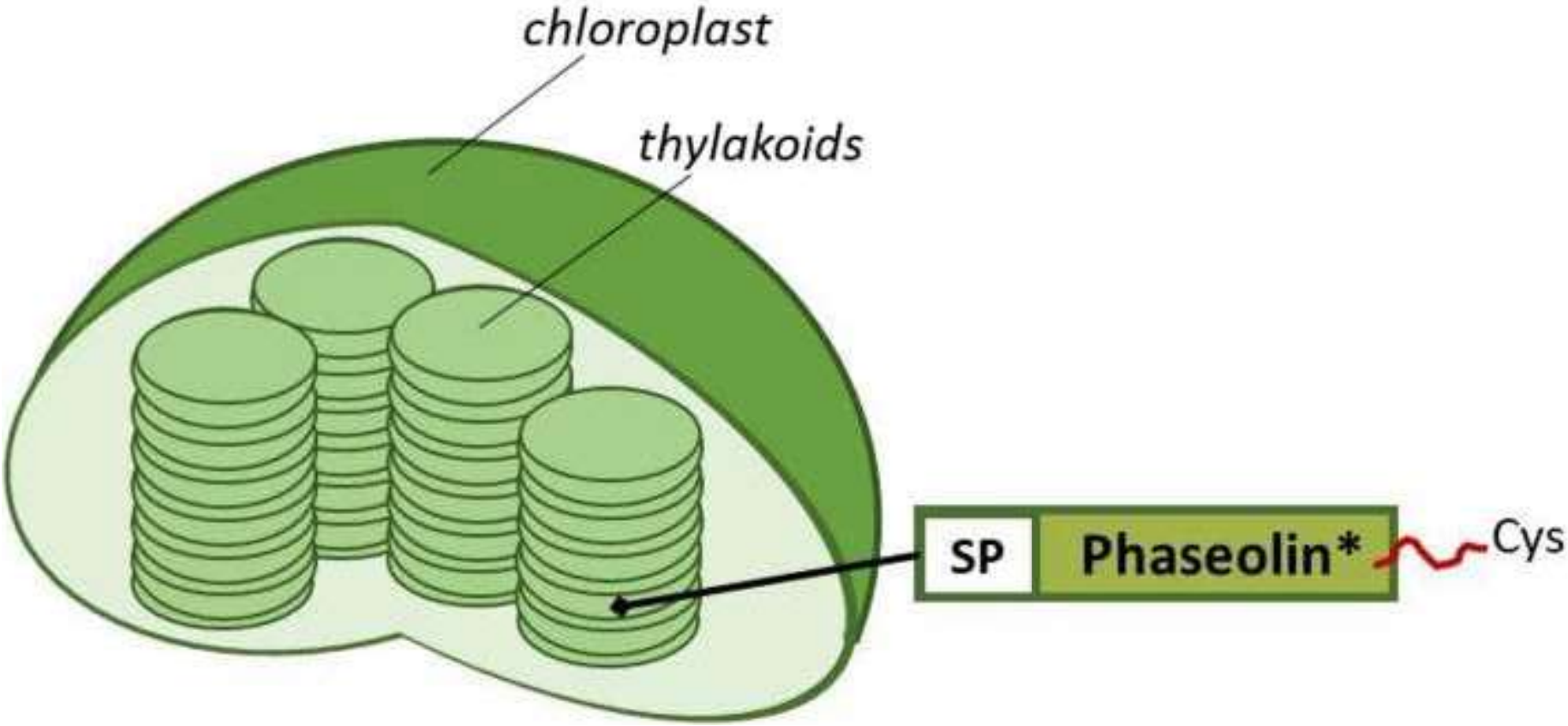
Figure 2*: (A) SDS-PAGE stained by Coomassie of P (0.5 μg) extracted from P. vulgaris (left) and Western blot of the same P (0.04 μg) sample using anti-phaseolin antiserum (right). Arrows indicate differentially glycosylated monomers of α- and β-phaseolin, (B) Total proteins (20 μg) extracted from leaves of transplastomic plants expressing P* and P (0.06 μg) extracted from P. vulgaris were analysed by Western blotting with anti-phaseolin antiserum. Due to the higher amount of protein loaded in the Western blot in (b) with respect to that in (a), phaseolin in the P sample migrated apparently as a single band, instead of three bands of α- and β-polypeptides as in (a). Black arrowhead indicates the mutated β-phaseolin monomer (P*) at 45 kDa, white arrowhead indicates P* dimers, and brace indicates P* high-molecular mass forms unresolved by SDS page. Numbers on the left of the figures indicate the positions of molecular mass markers in kDa.*

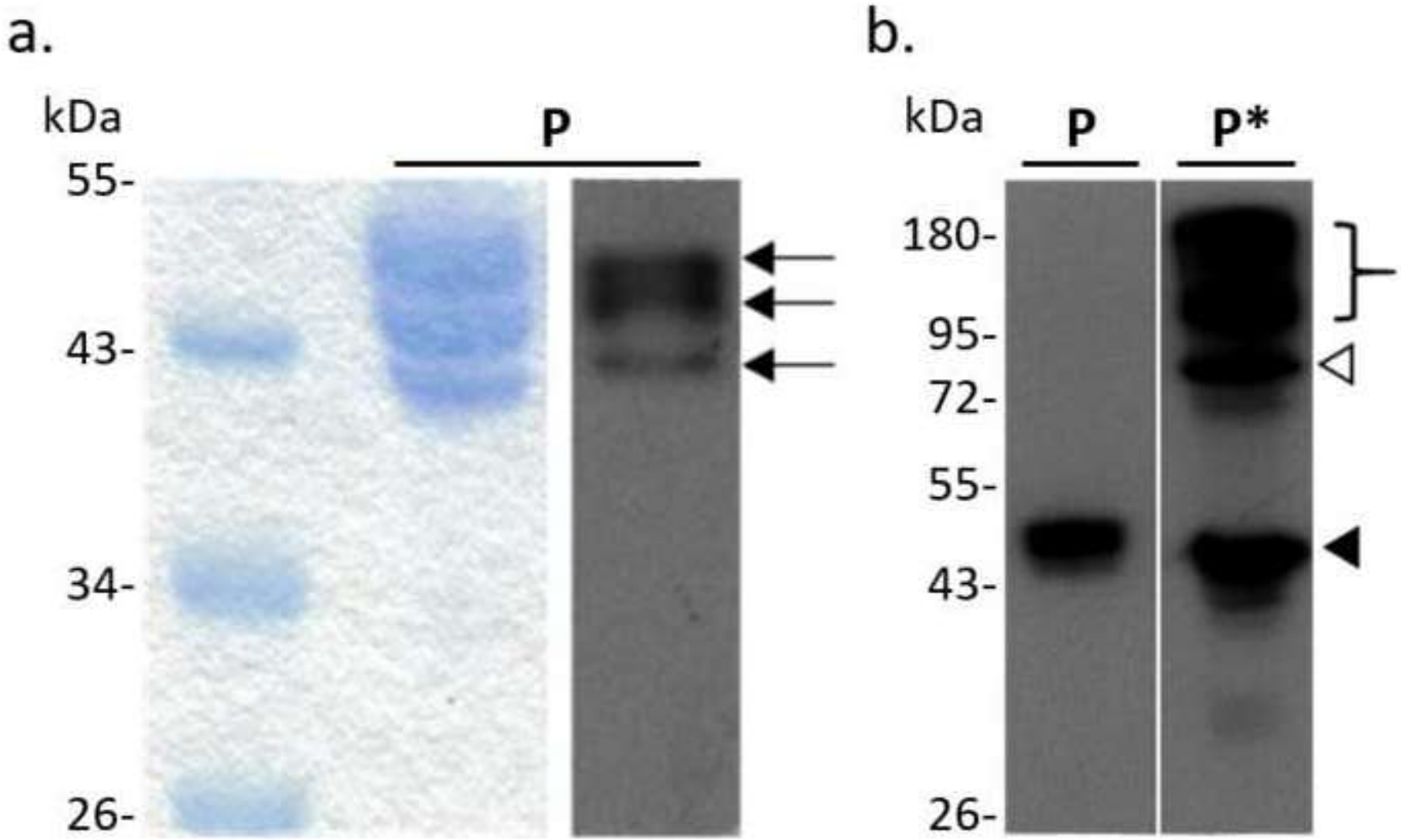
Figure 3*: (a) Schematic representation of P* purification that is described in detail in the Materials and methods section. (b) Each fraction and the pellet (indicated at the top) obtained from the velocity Suc gradient (a) analyzed by protein blot and visualized using anti-phaseolin antiserum. At the bottom, numbers indicate molecular mass of sedimentation markers (kDa) and the Suc concentration of the fractions is shown. Black arrowhead indicates P* monomers, white arrowhead indicates P* dimers, and brace indicates P* high-molecular mass forms.*

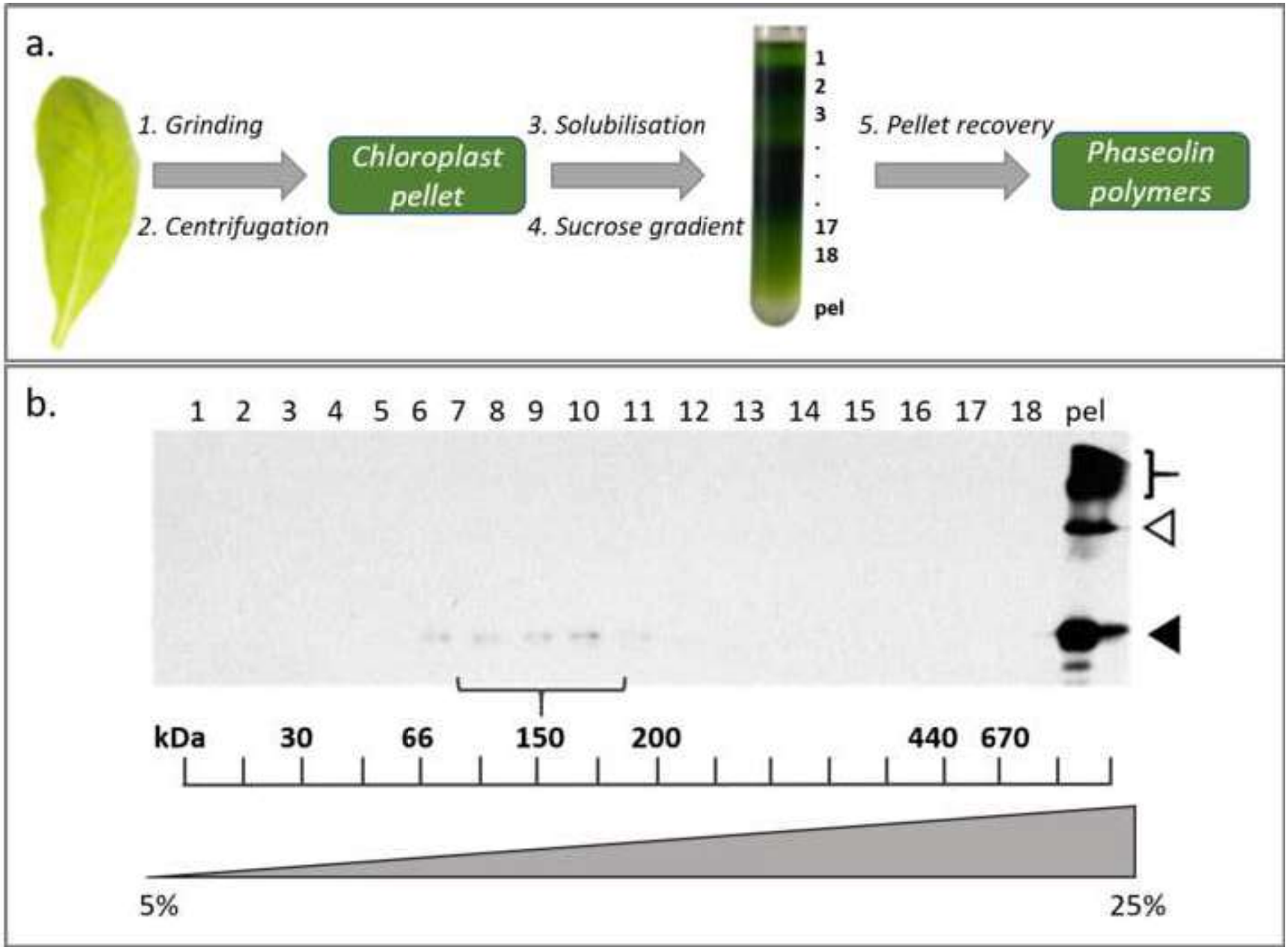
Figure 4. *Separation profile (grey, UV@280 nm) and molar mass calculation (green) of P characterization with the first method*

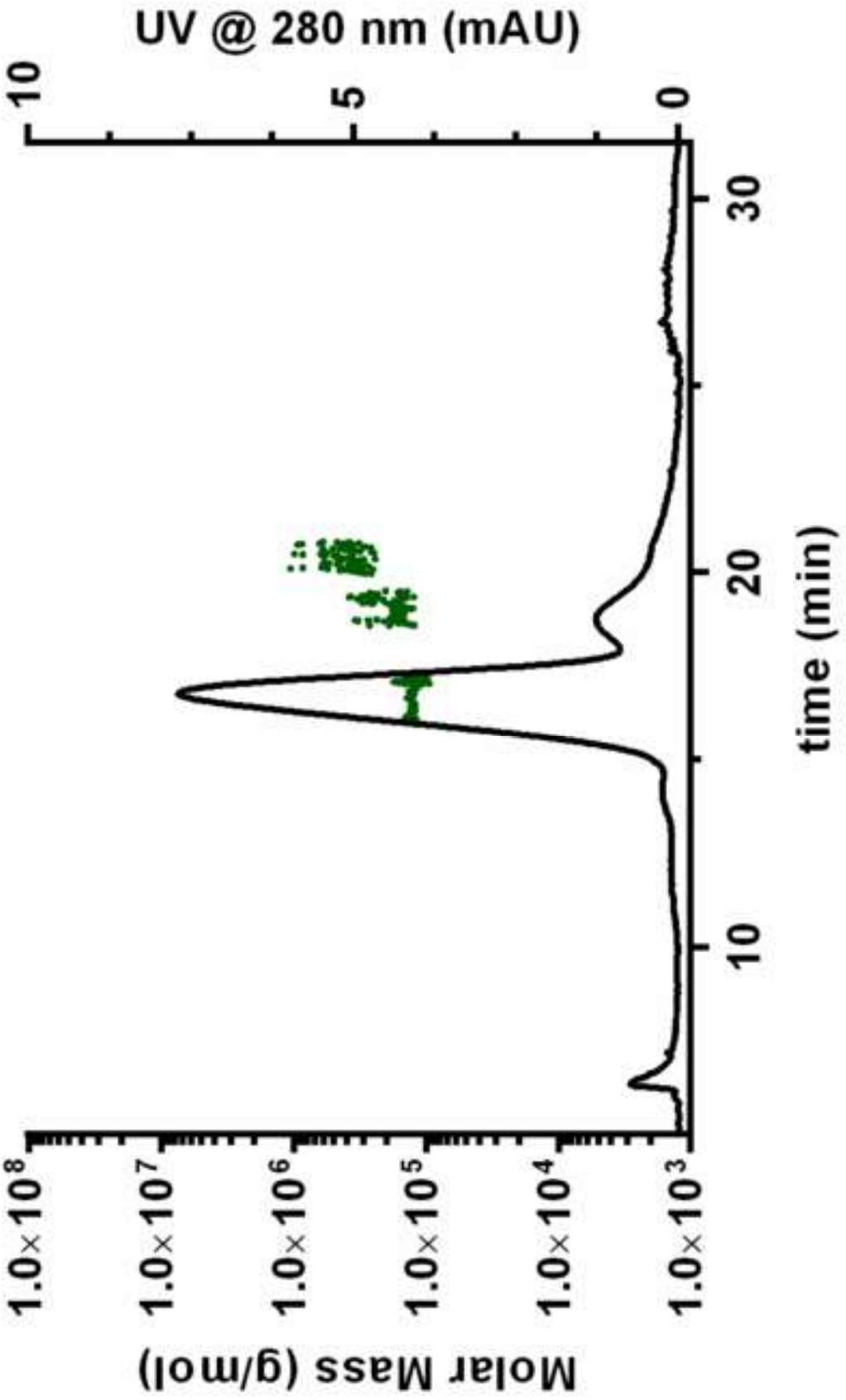
Figure 5. *Separation profile (grey, UV@280 nm) and molar mass calculation (green) of P* characterization with the first method*

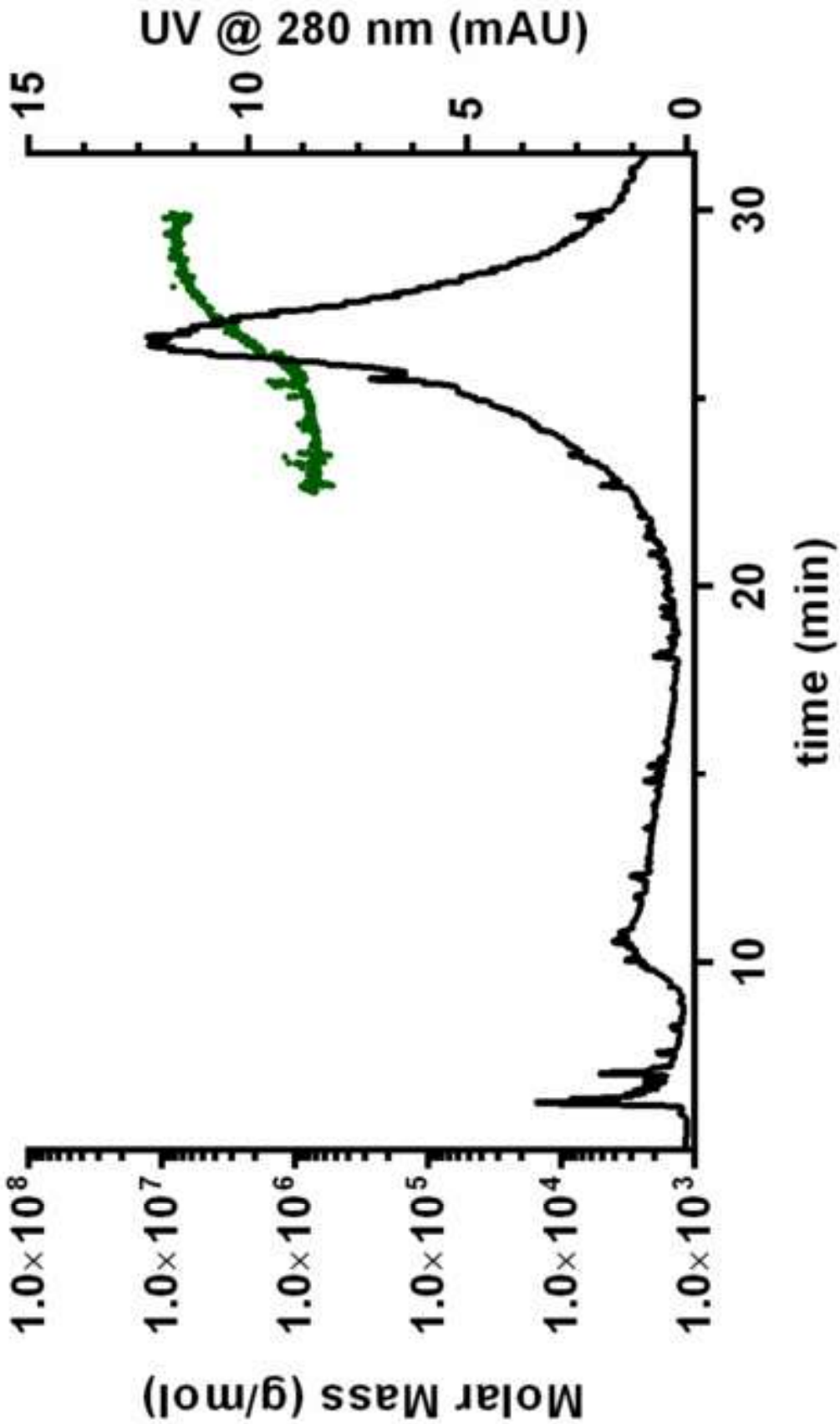
Figure 6*. Fractogram (grey, UV@280 nm) and molar mass calculation (green) of (a) P characterization and (b) P* characterization with the Method 2.*

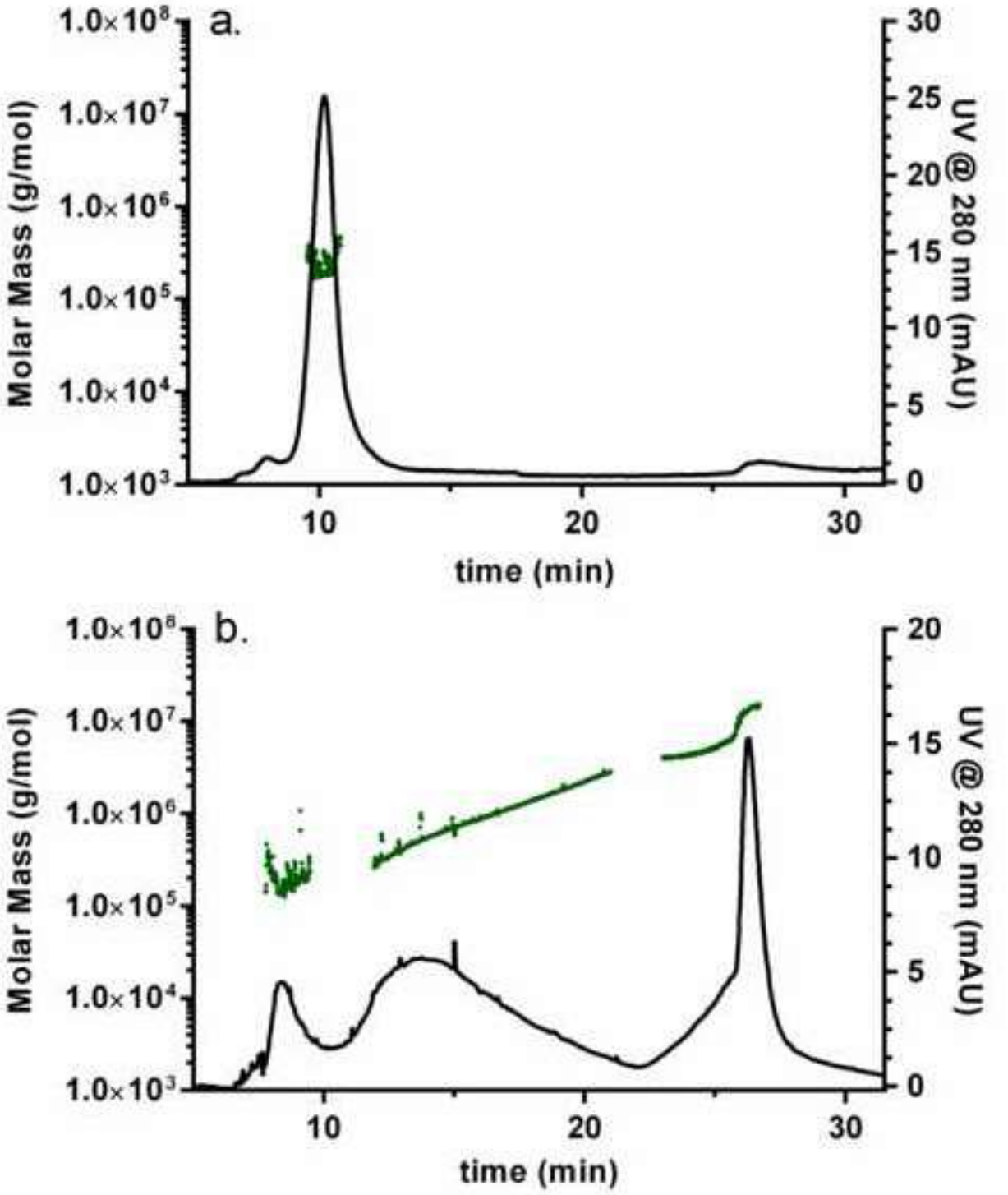












Steps → ↓Method	Focus (mL/min)	Focus-injection (mL/min)	Elution (mL/min)				Elution-Inject (mL/min)
Phaseolin characterisation	Vc=0.25 Vx=0.80 Time=0.5 min	Vc=0.25 Vx=0.80 Time=5 min	Vc=0.25 Vx=0.80 to 0.03 Time=20 min		Vc=0.25 Vx=0 Time=5 min		Vc=0.25 Vx=0 Time =3 min
Aggregates Characterisation	Vc=0.35 Vx=0.80 Time=0.5 min	Vc=0.35 Vx=0.80 Time=5 min	Vc=0.35 Vx=0.30 to 0.1 Time=5 min	Vc=0.35 Vx=0.1 Time=10 min	Vc=0.35 Vx=0.1 to 0.0 Time=5 min	Vc=0.35 Vx=0.0 Time=5 min	Vc=0.35 Vx=0 Time =3 min

Table 1. *Flow conditions for HF5 analyses*

Peaks → Characterisation ↓	**Min 14**	**Min 16.5**	**Min 19**	**Min 21**
% area	3%	78%	10%	9%
Peak attribution	2-mer	3-mer	6-mer	9-mer

Table 2. Phaseolin aggregation states

Shape factor of species (v value)	**Band 1 (9.5 to 11 minutes)**	**Band 2 (13 to 17 minutes)**	**Band 3 (23 to 26 minutes)**
Standard phaseolin	0.26 ± 0.01	-	-
Mutated P*	-	0.52 ± 0.09	0.8 ± 0.02

Table 3. Shape factors of phaseolin and mutated phaseolin (P) bands obtained from method 2.*

Authors Credit role

Valentina Marassi: formal analysis; methodology; data curation; Roles/Writing - original draft

Francesca De Marchis: Conceptualization; data curation; Roles/Writing - original draft; Writing - review & editing

Barbara Roda: Resources; Writing - review & editing

Michele Bellucci: Conceptualization; data curation; Roles/Writing - original draft; Writing - review & editing

Alice Capecchi: formal analysis; methodology; data curation

Pierluigi Reschiglian: Funding acquisition; Writing - review & editing

Andrea Pompa: Conceptualization; data curation; Roles/Writing - original draft; Writing - review & editing

Andrea Zattoni: Conceptualization; methodology; data curation; Roles/Writing - original draft; Writing - review & editing

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

V. Marassi, B. Roda, P. Reschiglian and A. Zattoni are associates of the academic spinoff company byFlow Srl (Bologna, Italy). The company mission includes know-how transfer, development, and application of novel technologies and methodologies for the analysis and characterization of samples of nano-biotechnological interest.