Pre-Processing Unsellable Tomatoes into Separated Streams with the Intention of Recovering Protein

Paul Baker (✉ paul.baker@bangor.ac.uk)
Bangor University BioComposites Centre  https://orcid.org/0000-0002-4226-0439

Dave Preskett
Bangor University BC: Bangor University BioComposites Centre

Dominik Krienke
GEA Westfalia Separator group

Kasper S. Runager
Danish Technological Institute: Teknologisk Institut

Anne C.S. Hastrup
Danish Technological Institute: Teknologisk Institut

Adam Charlton
Prifysgol Bangor: Bangor University

Research Article

Keywords: Kjeldahl, Bradford, carbohydrase, cellulase, pectinase, fibre

Posted Date: October 28th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-942516/v1

License: ☇ ① This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Purpose

A large proportion of the European Union's tomato crop is discarded during harvesting and there is a valorisation potential to recover proteins from this waste.

Methods

Cherry tomatoes were segregated into three separate components: juice, pomace (peels and skins), and seeds. The peels and skins, and seeds were separately hydrolyzed with carbohydrases to determine whether protein recovery could be increased. In addition, a strategy to fractionate the seeds using sequential washing of milled tomato seeds followed by low-speed centrifugation to remove the denser seed hulls and to collect the protein rich kernels remaining in suspension.

Results

The protein content of the seeds was highest with 27.4% while the peels and skins contained 7.6%. Carbohydrase mediated hydrolysis revealed a minor increase in protein recovery from seeds by 10% using Filta 02L (cellulase, xylananse and β-glucanase), and the quantity of protein recovered from peels and skins increased by 210% using Tail 157 (pectinase, hemicellulase). The strategy to separate the seeds into two fractions, revealed that a higher proportion of the fibre (65%) was associated with the hull fraction compared with the original seeds (47%). A significant proportion of the fibre in this fraction was composed lignin although the protein contents between both fractions was similar ranging from 27.4% to 29.9%.

Conclusions

These results reveal that carbohydrases were quite effective in protein extraction from peels and skins, but not from seeds. An alternative strategy was developed to remove the seed hulls from the milled seeds and to collect a crude protein fraction where the protein content could be further improved.

Statement Of Novelty

Large quantities of waste tomatoes are produced yearly that have valorisation potential. These waste tomatoes were separated into three different components: juice, peels and skins, and seeds. Previous studies have focused on solvent extraction of oils from seeds prior to protein extraction, whereas this study focused on developing a process to extract protein without the initial de-oiling step that would reduce solvent use.

Introduction
Tomatoes form a large proportion of the European Union’s fruit and vegetable crop harvest [1], with an annual production of 16.4 million tonnes. Approximately 0.5 million tonnes per annum is lost during the production stages [2], which is either used as animal feed or transferred to landfill sites. However, these losses may be based on indoor glasshouse production where pests, diseases and rates of growth can be controlled, whereas losses may be greater in some EU countries such as Spain that rely on outdoor growth [3]. There is an opportunity to valorize the unprocessed tomatoes or side residue streams generated from this crop, in order to isolate additional, value-added products. These include important compounds such as cutin found in the tomato skins [4] as a source of functional dietary fibre to reduce the risk of colorectal cancers [5] and carotenoids such as lycopene [6], with a range of medical related applications including in the prevention of certain cancers, cataracts and coronary diseases [7].

The current use of unsaleable tomatoes undergoes composting to form a fertilizer with a high protein content of up to 20% [8]. However, there is potential to upscale and this biomass to focus on tomato seeds. Previous studies have focused on the potential source of functional protein containing 23–34% of material and are separated from the tomato pulp by sedimentation [9, 10, 11]. The functional properties of these protein extracts showed higher oil absorbance compared with soy proteins, while the foaming and emulsion characteristics were similar [12]. In laboratory feeding trials on rats, tomato seed protein resulted in acceptable weight gains but to a lesser extent compared with the use of casein [13]. One of the benefits of tomato seeds is the higher proportion of the essential amino acid, lysine, that could be used to fortify foods low in lysine, whereas other plant proteins such as soy and corn contain lower amounts of this amino acid [11]. While the potential functional properties of tomato seed protein have been reported, the development of an industrial process that does not rely on de-oiling of the tomato seeds prior to protein extraction has not been explored. Such a process would provide further progression towards developing a biorefinery process and associated value-chain approach to improve the economic viability of tomato waste valorisation, thereby diverting material away from animal feed applications or landfill.

The aim of this study was to develop a protocol to separate tomatoes into juice, pomace (pulp and peel), and seeds. Compositional analysis was performed on each of these fractions to provide an insight into the valorisation potential and the presence of anti-nutritional compounds, the levels of which will impact on potential food applications areas because of reduced palatability. An assessment of the valorisation opportunities for these different fractions focused on proteins and involved determining the effectiveness of enzymatic hydrolysis to aid the separation process. Finally, a procedure was developed to recover crude protein from tomato seeds.

**Materials And Methods**

**Separation of tomato seeds**

Cherry tomatoes (4.5 kg) that were deemed unsaleable were supplied frozen by Annecoop Ltd., Spain which has one of the largest tomato farming operations in the EU and supplies fresh produce to customers across Europe. These tomatoes were heated at 50°C for 5 min to loosen the skins and then
passed through a juicing machine (Robot Coupe Automatic Sieve C200, Robot Coupe, France). Tomato juice was collected from one outlet and pomace was collected at the other. The pomace was placed into a 200 L pan, stirred and periodically passed through a sieve (5 mm diameter holes) and into a collection bucket. The pulp and skins collected within the sieve were emptied whenever this became full and the seeds possessing a gelatinous outer layer passed through the holes of the sieve. Any tomato pulp or skins that had passed through the sieve floated on the surface of the water in the collection bucket and these were skimmed off the surface. The collected seeds were then freeze dried to prevent degradation and ball milled.

**Compositional analysis of tomato components**

The moisture content of each fraction was determined on a 5 g sample using a moisture analyser at 105°C (Kern, Germany), until the evaporation rate was less than 2 mg of water/20 s. The total dry weights in the samples of the different tomato fruit fractions components were determined by multiplying the total wet weights by the percentage of dry weight that was determined using the moisture analyser. These dried samples were used to determine the fibre and phytic acid content. Fibre analysis was performed using a fibre analyser (Ankom Technology, USA) following the procedure described in the Ankom Technical Manual ([http://www.ankom.com](http://www.ankom.com)) for neutral detergent fibre (NDF) and acid detergent fibre (ADF), after extracting any residual oil in acetone by shaking ten times and incubating for 10 min. The acetone washed samples were air dried at room temperature and then oven dried at 103°C for 30 min. Neutral detergent fibre was determined using 2 L neutral fibre detergent, 20 g sodium sulphite and 4 ml amylase (Ankom Technology) for 1 h at 100°C and 10 psi and the fibre bags washed three times with 2.67 ml amylase (Ankom Technology) and 116.67 ml deionized water for 5 min each cycle. ADF was determined by treating the fibre bags with ADF solution for 1 h at 100°C and 10 psi followed by 3 x 5 min washes with hot distilled water. Lignin analysis was performed in the fibre bags in the Daisy Incubator (Ankom Technology) at room temperature (3 h in 500 ml 72% sulphuric acid) and then continuously washed in tap water until the pH was neutral as determined using a pH meter. The Ankom bags were soaked in acetone for 5 min and dried at 103°C for 3 h at end of each treatment with NDF, ADF and lignin analysis. The ash contents were determined on 0.5 g of sample, not previously used analysed for fibre content, at 600°C for 4 h in a furnace oven. Hemicellulose content was determined by subtracting ADF from NDF values, cellulose content was determined by subtracting lignin from ADF values, and lignin was determined by subtracting lignin from the weight of the empty bags. The ash contents in each of the respective fibre components were determined and subtracted from each component.

Protein determination, performed in duplicate, using the Bradford assay was carried out by weighing wet material to the equivalent of 1 g dry weight of material, as determined by moisture content. To the tubes containing pulp, 0.5 M NaOH (20 ml) was added, while 0.2 M NaOH (10 ml) was added to the seeds. The discrepancy in concentration and volume used reflected the higher moisture content associated with the pulp. The tubes were incubated at 50°C with end-over-end rotation (1 h) and then centrifuged at 1493 x g for 5 min. The supernatants were decanted into preweighed tubes and the weights of the supernatants were measured. Assuming that 1 g of supernatant was equivalent to 1 ml and the volumes were
determined. The concentration of protein in the supernatants were determined using the Bradford assay and each sample was diluted until a reading of 0.4 was obtained, using the microplate reader at 595 nm.

The presence of anti-nutrients occurring as phytic acid was determined using the assay kit (Phytic Acid Kit, Megazyme, Ireland), which involved measuring free phosphates and then enzymatically released phosphates bound as phytic acids. The phytic acids were solublized from 1 g sample in 20 ml 0.66 M HCl. The assay contains phytase that degrades any phytic acids present into myo-inositol polyphosphates and another enzyme, alkaline phosphatase, into free phosphates. The phosphates are detected by reacting them with ammonium molybdate and acid to form a blue coloured compound that can be measured in the plate reader at 655 nm. The free phosphate concentrations were performed beforehand and these values were subtracted from the phosphates indicating phytic acids.

**Enzyme mediated protein extraction from tomato seeds**

Freeze dried tomato seeds (20 g) were ball milled and samples of 100 mg were weighed into 2 ml Eppendorf tubes. To each tube, distilled water (1.65 ml) and carbohydrase (Table 1) diluted to one tenth (50 µl) was added and each treatment was performed in duplicate. The control consisted of distilled water (1.7 ml) without any enzyme. The milled seeds were resuspended in water and incubated at 50°C with end-over-end rotation at 40 rpm (3 h). The tubes were centrifuged at 15493 x g (10 min), the supernatants were collected noting the exact volume. The protein concentrations associated with each of the enzymes were determined and these values were subtracted from protein concentrations associated with each of the enzymes. The protein concentrations were determined in the supernatants using the Bradford method and measuring the absorbance at 595 nm until an absorbance value of ~0.4 was obtained.

**Enzyme mediated extraction of protein from tomato pulp**

Wet tomato pulp (1 g equivalent dry weight) was centrifuged at 5575 x g (10 min) and the supernatant was discarded. A minimal amount of protein was lost in the supernatant (0.057% of protein per g dry biomass). The pellets were resuspended in distilled water (20 ml) and the pH was adjusted from 4.2 to 4.7 ± 0.1 with 0.2 M NaOH. Each of the carbohydrases (Table 1) was separately added (50 µl) and the tubes were incubated with end-over-end rotation at 40 rpm for 3 h at the optimum temperature for the respective carbohydrase. The tubes were centrifuged at 5575 x g for 5 min and the supernatant was decanted into another tube. The weights of the supernatants were measured, and protein concentrations were determined in water using the Bradford assay. Controls were also prepared using only carbohydrases at the same concentrations used to determine the effectiveness of carbohydrases in protein extraction and these values were subtracted from protein concentrations determined with each carbohydrase.

**Laboratory scale protocol to extract proteins from tomato seeds**
Duplicate sets of milled seeds (2.5 g) were transferred into 50 ml Falcon tubes and distilled water (40 ml) added to each. The tubes were shaken for 20 s and centrifuged for 3 min at either 40 x g or 1493 x g. The supernatants were decanted into fresh tubes and aliquots of 0.5 ml were taken to determine protein content using the Bradford assay. Further extraction on the same samples involved adding distilled water (35 ml) to each pellet and the mixture was agitated (20 s), centrifuged (3 min at 40 x g or 1493 x g) and samples (0.5 ml) collected from the supernatant for protein determination. The remainder of the supernatant was decanted and pooled with the previous supernatants. This whole step from resuspension of the pellet in 35 ml distilled water, to pooling of supernatants was repeated three more times. Ethanol (1:1 (v/v)) was added to the pooled supernatants, left overnight at 4°C and the suspension was centrifuged at 5575 x g for 5 min to collect the protein pellet. The remaining seed hulls and the extracted protein were freeze-dried and weighed. The non-fibrous portion of the seeds was calculated by subtracting the weights of the seed hulls and seed protein from the original weight of seeds at the start of the experiment.

**Statistical analysis**

Levels of significant differences were calculated using Student's \( t \)-Test on each processed sample compared with the control. Each enzyme treatment and the controls were performed in duplicate.

**Gel electrophoresis and protein identification**

Protein samples were diluted (50 mg powder/ml) in 1x lithium dodecyl sulfate (LDS) sample buffer (Thermo Scientific) +/- 10 mM reducing agent (dithiothreitol, DTT, Sigma-Aldrich). Samples were then heated to 70°C for 30 minutes and centrifuged for 1 min at 10,000 \( \times \) g. Proteins were then separated using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) on 4-12% acrylamide gels (Genscript, Netherlands).

Protein bands of interest were excised with a scalpel and subjected to in-gel digestion using trypsin (Promega, USA). Extracted peptides were reduced and alkylated using DTT and iodoacetamide, respectively, and analyzed using matrix-assisted laser desorption ionization (MALDI) mass spectrometry (autoflex maX, Bruker Daltonics). The peptide mass list was searched against the NCBIProt database using Mascot search engine (Matrix Science, USA) with carbamidomethyl as a fixed modification of cysteine and oxidation of methionine as a variable modification.

**Results And Discussion**

**Tomato composition**

The whole tomato fruits were fractionated into juice, pomace (pulp and peel), and seeds (Fig. 1). It was impossible to separate the peel and pulp because at this stage the peel appeared as fragmented pieces. The juice formed the largest proportion of material in tomatoes in terms of dry biomass (60%), followed by the seeds (12%). However, a large proportion of these tomatoes was comprised of water and only 49 g
of tomato seeds were recovered from 4.5 kg fresh tomatoes. The processing of a larger quantity of round tomatoes (100 kg) resulted in a lower proportion of seeds being recovered (250 g) which would still be insufficient to perform pilot scale studies [Baker et al., unpublished]. Therefore, a different strategy was implemented to recover a larger quantity of seeds from a tomato juice processing plant [14]. However, the development of a strategy to recover crude protein from tomatoes that would be otherwise discarded is an important process that should be independently investigated.

Compositional analysis of the pomace and seeds indicated that the non-fibre was the major component in both fractions (Table 2), with a relatively high lignin content in the pomace (25.3 ± 1.5%). Whether the remaining material after fibre analysis was lignin is unclear and one possibility is that cutin, a macromolecule with properties similar to lignin, and that can form a structural component in mosses [15] may show acid resistant properties that are similar to lignin. In contrast, the lignin content in the seeds was considerably lower (9.51 ± 0.43%), with cellulose and lignin contents of the seeds having similar values to those reported previously [16]. There were significant quantities of cellulose in the pomace and both hemicellulose and cellulose in the seeds, indicating that the application of an enzymatic processing step using carbohydrases could have an effect in releasing proteins from both fractions.

The protein content was determined for each of the tomato fractions, with the milled tomato seeds containing the highest amounts (Table 2). The protein content in un-milled seeds was 0.17 ± 0.02% and so it was necessary to mill the seeds to reduce particle size and facilitate optimised, downstream protein recovery [17]. Kjeldahl analysis revealed that the protein concentrations in milled seeds (Table 2) were similar to those reported previously [10, 11], but there are no reported values for the protein content of tomato pulp. It is possible that some of the protein remained inaccessible even under the alkaline treatment, resulting in lower values determined using the Bradford assay. Similar findings were revealed in another study, where protein determination of tomato seeds under weak alkaline conditions, using the Bradford assay, was lower than values determined using Kjeldahl analysis [18]. The Bradford assay measures soluble protein although measurement with a buffer at a specific pH will reveal the majority of soluble proteins [19]. In contrast, Kjeldahl analysis will also include insoluble proteins such as globular proteins although very low concentrations of other nitrogenous compounds besides amino acids may also be included.

Three quarters of phosphorus in plants is stored as phytates, although it cannot be absorbed by ruminants and the presence of phytates in food can lead to the deficiency in other minerals such as iron and zinc [20]. Our results showed that 76% of the total phosphorus was in the form of phytic acid and the amount was within the expected range for most cereal grains (Megazyme, Ireland). A previous study has shown that the carryover of phytates into the protein extract is much less compared with the quantity found in the original seeds [21].

**Enzyme mediated protein extraction from seeds**

A range of different enzyme mixtures were assessed in order to determine any improvements in protein extraction, each with minor differences in optimum temperature and pH conditions (Table 1). Enzyme
mediated protein extraction from seeds indicated that the only carbohydrase which resulted in an increase in protein yield, following treatment, was Filta 02L. This was 10% higher and significantly different compared with the control (Fig. 2A) without enzyme (P = 0.008). Filta 02L possessed β-glucanase, xylanase and cellulase activities that would enable these carbohydrases to hydrolyse both the hemicellulose and cellulose present in the matrix. In contrast, Cellux 01L possessed β-glucanase and cellulase activities and did not show increased protein recovery. The other carbohydrases evaluated as part of this study, Tail 157 and Tail 01, did not possess cellulase activity and subsequently yielded lower quantities of protein compared with Filta 02L. Tail 113 which had similar carbohydrases to Filta 02L did not show increased levels compared with the control, perhaps due to lower levels of carbohydrase activity or possessed activity that targeted different carbohydrates. Therefore, carbohydrases that could completely degrade both hemicellulose and cellulose, such as xylanase, cellulase and β-glucanase appear to be necessary in order to increase the recovery of protein from tomato seeds. Some of the carbohydrases, Tail 157, Tail 113 and Cellux 01L, appeared to show a decrease in protein yield in comparison to the control without carbohydrase. However, only Tail 157 was significantly lower (P = 0.008) and it unclear why this resulted in a lower yield. However, this enzyme contained higher quantities of pectinases than the other enzymes and the release of galacturonate may form a gel-like substance in the presence of metallic ions [22], that may result in protein binding.

**Further separation of the milled seeds**

Another strategy was explored to separate the milled tomato seeds into two separate fractions, one containing predominately the seed hulls and the other fraction predominately containing the seed kernels. Tomato seeds are composed of ~50% non-fibrous material (Table 1), and it would appear, based on results from previous studies, that a significant amount of this soluble material could be tomato seed oil which has been found to comprise 31% of the total tomato seed weight [23]. Most studies have described the formation of tomato meal where the oils are removed through a defatting process by hexane extraction [11, 21, 24, 25, 26, 27] or by using pressing [23]. However, the use of large quantities of organic solvent in pilot scale studies presents significant flammability issues, along with increasing CAPEX and OPEX and it is desirable to develop a green process that limits organic solvent use. Therefore, a process was investigated that would postpone the use of organic solvents until a later stage where lower volumes would be required due to the decrease in biomass.

The total quantity of protein recovered after the first extraction and from the pooled sequential extraction mixtures was ~ 3% and 10% g of protein per g dry biomass, respectively. This indicated that a significant amount of protein could be recovered from subsequent washes, although this does require the use of large volumes of water and ethanol. Rather than using large quantities of ethanol, protein could be precipitated using citric acid and up to 500 mg/L chitin as a coagulating agent [28] or different pH acidification ranges [29]. The volume of water to weight of biomass ratio at 18:1 during the initial wash was similar to previously reported values, conducted under dilute alkaline conditions [24]. However, a considerable amount of residual protein remained in the seed hull fraction following processing (Table 4). The quantity of protein in the crude protein extracts as determined by the different methods, Bradford
and Kjeldahl analysis, were similar at ~30%, indicating that the majority of the protein was soluble and that a considerable proportion of other components were co-extracted with the protein. A hexane extraction revealed that 25% of the total crude protein was composed of oils and fats, indicating that the protein content would increase to ~40%.

It was determined that the sequential aqueous extraction of proteins revealed that use of higher centrifugation speeds resulted in a lower proportion of material remaining suspended and more protein was associated with the centrifuged pellet which was assumed to contain seed hulls fragments based on appearance (Table 4). Higher levels of protein were recovered when lower speeds were used but the precipitates appeared slighter browner compared with precipitates obtained with higher centrifugation speeds due to the carryover of smaller seed hulls fragments into this fraction. Nevertheless, there were minor differences between the crude fractions in terms of protein content determined using Kjeldahl analysis. The protein contents of the material left in the suspensions at different low speed centrifugation speeds were generally similar. Protein contents determined using Kjeldahl analysis or the Bradford assay where generally similar in contrast to protein contents determined in the milled seeds. This might indicate that the proteins in this separated fraction were more accessible to the reagents in the Bradford assay whereas the majority of those in the seeds were inaccessible. Fibre analysis of the residual hulls indicated that the lignin content had increased more than three-fold compared with the original seeds, whereas the levels of hemicellulose and cellulose remained unchanged, and the non-fibre content decreased (Fig. 4). This clearly indicated a separation of the tomato seed components.

An investigation into alkali extracted tomato seed proteins indicated that 61% were salt soluble globulins and 37% was glutenin and gliadins that were soluble in acetic acid and ethanol, respectively [27]. Similarly, analysis of the protein composition of the seed fraction containing mostly seed kernels using SDS-PAGE (Fig. 5) showed that the major proteins were globulins based on comparison with previously published data [30] showed that the protein bands migrating in the reducing lane (+DTT) at 49 kDa, 35 kDa and 20 kDa, represented globulins (Fig. 5). Another study showed that globulins precipitated at pH 3.8- 6.2, whereas those that were soluble proteins were precipitated at pH 3.5- 4.6 [31]. Therefore, it is possible that a higher proportion of soluble globulins were extracted under neutral pH conditions.

**Conclusions**

The carbohydrases showed higher protein efficiency of protein recovery in the seed fraction, through use of carbohydrases was lower compared with the activity on pulp and peel, despite the higher concentrations of hemicellulose in seeds. It would appear that under these conditions the use of carbohydrases might be ineffective from a technoeconomic perspective, due to the high production costs for these enzymes and when considering the minor increase in yield of protein in seeds and the lower quantity of protein associated with pulp and peel. However, a pre-process step was developed to remove a significant proportion of lignin from the milled seeds without the requirement for use of carbohydrases. This indicated that the other fraction contained a considerably lower lignin content and this milled seed fraction represented one third of the original weight of the seeds. Further processing to increase the
protein content and this approach could be investigated as part of the strategy for optimising and
developing a commercial process to extract protein from pulp and peel.

**Declarations**

Not applicable

**Acknowledgements**

The authors are grateful to Nicolas Vidal for co-ordinating the supply of whole cherry tomatoes from Annecoop Ltd (Spain) and to Simon Lausen from Tailorzymes Ltd. (Denmark) for provision of enzymes and sharing technical knowledge. The authors are also grateful to colleagues at the Food Technology Centre, Grŵp Llandrillo Menai, Gwynedd, Wales, for use of their processing facilities and provision of technical assistance. The authors would like to thank the Bio Based Industries Joint Undertaking for providing funding for the Pro-enrich project (Grant Agreement No. 792050), under the European Union's Horizon 2020 research and innovation programme.

**Data Availability**

Data is available on reasonable request

**References**

1. Union Européenne. Commission européenne, & EUROSTAT. (2015). Agriculture, forestry and fishery statistics. Publications office of the European Union. https://doi.org/10.2785/143455

2. Scherhaufer, S., Moates, G., Hartikainen, H., Waldron, K., & Obersteiner, G. (2018). Environmental impacts of food waste in Europe. *Waste Management*, 77, 98-113. https://doi.org/10.1016/j.wasman.2018.04.038

3. Epp, M. (2016) Tomato production varies in European countries. https://vegetablegrowersnews.com/article/tomato-production-varies-european-countries/

4. Osman, S. F., Irwin, P., Fett, W. F., O’Connor, J. V., & Parris, N. (1999). Preparation, isolation, and characterization of cutin monomers and oligomers from tomato peels. *Journal of Agricultural and Food Chemistry*, 47(2), 799-802. https://doi.org/10.1021/jf980693r

5. Mudgil, D., & Barak, S. (2019). Classification, technological properties, and sustainable sources. In *Dietary Fiber: Properties, Recovery, and Applications* (pp. 27-58). Academic Press. https://doi.org/10.1016/B978-0-12-816495-2.00002-2

6. Choksi, P. M., & Joshi, V. Y. (2007). A review on lycopene—extraction, purification, stability and applications. *International Journal of Food Properties*, 10(2), 289-298. https://doi.org/10.1080/10942910601052699
7. Baysal, T., Ersus, S. & Starmans, D. A. J. (2000). Supercritical CO$_2$ extraction of β-carotene and lycopene from tomato paste waste. *Journal of Agricultural and Food Chemistry, 48*(11), 5507-5511. https://doi.org/10.1021/jf000311t

8. Fritsch, C., Staebler, A., Happel, A., Cubero Márquez, M. A., Aguiló-Aguayo, I., Abadias, M., Gallur, M., Cigognini, I.M., Montanari, A., López, M.J. & Suárez-Estrella, F. (2017). Processing, valorization and application of bio-waste derived compounds from potato, tomato, olive and cereals: a review. *Sustainability, 9*(8), 1492. https://doi.org/10.3390/su9081492

9. Latlief, S. J., & Knorr, D. (1983b). Tomato seed protein concentrates: effects of methods of recovery upon yield and compositional characteristics. *Journal of Food Science, 48*(6), 1583-1586. https://doi.org/10.1111/j.1365-2621.1983.tb05036.x

10. Piyakina, G. A., & Yunusov, T. S. (1995). General characteristics of the proteins of tomato seed flour and tomato skin flour. *Chemistry of Natural Compounds, 31*(4), 495-499. https://doi.org/10.1007/BF01177421

11. Brodowski, D., & Geisman, J. R. (1980). Protein content and amino acid composition of protein of seeds from tomatoes at various stages of ripeness. *Journal of Food Science, 45*(2), 228-229. https://doi.org/10.1111/j.1365-2621.1980.tb02582.x

12. Liadakis, G. N., Tzia, C., Oreopoulou, V., & Thompoulos, C. D. (1998). Isolation of tomato seed meal proteins with salt solutions. *Journal of Food Science, 63*(3), 450-453. https://doi.org/10.1111/j.1365-2621.1998.tb15762.x

13. Sogi, D. S., Bhatia, R., Garg, S. K., & Bawa, A. S. (2005). Biological evaluation of tomato waste seed meals and protein concentrate. *Food Chemistry, 89*(1), 53-56. https://doi.org/10.1016/j.foodchem.2004.01.083

14. Miklavčič Višnjevec, A., Baker, P. W., Peeters, K., Schwarzkopf, M., Krienke, D., Charlton, A. (submitted) HPLC-DAD-qTOF compositional analysis of the phenolic compounds present in crude tomato protein extracts derived from food processing. *Molecules.*

15. Philippe, G., Sørensen, I., Jiao, C., Sun, X., Fei, Z., Domozych, D. S., & Rose, J. K. (2020). Cutin and suberin: assembly and origins of specialized lipidic cell wall scaffolds. *Current Opinion in Plant Biology, 55*, 11-20. https://doi.org/10.1016/j.pbi.2020.01.008

16. Dalimov, D. N., Dalimova, G. N., & Bhatt, M. (2003). Chemical composition and lignins of tomato and pomegranate seeds. *Chemistry of Natural Compounds, 39*(1), 37-40. https://doi.org/10.1023/A:1024128512801

17. Seikova, I., Simeonov, E., & Ivanova, E. (2004). Protein leaching from tomato seed—experimental kinetics and prediction of effective diffusivity. *Journal of Food Engineering, 61*(2), 165-171. https://doi.org/10.1016/S0260-8774(03)00083-9

18. Mechmeche, M., Kachouri, F., Ksontini, H., & Hamdi, M. (2017b). Production of bioactive peptides from tomato seed isolate by *Lactobacillus plantarum* fermentation and enhancement of antioxidant activity. *Food Biotechnology, 31*(2), 94-113. https://doi.org/10.1080/08905436.2017.1302888
19. Douglas Jr, F. W., Greenberg, R., Farrell Jr, H. M., & Edmondson, L. F. (1981). Effects of ultra-high-temperature pasteurization on milk proteins. *Journal of Agricultural and Food Chemistry, 29*(1), 11-15.

20. Raboy, V. (2020). Low phytic acid crops: Observations based on four decades of research. *Plants, 9*(2), 140. doi:10.3390/plants9020140

21. Sarkar, A., & Kaul, P. (2014). Evaluation of tomato processing by-products: a comparative study in a pilot scale setup. *Journal of Food Process Engineering, 37*(3), 299-307. https://doi.org/10.1111/jfpe.12086

22. Maire du Poset, A., Lerbret, A., Boué, F., Zitolo, A., Assifaoui, A., & Cousin, F. (2019). Tuning the structure of galacturonate hydrogels: external gelation by Ca, Zn, or Fe cationic cross-linkers. *Biomacromolecules, 20*(7), 2864-2872. https://doi.org/10.1021/acs.biomac.9b00726

23. Zuorro, A., Lavecchia, R., Medici, F., Piga, L. (2013). Enzyme-assisted production of tomato seed oil enriched with lycopene from tomato pomace. *Food and Bioprocess Technology, 6*(12), 3499-3509. https://doi.org/10.1007/s11947-012-1003-6

24. Mechmeche, M., Kachouri, F., Chouabi, M., Ksontini, H., Setti, K., & Hamdi, M. (2017a). Optimization of extraction parameters of protein isolate from tomato seed using response surface methodology. *Food Analytical Methods, 10*(3), 809-819. https://doi.org/10.1007/s12161-016-0644-x

25. Meshginfar, N., Mahoonak, A. S., Hosseinian, F., & Tsopmo, A. (2019). Physicochemical, antioxidant, calcium binding, and angiotensin converting enzyme inhibitory properties of hydrolyzed tomato seed proteins. *Journal of Food Biochemistry, 43*(2), e12721. https://doi.org/10.1111/jfbc.12721

26. Sogi, D. S., Arora, M. S., Garg, S. K., & Bawa, A. S. (2002a). Fractionation and electrophoresis of tomato waste seed proteins. *Food Chemistry, 76*(4), 449-454. https://doi.org/10.1016/S0308-8146(01)00304-1

27. Sogi, D. S., Garg, S. K., & Bawa, A. S. (2002b). Functional properties of seed meals and protein concentrates from tomato-processing waste. *Journal of Food Science, 67*(8), 2997-3001. https://doi.org/10.1111/j.1365-2621.2002.tb08850.x

28. Latlief, S. J., & Knorr, D. (1983a). Effect of chitin as coagulating aid on protein yield, composition and functionality of tomato seed protein concentrates. *Journal of Food Science, 48*(6), 1587-1590. https://doi.org/10.1111/j.1365-2621.1983.tb05037.x

29. Salikhova, S. R., Piyakina, G. A., & Yunusov, T. S. (1999). Proteins isolated from tomato seeds. *Chemistry of Natural Compounds, 35*(5), 572-574. https://doi.org/10.1016/j.foodhyd.2016.01.014

30. Sheoran, I. S., Olson, D. J. H., Ross, A. R. S., & Sawhney, V. K. (2005). Proteome analysis of embryo and endosperm from germinating tomato seeds. *Proteomics, 5*(14), 3752–3764. https://doi.org/10.1002/pmic.200401209

31. Piyakina, G. A., Maksimov, V. V., & Yunusov, T. S. (1998). Some properties of the protein fractions of tomato seeds. *Chemistry of Natural Compounds, 34*(4), 492-495. https://doi.org/10.1007/BF02329603
### Tables

#### Table 1 Carbohydrases and their main activities, and optimum temperatures and pH ranges

| Enzyme   | Activities                                | Operating Temperature | Operating pH  |
|----------|-------------------------------------------|-----------------------|---------------|
| Tail 157 | Pectinase / hemicellulase                 | 50 – 60°C             | 4.0 – 5.0     |
| Tail 113 | Pectinase / arabinase / hemicellulase / cellulase | 45 – 55°C             | 4.0 – 5.0     |
| Tail 01  | Pectinase / hemicellulase                 | 45 – 55°C             | 3.5 – 5.5     |
| Cellux 01L | Cellulase / β-glucanase                  | 50 – 60°C             | 4.5 – 6.0     |
| Filta 02L | Cellulase / xylananse / B-glucanase      | 40 – 65°C             | 4.0 – 6.5     |

#### Table 2 Composition of tomato components

| Component            | Juice | Pomace     | Seeds          |
|----------------------|-------|------------|----------------|
| Dry weight content   | 9.3   | 6.7        | 94.3           |
| Proportion in dry weight | 60.5  | 27.4       | 12.0           |
| Non-fibre            | ND    | 46.7 ± 0.4 | 53.1 ± 1.23    |
| Hemicellulose        | ND    | 5.4 ± 0.4  | 18.6 ± 1.0     |
| Cellulose            | ND    | 13.8 ± 1.6 | 18.4 ± 0.85    |
| Lignin               | ND    | 25.3 ± 1.5 | 9.5 ± 0.4      |
| Ash                  | ND    | 0.9        | 0.5            |
| Protein (Bradford)   | 3.0 ± 0.8 | 2.0 ± 0.03 | 10.0 ± 0.9     |
| Protein (Kjeldahl)   | ND    | 7.6 ± 0.2  | 27.4 ± 0.2     |
| Free phosphate       | ND    | ND         | 9.20x10^{-3} ± 1.07x10^{-4} |
| Phytic acid          | ND    | ND         | 0.29 ± 0.04    |

Table 3 is not available with this version

#### Table 4 The percentages of each fraction in terms of dry weight and protein contents determined using different methods
### Figures

#### Figure 1

The deseeding process for tomatoes
Figure 2

Enzyme mediated extraction of (A) protein from tomato seeds and (B) protein from pulp, compared to control without enzymes (none)
Figure 3

Protein extraction from tomato seeds using alkali compared with sequential extraction in water followed by precipitation of insoluble material at different centrifugal forces (based on 1 g as starting material).

Figure 4

Fibre analysis of tomato seed and after two sequential aqueous extractions.
**Figure 5**

SDS-PAGE gels of tomato seed protein extract under denaturing conditions with dithiothreitol (+DTT)

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Grahicalabstract.png