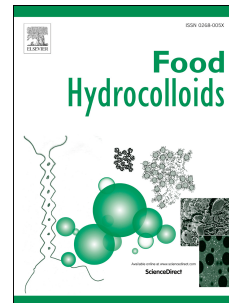# Journal Pre-proof

Production of soy protein concentrate with the recovery of bioactive compounds: From destruction to valorization

Ezequiel R. Coscueta, Luciana Pellegrini-Malpiedi, Maria Manuela Pintado, Bibiana B. Nerli

Please cite this article as: Coscueta, E.R., Pellegrini-Malpiedi, L., Pintado, M.M., Nerli, B.B., Production of soy protein concentrate with the recovery of bioactive compounds: From destruction to valorization, *Food Hydrocolloids* (2022), doi: https://doi.org/10.1016/j.foodhyd.2022.108314.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
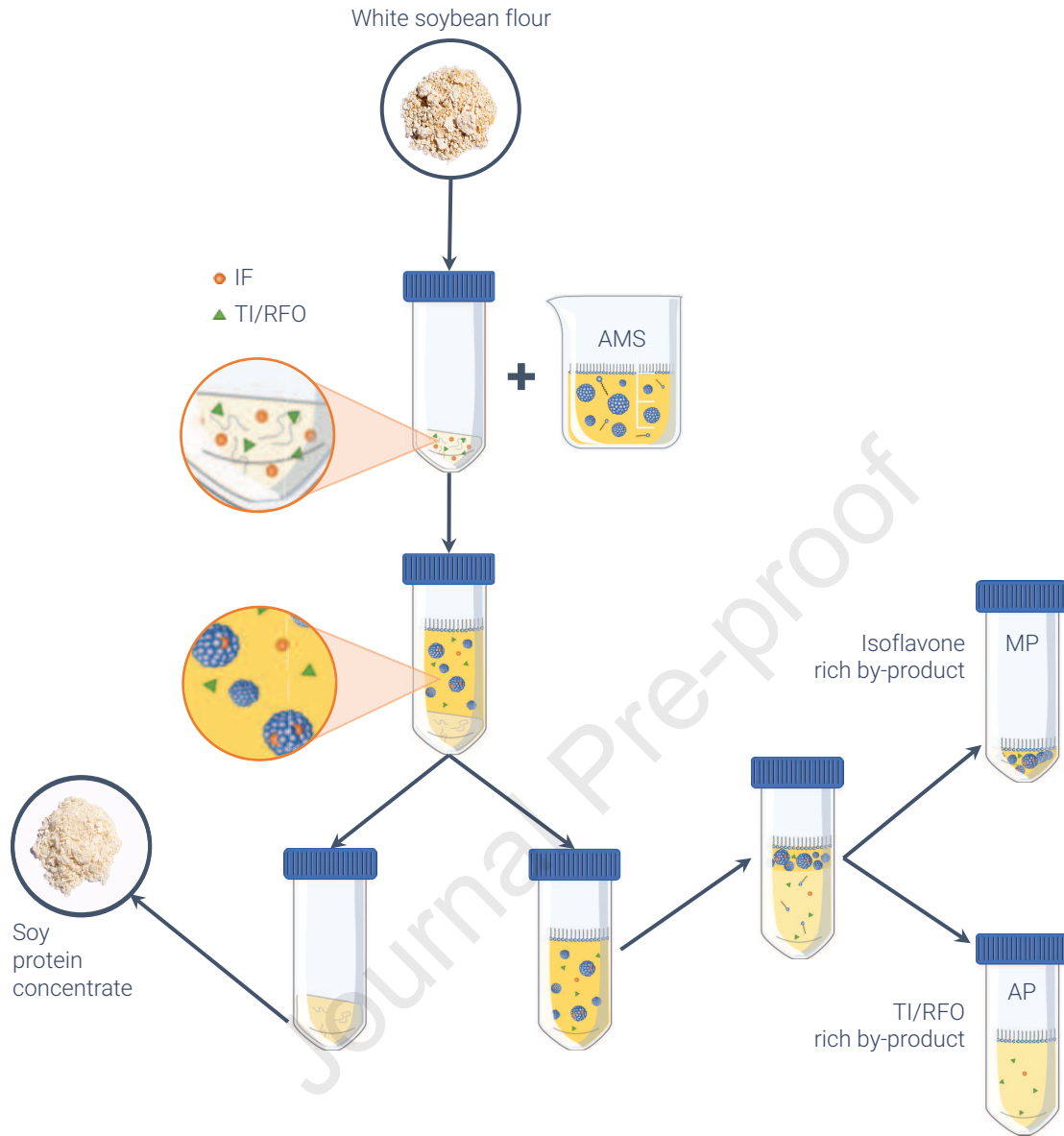
White soybean flour

- IF
- TI/RFO

AMS

+

Isoflavone
rich by-product

MP

Soy
protein
concentrate

TI/RFO
rich by-product

AP

1

2 Research paper

Production of soy protein concentrate with the recovery of

bioactive compounds: from destruction to valorization

5 Ezequiel R. Coscueta*[1,2], Luciana Pellegrini-Malpiedi[1], Maria Manuela Pintado[2] and

6 Bibiana B. Nerli[1]

7

8 [1] IPROBYQ (Instituto de Procesos Biotecnológicos y Químicos), UNR, CONICET,

9 Facultad de Ciencias Bioquímicas y Farmacéuticas (FCByF), Mitre 1998, 2000 Rosario,

10 Argentina.

11 [2] Universidade Católica Portuguesa, CBQF - Centro de Biotecnologia e Química Fina –

12 Laboratório Associado, Escola Superior de Biotecnologia, Rua Diogo Botelho 1327,

13 4169-005 Porto, Portugal.

14

15 *Correspondence:

16 Dr. Ezequiel R. Coscueta

17 E-mail: ecoscueta@ucp.pt

18 Tel.: +351 225 580 001 Ext 8047

19 Universidade Católica Portuguesa, CBQF - Centro de Biotecnologia e Química Fina –

20 Laboratório Associado, Escola Superior de Biotecnologia, Rua Diogo Botelho 1327,

21 4169-005 Porto, Portugal

22

Abstract

This work aimed to develop a novel methodology based on aqueous micellar systems (AMS), for producing soy protein concentrates (SPC) from soybean flour and recovering high-valuable bioactive compounds as by-products. Ethoxylated aliphatic alcohols Tergitol 15-S-7 and Tergitol 15-S-9, non-toxic and biodegradable surfactants, were selected to form the AMS. The methodology consisted of an extractive stage of soybean flour with AMS, which rendered both a pellet, i.e., the SPC, and a supernatant containing the extracted bioactive compounds. The latter was further heated above the cloud point temperature, thus resulting in a biphasic system formed by a micelle-rich phase (MP) and an aqueous phase (AP). Obtained SPC showed a noticeable loss (~90%) of trypsin inhibitor activity, a total protein content close to 60%, soluble protein amounts varying from 19% to 34%, and remarkable released (by simulated digestion) antioxidant and antihypertensive activities. Those indicators are similar to or even better than those corresponding to SPC from the classical acid-extraction method. The AMS also exhibited an enhanced efficiency for extracting antinutrients such as non-digestible oligosaccharides, trypsin inhibitors, and lectins mostly recovered at the AP and separated from isoflavones, which were concentrated and isolated at the MP. The recovery of all the mentioned bioactive compounds, whether beneficial or undesirable, broadens their uses in research, food, and pharmacological fields.

This successful performance, simplicity, scalability, and sustainability make the proposed AMS-based extraction a powerful tool for processing plant derivatives and valorizing their by-products.

48   **Keywords:** aqueous micellar two-phase system; isoflavones; trypsin inhibitors;

49   galactooligosaccharides; antioxidant activity; antihypertensive activity

50

1. Introduction

52   Soybean [*Glycine max* (L.) Merr.] is the most important legume crop

53   produced and consumed globally (Day, 2013; Jia et al., 2020). This legume and

54   its by-products are considered one of the primary alternative protein sources for

55   animal and human consumption (35-40%). Soybean has not only become an

56   increasingly popular food. However, it has also attracted much interest because of

57   the positive effect that its high intake produces on health, in particular in Asian

58   populations (Kulling et al., 2001). Soy-based foods also contain a wide range of

59   biologically active secondary metabolites, i.e., bioactive compounds, which can

60   confer either beneficial or undesirable effects. Among the former, antioxidant and

61   antihypertensive bioactivities have been widely reported (Balisteiro et al., 2013;

62   Coscueta et al., 2016; González-Montoya et al., 2016). Consumption of this

63   legume may reduce the risk of chronic diseases, such as cardiovascular diseases

64   and cancer, as well as reduce the risk of osteoporosis and relieve the symptoms of

65   menopause (Messina, 2014; Xiao et al., 2012). Phytochemicals responsible for

66   such protective activities include saponins, phytates, protease inhibitors, phenolic

67   acids, isoflavones, lecithin, lectin, and bioactive peptides (de Mejia et al., 2003;

68   C. C. Lee et al., 2017; Lule et al., 2015; Xu et al., 2015). However, some declared

69   beneficial activities, at specific doses and conditions, also have their dark side.

70   Soybean contains various antinutritional bioactive compounds that exhibit

71   undesirable physiological effects, such as preventing the absorption of nutrients.

72   The main antinutritional compounds comprise lipoxygenase, trypsin inhibitors,

73 lectin, and others in minor quantities, such as tannins, non-digestible

74 oligosaccharides, saponins, alkaloids, phenolic compounds, and phytates. Trypsin

75 inhibitors are the most critical components of antinutritional factors responsible

76 for growth retardation and digestive and metabolic diseases (Boisen & Eggum,

77 1991; Gatel, 1994). Therefore, the inactivation of trypsin inhibitors becomes a

78 requirement to improve the absorption of soy proteins in the digestive tract, thus

79 representing a challenge for the food industry and research development (Liu,

80 1997). At present, "detoxification," i.e., the inactivation of soybean antinutrients,

81 is carried out through various processes (Akande & Fabiyi, 2010; Newkirk,

82 2010). Denaturing the thermolabile antinutrients (trypsin inhibitors and lectins)

83 by heat is the most widespread classical inactivation process. It causes the loss of

84 their activity and results in improving protein digestibility. However, this process

85 also leads to the indiscriminate destruction of other essential nutrients and

86 bioactive molecules in the legume (Agrahar-Murugkar & Jha, 2010; Y. Chen,

87 2015; Jasti et al., 2015; Murugkar, 2015). Roasting in a rotary drum dryer or

88 conventional grain dryer (temperatures vary between 110-170 °C) can reduce

89 trypsin inhibitors up to 85%, while conventional drying with hot air at 100 °C for

90 two hours reduces trypsin inhibitors activity by 80% (Agrahar-Murugkar & Jha,

91 2010; Carvalho et al., 2013; Stewart et al., 2003). Other detoxification

92 methodologies include extrusion, micronization, sterilization by autoclaving,

93 dielectric thermal treatment technology, infrared, and enzymatic chemical

94 treatment. However, they are not applicable on a macro scale due to their high

95 cost (Vagadia et al., 2017). Besides, all the processes mentioned are destructive;

96 this deprives the opportunity to valorize those antinutrients that present beneficial

97 properties under certain conditions (Dang & Van Damme, 2015; Gomes et al.,

98   2011). As an initial step, extracting and isolating those compounds is necessary.

99   That would allow their toxicological and clinical evaluation application and their

100   subsequent commercialization as supplemental ingredients.

101   Soybean flour, which is obtained from grinding dehulled soybeans, is one

102   of the most widely marketed protein-rich food ingredients. Subjecting soybean

103   flour to a washing process with appropriate extractive liquids allows for obtaining

104   a soy protein concentrate (SPC). That is an ingredient with higher protein

105   content/quality and lower content of bioactive compounds, e.g., antinutrients,

106   which leach out in the washing liquids. This non-destructive detoxification

107   process allows for recovering bioactive compounds; however, it is not yet applied

108   for that purpose due to the lack of technologies capable of separating the different

109   phytochemicals remaining in the extracts in a viable and sustainable way.

110   Traditionally, SPC is obtained from defatted soybean flour by precipitating

111   proteins and discarding soluble sugar and minor constituents with an alcohol-

112   water mixture or a diluted acid solution in the pH range of 4.0-4.8 (Erickson,

113   1995). The yields of SPC for these conventional processes have been reported to

114   vary between 60-70% concerning the flour protein content (Erickson, 1995).

115   Thus, approximately two-thirds of the protein content of soy flour is recovered as

116   insoluble residues after the extraction. In contrast, the remaining one-third,

117   mainly containing protein antinutrients, is lost (Alibhai et al., 2006).

118   Recently, scientists have evaluated a wide range of new non-toxic, non-

119   flammable, and biodegradable solvents to develop sustainable and

120   environmentally friendly extraction methods (Bajkacz & Adamek, 2017). Certain

121   surfactants exhibit the mentioned properties and represent an economical

122   alternative to expensive and dangerous organic solvents. In aqueous media, they

123   form aggregates, i.e., micelles, capable of interacting with hydrophilic or

124   lipophilic molecules through hydrophobic, dipolar, and hydrogen bonding

125   interactions; thus, these AMS become useful for extractive purposes (Sharma et

126   al., 2015). Besides, they can separate into two phases, a micelle-poor one and a

127   micelle-rich one, when heated above a critical temperature (cloud point), thus

128   acquiring separating properties. AMS successfully recovered soybean

129   phytochemicals, such as isoflavones. A Genapol X-080 AMS performed

130   outstandingly in extracting daidzein from *Puerariae radix* (He et al., 2005).

131   Furthermore, previous works carried out by our team demonstrated that Triton X-

132   114, Genapol X-080, Tergitol 15-S-7 (Tg7), and Tergitol 15-S-9 (Tg9) AMS

133   were suitable for extraction of total soy isoflavones (Cordisco et al., 2016;

134   Coscueta et al., 2018). Genapol X-080 systems were suitable for extracting

135   antinutrients from soybean flour; however, a complete characterization (protein

136   content, antioxidant and antihypertensive activities) of the obtained SPC was not

137   performed (Haidar et al., 2018).

138   In this context, this work aimed to develop a soybean flour detoxifying

139   methodology, AMS-based, capable of producing protein concentrates with

140   similar or improved quality to those obtained by classical methods and recovering

141   high-valuable bioactive compounds as by-products. Thus, we proposed AMS

142   formed by the biodegradable surfactants Tg7 and Tg9 as extractive systems since

143   they represent a sustainable alternative industrially applicable. We considered

144   antioxidant and antihypertensive activities, digestibility, and trypsin inhibitory

145   activity as key parameters, other than protein content, to define the final SPC

146   quality. Regarding the leached bioactive compounds, we have considered the

147   molecular distribution pattern of the extracted proteins as one of the extraction

148 efficiency indicators, together with the recovery and distribution coefficients of

149 non-protein compounds (isoflavones and raffinose family oligosaccharides).

150 Finally, we compared the traditional methodology and the one proposed here

151 based on their advantages and disadvantages and provided future perspectives.

152

153 ## 2. Materials and methods

154 ### 2.1. Materials

155 The surfactants Tergitol 15-S-7 (Tg7) and Tergitol 15-S-9 (Tg9) were

156 supplied from Sigma-Aldrich (St. Louis, MO, USA). White soybean flour (i.e.,

157 non-thermal treated soybean flour) was supplied by the food processing company

158 Molinos Río de la Plata SA (San Lorenzo, Argentina). Crystallized salt-free

159 bovine trypsin, porcine pepsin (800-1000 U mg^{-1} protein), pancreatin (4xUSP),

160 angiotensin-I converting enzyme (peptidyl-dipeptidase A, EC 3.4.15.1, 5.1 U

161 mg^{-1}), crystalline α-N-benzoyl-DL-arginine-p-nitroanilide (BAPNA), bile salts

162 and Trolox (6-hydroxy-2,5,7,8-tetramethyl-chroman-2-carboxylic acid) were

163 obtained from Sigma-Aldrich (St. Louis, MO, U.S.A.) and used without further

164 purification. Fluorescein [3',6'-dihydroxyspiro (isobenzofuran-1[3H],9'[9H]-

165 xanten)-3-one] was purchased from Fisher Scientific (Hanover Park, IL). AAPH

166 [2,2'-azobis (2-amidi-nopropane) dihydrochloride] was purchased from Aldrich

167 (Milwaukee, WI). The tripeptide Abz-Gly-Phe(NO$_2$)-Pro was obtained from

168 Bachem Feinchemikailen (Bubendorf, Switzerland). Tris [tris (hydroxymethyl)

169 aminomethane] was obtained from Fluka (Gmbh, Germany). Isoflavone standards

170 (Sigma-Aldrich, St. Louis, MO, USA) were dissolved in pure methanol until

171 obtaining the following concentrations: daidzin 1.5-80.0 µg mL^{-1}; daidzein 0.3-

172   17.0 µg mL^{-1}; genistin 0.7-40.0 µg mL^{-1} and genistein 0.2-13.0 µg mL^{-1}. All the

173   other reagents were of analytical grade and used without further purification.

174   Deionized water was used to prepare all the solutions.

175   **2.2. Methods**

176   White soybean flour (henceforth called soybean flour) was treated following

177   traditional and AMS-based extraction processes. Samples from the resulting

178   supernatants, treated solids, permeates, and extracts were subjected to different

179   analytical procedures to determine performance indicators. Simultaneously, non-treated

180   soybean flour was also evaluated for comparison purposes. Figure 1, provided for a

181   more precise understanding, shows a schematic diagram of all the developed processes

182   and techniques whose descriptions are given below.

183   *2.2.1. Production of soy protein concentrates*

184   Soy protein concentrates (SPC) were prepared according to the classical

185   acid-washing method (Sair, 1959). Briefly, a suspension of 20.00% m/V soybean

186   flour and 25.00 mM citrate buffer pH 4.5 (3.00 g of flour in 60.00 mL of buffer)

187   was incubated in a bath at 40 °C and continuously stirred (150 rpm) for 45 min.

188   Then, the system was centrifuged (4000 rpm) at 15 °C for 15 min, thus obtaining

189   two fractions: a supernatant and an insoluble fraction, i.e., treated solid (SPCa).

190   SPC were also obtained by applying AMS as extractive solvents. AMS

191   formed by two non-ionic surfactants, Tg7 and Tg9, were evaluated. Appropriate

192   amounts of surfactants were dissolved in 50.00 mM sodium citrate (NaCit) at pH

193   4.5 until a final concentration of 5% m/m. This surfactant concentration was

194   selected according to previous optimization tests (Coscueta et al., 2022). Higher

195   concentrations were discarded to avoid high viscosity conditions representing a

196   technical disadvantage. Each AMS was prepared by mixing 3.00 g of a given

197   surfactant (Tg7 or Tg9) with 50.00 mM NaCit buffer pH 4.5 until a final system

198   volume of e 60.00 mL. This mixture and a sample of soybean flour (3.00 g) were

199   incubated separately in a thermostatic bath for 20 min to reach the extraction

200   temperature (45 °C). After the incubation, both components (flour and surfactant

201   solution) were placed into an Erlenmeyer, mixed, and stirred for 45 min inside

202   the thermostatic bath, maintaining the temperature at 45 °C. Then, the system was

203   centrifuged (4000 rpm) at room temperature for 10 min, thus obtaining a pellet

204   (SPC) and a supernatant, separated by decantation. The supernatant was

205   incubated again in a thermostatic bath and allowed to stand until total phase

206   separation (approximately 15 min). A phase of low micelle concentration, i.e.,

207   aqueous phase (AP) and a phase rich in micelles (MP), were obtained by

208   decanting. The phase separation temperatures, 45 °C for Tg7 and 60 °C for Tg9,

209   were selected for each surfactant according to the phase diagram previously

210   determined. These temperatures were a few degrees above the respective cloud

211   point to obtain similar volumes of MP. Experiments were performed in triplicate

212   for each surfactant.

213   All the resulting SPC (from traditional and AMS-based extractions) were

214   neutralized (final pH 6.5) by adding 20.00 mL of 25.00 mM phosphate buffer

215   solution at pH 7.0 and 0.50 mL of 1.00 M NaOH solution and then lyophilized.

216   Finally, the Kjeldahl reference method analyzed the total protein content of each

217   lyophilized SPC (SPCa, SPC Tg7, SPC Tg9).

218   **2.2.2. FTIR spectrometry**

219   Fourier transform infrared (FTIR) spectrometry characterized the soybean

220   flour and the SPC. The spectra were acquired by averaging 30 measurements at

221   wavenumbers from 500 to 4000 cm^{-1} with a resolution of 4 cm^{-1}. The spectra

222   were normalized for the maximum absorbance. The amide-I band's second

223   derivative was used to identify the different spectral components of soybean flour

224   and SPC. The subsequent spectral deconvolution was performed by applying a

225   Gaussian fitting (Arrondo et al., 1993; Zana et al., 1998). Measurements were

226   carried out in an infrared spectrometer, model ABB MB3000 (ABB,

227   Switzerland), equipped with a deuterated triglycine sulfate detector and provided

228   with a horizontal reflection accessory MIRacle^{TM} (PIKE Technologies, USA), for

229   attenuated total reflectance, with a diamond crystal plate/Se.

230   *2.2.3. Extraction of trypsin inhibitors and soluble protein*

231   Reference methods globally accepted for estimating trypsin inhibitory

232   activity in soybean-derived products share the same principle: trypsin inhibitors

233   are extracted from a weighed sample under the alkaline condition at which soy

234   protein solubility is enhanced (Liu, 2021). In this work, the extraction/lixiviation

235   was carried out by mixing 1.00 g of soybean flour/SPC with 50.00 mL (V_E) of

236   10.00 mM NaOH and stirring at room temperature for three hours, according to

237   the standard method proposed by Kakade et al. (1974) and later improved

238   (AOCS, 2009; Kakade, 1974). A final centrifugation step (3500 rpm) for 10 min

239   was carried out to recover the supernatant, which was conveniently diluted (with

240   a dilution factor F_D) to determine the trypsin inhibitory activity (TIA) and soluble

241   protein content.

242   *2.2.4. Trypsin inhibitory activity (TIA)*

243   TIA was measured according to the modified methodology proposed by

244   Coscueta et al. (2017). The procedure was adapted to reduce the working

245   volumes and develop the assay in a 96-well microplate, as shown in Table 1. The

246   enzymatic reaction progress was monitored by absorbance measurements (400

247   nm) for 2 min. All the determinations were performed on the Multiskan GO

248   (Thermo Fisher Scientific Corporation) spectrophotometer operated by the SkanIt

249   3.2 software (Thermo Fisher Scientific Corporation). The TIA was calculated as

250   follows:

$$TIA' = 100 \; \frac{0.350 \; (m_{control} - m_{sample})}{0.040} \; F_D V_E \qquad (1)$$

251   where 100 is a conversion factor (to convert 0.01 u. Abs in trypsin inhibition

252   units); $m_{control}$-m_{sample}, the difference between the slopes corresponding to the

253   absorbance vs. time curves, in the absence (control) and presence of trypsin

254   inhibitors (sample); F_D, the dilution factor of supernatant from extraction; V_E is

255   the extraction volume of 0.01 M NaOH solution (50 mL, see section 2.2.3), used

256   for 1 g of soybean flour; 0.040, the aliquot (mL) of supernatant dilution used in

257   the continuous assay; and 0.350, the final reaction volume (mL) in the microplate

258   well. Finally, the TIA of a given sample was expressed as a percentage of the

259   total TIA present in soybean flour (TIA=TIA'$_{sample}$/TIA'$_{soybean\ flour}$ x 100) to

260   facilitate the comparison of extractive efficiencies exhibited in the different

261   treatments.

262 **Table 1.** Continuous method to determine trypsin inhibitory activity adapted for use in a

263 microplate reader.

Reagents	Control[†]	Sample[†]
Tris buffer 0.050 M, pH 8.20	140	100
Trypsin working solution	70	70
Diluted supernatant	---	40
Incubate at 37 °C inside the microplate reader for 2 min, then add		
BAPNA working solution (pre-heated at 37 °C)	140	140

[†]Volumes expressed in µL.

264

265 ### 2.2.5. *Determination of soluble protein content*

266 We determined the soluble protein content by the bicinchoninic acid

267 method (Smith et al., 1985), adapted for use in a 96-well microplate. This method

268 measured the protein content in extracts from alkaline lixiviation of SPC and

269 permeates from *in vitro* gastrointestinal simulation (see section 2.2.6). A stock

270 bicinchoninic acid (BCA) solution was prepared with this composition: BCA

271 1.00% (m/V), sodium tartrate 0.16% (m/V), Na_2CO_3 2.00% (m/V), NaOH 0.40%

272 (m/V), and $NaHCO_3$ 0.95% (m/V), the final pH being 11.2. A stock $CuSO_4$

273 solution of 4.00% (m/V) was also prepared. The working BCA reagent was

274 prepared by mixing the BCA stock solution (diluted 1:10) with the stock $CuSO_4$

275 solution to a 50:1 ratio. A calibration curve made with bovine serum albumin

276 (BSA) with concentrations of 50-1000 µg mL^{-1} was used. The protocol was

277 carried out by placing 25 µL of sample in each well and then adding 200 µL of

278 working BCA reagent simultaneously in all the wells. The microplate was

279 incubated at 37 °C for 30 min before obtaining the absorbance values at 562 nm.

280 Incubation and readings were performed on the Multiskan GO (Thermo Fisher

281   Scientific Corporation) spectrophotometer operated by the SkanIt 3.2 software

282   (Thermo Fisher Scientific Corporation).

283   *2.2.6. In vitro simulated gastrointestinal digestion*

284   The *in vitro* simulated gastrointestinal digestion (SGI) was carried out for

285   all SPC and soybean flour. The digestive process in the mouth and esophagus,

286   where carbohydrates are mainly affected, was not simulated since this work

287   focused on the digestion of proteins and phenolic compounds. Replicates (300

288   mg) from each applied methodology were adequately pooled and homogenized

289   into one sample. A given amount of the pooled sample (280 mg) was mixed with

290   6.00 mL of acidified water (pH 2.0). The pH was adjusted to 2.0 with 1.00 N

291   HCl, and the final volume was completed to 7.00 mL with the same acidified

292   water. The mixture was incubated at 37 °C and shaken at 130 rpm for 20 minutes

293   to temper the digestive process. Before starting the SGI, 1.00 mL of supernatant,

294   representing the sample at the initial time (T1), was removed from each

295   experiment. The initial stage of the SGI began with the stomach digestion step,

296   the gastric juice being simulated with pepsin 25 mg mL^{-1}, prepared in 0.10 N HCl

297   (Aura, 2005). 0.3 mL of this "gastric juice" was added, left at 37 °C, and shaken

298   at 130 rpm for 60 min. Then, the gastric stage was terminated by increasing the

299   pH to 6.5 with 100.00 mM NaHCO$_3$ solution. An aliquot of 1.00 mL was then

300   withdrawn, thus representing the sample corresponding to the end of the

301   mentioned stage (T2). For the intestinal step, pancreatic juices were simulated

302   with a solution of pancreatin 2 mg mL^{-1} and bile salts 12 mg mL^{-1} diluted in a

303   solution of 100.00 mM NaHCO$_3$ (Laurent et al., 2007). 1.50 mL of pancreatic

304   solution was added to the system. The temperature was returned to 37 °C and

305   stirring decreased to 45 rpm. This stage was extended for 90 min (T3) and then

306   stopped by freezing at -30 °C. Three independent SGI experiments were

307   performed for each pool. Finally, all the samples from the SGI process were

308   nano-filtered in 3 kDa pore membranes (Amicon® Ultra-4, Millipore), the

309   resulting permeates (T1, T2, and T3) being subsequently analyzed (bioactive

310   peptides and phenolic compounds). Enzymatic solutions were freshly prepared

311   and sterilized by filtration with 0.22 μm membrane filters (Millipore, Billerica,

312   MA, USA). After being sterilized, the solutions were kept in an ice bath to avoid

313   enzymatic self-degradation. A thermostatic water bath at 37 °C was used to

314   simulate the physiological temperature of the human body. Mechanical agitation

315   (parallel peristaltic movements) was implemented, with intensities emulating

316   those reached in each digestive compartment.

2.2.7. Antioxidant activity

318   The oxygen radical scavenging capacity (ORAC) was determined by the

319   method proposed by Coscueta, Brassesco, & Pintado (2021). The antioxidant activity of

320   each sample (permeates T1, T2, and T3) was evaluated in triplicate and expressed

321   as μmol Trolox Equivalent per gram of initial digested sample (SPC or soybean

322   flour) on a dry basis.

2.2.8. ACE inhibitory activity

324   The ACE inhibitory activity was carried out using the fluorometric assay

325   described by Coscueta et al. (2021). The inhibitory activity on ACE (iACE) of

326   each sample (permeates T1, T2, and T3) was evaluated in triplicate and expressed

327   as the concentration capable of inhibiting 50% of the enzymatic activity (IC_{50}). A

328   non-linear model calculated the IC_{50} values. The results were expressed as μL

329   mL^{-1} to inhibit 50% of the enzymatic activity.

14

2.2.9. Size exclusion chromatography

The molecular mass distribution of the protein components in the supernatant from traditional extraction and the aqueous phase (AP) from AMS-based extraction was analyzed by size exclusion chromatography. The chromatographic runs were performed at a flow of 0.5 mL min^{-1} with 25 mM phosphate buffer (pH 7) containing 150 mM NaCl and 0.2 g L^{-1} NaN$_3$. Standard proteins with known molecular masses (Thyroglobulin, 669 kDa; Ferritin, 440 kDa; Aldolase, 158 kDa; Conalbumin, 75 kDa; Ovalbumin, 43 kDa; Carbonic anhydrase, 29 kDa; Ribonuclease A, 13.7 kDa; Aprotin, 6, 5 kDa) were used to calibrate the system. The AKTA pure 25 L system (GE Healthcare Life Sciences, Freiburg, Germany) was used in a configuration consisting of two high-performance piston pump systems, a pressure monitoring system for column protection, a mixing chamber, a V9-IA injection valve, a Superdex® 200 10/300 GL column connected in series to a Superdex Peptide 10/300 GL column (GE Healthcare Life Sciences, Freiburg, Germany), and a length U9-L UV detector fixed wave at 280 nm. The system was controlled by UNICORN software.

2.2.10. Determination of isoflavones by HPLC

The identification and quantification of extracted isoflavones on supernatant from traditional extraction and both phases (AP and MP) of micellar systems were carried out using high-performance liquid chromatography (HPLC). Initially, each analyzed sample was conveniently diluted (1:3 with pure MetOH) and filtered on a 0.45 μm filter. Chromatography was carried out on a reversed-phase column (COSMOSIL 5C18-AR-II Packed Column - 4.6 mm D.I. x 250 mm) with two mobile phases. Mobile phase A consisted of ultra-pure water

15

354 acidified with glacial acetic acid (0.1% V/V), while phase B was constituted by

355 acetonitrile acidified with glacial acetic acid (0.1 % V/V). The operation involved

356 injecting 20 μL of sample and eluting it with a linear gradient from 80 to 0% of

357 phase A over 25 min, at a constant flow of 0.8 mL min^{-1} and a column

358 temperature of 25 °C. Each isoflavone was identified by considering the

359 absorbance spectrum and the retention time of standards. The concentration of

360 each isoflavone was determined by measuring the area under the peak and

361 interpolating it on the corresponding calibration curve. The results were

362 expressed in μg of isoflavone per gram of soybean flour on a dry basis. The

363 Waters e2695 modular separation system was used, with a UV/Vis photodiode

364 array detector (PDA 190-600 nm). The acquisition of the data and the analysis

365 were carried out using the Empower 3 software.

366 *2.2.11. Determination of non-digestible oligosaccharides by HPLC*

367 The content of raffinose family oligosaccharides in the supernatant from

368 traditional extraction and both phases (AP and MP) of micellar systems was

369 determined by HPLC. Aliquots (30 μL) of the undiluted liquid samples were

370 analyzed by chromatography. The mobile phase was a 13.00 mM H_2SO_4 solution,

371 with a flow of 0.8 mL min^{-1}, at isocratic elution. HPLC quality standards of

372 raffinose, stachyose, glucose, and sucrose were used to identify the mentioned

373 compounds in the samples by considering their retention time. The peaks for

374 raffinose and stachyose were so close together that we had to measure them as

375 one, which we called raffinose family oligosaccharides. The values were

376 expressed in relative units (RU), given by the product between the integrated area

377 and the mass of liquid obtained for each fraction, referred to as 1 gram of soybean

378 flour (RU g^{-1}). An HPLC system prepared for the identification and quantification

379   of sugars was used, consisting of a Knauer WellChrom Pump K-1001 module

380   (Knauer GmbH, Germany) and a differential refractive index (RI) detector K-

381   2301 (Knauer GmbH, Germany). The separation was carried out on an Aminex

382   HPX-87H 300 x 7.8 mm column (Bio-Rad, Hercules, USA) coupled to a Micro-

383   Guard Cation H^+ precolumn (Bio-Rad, Hercules, USA), maintained at a

384   temperature of 42 °C by an Eldex CH-150 column oven (Eldex Laboratories,

385   Napa, CA). The data were acquired and analyzed using Clarity v.5.0.5.98

386   software (DataApex Ltd, Prague, Czech Republic).

387   *2.2.12. Extraction efficiency parameters*

388   The distribution coefficients (Kd) and phase recovery performance (Pr) in

389   the different AMS were determined for both isoflavones and raffinose family

390   oligosaccharides. The Kd was estimated as:

$$Kd = \frac{C_{MP}}{C_{AP}} \tag{2}$$

391   where C_{MP} and C_{AP} are the concentrations of the analyte (isoflavones or raffinose

392   family oligosaccharides) in the MP (micellar phase) and AP (aqueous phase),

393   respectively. On the other hand, the Pr was determined:

$$Pr(\%) = \frac{C_{MP/AP}V_{MP/AP}}{(C_{MP}V_{MP} + C_{AP}V_{AP})} \, 100 \tag{3}$$

394   where $C_{MP/AP}$ is the concentration of the analyte in either the MP or AP, selected

395   according to the phase of the highest recovery, these $V_{MP/AP}$ is the volume of the

396   corresponding phase. Equation (3) is only valid when analytes distribute entirely

397   between the two phases without precipitating at the interphase.

398   *2.2.13. Statistical analysis*

399   All the experiments were carried out in triplicate, except for specific cases

400   where something different was indicated, and the results were expressed as the

401   mean value with their standard deviation (SD). Before any comparative statistical

402   analysis, an exploratory study was carried out for all the data sets to determine

403   the compliance of normality and homoscedasticity. Then the means were

404   analyzed statistically by analysis of variance (ANOVA) followed by a posthoc

405   test using Tukey (Tukey, 1949). A significance level of 5% was considered. All

406   statistical analysis was carried out with the aid of RStudio V 1.0.143.

407   ## 3. Results and discussion

408   According to previous results (Coscueta et al., 2022), we designed a new AMS-

409   based process applied to soybean flour for producing SPC and recovering valuable

410   bioactive compounds as by-products. A graphical scheme and technical details of the

411   proposed methodology are presented in Figure 2 to facilitate reading this article and

412   interpreting the results.

413   Figure 2 shows an initial micellar extraction performed by applying AMS on

414   soybean flour. Then, a phase separation produced a MP and an AP by heating the

415   extract (supernatant of the first extraction) to a temperature higher than the cloud point.

416   It is important to note that the micellar extraction for Tg7 is considered a "cloud point"

417   extraction since the extraction temperature (45 °C) is above the Tg7 CP (39 °C). This is

418   not so for Tg9 since its cloud point (61 °C) is higher than the working temperature.

419   **3.1. Producing soy protein concentrates: characterization of the product**

420   *3.1.1. FTIR spectroscopy*

421   Although the literature has reported the non-ionic Tg7 and Tg9 as mild

422   surfactants, it was necessary to evaluate their effect on soy protein structure since

423   it might affect its functional properties (Nadar et al., 2017; Vicente et al., 2017).

424   The spectroscopic analysis of polymeric molecules is complex due to the

425   molecular vibrations that arise from numerous atoms. FTIR is an advantageous

426   technique for the study of protein systems. There are several easily identifiable

427   regions in the middle infrared spectrum, having nine characteristic bands called

428   amide-A, B, I, II ... VII; the amide-I, amide-II, and amide-III being the most

429   widespread ones used in structural studies of protein

430   Figure 3A shows the FTIR spectra for flour (starting material) and SPC

431   (products). The amide-A (3500-3200 cm^{-1}) and amide-B (3100-2500 cm^{-1}) bands come

432   from a Fermi resonance between the first harmonic of amide-II and the vibration of the

433   N-H stretch. Notably, the spectral band of amide-B showed an intense alteration after

434   the process, which indicates a conformational change in the secondary structure of the

435   matrix of the SPC (Chang & Tanaka, 2002). This band's intensity increased drastically for

436   the SPCa and the SPC Tg9 compared to the soybean flour. However, only slight

437   changes were noticeable for the SPC corresponding to Tg7. The amide-I and II bands

438   are the two central regions of the infrared spectrum for protein. The amide-I (between

439   1700 and 1600 cm^{-1}) is mainly associated with the stretching vibration C=O (70-85%)

440   and is directly related to the conformation of the main peptide chain (Chang & Tanaka,

441   2002). Amide-II (1600-1500 cm^{-1}) results from the N-H bending vibration (40-60%) and

442   the C-N stretch vibration (18-40%). When structural alterations such as denaturation or

443   aggregation occur, these two bands significantly decrease in intensity (Nishinari et al.,

444   2014). Figure 3B shows the highest and lowest amide-I and II signals for the flour and

445   SPCa, respectively. In contrast, it shows intermediate intensities for the SPCs obtained

446   by the micellar systems. The SPC Tg7 spectrum was like the flour, while the SPC Tg9

447   spectrum was higher than the SPCa. Amide-III (between 1300 and 1200 cm^{-1}) is a very

448   complex band that results from a mixture of several coordinate shifts; it does not exhibit

449   noticeable differences between SPCs and flour (Figure 3B).

450   Even if the primary protein structures were the same, the secondary and the

451   tertiary structures sometimes are not. The amide-B allowed us to analyze what

452   happened to the secondary structures of the proteins. The most interesting regions in the

453   FTIR spectra are the amide-I and II regions since changes in these regions were

454   observed for other proteins (Arrondo et al., 1993; Susi & Byler, 1986). To analyze the

455   secondary structures, we applied a deconvolution procedure of the amide-I band (Figure

456   3C-F) (Fabian et al., 1993). The second derivative curve allowed the identification of

457   the different spectral components used in a deconvolution procedure with Gaussian

458   curve fitting. Table 2 presents the contribution percentages of the various secondary

459   structure types, calculated as the ratios between the areas of the curves corresponding to

460   the different component bands and the total area below the spectral curve.

461   **Table 2.** Protein secondary structure in soybean flour and SPC by FTIR analysis.

Product	Secondary structure (%) [†]			
	α-helix	β-sheet	Random coil	β-turns
Soybean flour	62	38	0	0
SPCa	42	44	0	14
SPC Tg7	12	24	43	20
SPC Tg9	67	28	0	4

[†] Structural composition of the amide-I, obtained from the area under each deconvoluted band and expressed as a percentage of the total area.

462

463   The data revealed that amide-I for all SPCs consists of nine or ten main

464   components, unlike flour, which exhibits only two. We found the main band of the

465   soybean flour and SPCa FTIR spectra in the amide-I region between 1660-1650 cm^{-1}

20

466 (62% and 42% of the total area, respectively, Table 2). This corresponds to the

467 vibrational movements of the amide residues of the main chain in a helical

468 conformation. The soybean flour also presented another large band between 1650 and

469 1600 cm-1 corresponding to the β-sheet structure, which is in the minority (38%).

470 Meanwhile, the SPCa also had smaller bands at 1637-1600 cm^{-1} and 1700-1682 cm^{-1},

471 related to β-sheet structures (44%), which slightly exceeded the content of α-helix

472 (42%). Besides, it presented two bands corresponding to β-turn structures (1674 cm^{-1}

473 and 1666 cm^{-1}). Both β-sheet and β-turn structures indicate a greater union of the amide

474 protons by forming hydrogen bonds (FAO & Berk, 1992). We did not observe a common

475 pattern in the FTIR spectra of the Tg7 and Tg9 SPCs (Table 2). The SPC Tg7 showed

476 the principal peak in the region corresponding to the unordered structure (1639-1638

477 cm^{-1}), thus representing the highest structural percentage (43%). The second main

478 structure was the β-sheet type, contributing about 24%. Finally, the SPC Tg9 had a

479 structural distribution like the flour, with 67% of α-helix and 28% of β-sheet.

480   The structural alterations suffered by the SPC protein are related to the process

481 conditions, such as temperature, surfactant concentration, and the medium's pH. It

482 should be noticed that those conditions may affect inter/intramolecular interactions

483 (formation/rupture of hydrogen bonds, hydrophobic effects) of protein molecules, thus

484 resulting in conformational changes. Notably, the applied pH (pH 4.5) corresponds to

485 the isoelectric point of most of the matrix proteins. This pH facilitates aggregation by

486 non-electrostatic forces, thus decreasing protein solubility in the final concentrate

487 (Derringer & Suich, 1980). We must not ignore that although the applied methodologies

488 allow concentrate proteins, they cause the loss of soluble carbohydrates and a

489 considerable amount of soluble protein. That enriches the final product in proteins

490 already insoluble in the starting soybean flour, either by extracting the oil with hexane

491 or the solvent removal step in the previous soybean processing. It should be noticed that

492 changes in the secondary structure of soy proteins caused by a given treatment might

493 affect not only their conformation but also functional properties such as gel

494 transparency, solubility, surface hydrophobicity, and emulsifying capacity (X. Chen et

495 al., 2013; Zhao, Chen, Chen, et al., 2008). Therefore, the information introduced here

496 could be critical for further understanding and correlating the SPC applications.

3.1.2. Protein content and trypsin inhibitory activity

498 The SPC produced by the different methodologies were analyzed and compared

499 with each other and the starting soybean flour. Initially, we analyzed the total protein

500 content by the Kjeldahl reference method (section 2.2.1). Then, we performed aqueous

501 extractions to evaluate the trypsin inhibitory activity (TIA) and the amount of soluble

502 protein by the bicinchoninic acid method (section 2.2.5). The results reported in Table 3

503 indicated that the total protein varied according to the type of surfactant used in the

504 proposed technique. SPC Tg9 presented a total protein value comparable to SPCa, while

505 SPC Tg7 showed a total protein like the starting soybean flour. When comparing the

506 surfactants, Tg7 was demonstrated to be less efficient in concentrating proteins than

507 Tg9, thus evidencing a higher protein extractive power.

508 **Table 3.** Total protein content, soluble protein, and trypsin inhibitory activity analysis for SPC
509 and soybean flour.

Product	Pt [†, 1]	TIA [†, 2]	Ps [†, 3]
Soybean flour	54.6 ± 2.1 [a]	100 ± 7	84 ± 4
SPCa	64.0 ± 1.5 [b]	11 ± 1 [a]	29 ± 4 [a, b]
SPC Tg7	57.4 ± 0.7 [a, c]	10 ± 2 [a]	34 ± 7 [b]
SPC Tg9	60.7 ± 2.1 [b, c]	12 ± 3 [a]	19 ± 4 [a]

† Values are expressed as mean ± SD on a dry basis. [1] Total protein content (Pt) expressed as g proteins (100 g product) [-1]. [2] Trypsin inhibitory activity (TIA) expressed as a percentage from the starting soybean flour. [3] Soluble protein content (Ps) expressed as g BSA (100 g total proteins) [-1]. [a-d] Values from the same column that share a superscript do not present statistical differences, given the analysis of their variances (ANOVA) with Tukey's posthoc test for a significance level of 0.05.

510

511   As shown in Table 3, the inhibitory activity of trypsin was considerably reduced,

512   around 90%, in all the SPC, thus reaching satisfactory levels that agreed with those

513   predicted in a previous optimization work (Coscueta et al., 2022). In this regard, none of

514   the methodologies differed significantly. However, this drastic reduction in the TIA

515   resulted in a significant decrease in the protein solubility in 10 mM NaOH. That is a

516   well-known disadvantage of protein concentration methods, i.e., the resulting SPC

517   present low protein solubility after rehydration and poor functional properties. Protein

518   exposure to extreme conditions may cause this low solubility (Alibhai et al., 2006;

519   Fisher et al., 1986). We mean alcohol or acid extraction, heat treatment, precipitation, or

520   centrifugation by extreme conditions. Besides, the reduction of protein solubility could

521   also be the result of the enrichment of the final product is already insoluble proteins

522   before the concentrating process, as we mentioned previously. Different treatments that

523   allow for significant solubility recovery can mitigate this disadvantage's impact

524   (Johnson, 1999). Also, particularly for the AMS-based extractions, the Tg9 led to the

525   most significant protein solubility loss.

526   *3.1.3. Digestibility and bioactivity*

527   To analyze the performance of different SPC at *in vitro* simulated

528   gastrointestinal digestion (SGI), we measured soluble protein, antioxidant activity, and

529   antihypertensive activity at each stage (T1, T2, and T3). We measured soluble protein

530   considering protein compounds with MW less than 3 kDa (enriched peptide fraction),

531 the antioxidant activity by ORAC, and the antihypertensive activity from the inhibitory

532 capacity on ACE (iACE). The response variable was recorded three times on the same

533 experimental unit (pooled sample); therefore, we considered the collected data "repeated

534 measures." So, we analyzed through the "Generalized Linear Models" procedure. Table

535 A1 (see Annex) reports the results of fitting the general linear statistical models. These

536 models relate the response variables soluble protein, antioxidant activity, and

537 antihypertensive activity with three categorical predictive factors: experimental units (I),

538 substrate (M: starting soybean flour or SPC from a given extractive methodology), and

539 SGI's stage (T). Relationships between the observed responses and the predictor

540 variables were statistically significant since, for each model, the P-value was less than

541 0.05. Besides, the determination coefficients (R^2) indicated that the adjusted models

542 explained between 99.6% and 99.7% of the respective variability.

543 For the general digestion process (Table A2 in Annex), irrespective of the

544 sample considered, the three responses adopted significantly different values according

545 to the stage of the SGI. Interestingly, T2 showed higher soluble protein than T3. This

546 finding is attributable to the analytical technique used for protein quantification, the

547 bicinchoninic acid method. That assay depends on the ability of proteins to reduce Cu^{+2}

548 to Cu^{+1} in an alkaline solution (biuret reaction), thus resulting in a purple product.

549 Cysteine, tyrosine, and tryptophan residues are mainly responsible for reducing copper.

550 However, unlike Coomassie dye-binding methods, the peptide skeletons also contribute

551 to color formation, helping to minimize the variability derived from differences in

552 protein composition. The peptide skeleton loss resulting from the release of peptides

553 during stomach digestion (T2) may cause the decreasing signal. These peptides can be

554 further degraded to tripeptides, dipeptides, and free amino acids in the intestinal stage

555 (T3). On the other hand, the ORAC values increased during the gastric phase and

24

556   practically kept constant in the intestinal phase. About iACE, the IC_{50} decreased

557   significantly at each stage of the process, thus indicating a sustained increase in the

558   inhibitory activity.

559   When considering the samples separately (Table A3 in Annex), we identified

560   three homogeneous groups concerning ORAC and two for soluble protein and iACE.

561   The mean soluble protein released during the digestion of the SPC Tg7 was the lowest

562   one, while the remaining samples did not show differences. Concerning ORAC, the

563   SPC Tg9 and the soybean flour did not differ significantly and surpassed the other

564   samples, the SPC Tg7 showing the lowest level of this bioactivity. On the other hand,

565   the maximum iACE was for the SPC Tg7 and the minimum for SPCa and SPC Tg9,

566   while the soybean flour did not differ from the mentioned groups.

567   Figure 4 shows the soluble protein content, ORAC, and iACE profiles as a

568   function of each digestive stage. The lines represent each I and the color, each level of

569   M. For soluble protein (Figure 4A), the SPCa presented a different behavior than the

570   others. As explained above, it showed a marked peptide release during the gastric phase

571   and a later decrease in the signal during the intestinal phase. That could be indicative of

572   more significant peptide degradation. The SPC Tg9 exhibited an increase in the peptide

573   content during the first stage, keeping it constant until the end of digestion. About the

574   soybean flour and SPC Tg7, the peptide release was almost continuous throughout the

575   digestive process. The results for ORAC (Figure 4B) differed, with SPC Tg7 and SPC

576   Tg9 being markedly different from the other samples. SPC Tg7 showed constant

577   antioxidant activity during the gastric phase, which increased during intestinal

578   digestion. At the same time, SPC Tg9 evidenced a drastic increase and a later reduction

579   in the gastric and intestinal stages, respectively. Unlike the other two responses, the

580   iACE pattern (Figure 4C) was similar for all samples throughout the SGI, with a greater

581   increase in inhibitory activity during the gastric phase.

582   At the end of the digestion, the soluble protein was similar (Table 4) for all the

583   products. At the same time, the antioxidant activity of both SPCa and SPC Tg7 was

584   practically 30% lower than that corresponding to the soybean flour and SPC Tg9.

585   Concerning iACE, all the products achieved high bioactivities.

586   **Table 4.** Multiple comparisons for each response by product.

Substrate	Means observed at the end of the SGI [†]		
	Ps [1]	ORAC [2]	iACE [3]
Soybean flour	8.5 ± 0.8 [a]	737 ± 51 [a]	47 ± 6 [a]
SPCa	7.2 ± 0.2 [a]	518 ± 20 [b]	53 ± 6 [a]
SPC Tg7	7.4 ± 0.8 [a]	437 ± 34 [b]	58 ± 4 [a]
SPC Tg9	8.8 ± 0.6 [a]	672 ± 65 [a]	49 ± 6 [a]

[†] Values expressed mean $\pm$ SD on a dry basis. [1] Soluble protein content (Ps) in mg BSA (g of substrate)$^{-1}$, [2] ORAC in µmol TE (g of substrate)$^{-1}$, [3] iACE as IC$_{50}$, µg mL^{-1}. [a, b] Values from the same column that share a superscript do not present statistical differences, given the analysis of their variances (ANOVA) with Tukey's posthoc test for a significance level of 0.05.

587   It is expectable that soybean flour with higher protein solubility will release a

588   greater number of peptides from its digestion and even higher antioxidant and

589   antihypertensive activities. However, we did not observe this behavior in this work,

590   probably due to its content of protease inhibitors. Protease inhibitors reduce proteolysis,

591   releasing a larger proportion of peptides with more than 20 amino acids during flour

592   digestion (Capriotti et al., 2015). These peptides probably play a nutritional role since

593   their size belongs to the range (3-51 amino acids) compatible with the absorption

594   through the intestinal epithelium. However, it should be noted that peptides that exceed

595   20 amino acids are known to exhibit minimal bioactivity properties (Roberts et al.,

596   1999). For this reason, we previously filtrated with 3 kDa pore membranes samples

597   used to determine ORAC and iACE bioactivities. Consequently, the retention of most

598　peptides released during the filtration step could cause the reduced bioactivity values

599　found for samples from soybean flour digestion.

600　**3.2. Recovering bioactive compounds: characterization of by-products**

601　*3.2.1. Protein characterization by size*

602　After applying both methodologies to prepare protein concentrates, we

603　characterized the protein size profile of the supernatant from the classical

604　extraction (acid extraction) and the AP from AMS-based extraction, presumably

605　rich in antinutritional factors. For that, we used size exclusion chromatography

606　(SEC).

607　We identified three main regions when analyzing the chromatograms

608　(Figure 5). The first region (I) corresponds to large proteins and polypeptides, the

609　second one (II) to medium/small peptides, while the third (III) contains free

610　amino acids and small molecules capable of absorbing at 280 nm. We focused

611　our principal interest on the region I since it corresponds to the protein

612　antinutritional factors in soybean flour. Even though they are unwanted

613　compounds for nutritional purposes, they have biological capacities in the health

614　area, making it interesting to recover them intact.

615　Within region I, peaks 1 and 2, with approximate average MW of ~690

616　kDa and ~296 kDa, respectively, may include protein aggregates and the native

617　quaternary structure of glycinin. Peak 3 (~120 kDa MW) may correspond to β-

618　conglycinin (native and denatured) and denatured glycinin. Peak 4 (~72 kDa

619　MW) may include subunits of β-conglycinin (α, α', and β) and acid subunits of

620　glycinin. On the other hand, peak 5 (~30 kDa MW) may comprise the basic

621 subunits of glycinin, lectin, and the Kunitz-type trypsin inhibitor. The Bowman-

622 Birk type trypsin inhibitor was in peak 6 (~13 kDa MW). The soybean's principal

623 antinutrients are lectin, the Kunitz, and Bowman-Birk trypsin inhibitors, all being

624 the extraction's target molecules. Finally, the peak 7 (~4 kDa MW) may contain a

625 peptide of high commercial value and currently growing interest, lunasin (Seber

626 et al., 2012; Serra et al., 2016; Singh et al., 2017). We fractionated both peaks 5

627 and 6 prior to analysis by mass spectrometry (MALDI-TOF/TOF), thus

628 confirming the presence of the lectin and the trypsin inhibitors.

629   Table A4 (see Annex) shows the areas calculated for each peak in region I

630 and the total area for regions II and III. Neither the micellar extraction with Tg9

631 nor the classical method extracted protein aggregates or glycinin, while the Tg7

632 extracted them. For peak 3, the AP Tg9 differed from the others showing the

633 lowest value, while for peak 4, it only differed from the AP Tg7. Notably, the

634 Tg7 system exhibited the highest leaching ability when the first four peaks were

635 analyzed globally. Concerning antinutritional factors, both the classic method and

636 the AP Tg7 showed to be the best ones to extract Kunitz trypsin inhibitors and

637 lectins. At the same time, the AP Tg9 exhibited the best performance in removing

638 the Bowman-Birk trypsin inhibitor.

639   Additionally, it is worth mentioning that Tg9 evidenced a marked

640 extractive capability for peak 7, which could contain the lunasin. This is a

641 remarkable feature given the high value that this peptide has acquired for its

642 proven beneficial properties for health, namely antioxidant, anti-inflammatory,

643 and chemopreventive (de Mejia et al., 2021; Dong et al., 1995; French et al.,

644 2004; Hao et al., 2020; Pivato et al., 2012; Zhao, Chen, Xue, et al., 2008).

645 Finally, from inspection of regions II and III, the AP Tg9 presented the highest

646   content of small peptides, free amino acids, and small molecules that absorb at

647   280 nm relative to the classic and Tg7 AP.

648   ### 3.2.2. *Extraction and purification of isoflavones*

649   The potential ability to extract and isolate isoflavones of high commercial

650   value is one of the proposed methodology's attractive properties. Table 5 reports

651   isoflavone amounts leached by each system, determined by HPLC. The AMS

652   differed markedly from the control (classical process), exhibiting a higher

653   extractive ability of total isoflavone content. A more detailed analysis showed

654   that the AMS Tg7 stood out from the classical method when extracting the

655   glycosylated forms (daidzin and genistin). However, for the aglycones, those two

656   systems were not different. On the other hand, Tg9 highly exceeded the control at

657   extracting daidzin. However, it led to achieving similar or even lower genistein

658   and daidzein extraction, respectively, probably due to its more hydrophilic nature.

659   Table 5 also reveals that all the extracts have different compositions of

660   isoflavones, with the glycosylated forms being the predominant ones in all cases.

661   These results agree with the higher content of β-glycosides concerning the

662   aglycones in soybean and derivatives. For example, it is well known that the

663   aglycones of soybean flour are approximately 2-3% of the total isoflavone

664   content, while the unconjugated β-glycosides represent 34% of the total

665   isoflavone (Andrade et al., 2016). Table 5 also includes the distribution

666   coefficients (Kd) of isoflavones between the micelle-rich/-poor phases (MP and

667   AP). All the Kd measured were >>1, thus confirming a marked isoflavones

668   preference for the MP. Table 5 also shows the phase recovery performance at the

669   micellar phase (Pr). The calculation of this efficiency parameter needed the phase

670   volumes in each AMS. The mean MP volumes of the Tg7 and Tg9 systems were

671   10.3 and 9.1 mL, respectively, while their corresponding mean AP volumes were

672   44.7 and 40.9 mL. The differences in the total volumes recovered were due to the

673   liquid retained by the solid in each case. Note that the reduced volume of MP

674   regarding that of AP (four times lower) and Kd values higher than 1 results in

675   increased isoflavone concentration in the MP. On the other hand, the high found

676   Pr values demonstrate the remarkable effectiveness of AMS at recovering the

677   isoflavones in the MP. Considering the extractive and concentrating properties,

678   the AMS Tg7 was the best system to recover all the isoflavones. It allowed not

679   only to extract more isoflavones than the traditional method but also to achieve

680   four to five times higher concentrations. This finding is technologically

681   promising since it involves a considerable reduction of volumes to be handled,

682   thus leading to industrially advantageous processes.

683   **Table 5.** Comparison of the amount of isoflavones extracted by the different methodologies.

Isoflavone			Methodology		
			Micellar extraction		Classic extraction
			Tg7	Tg9	
Aglycones [†]	Daidzein	Amount [1]	62.0 ± 3.6 [a, b]	61.1 ± 1,4 [a]	70.0 ± 3.1 [b]
		Kd [2]	23.8 ± 2.5	6.1 ± 0,6 [a]	NA
		Pr [3]	85.8 ± 3.1	57.6 ± 2.4	NA
	Genistein	Amount [1]	40.2 ± 2.0 [a]	30.3 ± 1.8 [a]	35.0 ± 2.4 [a]
		Kd [2]	14.3 ± 2.7 [a]	14.2 ± 2.8 [a]	NA
		Pr [3]	78.7 ± 6.5 [a]	78.2 ± 7.0 [a]	NA
Glycones [†]	Daidzin	Amount [1]	490.1 ± 20.3 [a]	513.9 ± 22.6 [a]	183.0 ± 15.8
		Kd [2]	4.2 ± 0.2	6.0 ± 0.3	NA
		Pr [3]	49.3 ± 1.1	57.2 ± 1.3	NA
	Genistin	Amount [1]	770.4 ± 21.4	461.2 ± 4.3 [a]	418.7 ± 15.3 [a]
		Kd [2]	7.4 ± 0.4	11.2 ± 1.1	NA
		Pr [3]	63.0 ± 1.3	71.4 ± 2.0	NA
Total [†]		Amount [1]	1362.8 ± 47.2	1066.5 ± 30.1	706.7 ± 36.7
		Kd [2]	6.4 ± 0.4	8.0 ± 0.7	NA

Pr [3]	59.5 ± 1.5 [a]	64.0 ± 1.9 [a]	NA

[†] Values are expressed as mean ± SD. [1] Amount in μg (g of soybean flour)$^{-1}$. [2] Distribution coefficient (Kd). [3] Phase rich in micelles recovery performance (Pr) expressed in %. [a, b] Values from the same row that share a superscript do not present statistical differences, given the analysis of their variances (ANOVA) with Tukey's posthoc test for a significance level of 0.05.

684

685   When analyzing our results and those corresponding to other
686   methodologies (Table 6), we found that the amounts of aglycones recovered with
687   the proposed processes are similar or superior to most of them. The extraction
688   with 70% ethanol achieved the highest amount of genistein and daidzein, but the
689   extracted matrix was another soy derivative, such as okara (Jankowiak et al.,
690   2014). When comparing the yields from ethanol and pure water extractions, water
691   alone cannot extract the less hydrophilic isoflavones. However, adding
692   surfactants can revert that (Cao et al., 2012). It is also important to note that direct
693   micelles improve extractive performance. AMS here proposed (Tg7/Tg9) and
694   Triton X-114 AMS (Cordisco et al., 2016) extracted more isoflavones (aglycones
695   and β-glycosides) than those obtained by reverse micelles (Zhao et al., 2010).
696   That suggests that the use of direct micelles improves extractive performance.
697   When MP concentrates the target compounds, the surfactant presence may
698   compromise the applicability of the final product. To solve this problem, in
699   previous work, our research group applied a second extractive step, i.e., a back
700   extraction, to make isoflavones displace from the AM to the AP (Cordisco et al.,
701   2016). It is worth mentioning that the decision to include this step in the process
702   will depend on the final use of isoflavones, i.e., as cosmetics or food ingredients.

703

704 **Table 6.** Soy isoflavones extracted with different methodologies.

Methodology	Source	Amount [†]				Reference
		Daidzein	Genistein	Daidzin	Genistin	
AMS Tg7	Flour	62.00 ± 3.60 [a]	40.20 ± 2.00 [a]	490 ± 20 [a]	770.40 ± 21.40 [a]	**
AMS Tg9	Flour	61.10 ± 1.40 [a]	30.30 ± 1.80 [b, c]	514 ± 23 [a]	461.20 ± 4.30 [b]	**
AMS Triton X-114	Flour	142.60 ± 2.60	69.60 ± 0.20	350 ± 8 [b]	736.50 ± 3.30 [a]	(Cordisco et al., 2016)
Water/acetone/ethanol	Flour	NA	38.00 ± 0.01 [a]	382 ± 12 [b]	467.00 ± 8.00 [b]	(Andrade et al., 2016)
Acetonitrile 80%	Bean	36.60 ± 2.00 [b]	36.30 ± 0.40 [a, b]	670 ± 18 [c]	644.70 ± 22.40	(M. J. Lee et al., 2015)
Methanol 80%	Bean	22.10 ± 2.50 [c]	26.40 ± 1.30 [c]	631 ± 15 [c]	587.60 ± 2.20	(M. J. Lee et al., 2015)
Water	Okara	22.00 ± 10.00 [c]	3.00 ± 4.00 [d]	108 ± 22	71.00 ± 16.00 [c]	(Jankowiak et al., 2014)
Ethanol 70%	Okara	173.00 ± 8.00	194.00 ± 14.00 [*]	160 ± 7 [d]	171.00 ± 7.00	(Jankowiak et al., 2014)
Reverse AMS Triton X-100	Flour	NA	79.03 ± 4.31	168 ± 5 [d]	118.24 ± 4.98	(Zhao et al., 2010)
Supercritical CO_2	Flour	30.93 ± 0.04 [b, c]	1.71 ± 0.01 [d]	–	53.64 ± 0.06 [c]	(Rostagno et al., 2002)

[†] Amount of isoflavones expressed as mean $\pm$ SD, in µg (g of soybean flour) [-1]. [a-d] Values from the same column that share a superscript do not present statistical differences, given the analysis of their variances (ANOVA) with Tukey's posthoc test for a significance level of 0.05. [*] Values not included in the analysis due to being out of homoscedasticity (Bartlett's test). [**] Values corresponding to the present work.

705

706

707 ### *3.2.3. Extraction of raffinose family oligosaccharides*

708   Soy carbohydrates are largely undesirable due to their low digestibility. In

709 this regard, the process evaluated here could extract oligosaccharides belonging

710 to the raffinose family. Table 7 shows that the amount of leached oligosaccharide

711 by the AMS was 48-53% higher than that of the traditional methodology. So,

712 both AMS Tg7 and AMS Tg9 had a better extracting performance, i.e., higher

713 detoxifying capacity. According to the Kd values lower than 1, those

714 oligosaccharides slightly prefer the AP, probably due to their hydrophilic nature.

715 This finding, together with the larger volumes of the AP (3.3-4.5 times the MP

716 volume), are responsible for the high obtained Pr, close to 85% for both

717 surfactants.

718 **Table 7.** Comparison of the amount of raffinose family oligosaccharides extracted by the
719 different methodologies.

Methodology		Raffinose family oligosaccharides [†]		
		Amount [1]	Kd [2]	Pr [3]
Micellar extraction	Tg7	17.9 ± 1.0 [a]	0.76 ± 0.05 [a]	85.1 ± 1.3 [a]
	Tg9	18.5 ± 1.3 [a]	0.75 ± 0.02 [a]	85.7 ± 0.9 [a]
Classic extraction		12.1 ± 0.8	NA	NA

[†] Values are expressed as mean ± SD. [1] Amount in RU (g of soybean flour) [-1]. [2] Distribution coefficient (Kd). [3] Aqueous phase recovery performance (Pr) expressed in %. [a] Values from the same column that share a superscript do not present statistical differences among themselves, given the analysis of their variances (ANOVA) with Tukey's posthoc test for a significance level of 0.05.

720

721   Considering that it is also possible to separate most of the raffinose family

722 oligosaccharides from the isoflavones by simply decanting the phases, the

723 proposed methodology constitutes a better alternative than those currently used.

724 Besides, soy galactooligosaccharides represent an alternative and economical

725 carbon source for microbial conversion to value-added products such as succinic

726 acid and ethanol (Thakker et al., 2014). Consequently, the applied AMS-based

727 extraction allows obtaining a by-product for further uses in bio-conversion

728 processes.

4. Conclusions

730 This study, aimed to perform a novel AMS-based extracting methodology,

731 was driven by the following research questions: -Are SPC from the traditional

732 and proposed methodology of comparable quality? -What about their nutritional

733 and bioactive properties? -Compared to the classical methodology, how is the

734 AMS performing in extracting soluble proteins, antinutrients, and isoflavones? -Is

735 it possible to recover these components separately from the extracts to valorize

736 them? In light of all presented results, it is possible to answer all the questions

737 and conclude:

738 -The AMS-based extractive methodology produced SPC noticeably

739 depleted in trypsin inhibitor content (TIA loss of 90%) with total protein close to

740 60% and variable protein soluble amounts ranging from 19% to 34%, these

741 parameters being similar to those corresponding to SPC from classical methods.

742 The secondary structure content, determined as a fingerprint of the obtained

743 product, showed either alpha helix (SPC Tg7) or random coils (SPC Tg9)

744 increases concerning SPC control (SPCa). This finding might affect their

745 functional properties (emulsifying, solubility, etc.) and open potential new

746 applications. Additionally, remarkable antioxidant and antihypertensive activities,

747 other than high soluble protein content, were detected in peptides released during

748 *in vitro* digestibility assays, thus confirming the high quality of the obtained

749 product.

750 -The AMS-based methodology also exhibited a markedly enhanced extractive

751 performance of raffinose family oligosaccharides (~50% higher than the

752 traditional one), thus resulting in a successful soybean flour detoxifying strategy.

753 Due to its non-destructive character, it allowed recovering raffinose family

754 oligosaccharides and other valuable by-products, such as Kunitz and Bowman-

755 Birk trypsin inhibitors, lectins, and lunasin, potentially applicable in research and

756 fermentation fields. Additionally, these compounds, primarily recovered in the

757 aqueous phase of AMS, were separated from isoflavones (isolated and

758 concentrated in the micelle phase). Regarding this issue, AMS demonstrated

759 improved efficiency in extracting isoflavones, thus exhibiting both selectivity for

760 aglycone/glycone isoflavone forms and high yields, e.g., almost twice as large as

761 that of the traditional method and even higher than those obtained with organic

762 solvents such as methanol. These advantages make AMS a powerful tool for the

763 food and cosmetic industries.

764 Finally, it is possible to affirm that the proposed AMS-based extraction

765 constitutes a sustainable, simple, and scalable alternative that deserves

766 consideration among the new technologies for processing plant derivatives and

767 obtaining more valuable products and by-products.

768

Acknowledgments

770 Author E.R. Coscueta acknowledges Hugo Osório and IPATIMUP (Instituto de

771 Patologia e Imunologia Molecular da Universidade do Porto) for their contribution to

772 the analysis of MALDI-TOF/TOF.

773 **Funding**

779   **References**

780   Agrahar-Murugkar, D., & Jha, K. (2010). Effect of drying on nutritional and functional quality
781   and electrophoretic pattern of soyflour from sprouted soybean (Glycine max). *Journal of*
782   *Food Science and Technology*, *47*(5), 482–487. https://doi.org/10.1007/s13197-010-0082-
783   5

784   Akande, K. E., & Fabiyi, E. F. (2010). Effect of processing methods on some antinutritional
785   factors in legume seeds for poultry feeding. *International Journal of Poultry Science*,
786   *9*(10), 996–1001. https://doi.org/10.3923/ijps.2010.996.1001

787   Alibhai, Z., Mondor, M., Moresoli, C., Ippersiel, D., & Lamarche, F. (2006). Production of soy
788   protein concentrates/isolates: traditional and membrane technologies. *Desalination*,
789   *191*(1–3), 351–358. https://doi.org/10.1016/j.desal.2005.05.026

790   Andrade, J. C., Mandarino, J. M. G., Kurozawa, L. E., & Ida, E. I. (2016). The effect of thermal
791   treatment of whole soybean flour on the conversion of isoflavones and inactivation of
792   trypsin inhibitors. *Food Chemistry*, *194*, 1095–1101.
793   https://doi.org/10.1016/j.foodchem.2015.08.115

794   AOCS. (2009). Trypsin inhibitor activity. In *Sampling and analysis of oilseed by-products*.
795   American Oil Chemist's Society.

796   Arrondo, J. L. R., Muga, A., Castresana, J., & Goni, F. M. (1993). Quantitative studies of the
797   structure of proteins in solution by fourier-transform infrared spectroscopy. *Progress in*
798   *Biophysics*, *59*, 23.

799   Aura, A. M. (2005). In vitro digestion models for dietary phenolic compounds. In *VTT*
800   *Publications* (Issue 575, pp. 1–107). VTT Technical Research Centre of Finland.

801   Bajkacz, S., & Adamek, J. (2017). Evaluation of new natural deep eutectic solvents for the
802   extraction of isoflavones from soy products. *Talanta*, *168*, 329–335.
803   https://doi.org/10.1016/j.talanta.2017.02.065

804   Balisteiro, D. M., Rombaldi, C. V., & Genovese, M. I. (2013). Protein, isoflavones, trypsin
805   inhibitory and in vitro antioxidant capacities: Comparison among conventionally and
806   organically grown soybeans. *Food Research International*, *51*(1), 8–14.
807   https://doi.org/10.1016/j.foodres.2012.11.015

808   Boisen, S., & Eggum, B. O. (1991). Critical evaluation of in vitro methods for estimating
809   digestibility in simple-stomach animals. *Nutrition Research Reviews*, *4*(1), 141–162.
810   https://doi.org/10.1079/NRR19910012

811   Cao, Y., Xing, H., Yang, Q., Bao, Z., Su, B., Yang, Y., & Ren, Q. (2012). Separation of soybean
812   isoflavone aglycone homologues by ionic liquid-based extraction. *Journal of Agricultural*
813   *and Food Chemistry*, *60*(13), 3432–3440. https://doi.org/10.1021/jf3003009

814   Capriotti, A. L. A. A. L., Caruso, G., Cavaliere, C., Samperi, R., Ventura, S., Zenezini Chiozzi, R.,
815   Laganà, A., Group, P., Ventura, S., Zenezini Chiozzi, R., & Laganà, A. (2015). Identification
816   of potential bioactive generated by imulated gastrointestinal digestion of soybean seeds
817   and soy milk proteins. *Journal of Food Composition and Analysis*, *44*, 205–213.
818   https://doi.org/10.1016/j.jfca.2015.08.007

819   Carvalho, A. W. de, Natal, D. I. G., Silva, C. O. da, Dantas, M. I. de S., Barros, E. G. de, Ribeiro, S.
820   M. R., Costa, N. M. B., & Martino, H. S. D. (2013). Heat-treatment reduces antinutritional
821   phytochemicals and maintains protein quality in genetically improved hulled soybean
822   flour. *Food Science and Technology (Campinas)*, *33*(2), 310–315.
823   https://doi.org/10.1590/S0101-20612013005000048

824   Chang, M. C., & Tanaka, J. (2002). FT-IR study for hydroxyapatite/collagen nanocomposite
825   cross-linked by glutaraldehyde. *Biomaterials*, *23*(24), 4811–4818.
826   https://doi.org/10.1016/S0142-9612(02)00232-6

827   Chen, X., Ru, Y., Chen, F., Wang, X., Zhao, X., & Ao, Q. (2013). FTIR spectroscopic
828   characterization of soy proteins obtained through AOT reverse micelles. *Food*
829   *Hydrocolloids*, *31*(2), 435–437. https://doi.org/10.1016/j.foodhyd.2012.11.017

830   Chen, Y. (2015). *Effects of Micronization, Ethanol Washing, and Enzymatic Hydrolysis*
831   *Processing Alone or in Combination on Trypsin Inhibitors, Lipoxygenase Activities and*
832   *Selected "Beany" Flavour Related Compounds in Soybean Flour*. University of Manitoba.

833   Cordisco, E., Haidar, C. N., Coscueta, E. R., Nerli, B. B., & Malpiedi, L. P. (2016). Integrated
834   extraction and purification of soy isoflavones by using aqueous micellar systems. *Food*
835   *Chemistry*, *213*, 514–520. https://doi.org/10.1016/j.foodchem.2016.07.001

836   Coscueta, E. R., Amorim, M. M., Voss, G. B., Nerli, B. B., Picó, G. A., & Pintado, M. E. (2016).
837   Bioactive properties of peptides obtained from Argentinian defatted soy flour protein by
838   Corolase PP hydrolysis. *Food Chemistry*, *198*, 36–44.
839   https://doi.org/10.1016/j.foodchem.2015.11.068

840   Coscueta, E. R., Brassesco, M. E., Pellegrini Malpiedi, L., & Nerli, B. B. (2022). Non-Ionic
841   Aqueous Micellar Extraction of Trypsin Inhibitors and Isoflavones from Soybean Meal:
842   Process Optimization. *Preprints*. https://doi.org/10.20944/preprints202201.0243.v1

843   Coscueta, E. R., Brassesco, M. E., & Pintado, M. (2021). Collagen-Based Bioactive Bromelain
844   Hydrolysate from Salt-Cured Cod Skin. *Applied Sciences*, *11*(18), 8538.
845   https://doi.org/10.3390/app11188538

846   Coscueta, E. R., Pellegrini Malpiedi, L., & Nerli, B. B. (2018). Micellar systems of aliphatic
847   alcohol ethoxylates as a sustainable alternative to extract soybean isoflavones. *Food*
848   *Chemistry*, *264*, 135–141. https://doi.org/10.1016/j.foodchem.2018.05.015

849   Coscueta, E. R., Pintado, M. E., Picó, G. A., Knobel, G., Boschetti, C. E., Malpiedi, L. P., & Nerli,
850   B. B. (2017). Continuous method to determine the trypsin inhibitor activity in soybean
851   flour. *Food Chemistry*, *214*, 156–161. https://doi.org/10.1016/j.foodchem.2016.07.056

852 Dang, L., & Van Damme, E. J. M. (2015). Toxic proteins in plants. *Phytochemistry*, *117*(1), 51–
853   64. https://doi.org/10.1016/j.phytochem.2015.05.020

854 Day, L. (2013). Proteins from land plants - Potential resources for human nutrition and food
855   security. *Trends in Food Science and Technology*, *32*(1), 25–42.
856   https://doi.org/10.1016/j.tifs.2013.05.005

857 de Mejia, E. G., Bradford, T., & Hasler, C. (2003). The anticarcinogenic potential of soybean
858   lectin and lunasin. *Nutrition Reviews*, *61*(July), 239–246.
859   https://doi.org/10.131/nr.2003.jul.239

860 de Mejia, E. G., Castañeda-Reyes, E. D., Mojica, L., Dia, V., Wang, H., Wang, T., & Johnson, L. A.
861   (2021). Potential health benefits associated with lunasin concentration in dietary
862   supplements and lunasin-enriched soy extract. *Nutrients*, *13*(5), 1618.
863   https://doi.org/10.3390/nu13051618

864 Derringer, G., & Suich, R. (1980). Simultaneous Optimization of Several Response Variables.
865   *Journal of Quality Technology*, *12*(4), 214–219.
866   https://doi.org/10.1080/00224065.1980.11980968

867 Dong, A., Prestrelski, S. J., Allison, S. D., & Carpenter, J. F. (1995). Infrared spectroscopic studies
868   of lyophilization- and temperature-induced protein aggregation. *Journal of
869   Pharmaceutical Sciences*, *84*(4), 415–424. https://doi.org/10.1002/jps.2600840407

870 Erickson, D. R. (1995). *Practical handbook of soybean processing and utilization*. AOCS Press.

871 Fabian, H., Choo, L. P. I., Szendrei, G. I., Jackson, M., Halliday, W. C., Otvos, L., & Mantsch, H. H.
872   (1993). Infrared spectroscopic characterization of Alzheimer plaques. *Applied
873   Spectroscopy*, *47*(9), 1513–1518. https://doi.org/10.1366/0003702934067469

874 FAO, & Berk, Zeki. (1992). *Technology of production of edible flour and protein products from
875   soybeans*.

876 Fisher, R. R., Glatz, C. E., & Murphy, P. A. (1986). Effects of mixing during acid addition on
877   fractionally precipitated protein. *Biotechnology and Bioengineering*, *28*(7), 1056–1063.
878   https://doi.org/10.1002/bit.260280716

879 French, D. L., Arakawa, T., & Li, T. (2004). Fourier Transformed Infrared Spectroscopic
880   Investigation of Protein Conformation in Spray-Dried Protein/Trehalose Powders.
881   *Biopolymers*, *73*(4), 524–531. https://doi.org/10.1002/bip.10558

882 Gatel, F. (1994). Protein quality of legume seeds for non-ruminant animals: a literature review.
883   *Animal Feed Science and Technology*, *45*(3–4), 317–348. https://doi.org/10.1016/0377-
884   8401(94)90036-1

885 Gomes, M. T. R., Oliva, M. L., Lopes, M. T. P., & Salas, C. E. (2011). Plant proteinases and
886   inhibitors: an overview of biological function and pharmacological activity. *Current
887   Protein & Peptide Science*, *12*(5), 417–436.
888   https://doi.org/10.2174/138920311796391089

889 González-Montoya, M., Ramón-Gallegos, E., Robles-Ramírez, M. del C., & Mora-Escobedo, R.
890   (2016). Evaluation of the antioxidant and antiproliferative effects of three peptide
891   fractions of germinated soybeans on breast and cervical cancer cell lines. *Plant Foods for
892   Human Nutrition*, *71*(4), 368–374. https://doi.org/10.1007/s11130-016-0568-z

893   Haidar, C. N., Coscueta, E., Cordisco, E., Nerli, B. B., & Malpiedi, L. P. (2018). Aqueous micellar
894   two-phase system as an alternative method to selectively remove soy antinutritional
895   factors. *LWT*, *93*, 665–672. https://doi.org/10.1016/j.lwt.2018.04.025

896   Hao, Y., Fan, X., Guo, H., Yao, Y., Ren, G., Lv, X., & Yang, X. (2020). Overexpression of the
897   bioactive lunasin peptide in soybean and evaluation of its anti-inflammatory and anti-
898   cancer activities in vitro. *Journal of Bioscience and Bioengineering*, *129*(4), 395–404.
899   https://doi.org/10.1016/j.jbiosc.2019.11.001

900   He, J., Zhao, Z., Shi, Z., Zhao, M., Li, Y., & Chang, W. (2005). Analysis of isoflavone daidzein in
901   Puerariae radix with micelle-mediated extraction and preconcentration. *Journal of
902   Agricultural and Food Chemistry*, *53*(3), 518–523. https://doi.org/10.1021/jf048545q

903   Jankowiak, L., Kantzas, N., Boom, R., & Van Der Goot, A. J. (2014). Isoflavone extraction from
904   okara using water as extractant. *Food Chemistry*, *160*, 371–378.
905   https://doi.org/10.1016/j.foodchem.2014.03.082

906   Jasti, L. S., Lavanya, K., & Fadnavis, N. W. (2015). Adsorption induced denaturation: application
907   to denaturation of soybean trypsin inhibitor (SBTI) and lipoxygenase (LOX) in soymilk.
908   *Biotechnology Letters*, *37*(1), 147–151. https://doi.org/10.1007/s10529-014-1659-2

909   Jia, F., Peng, S., Green, J., Koh, L., & Chen, X. (2020). Soybean supply chain management and
910   sustainability: A systematic literature review. *Journal of Cleaner Production*, *255*, 120254.
911   https://doi.org/10.1016/J.JCLEPRO.2020.120254

912   Johnson, L. A. (1999). Process for Producing Improved Soy Protein Concentrate From
913   Genetically-Modified Soybeans (Patent No. 5936069). In *US Patent 5936069* (No.
914   5936069).

915   Kakade. (1974). Determination of Trypsin Inhibitor Activity of Soy Products: A Collaborative
916   Analysis of an Improved Procedure. In *Cereal Chemistry*.

917   Kulling, S. E., Honig, D. M., & Metzler, M. (2001). Oxidative metabolism of the soy isoflavones
918   daidzein and genistein in humans in vitro and in vivo. *Journal of Agricultural and Food
919   Chemistry*, *49*(6), 3024–3033. https://doi.org/10.1021/jf0012695

920   Laurent, C., Besançon, P., & Caporiccio, B. (2007). Flavonoids from a grape seed extract
921   interact with digestive secretions and intestinal cells as assessed in an in vitro
922   digestion/Caco-2 cell culture model. *Food Chemistry*, *100*(4), 1704–1712.
923   https://doi.org/10.1016/j.foodchem.2005.10.016

924   Lee, C. C., Dudonné, S., Dubé, P., Desjardins, Y., Kim, J. H., Kim, J. S., Kim, J. E., Park, J. H. Y.,
925   Lee, K. W., & Lee, C. Y. (2017). Comprehensive phenolic composition analysis and
926   evaluation of Yak-Kong soybean (Glycine max) for the prevention of atherosclerosis. *Food
927   Chemistry*, *234*, 486–493. https://doi.org/10.1016/j.foodchem.2017.05.012

928   Lee, M. J., Chung, I. M., Kim, H., & Jung, M. Y. (2015). High resolution LC-ESI-TOF-mass
929   spectrometry method for fast separation, identification, and quantification of 12
930   isoflavones in soybeans and soybean products. *Food Chemistry*, *176*, 254–262.
931   https://doi.org/10.1016/j.foodchem.2014.12.073

932   Liu, K. (1997). Chemistry and Nutritional Value of Soybean Components. In *Soybeans* (pp. 25–
933   113). Springer US. https://doi.org/10.1007/978-1-4615-1763-4_2

934   Liu, K. (2021). Comparison of ISO14902:2001 with AOCS Ba 12a-2020 for determining trypsin
935   inhibitor activity in protein products. *Journal of the American Oil Chemists' Society*,
936   *98*(12), 1115–1129. https://doi.org/10.1002/aocs.12542

937   Lule, V. K., Garg, S., Pophaly, S. D., Hitesh, & Tomar, S. K. (2015). "Potential health benefits of
938   lunasin: A multifaceted soy-derived bioactive peptide." *Journal of Food Science*, *80*(3),
939   C485–C494. https://doi.org/10.1111/1750-3841.12786

940   Messina, M. (2014). Soy foods, isoflavones, and the health of postmenopausal women.
941   *American Journal of Clinical Nutrition*, *100*(SUPPL. 1).
942   https://doi.org/10.3945/ajcn.113.071464

943   Murugkar, D. A. (2015). Effect of different process parameters on the quality of soymilk and
944   tofu from sprouted soybean. *Journal of Food Science and Technology*, *52*(5), 2886–2893.
945   https://doi.org/10.1007/s13197-014-1320-z

946   Nadar, S. S., Pawar, R. G., & Rathod, V. K. (2017). Recent advances in enzyme extraction
947   strategies: A comprehensive review. *International Journal of Biological Macromolecules*,
948   *101*, 931–957. https://doi.org/10.1016/J.IJBIOMAC.2017.03.055

949   Newkirk, R. (2010). SoyBean, Feed Industry Guide. In *Feed Industry Guide 1st Editin, 2010* (pp.
950   1–48).

951   Nishinari, K., Fang, Y., Guo, S., & Phillips, G. O. (2014). Soy proteins: A review on composition,
952   aggregation and emulsification. *Food Hydrocolloids*, *39*, 301–318.
953   https://doi.org/10.1016/j.foodhyd.2014.01.013

954   Pivato, M., De Franceschi, G., Tosatto, L., Frare, E., Kumar, D., Aioanei, D., Brucale, M., Tessari,
955   I., Bisaglia, M., Samori, B., de Laureto, P. P., & Bubacco, L. (2012). Covalent α-Synuclein
956   Dimers: Chemico-Physical and Aggregation Properties. *PLoS ONE*, *7*(12).
957   https://doi.org/10.1371/journal.pone.0050027

958   Roberts, P. R., Burney, J. D., Black, K. W., & Zaloga, G. P. (1999). Effect of chain length on
959   absorption of biologically active peptides from the gastrointestinal tract. *Digestion*, *60*(4),
960   332–337.

961   Rostagno, M. A., Araújo, J. M. A., & Sandi, D. (2002). Supercritical fluid extraction of isoflavones
962   from soybean flour. *Food Chemistry*, *78*(1), 111–117. https://doi.org/10.1016/S0308-
963   8146(02)00106-1

964   Sair, L. (1959). *Proteinaceous soy composition and method of preparing* (Patent No. 2881076).

965   Seber, L. E., Barnett, B. W., McConnell, E. J., Hume, S. D., Cai, J., Boles, K., & Davis, K. R. (2012).
966   Scalable purification and characterization of the anticancer lunasin peptide from
967   soybean. *PLoS ONE*, *7*(4). https://doi.org/10.1371/journal.pone.0035409

968   Serra, A., Gallart-Palau, X., See-Toh, R. S.-E., Hemu, X., Tam, J. P., & Sze, S. K. (2016).
969   Commercial processed soy-based food product contains glycated and glycoxidated
970   lunasin proteoforms. *Scientific Reports*, *6*, 26106. https://doi.org/10.1038/srep26106

971   Sharma, S., Kori, S., & Parmar, A. (2015). Surfactant mediated extraction of total phenolic
972   contents (TPC) and antioxidants from fruits juices. *Food Chemistry*, *185*, 284–288.
973   https://doi.org/10.1016/j.foodchem.2015.03.106

974 Singh, B. P., Yadav, D., & Vij, S. (2017). Soybean Bioactive Molecules: Current Trend and Future
975   Prospective. In J.-M. Mérillon & K. G. Ramawat (Eds.), *Bioactive Molecules in Food* (pp. 1–
976   29). Springer International Publishing. https://doi.org/10.1007/978-3-319-54528-8_4-1

977 Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D.,
978   Fujimoto, E. K., Goeke, N. M., Olson, B. J., & Klenk, D. C. (1985). Measurement of protein
979   using bicinchoninic acid. *Analytical Biochemistry*, *150*(1), 76–85.
980   https://doi.org/10.1016/0003-2697(85)90442-7

981 Stewart, O. J., Raghavan, G. S. V, Orsat, V., & Golden, K. D. (2003). The effect of drying on
982   unsaturated fatty acids and trypsin inhibitor activity in soybean. *Process Biochemistry*, *39*,
983   483–489. https://doi.org/10.1016/S0032-9592(03)00130-4

984 Susi, H., & Byler, D. M. (1986). [13] Resolution-Enhanced Fourier Transform Infrared
985   Spectroscopy of Enzymes. *Methods in Enzymology*, *130*(C), 290–311.
986   https://doi.org/10.1016/0076-6879(86)30015-6

987 Thakker, C., San, K.-Y., & Bennett, G. N. (2014). Soybean Carbohydrates as a Renewable
988   Feedstock for the Fermentative Production of Succinic Acid and Ethanol - ACS Symposium
989   Series (ACS Publications). *ACS Symposium Series*, 81–107. https://doi.org/10.1021/bk-
990   2014-1178.ch004

991 Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, *5*(2),
992   99–114. https://doi.org/10.2307/3001913

993 Vagadia, B. H., Vanga, S. K., & Raghavan, V. (2017). Inactivation methods of soybean trypsin
994   inhibitor – A review. *Trends in Food Science and Technology*, *64*, 115–125.
995   https://doi.org/10.1016/j.tifs.2017.02.003

996 Vicente, F. A., Cardoso, S., Sintra, nia E., Lemus, J., Marques, E. F., M Ventura, nia P., &
997   Coutinho, A. P. (2017). *Impact of Surface Active Ionic Liquids on the Cloud Points of*
998   *Nonionic Surfactants and the Formation of Aqueous Micellar Two-Phase Systems*.
999   https://doi.org/10.1021/acs.jpcb.7b02972

1000 Xiao, C. W., Wood, C. M., Robertson, P., & Gilani, G. S. (2012). Protease inhibitor activities and
1001   isoflavone content in commercial soymilks and soy-based infant formulas sold in Ottawa,
1002   Canada. *Journal of Food Composition and Analysis*, *25*(2), 130–136.
1003   https://doi.org/10.1016/j.jfca.2011.10.001

1004 Xu, M. L., Liu, J., Zhu, C., Gao, Y., Zhao, S., Liu, W., & Zhang, Y. (2015). Interactions between soy
1005   isoflavones and other bioactive compounds: a review of their potentially beneficial health
1006   effects. *Phytochemistry Reviews*, *14*(3), 459–467. https://doi.org/10.1007/s11101-015-
1007   9398-0

1008 Zana, R., Lévy, H., & Kwetkat, K. (1998). Mixed Micellization of Dimeric (Gemini) Surfactants
1009   and Conventional Surfactants. I. Mixtures of an Anionic Dimeric Surfactant and of the
1010   Nonionic Surfactants C12E5 and C12E8. *Journal of Colloid and Interface Science*, *197*(2),
1011   370–376. https://doi.org/10.1006/jcis.1997.5248

1012 Zhao, X., Chen, F., Chen, J., Gai, G., Xue, W., & Li, L. (2008). Effects of AOT reverse micelle on
1013   properties of soy globulins. *Food Chemistry*, *111*(3), 599–605.
1014   https://doi.org/10.1016/j.foodchem.2008.04.026

1015 Zhao, X., Chen, F., Xue, W., & Lee, L. (2008). FTIR spectra studies on the secondary structures
1016 of 7S and 11S globulins from soybean proteins using AOT reverse micellar extraction.
1017 *Food Hydrocolloids*, *22*(4), 568–575. https://doi.org/10.1016/j.foodhyd.2007.01.019

1018 Zhao, X., Wei, Z., Du, F., & Zhu, J. (2010). Effects of surfactant and salt species in reverse
1019 micellar forward extraction efficiency of isoflavones with enriched protein from soy flour.
1020 *Applied Biochemistry and Biotechnology*, *162*(7), 2087–2097.
1021 https://doi.org/10.1007/s12010-010-8984-2

1022

1023 **Figure Captions**

1024 **Figure 1. Methodology outline.** Flow diagram of the complete methodology, including

1025 the raw material (white soybean flour), the processes applied, the products and by-

1026 products obtained, and the analyses performed. For the processes and analyses, the

1027 subsection number corresponding to their description in the Materials and Methods

1028 section was also indicated.

1029 **Figure 2. Process outline.** Scheme of the proposed methodology to produce SPC with

1030 SMA. ME: micellar extraction; MP: micellar phase; AP: aqueous phase.

1031 **Figure 3. FTIR.** Standardized ATR-FTIR absorbance spectra of soybean flour and

1032 protein concentrates (SPCs). **A)** Full spectrum from 4000 to 800 cm^{-1}. **B)** Most

1033 important characteristic amide regions (I, II and III). Comparison of the amide-I region

1034 of the deconvoluted ATR-FTIR absorbance spectra for: **C)** soybean flour; **D)** SPCa; **E)**

1035 SPC Tg7; **F)** SPC Tg9. The curve fit was performed with Gaussian lines. The peak

1036 position of each component corresponding to the amide band was deduced from the

1037 second derivative spectra. The sum of the fitted curves is shown as a solid line, closely

1038 overlaid the original spectrum, shown as a dashed line (C) or colored line (D, E, and F).

1039 **Figure 4. SGI.** Levels of each response depending on the stage of the SGI. The lines

1040 display the observations for each experimental unit for: **A**) soluble protein (Ps); **B**)

1041 ORAC; **C**) iACE. Stages of the process: T1, beginning of the gastric stage; T2, end of

1042 gastric stage; T3, end of the intestinal stage.

1043 **Figure 5. Size exclusion chromatography.** Chromatograms (solid lines) of molecular

1044 size distribution of the APs obtained by the classical method and the AMSs of Tg7 and

1045 Tg9. The horizontal dotted lines represent the baseline of each chromatogram, and the

1046 vertical dashed ones divide regions: I, II and III. The numbers represent the peaks

1047 corresponding to region I.

Annex

Production of soy protein concentrate with the recovery of bioactive compounds: from destruction to valorization

Ezequiel R. Coscueta, Luciana Pellegrini Malpiedi, Maria Manuela Pintado and Bibiana B. Nerli

Table A1. Analysis of variance for fitted general linear models.

Response	Source	SS[*]	DF[**]	MS[***]	F	P	R^2
Ps	Model	687	24	28.604	228	0.000	0.996
	Residue	2,510	20	0.126			
	Total	689	44				
ORAC	Model	3557270	24	148219	276	0.000	0.997
	Residue	10759	20	538			
	Total	3568030	44				
iACE	Model	7580240	24	315843	251	0.000	0.997
	Residue	25209	20	1260			
	Total	7605440	44				

[*]SS, sum of squares; [**]DF, degrees of freedom; [***]MS, mean square.
P < 0,05 is considered significant.

Table A2. Multiple comparisons for each response by T.

T	Least squares mean [†]		
	Ps [1]	ORAC [2]	iACE [3]
T1	2.59 ± 0.09	51 ± 6	1015 ± 9
T2	9.63 ± 0.09	460 ± 6	283 ± 9
T3	8.37 ± 0.09	577 ± 6	57 ± 9

[†] Values are expressed on dry basis, as mean ± SEM calculated by least squares. [1] Ps in mg (g of product)$^{-1}$, [2] ORAC in µmol TE (g of product)$^{-1}$, [3] iACE as IC_{50}, µg mL^{-1}. All the values in the same column present statistical differences among themselves, given the analysis of their variances (ANOVA) with Tukey's posthoc test for a significance level of 0.05.

Table A3. Multiple comparisons for each response by product.
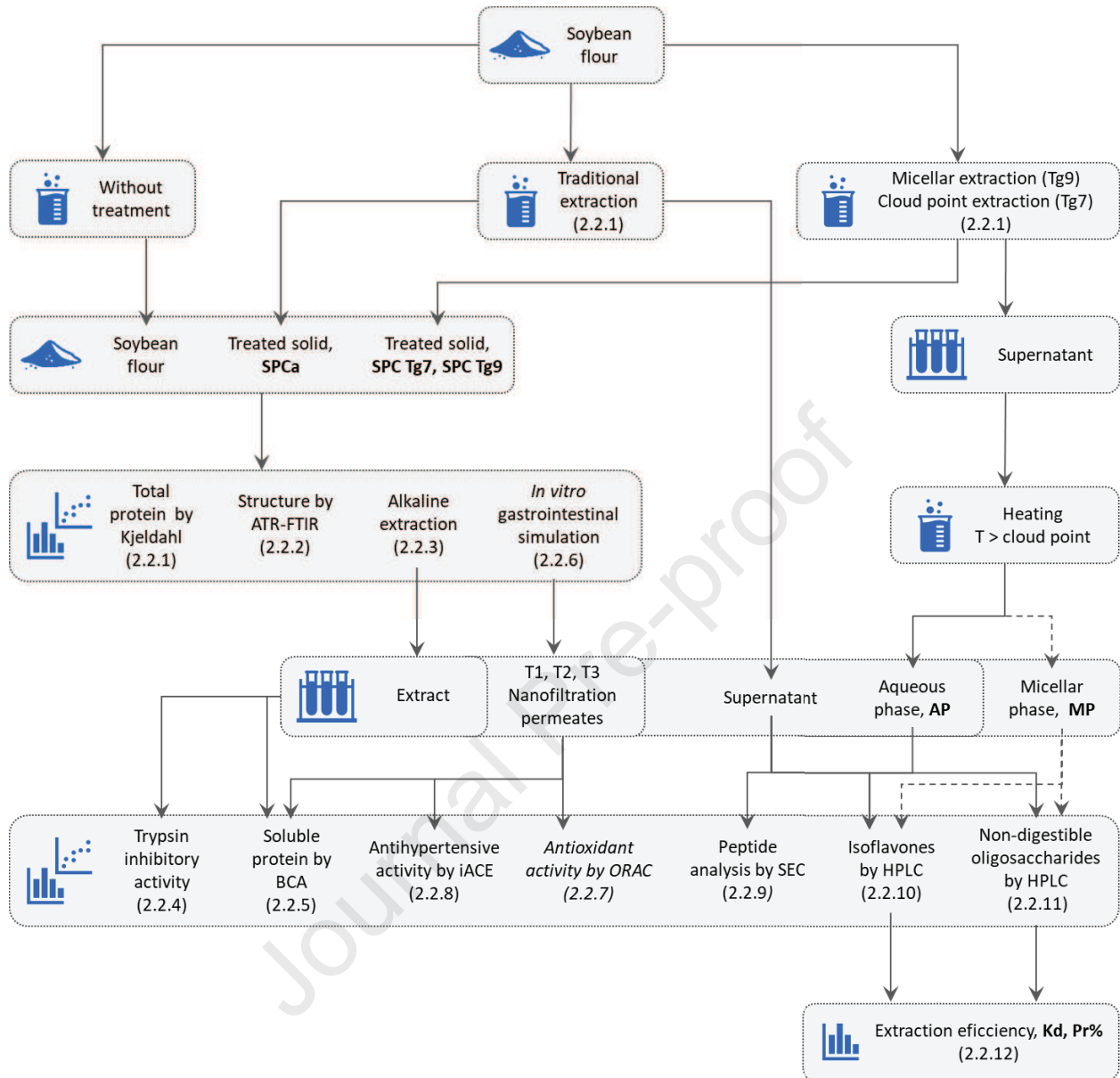
Product	Least squares mean [†]

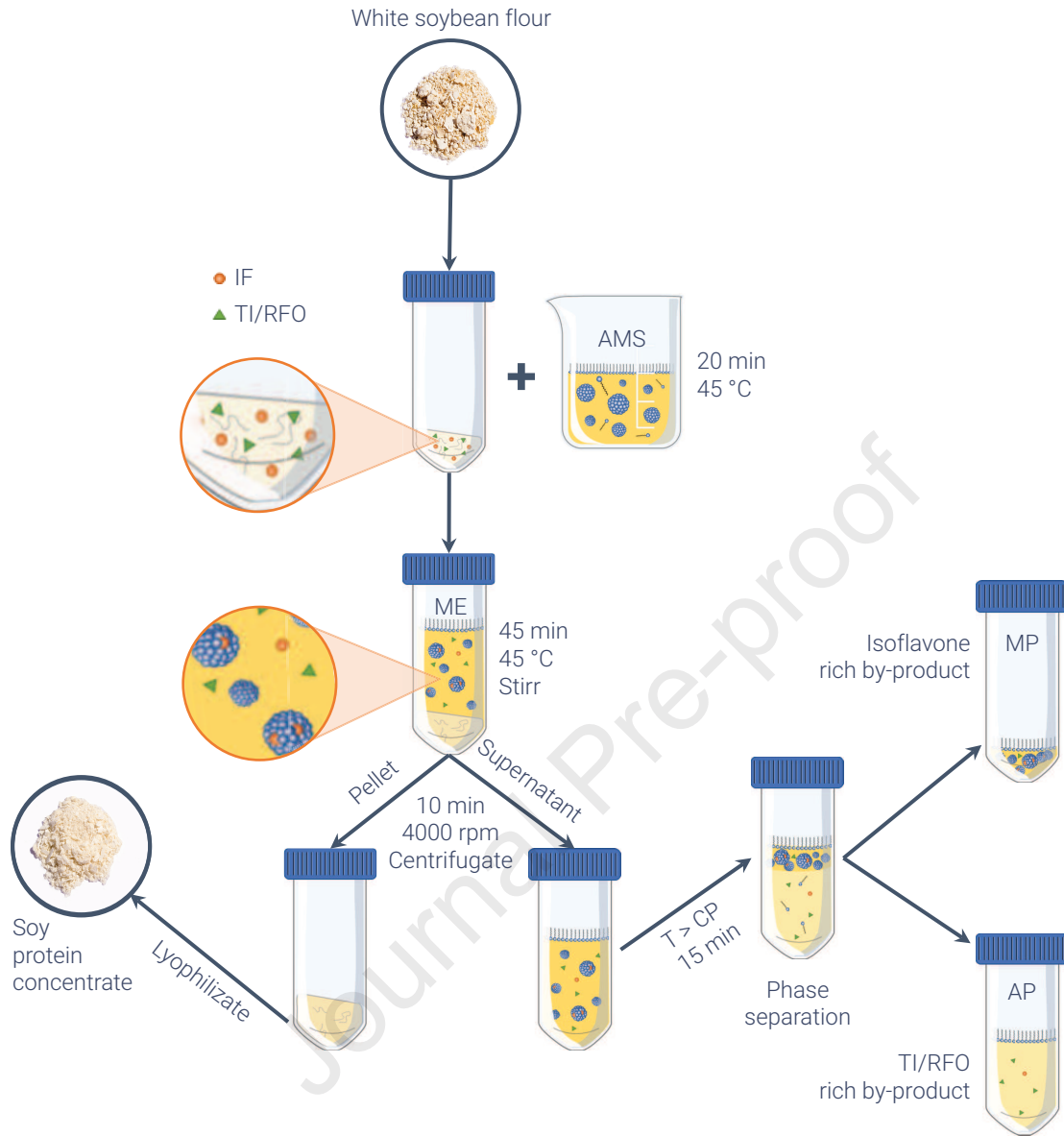|	Ps [1]	ORAC [2]	iACE [3]
Flour	6.3 ± 0.3 [a]	473 ± 13 [a]	452 ± 12 [a, b]
SPCa	7.5 ± 0.3 [a]	326 ± 13	428 ± 12 [a]
SPC Tg7	4.8 ± 0.3	178 ± 13	489 ± 12 [b]
SPC Tg9	6.9 ± 0.3 [a]	534 ± 13 [a]	431 ± 12 [a]

[†] Values are expressed on a dry basis, as mean $\pm$ SEM calculated by least squares. [1] Ps in mg (g of product) $^{-1}$, [2] ORAC in µmol TE (g of product) $^{-1}$, [3] iACE as IC_{50}, µg mL^{-1}. [a, b] Values from the same column that share a superscript do not present statistical differences among themselves, given the analysis of their variances (ANOVA) with Tukey's posthoc test for a significance level of 0.05.
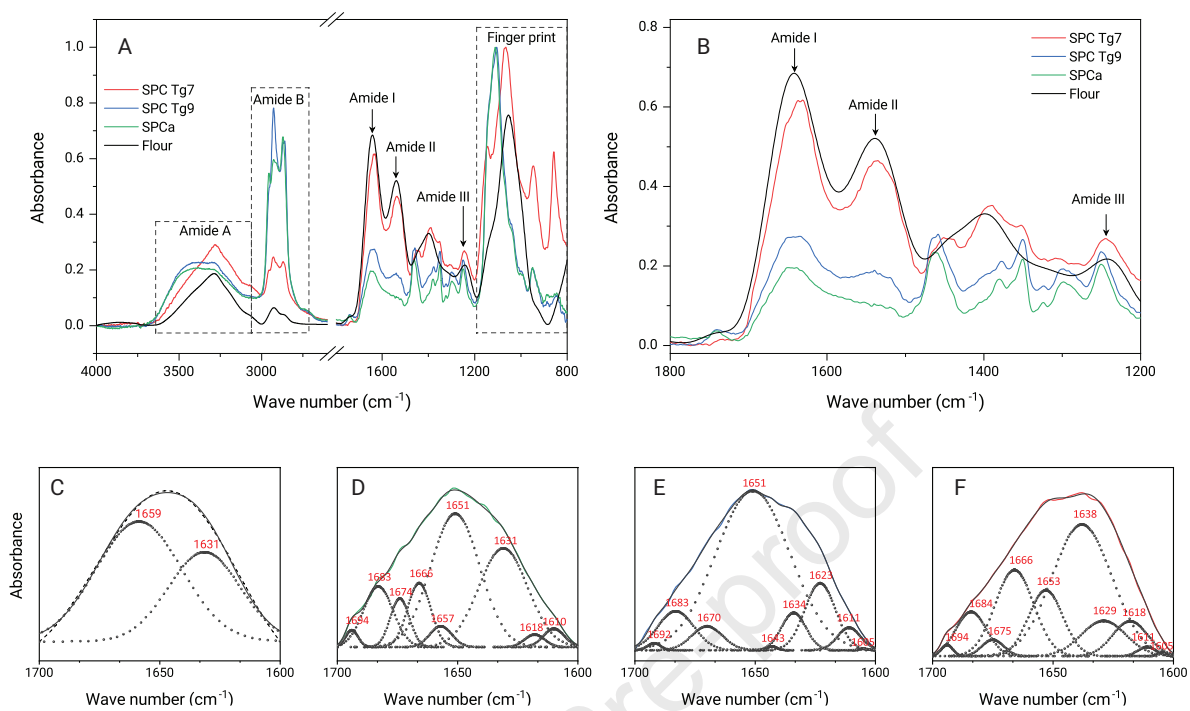
Table A4. Peak areas belonging to the region I and total areas of regions II and III.
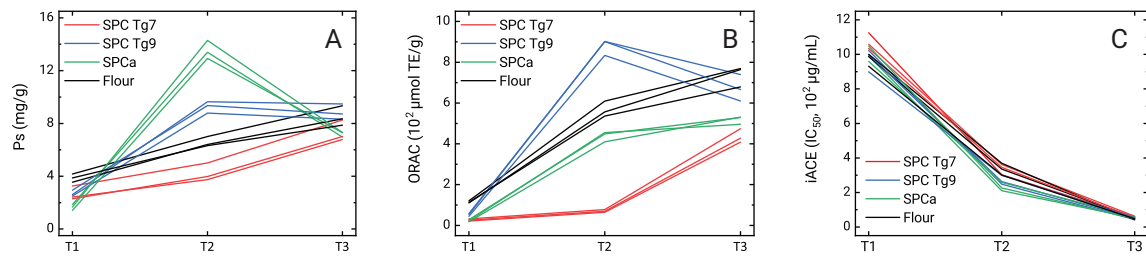
Region	Peak	Area (mL*mAU) [†]		
		Tg7 AP	Tg9 AP	Classic AP
I	1	1.6 ± 0.1	NA	NA
	2	2.5 ± 0.2	NA	NA
	3	7.3 ± 0.5 [a]	5.4 ± 0.5	7.3 ± 0.6 [a]
	4	4.7 ± 0.3 [a]	4.5 ± 0.3 [a]	4.8 ± 0.4 [a]
	5	11.2 ± 0.8 [a]	7.3 ± 0.5	14.4 ± 1.0 [a]
	6	4.0 ± 0.2	5.8 ± 0.5	3.1 ± 0.2
	7	2.4 ± 0.1	5.1 ± 0.3	1.1 ± 0.1
II		49.6 ± 4.2 [a]	77.1 ± 3.9	49.2 ± 3.2 [a]
III		66.3 ± 4.0	94.3 ± 7.1	51.3 ± 2.3

[†] Values are expressed as mean $\pm$ SD. [a, b] Values from the same row that share a superscript do not present statistical differences among themselves, given the analysis of their variances (ANOVA) with Tukey's posthoc test for a significance level of 0.05.
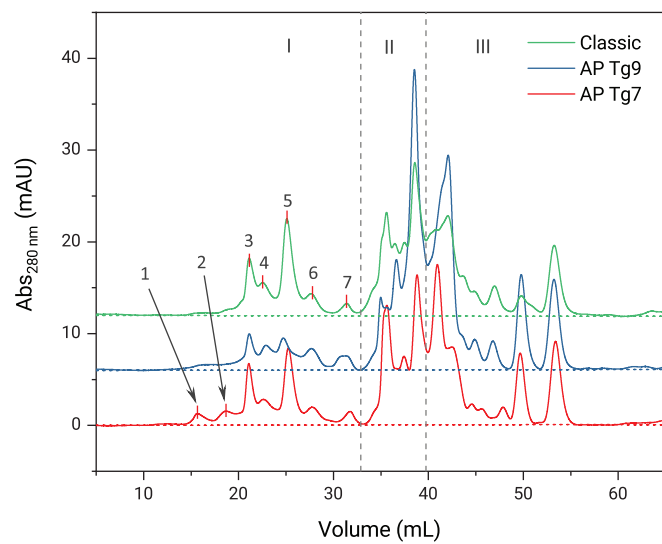
White soybean flour

- IF
- TI/RFO

AMS
20 min
45 °C

+

ME
45 min
45 °C
Stirr

Pellet

Supernatant

10 min
4000 rpm
Centrifugate

Soy
protein
concentrate

Lyophilizate

T > CP
15 min

Phase
separation

Isoflavone
rich by-product

MP

AP

TI/RFO
rich by-product

Highlights

Aqueous micellar systems (AMS) allowed producing soy protein concentrates (SPC)

Quality of SPC from AMS was similar to that of SPC from classical acid washing method

SPC from AMS showed remarkable antioxidant and antihypertensive activities

Trypsin inhibitors and non-digestible oligosaccharides were recovered as by-products

Isoflavones were concentrated and separated from antinutrients

CRediT author statement

Ezequiel R. Coscueta: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Luciana Pellegrini Malpiedi:** Methodology, Formal analysis, Writing - Review & Editing. **Maria Manuela Pintado:** Writing - Review & Editing, Supervision, Project administration, Funding acquisition. **Bibiana B. Nerli:** Conceptualization, Resources, Writing - Review & Editing, Supervision, Project administration.

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: