Amonsou Eric Oscar (Orcid ID: 0000-0001-6299-0466)
Wokadala Obiro Cuthbert (Orcid ID: 0000-0002-0292-400X)

Influence of amino acid profiles, and secondary structure on nutritional and functional properties of *Trichilia emetica* and *Trichilia dregeana* protein concentrates

Gugu Felicity Tsomele[1,6], Belinda Du Plessis[2], Tonna Ashim Anyasi[1], Victor Mlambo[3], Eric Amonsou[4], Sello Presley Lepule[5], Muthulisi Siwela[6], Obiro Cuthbert Wokadala[7*]

[1]*Agro-Processing and Postharvest Technologies Division, Agricultural Research Council Tropical and Subtropical Crops, Mbombela, 1200, Mpumalanga, South Africa*

[2]*Department of Biotechnology and Food Technology, Tshwane University of Technology, Private Bag X680, Pretoria, 0083, South Africa*

[3]*School of Agricultural Sciences, Faculty of Agriculture and Natural Sciences, University of Mpumalanga, Private Bag X11283, Mbombela, 1200, South Africa*

[4]*Department of Biotechnology and Food Science Technology, Durban University of Technology, Durban, 4001, South Africa*

[5]*Department of Chemistry, Tshwane University of Technology, Private Bag X680, Pretoria, 0001, South Africa*

[6]*School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Scottsvile 3209, Pietermaritzburg, South Africa*

[7]*Postharvest Technology Division, School of Agricultural Sciences, Faculty of Agricultural and Natural Sciences, University of Mpumalanga, Nelspruit, 1200, South Africa*

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/ijfs.16425

Abstract

The prevalent problem of protein energy malnutrition (PEM) has been a major concern in Africa. *Trichilia* species is one of the underutilized oilseeds. This study determined and compared the amino acid profile, functional properties, and the secondary structure of the proteins from *Trichilia emetica and Trichilia dregeana* oilseeds. The results showed that the essential amino acid levels of the two *Trichilia* seed proteins were above the recommended Food and Agricultural Organisation standard for adults, with E/T (Essential-to-Total amino acid ratios (*T. dregeana*-52.8%; *T. emetica* 53.4%) and chemical scores (100% for both) not significantly different ($p>0.05$) from those of soy proteins (53.0% and 100% respectively). The amino acid profile results indicated that *Trichilia* oilseeds can serve as candidates for addressing protein malnutrition in the predominantly economically disadvantaged sub-Saharan Africa. The secondary structure showed that the *Trichilia species*' proteins had significantly ($p<0.05$) more β-conformations compared to soybean proteins. Further, the functional properties results suggested that *Trichilia* oilseed proteins can be used as ingredients in food and feed formulations due to their versatile functional properties.

Keywords: *Trichilia species*; plant proteins, oilseeds; secondary structure; water holding capacity; oil holding capacity; foaming capacity; globular structure; amino acid composition

1. Introduction

Trichilia species are indigenous in Sub-Saharan and their oilseeds have a potential to contribute significantly to food security in this region. However, commercial production and the use of the oilseeds as a food source are still limited or not known (Van Wyk et al., 2000 &; Orwa et al., 2009). In traditional African culinary practice, the dehulled/decorticated Trichilia seeds are eaten raw, or ground and the extracted milky potable liquid mixed with vegetables such as spinach (Orwa et al., 2009 &; Uamusse et al., 2016). There are no studies reported on the presence antinutrient factors in the decorticated/dehulled seeds of the two Trichilia seeds. The protein content of *T. emetica* and *T. dregeana* oilseed (25.65% and 17.3%, respectively) is comparable to other oilseeds, which are utilized as a source of protein concentrates, such as soybean, peanuts, castor and sunflower seeds (Tsomele et al., 2021).

Understanding the secondary structure of the Trichilia oilseed protein can help to understand the relationships between structure and techno-functional properties, which are important to determine the utilization of the protein concentrates in the food industry. The primary structure determines the formation of secondary, tertiary, and quaternary structures, which as a whole will determine the shape of the molecule and the existence of hydrophilic zones and hydrophobic patches on the protein surface. The amino-acid side chains influence the three-dimensional shape and overall hydrophobicity of the protein. A relatively high proportion of amino acids with hydrophilic side chains tend to result rod-like shape of the protein molecules, which causes an increase in water holding and foaming capacity of the protein (Hettiarachchy et al., 2012). Secondary structures with relatively high proportions of charged (negative or positive) amino acid resides are related to the electrostatic repulsion and ionic hydration, which promote the solubilization of protein and influence functional properties, such as gelation, emulsification and foaming when exposed to moderately pH

values above or below the protein isoelectric point (Hettiarachchy et al., 2012; Day, 2013; Shevkani et al., 2015; Tang & Sun, 2011).

The secondary structure and major functional groups of the storage protein can be determined using *Fourier*-transform infrared spectroscopy (FTIR) (Zhang et al., 2015), Raman spectroscopy (Li et al., 2014), Circular dichroism (CD) (Chandrapala et al., 2012), X-ray diffraction (XRD) (Jenkins et al., 2013), Nuclear magnetic resolution (NMR) (Mao et al., 2014) and UV-spectroscopy (Barrios-Peralta et al., 2012). These methods provide valuable ways for studying conformational changes of protein in solid, crystal and solution states with high quality spectra of about 1600-1700 cm^{-1}. The protein materials usually studied include isolates, concentrates and meals; these extracts are composed of different individual proteins rather than purified fractions (Achouri et al., 2012). The Raman spectroscopy is a convenient procedure and provides effective information about the molecular vibrations of protein related to the secondary structure and microenvironment of the protein side chains. However, use of The Raman spectroscopy is limited by strong disturbance of biological fluorescence, which can hamper the collection of high-resolution spectra (Wang et al., 2017). While CD can be used to determine the secondary structure, folding and binding properties of protein at lower protein concentration, it is relatively time-consuming and provides low-resolution secondary structural information (Achouri et al., 2012). The secondary structure is dependent on the amino acid profile and strongly influences the functional properties of protein concentrates.

Extensive work has been done on secondary structure, amino acid profile and functional properties of protein concentrates from soybean (Rao et al., 2002), hemp (Wang et al., 2008), rapeseed/ canola (Gerzhova et al., 2015), peanut (Lui et al., 2019), sunflower (Ishii et al., 2021)

and walnut (Mao & Hua, 2012) protein concentrates. To the best of our knowledge, the amino acids profile, functional properties, and secondary structure of *Trichilia* oilseed protein is not known. Hence, the objective of the current study was to determine the secondary structure, amino acid profile and functional properties of protein concentrates obtained from the two *Trichilia species* (*T. emetica* and *T. dregeana*) in comparison to soybean protein, a standard plant protein.

2. Materials and Methods

2.1. Materials

The *Trichilia* (*Trichilia emetica* and *Trichilia dregeana*) oilseeds were gathered in 2021 from various trees growing at the Agricultural Research Council-Tropical and Subtropical Crops (ARC-TSC) (Mbombela, South Africa, -25.45127 S 30.96919 E) between July and November. The Soybean (*Glycine max* (L.) Merrill) seeds used as a standard in this work were obtained from Agricultural Research Council- Grain Crops (ARC-GC) (Potchefstroom, South Africa; 26.72866 S 27.07972 E). The seeds were dried in an oven (AB 3000 Agridrier, Dryer for Africa, South Africa) at 50°C for 16 h and stored in a refrigerator at 4°C; in air-tight zip-lock plastic bags.

2.2. Methods

2.2.1. Preparation of protein concentrates from oilseeds of *Trichilia species* and soybean

The protein concentrates were prepared through isoelectric precipitation according to Chove et al. (2001) with modifications. The dried seeds of *T. emetica*, *T. dregeana* and soybean (1000 g) were dehulled manually to remove the seedcoat. The dehulled seeds of *T.emetica* and *T. dregeana* and soybean were milled to flour using a Drotsky S1 hammer miller (Alberton,

Gauteng, South Africa) with a 0.8 mm sieve. The flour was stored in air-tight plastic containers at 4°C until further analysis. The *T. emetica*, *T. dregeana* and soybean flour were defatted to constant weight using diethyl ether (1:5) at 25°C for 72 h with constant stirring and three solvent changes (24 h each). The protein in the flour samples (*T. emetica*, *T. dregeana* and Soybean) were solubilized using 2 N NaOH containing sodium azide (0.005 M) for 2 h at 25°C with continuous stirring shaking in an Incoshake© incubator (Labotec, South Africa) at (120 g (25°C). The slurry/mixture was filtered in 6 folds cheese cloth and then liquids centrifuged at 2000 g for 5min at 5°C. The protein in the supernatants were precipitated by isoelectric precipitation using 2M HCL at pH 4.25 (for Soybean and *T. emetica*) and 4.35 (*T. dregeana*) for 1h. The isoelectric points were estimated through centrifugation weight yield using pH of 3.5, 4.0, 4.5, 5.0, 5.0 and 6.0 according to Chove et al. (2001) and Boye et al. (2010). The precipitates were recovered by centrifugation at 3000 g for 10 min at 5°C and the supernatants discarded. The precipitates were re-suspended in distilled water and were dialyzed against distilled water for 48 h at 25°C (with three changes of water) and then recovered by centrifugation (3000 g for 10 min at 5°C). The dialyzed extracts were freeze dried and the resulting powder stored in airtight containers at 5°C until further analysis. The extraction was done in triplicate for each seed type.

2.2.2. Amino acid profile determination of *Trichilia emetica* and *T. dregeana* protein concentrate

The amino acid content of the protein concentrates was determined using Waters Acquity Ultra Performance Liquid Chromatograph (UPLC) AccQ-Tag system (Waters Corporation, Milford, Massachusetts, 01757, United States) with MassLynx software fitted with a Photodiode array (PDA) detector (Waters, Millipore Corp, Milford, MA) according to

Bidlingmeyer et al. (1984) with modification. The amino acid determination principle is based on reverse-phase chromatography with Phenylthiocarbamyl precolumn derivatization following acid digestion. Samples were hydrolysed with 6 M HCL at 110°C for 24 h in a sealed tube prior chromatographic analysis. The cysteine and Met content were determined after performic acid oxidation (Gehrke et al., 1985) while Trp content was determined after alkaline hydrolysis (Landry & Delhaye, 1992). The digests were separated onto a Waters Ultra Tag C_{18} column, USA (2.1 X 50 mm x 1.7 µm column) using a gradient of sodium citrate buffers at 0.70 ml/min flow rate.

2.2.3. *Fourier*-transform infrared spectroscopy (FTIR) of *Trichilia* and soybean protein concentrates

The FTIR analysis of *Trichilia* and soybean protein concentrates was performed according to Byler & Susi (1986); Choi & Ma (2005); and Sadat & Joye (2020), with modification, using a Perkin Elmer FTIR Spectrometer (Spectrum Two, Version 10.5.4; Perkin Elmer Inc, Llantrisant, UK). The spectra were taken in the spectral range of 4000 cm^{-1} to 700 cm^{-1} with a 4.0 cm^{-1} resolution, by accumulation of 32 scans. The data was baseline corrected using the Perkin Elmer Spectra100 software. The second derivative of the spectral region 1700 cm^{-1} to 1600 cm^{-1} was used to empirically identify the positions of the hidden constitutive Amide I band peaks using Origin Pro-software (OriginLab Corporation). The areas of the identified constitutive peaks were determined by deconvolution of the amide I band with fitting to Lorentzian distributions. The major component peaks of the amide I band for the *Trichilia* (*emetica* and *dregeana*) and soybean proteins were assigned as; 1630 cm^{-1} for antiparallel β-sheets, 1682 cm^{-1} for parallel β-sheet, 1580-1615 cm^{-1} for side-chains, 1642-1657cm^{-1} for random coils, 1662-1686 cm^{-1} for β-Turns, 1643–1645 cm-1 for random coils and 1648-1657

cm^{-1} for α-helix, based on Shevkani et al. (2019; Achouri et al. (2012); and Barth (2007). The peak area of each secondary structure was reported as a percentage of the total peaks' area.

2.2.4. Functional properties of *T. emetica, T. dregeana* and soybean protein concentrate

The functional properties of the protein concentrate of the *Trichilia spp* in comparison with soybean protein concentrates were determined in terms of water holding capacity (WHC), oil holding capacity (OHC), emulsion activity index (EAI), foaming capacity (FC) and foam stability (FS). The WHC and OHC were determined according to Chakraborty, (1986) with modifications. For WHC, accurately, 1 g of a sample was dispersed in 10 ml distilled water and vortex for 1 min at high speed. The suspension was centrifuged (2200 g for 30 min at 25°C) and the WHC was calculated using the formula:

$WHC= (W_2-W_1)/W_0 X100$

Where W_0 is the weight of the dry sample (g)

W_1 is the weight of the dry sample with the tube (g)

W_2 is the weight of the sediment with the tube (g)

For OHC, 1 g of the sample (W_0) was mixed with 10 ml vegetable oil or corn oil with vortexing every 1 min for 10 min. The mixture was centrifuged (1600 g for 10 min) and free oil was decanted by venting the tube for 1 h. The OHC was calculated using the formula: $OHC= (W_2-W_1)/W_0 \times 100$

Where W_0 is the weight of the dry sample (g)

W_1 is the weight of the dry sample with the 15 mL tube (g)

W_2 is the weight of the sediment with the tube (g)

The EAI was determined according to Pearce & Kinsella, (1978) with modifications. Samples of about 0.8 g was dispersed into 40 ml distilled water and homogenized (Polytron, PT 1300D, 1000 g) for 2 min. A vegetable oil (40 ml) was added to the suspension and further mixed using a homogenizer for 1 min. The mixture was then centrifuged for 5 min at 1200 g. EAI was calculated using the formula:

$$\text{EAI } (\%) = \frac{\text{Height of emulsification layer}}{\text{Height of the content of the tube}} X100$$

The FC and FS were done according to Sathe & Salunkhe, (1981). One gram sample was dispersed into 100 ml distilled water to make 1% (w/v) solution. The pH was adjusted using 0.1 N NaOH and the mixture was homogenized at 1000 g for 2 min. The volume of the foam was measured at 0 min, 15 min, 30 min and 60 min.

$$\text{FC } (\%) = \frac{\text{Vol after homogenization} - \text{Volume before homogenization}}{\text{Volume before homogeniz}ation} X100$$

$$\text{FS} = \frac{\text{Vol at set time}}{\text{Initial vol of foam}} \ X100$$

2.2.5 Statistical analysis

Amino acid profile, secondary structure and functional properties analysis data were analysed by one-way analysis of variance (ANOVA) using Statistica® version 8 (Statsoft Inc, Tulsa). The mean values of amino acid profile, secondary structure and functional properties analyses of

T. emetica, *T. dregeana* and soybean protein were compared by the Fishers Least Significant Difference test (LSD) with a 95% confidence interval.

3. Results and Discussion

3.1. Amino acid profile of *Trichilia emetica* and *Trichilia dregeana* protein concentrates

The amino acid content of the protein concentrates were reported a percentage of the protein content, which was 60.9, 54.0 and 69.4 % w/w for *T. dregeana*, *T. emetica* and soybean respectively (Table 1). The hydrophobic amino acids (HAA), negatively charged amino acids (NCAA), aromatic amino acid (AAA) and branched chain amino acid (BCAA) were not significantly different ($p>0.05$) between the two Trichilia proteins (Table 2). The values of the HAA, NCAA, PCAA, AAA and BCAA proportions for the two oil seeds were similar at various extents to those reported in other oils seeds such chia, sunflower and flaxseed to various extents (Coelho et al., 2018; Bautista et al., 1996; Nwachukwu & Aluko, 2018). The *Trichilia* spp proteins had significantly higher ($p<0.05$) BCAA than soy proteins while their NCAAs were not significantly different ($p>0.05$) from those of soy proteins (Table 2). The PCAA was significantly ($p<0.05$) different between the two Trichilia proteins with *T. dregeana* protein PCAA being similar to that of soy proteins. The *Trichilia spp* proteins contained substantial amounts of sulphur-containing amino acid which are limiting amino acids in other legumes. The SCAAs were not significantly ($p>0.05$) different between *T. dregeana* and *T. emetica* proteins but the SCAAs for both were significantly ($p< 0.05$) lower than for soy proteins (Table 2), due to lower methionine levels (Table 1). The present SCAA results for *T. dregeana* and *T. emetica* were comparable to those of sunflower (3.75 g/100g protein), chia (4.2-4.6 g/100g)

and flaxseed (2.4-3.7 g/100g) protein (Bautista et al., 2018; Coelho et al., 2018; Nwachukwu & Aluko, 1996).

Essential amino acids (EAA) are amino acids that must be obtained from the diet or supplements and non-essential amino acids are amino acids that are synthesized by the human body. The proportion of essential amino acids (E) to total (T) amino acids was not significantly different ($p > 0.05$) between both Trichilia proteins and soy protein (Table 2). These results implied that the *Trichilia* seed proteins could serve as substitutes for soy proteins in both food and animal feed applications. *In-vivo* feeding studies are however required to verify the substitution effects.

Although the total essential amino acids *T. emetica* and *T. dregeana* were above the chia, flax, sunflower protein concentrate and sunflower (Bautista et al., 1996; Coelho et al., 2018; Nwachukwu & Aluko, 2018), like soy, they are below the FAO/WHO dietary recommendation for children (Table 3). Essential amino acids help in calcium absorption, boost muscle growth and play an important role in collagen formation (Civitelli et al., 1992; Bifari & Nisoli, 2017). Like soybean, and *Trichilia species* protein concentrate had lower amino acids score for the essential amino acids (Thr, Lys, Met, Val, Ile, Leu, Trp and Phe+ Tyr) than the recommended the FAO score for children (FAO/WHO, 2007) (Table 3). For children, the *Trichilia* protein may have to be composited with other protein sources to improve its nutritional quality.

For adults, the essential amino acids scores for *Trichilia species* and soybean protein concentrate for adults were higher than the FAO requirements for adults (FAO/WHO, 2007) (Table 2). The essential amino acids scores of the Trichilia seed proteins for both children and adults (His, Ile, Leu, Val, Lys, Phe and Tyr) were higher than chia, sunflower and flaxseed proteins (Hughes et al., 2011; Bautista et al., 1996; Coelho et al., 2018; Nwachukwu & Aluko,

2018). However, like soy protein, Threonine (Thr) was the limiting amino acid for both *Trichilia* protein (Table 3) unlike in chia, sunflower and flaxseed limited in lysine was the limiting amino acid for children and adults (Table 3). The associated chemical scores based on the Thr were 100% (AAS×100, truncated to 100% if >100) and were not significantly different ($p>0.05$) between the two Trichilia seed proteins. Given the soy protein chemical score was also 100%, is deemed that the two Trichilia proteins had a similar chemical score to soy proteins although the AAS was significantly lower than that of soy proteins. This further elucidated the potential of the two Trichilia proteins in replacement of soy proteins in food and feed applications.

Non-essential amino acids (NEAA) can be deemed as essential amino acids during illness or stress and are called conditional essential amino acids (Wu, 2009). The non-essential amino acids (NEAA) for *T. dregeana* were not significantly ($p>0.05$) different from those of *T. emetica* except Arg and Trp. The NEAA content of the two Trichilia spp proteins were comparable to those of soybean proteins to various extents (Table 4) with only Glutamic acid being significantly higher in soybean proteins alone. The results for the non-essential amino acids followed the similar trends as chia and sunflower protein in terms (Coelho et al., 2018; Nwachukwu & Aluko, 2018).

3.2. *Fourier*-transform infrared spectroscopy of *Trichilia emetica* and *Trichilia dregeana*

The FTIR spectra for *T. emetica*, *T. dregeana* and soybean protein concentrate is shown in Fig 1. A peak at lower wavenumber (1060 cm^{-1}) was observed for *Trichilia* and Soybean protein concentrate IR spectra (Fig 1). This peak is associated with intermolecular sheet interactions which support the oligomeric nature of the protein structure (Aider & Barbara, 2011). However, the peak for *T. emetica* protein concentrate intermolecular sheet interaction

showed a higher peak intensity compared with *T. dregeana* and soybean protein concentrate. The increase in peak intensity may be attributed with the higher bond polarity, repetition of the same functional group which leads to larger and intense peak; and the relative of two overlapped bonds. The *Trichilia* protein concentrate (*T. emetica* and *T. dregeana*) intermolecular sheet interaction showed higher peak intensity compared with the soybean protein concentrate. The high peak intensity of *Trichilia* protein concentrate intermolecular sheet interactions than soy protein concentrate may suggest that the *Trichilia* protein concentrate may have a more stable oligomeric nature of the protein than soybean.

There were other five typical peaks for *Trichilia* and soybean protein concentrates IR spectrum which were observed. These peaks were attributed with Amide I, II, III, A and B bands which derived from bending and stretching vibration of the molecular bonds (Byler & Susi, 1986). Generally, the Amide I band (1631 cm^{-1}) showed an increase in peak intensity for *Trichilia* and Soybean protein concentrates compared with other Amide band II, III, A and B (See Fig 1). This band is due to the C=O bonds stretching vibrations (Litvinov et al., 2012 & Riaz et al., 2018). The *T. emetica* protein concentrate showed an increase in peak intensity compared with soybean and *T. dregeana* protein concentrate. However, *T. dregeana* showed an increased in peak intensity than Soy protein concentrate (Fig 1). Amide II (1488 cm^{-1} and 1535 cm^{-1}) observed in this study may be due to the bending vibrations of N-H bonds coupled with C-N stretching vibrations (Litvinov et al., 2012 & Riaz et al., 2018) in the *Trichilia* and Soybean protein concentrate, Amide III (1233 cm^{-1} and 1393 cm^{-1}) for *Trichilia* and Soybean protein may also be due to the stretching vibrations of C-N coupled with N-H bending vibrations with weak C-C stretching and C=O bending vibrations. The chia protein concentrates had previously shown similar Amide II (Coelho et al., 2018). Amide A (3287 cm^{-1} for the current

protein) reported in this study may be caused by the vibration of O-H group with a carbonyl group of the peptide chain and the Amide B (2843 and 2915 cm^{-1}) which may be related to the stretching vibrations of CH_2 (Litvinov et al., 2012; Riaz et al., 2018; Jackson & Mantsch, 1995).

The Amide band I for *T. emetica*, *T. dregeana* and Soybean had a strong and broad peak (Fig 1). This increase in peak intensity may be attributed to a greater concentration of the amine groups for the *T. emetica*, *T. dregeana* and Soybean protein concentrate. Bambara and cowpea protein concentrates had previously shown high Amide I band of about 1633 cm^{-1} (Mune-Mune & Sogi, 2016) which agreed with the current results. However, the chia, sunflower, bambara and cowpea protein concentrate had previously shown Amide II bands with peak intensity of about 1549 cm^{-1}- 1551 cm^{-1}; 1580-1480 cm^{-1} and 1550 cm^{-1} (Coelho et al., 2018; Ishii et al., 2021; Mune-Mune & Sogi, 2016) which were within range with the *Trichilia* and Soybean protein concentrate. The presence of Amide B and II adsorption may indicate that some of the peptide bonds Hydrogen atoms have not completely been exchange deuterium. This may be due to the compact inaccessible form of some native folded protein molecules (Clark et al., 1981 & Kavanagh et al., 2000). Both C=O and N-H bonds are involved in the H-bonding which takes place between the different elements of the secondary structure. Therefore, the locations of both the Amide I (C=O) and Amide II (N-H) bands are sensitive to the secondary structure content of a protein (Goormaghtigh et al., 2006). However, due to the complex vibrations of multiple functional groups, the Amide II bands are less useful for the protein structure prediction (Jackson & Mantsch, 1995).

The deconvoluted Amide I band IR spectra for the *Trichilia spp* and soybean protein concentrate with several bands were shown in Fig 2. The data revealed that the Amide I band for *T. emetica*, *T. dregeana* and soybean protein Amide I band consisted of nine to eleven major component peaks which were found in all the IR spectra (Fig 2). The β-sheet and α-helices provide strength and rigidity to the protein structure (Carbonaro et al., 2012; Gropper et al., 2009 &; Shivu et al., 2013). The intra-molecular β-sheet holds the 3-dimension tertiary structure of the protein together; inter-molecular β-sheet stabilizes the secondary structure into a three-dimensional fold, whilst the side chain plays a crucial role in designing the three-dimensional conformation of the protein (Hettiarachchy et al., 2012).

The relative proportion of the different structures in the *T. emetica*, *T. dregeana* and soy protein concentrates is shown in Fig 3. The secondary structure of proteins plays an important role in understanding the nutritional quality of the protein and functional properties such as texture, protein availability and protein digestibility (Zhu et al., 2022). The secondary structure of *T. emetica* protein had a relatively greater proportion of the β-sheet than the secondary structure of *T. dregeana*. However, the *T. dregeana* had a significantly ($p< 0.05$) lower proportion of the β-sheet in its secondary compared with the secondary structure of soy protein concentrates (Fig 3). The lower proportion of the β-sheet in the secondary structure of *T. dregeana* protein may suggest that the protein was partially unfolded during extraction. The unfolding of the protein could cause a decrease in the foaming capacity (FC) and an increase in the emulsion activity index (EAI). Mune-Mune & Sogi (2016) previously reported that different cowpea and Bambara proteins dried using different methods caused partial unfolding which lead to a decrease in β-sheet aggregates. The results of the relative proportion of the β-sheet in the secondary structure of *Trichilia spp* protein and soybean

protein are in agreement with what had been reported for cowpea (41-42%) (Mune-Mune & Sogi, 2016) and Bambara (40-42%) (Mune-Mune & Sogi, 2016). However, sunflower protein (30.7%) and walnut protein (11%) had a lower proportion of the β-sheet. On the other hand, canola protein had more proportion of the β-sheet compared with the current results (Gerzhova et al., 2015; Ishii et al., 2021; Mao & Hau, 2014).

The α-helices for the *Trichilia* protein were not significantly different ($p > 0.05$) from those of soy protein (Fig 3). The proportion of α-helices in walnut protein was previously observed to be more than the current results for the *Trichilia spp* and soybean protein (Mao & Hau, 2014). However, canola protein and sunflower protein had a lesser proportion of the α-helices compared with the *Trichilia spp* protein (Gerzhova et al., 2015 &; Ishii et al., 2021).

The relative proportion of the β-turns in the secondary structure of *T. emetica* was significantly ($p < 0.05$) lower than in the secondary structure of the *T. dregeana* protein. *T. dregeana* β-turns were significantly ($p < 0.05$) more than the soy too (Fig 3). Other oilseeds such as walnut (23.3%), sunflower (14.4-14.6%), cowpea (17-25%) and Bambara (18-24%) protein concentrate had more β-turns compared with the current results for *Trichilia spp* and soy (Mao & Hau, 2014; Ishii et al., 2021; Mune-Mune & Sogi, 2016).

The random coil and side chains for *T. emetica* and *T. dregeana* were not significantly different ($p > 0.05$) (Fig 3). However, soy protein concentrate had a significantly ($p<0.05$) lower random coil compared with the *Trichilia spp*. The side chain for soy protein was significantly ($p < 0.05$) higher than the *Trichilia spp* (Fig 3). Mao & Hua, (2014) reported walnut protein concentrate had a higher proportion of the α-helix (34.9%) and random coil (32%) which was different from the current observation for the *Trichilia spp* and soybean protein concentrates. Sunflower, canola, Bambara and cowpea protein concentrate had previously been observed

to have a more β-sheet proportion (Ishii et al., 2021; Gerzhova et al., 2015; Mune-Mune & Sogi, 2016) which shows a similar trend to current results observed for the *Trichilia spp* and soy protein. The present results indicate that the *T. emetica* and *T. dregeana* protein like most of the other oilseed (rapeseed, sunflower and soybean) protein concentrate had more β-conformations proportion than the α-helix with a more ordered secondary structure. The β-conformation is associated with the protein globular structure; this may imply that the *Trichilia emetica* and *T. dregeana* have a globular structure which may cause a lower digestibility and improved functional properties such as water holding capacity and foaming capacity.

3.3. Functional Properties of *T. emetica* and *T. dregeana* in comparison with soybean protein concentrates

The water holding capacity (WHC), oil holding capacity (OHC), emulsion activity index (EAI) foaming capacity (FC) and foaming stability (FS) results for *T. emetica* and *T. dregeana* in comparison with soybean protein concentrates are shown in Table 5. High WHC and OHC contribute to flavour retention, desirable texture and mouthfeel; and thereby enhance the organoleptic acceptability of the food product (Hutton & Campbell, 1981). The WHC is essential in viscous foods such as soup, dough, baked foods and custard (Sreerama et al., 2012), while the high OHC is required in baked foods, meat replacers, emulsions (e.g., ice cream) and extenders in which the oil contributes to the texture of the finished food product (Shevkani et al., 2019).

The *T. emetica* had a significantly ($p< 0.05$) higher WHC compared with *T. dregeana* followed by soy protein concentrate. This observation is similar to that of HAA in *T. emetica* (See Section 3.1). The WHC results for the current study are in agreement with the results reported for walnut, canola, Cowpea and Bambara groundnut protein (Mao & Hua, 2014; Gerzhova et al., 2015; Mune-Mune & Sogi, 2016). However, previous studies showed that hemp, soy and sesame protein concentrate had a higher WHC than the samples of the current study (Malomo & Aluko, 2015; Mao & Hua, 2014; Teh et al., 2014; Gerzhova et al., 2015; Adeleke et al., 2018 & Liu et al., 2019). The higher WHC of *T. emetica* protein may be due to the presence of exposed surface hydrophilic groups in the amino acid side chains. The *T. emetica* protein had a significantly ($p< 0.05$) lower OHC compared with *T. dregeana* protein. Soybean protein had a significantly ($p> 0.05$) higher OHC compared with *T. emetica* and *T. dregeana* (Table 5). The higher OHC of soy protein concentrate could be due to a higher proportion of hydrophobic group and polar amino acids on the surface, which bind the aliphatic chains of the oil (Kumar et al., 2022; Kaur & Singh, 2005; Shevkani et al., 2019). The current results are in agreement with the amino acids profile results, the soybean protein had a higher proportion of hydrophobic amino acids, such as Leu, Ileu, Pro, Gly, Val and Phe than the *Trichilia species* (See Section 3.1). Peanut, canola, sesame and chia protein had a higher OHC which is similar to the results reported for *T. dregeana* (Yu et al., 2007; Gerzhova et al., 2015; Escamilla-Silva et al., 2003 & Coelho et al., 2018).

The *T. emetica* and *T. dregeana* protein had shown no significantly different ($p> 0.05$) for EAI. However, soy protein concentrates had significantly ($p< 0.05$) higher EAI compared with *Trichilia spp* (Table 5). The high EAI may be related to the high presence of hydrophobic amino acids in the soy protein concentrates (See Section 3.1). This current result for high EAI is in

agreement with other oilseeds such as soy protein concentrate and peanuts (Rao et al., 2002; Liu et al., 2019 &; Yu et al., 2007).

Foaming capacity relay on the ability of the protein to diffuse to the interface and form a viscous film without excessive aggregation whilst foaming stability is the ability of the protein to maintain the foam (Shevkani et al., 2019). *Trichilia emetica* had a significantly ($p< 0.05$) higher foaming capacity (FC) and stability (FS) compared with *T. dregeana*. Soy protein concentrate had significantly ($p< 0.05$) lower FC and FS compared with the *Trichilia spp* (Table 5). The high FC for *Trichilia protein* is not in agreement with the hydrophobic amino acids data, however, in agreement with the compact structure for the *Trichilia protein*. The high FC and FS for the *Trichilia* protein concentrate could be due to the balanced hydrophobic and hydrophilic side chains in the amino acids of the protein concentrate. The results from this current study indicate that the *Trichilia spp* protein concentrate could be used as a food ingredient to improve foaming capacity requirements in products such as baked products, and dairy products (egg ice cream).

Principal Component Analysis showing the relationship between the amino acid content, secondary structure and functional properties of *Trichilia* seeds and soybean protein

A correlation based PCA was done to investigate the relationship between protein chemical structure, amino acids and functional properties of *Trichilia emetica*, *Trichilia dregeana* and soybean protein (Fig 4). The first two PC score plots explained 92.5% of the total variation with PC1 explained 56% of the total variation. Thus, this suggests that the random coil, BCAA and β-turns had a large positive loading on principal component 1 (PC1) which imply that this

component was effective in defining the chemical structure of the protein. However, OHC, SCAA, HAA and side chain had larger negative loading on PC2 (Fig 4). Therefore, this could imply that the component measured the chemical structure, amino acids and functional properties of the protein concentrates.

The clustering of variables on a correlation PCA plot indicates positive correlation between the variables. The oil holding capacity (OHC) had a positive correlation with side chain structures, SCAA and the HAA (left-side of PC1, Fig 4). This indicated that the protein structure may contain more of the polar amino acids on the side chain which could bind with the protein. In this current study, the PCA results agreed with the soy protein amino acids and chemical structure which was found to have significantly ($p< 0.05$) higher HAA, SCAA and side chains.

The EAI correlated with the AAA (Fig 4). The EAI was previously observed to be positively correlated with surface hydrophobicity (Shevkani et al., 2015). Thus, surface hydrophobicity would be due to increased exposure of hydrophobic amino acids during the partial unfolding of the protein concentrates (Shevkani et al., 2015; Mune-Mune & Sogi, 2016). As has been described before soy protein had higher EAI with an increase in β-sheet and AAA compared with *Trichilia* protein.

The OHC had a positive correlation with the side chain of the protein structure, SCAA and the HAA on the left-side of PC1 (Fig 4). This indicated that the HAA and SCAA side chains were probably anchored in the hydrophobic oil phase and facilitated the binding of the oil. This agreed with the soy protein amino acid content (Table 2) and secondary structure (Fig 3) which showed significantly ($p> 0.05$) higher HAA, SCAA and side chains respectively, compared to the *Trichilia spp* proteins (Fig 4).

The PCA showed that the random coil and BCAA were positively correlated (Fig 4). This could imply that the random coil segments could have been mostly made up of branched amino acids. The random coil and BCAA were opposite to EAI and OHC on PC1 (Fig 4). This indicated that the random coil and BCAA were negatively correlated with EAI and OHC. The BCAA containing random coil segments lead to increased segmental flexibility (Mune-Mune & Sogi, 2016), which probably limited interphase dispersion of the protein at expense of intra-phase dispersions.

The EAI was closely related with the AAA (Fig 4). The EAI was previously observed to be positively correlated with surface hydrophobicity (Shevkani et al., 2015). Thus, surface hydrophobicity would be due to increased exposure of hydrophobic amino acids during the partial unfolding of the protein concentrates (Shevkani et al., 2015; Mune-Mune & Sogi, 2016). As has been described before that soy protein had higher EAI with an increase in β-sheet and AAA compared with *Trichilia* protein.

Water holding capacity (WHC) was positively correlated with anti-parallel β-sheet aggregates (Fig 4). This implies that the protein structure had decreased in partial unfolding which lead to a decrease in flexibility and buried hydrophobic amino acids. However, negatively related to the α-helix and β-turns on the far end opposite of PC2 (Fig 4). Shevkani et al. (2015) had previously observed that the WHC was positively related to the protein aggregates structure of the protein.

Foaming capacity (FC) was diagonally opposite to the NCAA and PCAA on the far opposite end of PC2. This implies that FC may be negatively related to the NCAA and PCAA. Mune-Mune & Sogi (2016) observed a negative correlation between the unordered structure (random coil) and FC. An increase in random coils probably causes a decrease in FC due to increase in

segmental flexibility of the protein which probably led to decreased ability to form durable foams. The samples which include soy, *T. dregeana* (TD) and *T. emetica* (TE) were widely separated. This could imply that the samples were different in protein chemical structure arrangement, amino acids sequencing and composition hence the samples had different effects on the protein functionality. *Trichilia emetica* was closely related to the FC whilst *T. dregeana* was related to the β-turns and soy with the side chains and intra-molecular β-sheet.

Variables modeling power closer to one is regarded as more relevant to the model. For PC2, the variables power values of the chemical structure, amino acids, and functional properties in the decreased order: FC (0.941)> anti-parallel β-sheets aggregates (0.937)> WHC (0.922)> EAI (0.920)> AAA (0.905)> intra-molecular β-sheet (0.505). The EAI, FC and WHC were closely related to the AAA and anti-parallel β-sheets aggregates. The PCA indicated that the EAI, FC and WHC increase with the increase in AAA and more anti-parallel β-sheets aggregates; however, an increase in intra-molecular β-sheets could lead to a decrease in EAI, FC and WHC.

4. Conclusion

The present study elucidated the amino acids profiles, functional properties, and secondary structure of *T. emetica and T. dregeana* oilseed proteins. The essential amino acid contents of the *Trichilia* seed proteins were three to four times greater than the recommended FAO/WHO standard for adults. The results from this study demonstrated a similarity in the protein chemical structure between the two *Trichilia* seed proteins, and soybean proteins. The *Trichilia* protein secondary structure results suggest that the two species had more β-conformations than α-helix as in soybean protein. The presented research will facilitate the utilization of the Trichilia seed proteins by food and feed processors, and in nutritional

intervention . Given the amino acid profile impacts the nutritional quality of the proteins and digestibility, future research must be conducted to determine digestibility and anti-nutritional factors which are important from a nutritional perspective.

Acknowledgements

The authors would like to thank the Agricultural Research Council-Tropical and Subtropical South Africa and University of KwaZulu-Natal for the support and funding of the research.

Funding: This study was funded by National Research Foundation (NRF) (Grant Number 98687), Agricultural Research Council Tropical and Subtropical crops (ARC-TSC).

Code of availability: No code available.

Data Availability statement

Data available on request from the authors.

Authors Contributions

All authors contributed to the study as follows: Gugu Tsomele (experimental runs, data analysis, manuscript writing), Belinda Du Plessis, Dr Tonna Anyasi, Prof Victor Mlambo, Prof Eric Amonsou, Sello Lepule, Prof Mthulisi Siwela (student supervision, manuscript review), Dr Obiro Wokadala (research conceptualization, experimental design, student supervision, manuscript review).

Ethics and Integrity

Ethics approval was not required for this research.

Conflict of Interest

The authors have no conflict of interest to declare.

References

Achouri, A., Nail, V., and Boye, J.I. (2012). Sesame protein isolate: Fractionation, secondary structure and functional properties. *Food Research International Journal*, 46, 360-369

Adeleke, O.R., Adiamo, O.Q., & Fawale, O.S. (2018). Nutritional, physicochemical and functional properties of protein concentrate and isolate of newly-developed Bambara groundnut cultivars. *Food Science and Nutrition*, 6, 229-242

Aider, M., & Barbana, C. (2011). Canola proteins: Composition, extraction, functional properties, bioactivity: Applications as a food ingredient and allergenicity- a practical and critical review. *Trends in Food Science and Technology*, 22, 21-39

Barrios-Peralta, P., Perez-Won, M., Tabilo-Munizaga, G., & Briones-Labarca, V. (2012). Effect of high pressure on the interactions of myofibrillar proteins from abalone (*Haliotis rufencens*) containing several food additives. *LWT-Food Science and Technology*, 49 (1), 28-33

Barth, A., 2007. Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1767(9), 1073-1101.

Bautista, J., Hernandez-Pinzon, I., Alaiz, M., Parrado, J., and Millan, F. (1996). Low molecular weight sunflower protein hydrolysate low concentrate in aromatic amino acids. *Journal of Agriculture Food Chemistry*, 44 (4): 967-971

Bidlingmeyer, B.A., Cohen, S.A., and Tarvin, T.L. (1984). Rapid analysis of amino acids using pre-column derivatization. *Journal of Chromatography*, 336: 93-104

Bifari, F., & Nisoli, E. (2017). Branched-chain amino acids differently modulate catabolic and anabolic states in mammals: A pharmacological point of view. *British Journal of Pharmacology*, 174 (1), 1366-1377

Boye, J.I., Aksay, S., Roufik, S., Ribereau, S., Mondor,M., Farnworth, E. (2010). Comparison of the functional properties of pea, chickpea and lentil protein concentrates processed using ultrafiltration and isoelectric precipitation techniques. *Food Research International*, 43, 537-546

Byler, D.M., & Susi, H. (1986). Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers Journal*, 25, 469-487

This reference was one of the first to provide a foundation for analysis of protein secondary structure using FTIR. The article gives guidelines on deconvolution of the Amide I band. The analysis of the secondary structure of the proteins was essential to the work in the present manuscript.

Carbonaro, M., Maselli, P., & Nucara, A. (2012). Relationship between digestibility and secondary structure of raw and thermally treated legume protein: A Fourier transform infrared (FTIR) spectroscopic study. *Amino acids Journal*, 43, 911-921

Chakraborty, P. (1986). Coconut protein isolate by ultrafiltration. *Food engineering and process applications*, 226-242

Chandrapala, J.,Zisu, B., Kentish, S., & Ashokkumar, M. (2012). The effects of high-intensity ultrasound on the structural and functional properties of α-Lactalbumin, β-Lactoglobulin and their mixtures. *Food Research International Journal*, 48 (2), 940-943

Choi, S.M., & Ma, C. –Y. (2005). Conformational study of globulin from common buckwheat (*Fagopyrum esculentum Moench*) by Fourier transform infrared spectroscopy and differential scanning calorimetry. *Journal of Agricultural and Food Chemistry*, 53, 8046-8053

Chove, B.E., Grandison, A.S., and Lewis, M.J., (2001). Emulsifying properties of soy protein isolate fractions obtained by isoelectric precipitation. *Journal of the Science of Food and Agriculture*, 81, 759-763

This article provided a reference method for isolation of proteins form oil seeds by isoelectric precipitation. The reference is used specifically because it deals with high oil seeds like soybean which is used as a reference seed in the present manuscript.

Civitelli, R., Villareal, D.T., Agnusdei, P., Nardi, P., Avioli, L.V., Gennari, C. (1992). Dietary L-lysine and calcium metabolism in humans. *Nutrition*, 8 (6), 400-405

Clark, A.H., Saunderson, D.H.P., & Suggett, A. (1981). Infrared and laser Raman spectroscopic studies of thermally induced globular protein gels. *International Journal of Peptide and Protein Research*, 17, 353-364

Coelho, M.S., Salas-Mellado, M.M. (2018). How extraction method affects the physicochemical and functional properties of chia protein. *LWT-Food Science and Technology*, 96, 26-33

This article studied chia seed protein. Chia seed protein is an appropriate oil seed protein for comparison with the Trichilia spp seed proteins in the present research.

Day, L. (2013). Proteins from land plants-potential resources for human nutrition and food security. *Trends in Food Science and Technology*, 32, 25-42

Day, L., Cakebread, J.A., & Loveday, S.M. (2022). Food proteins from plants and animals: Differences in nutritional and functional properties. *Trends in Food Science and Technology*, 119, 428-442

Elias, R.J., Mcclements, D.J., and Decker, E.A. (2005). Antioxidant activity of cysteine, tryptophan and methionine residues in continuous phase β-lactoglobulin in oil-in-water emulsion. *Journal of Agricultural and Food Chemistry*, 53 (26), 10248-10253

Escamilla-Silva, E.M., Guzman-Maldonado, S.H.,Cano-Medinal, A., & Gonzalez-Alatorre, G. (2003). Simplified process for the production of sesame protein concentrate: Differential scanning calorimetry and nutritional, physicochemical and functional properties. *Journal of the Science of Food and Agriculture*, 83 (9), 972-979

FAO/WHO. Protein and amino acid requirements in human nutrition, Report of a Joint WHO/FAO/UNU Expert Consultation, WHO technical report series 935. *Food and Agriculture Organization/ World Health Organization*, Geneva, Switzerland, 2007.

Gehrke, C.W., Sr, W. L. L., Absheer, J.S. (1985). Sample preparation for chromatography of amino acids: Acid hydrolysis of protein. *Journal- Association of Official Analytical Chemists*, 68 (5): 811-821

Gerzhova, A., Mondor, M., Benali, M., & Aider, M. (2015). Study of the functional properties of canola protein concentrates and isolates extracted by electro-activated solutions as non-invasive extraction method. *Food Bioscience*, 12, 128-138

Goormaghtigh, E., Ruysschaert, J.M., and Raussens, V. (2006). Evaluation of the Information Content in Infrared Spectra for Protein Secondary Structure Determination. *Biophysical Journal*, 90, 2946-2957

Gropper, S.S., Smith, J.L., & Groff, J.L. (2009). Advanced nutrition and human metabolism. 6th Edition. Belmont, CA: Wadsworth Cengage Learning, Pp 183-209

Hadnadev, M.S., Hadnadev, T.R.D., Pojic, M.M., Saric, B.M., Misan, A.C., Jovanov, P.T., & Sakac, M.B. (2018). Progress in vegetable proteins isolation techniques: A review. *Food and Feed Research*, 44 (1), 11-21

Herman, M.A., She, P., Peroni, O.D., Lynch, C.J., & Kahn, B.B. (2010). Adipose tissue branched-chain amino acids metabolism modulates circulating BCAA levels. *Journal of Biological Chemistry*, 285,11348-11356

Hettiarachchy, N.S., Sato, K., Marshall, M.R., & Kannan, A. (2012). Food Proteins and peptides: Chemistry, functionality, interactions, and commercialization. CRC Press

Hughes, G.J., Ryan, D.J., Mukherjea, R., and Schasteen, C.S. (2011). Protein digestibility-corrected amino acids scores (PDCAAS) for soy protein isolate and concentrate: Criteria for evaluation. *Journal of Agricultural and Food Chemistry*, 59, 12707-12712

Hutton, C.W., & Campbell, A.M. (1981). Water and fat absorption protein functionality in foods, Vol 147, USA: *American Chemistry Society*

Ishii, A.K., Pacioles, C.T., & Were, L. (2021). Colour and structural modifications of alkaline extracted sunflower protein concentrates and isolates using L-cysteine and glutathione. *Food Research International*, 147 (110574), 1-12

Jackson, M., & Mantsch, H.H. (1995). The Use and Misuse of FTIR Spectroscopy in the Determination of Protein Structure. *Critical Reviews in Biochemistry and Molecular Biology Journal*, 30 (2), 95-120

Jenkins, J.E., Sampath, S., Butler, E., Kim, J., Henning, R.W., Holland, G.P., et al., (2013). Characterizing the secondary protein structure of black widow dragline silk using solid-state NMR and X-ray diffraction. *Biomacromolecules Journal*, 14 (10), 3472-3483

Kaur, M., & Singh, N. (2005). Studies on functional, thermal and pasting properties of flours from different chickpea (*Cicer arietinum L*.) cultivars. *Food Chemistry*, 9, 403-411

Kavanagh, G.M., Clark, A.H., & Ross-Murphy, S.B. (2000). Heat-induced gelation of globular proteins: Part 3. Molecular studies on low pH β-lactoglobulin gels. *International Journal of Biological Macromolecules*, 28, 41-50

Kumar, M., Tomar,M., Potkule, J., Reetu, Punia, S., Dhakane-Lad, J., Singh, S., Dhumal, S., Pradhan, P.C., Bhushan, B., Anitha, T., Alajil, O., Alhariri, A., Amarowicz, R., Kennedy, J.F.

(2022). Functional characterization of plant-based protein to determine its quality for food applications: Review. *Food Hydrocolloids*, 123 (106986), 1-20

Landry, J., & Delhaye, S. (1992). Simplified procedure for the determination of tryptophan of foods and feedstuffs from burytic hydrolysis. *Journal of Agricultural and Food Chemistry*, 40 (5), 776- 779

Li, K., Kang, Z. –L., Zhao, Y. –Y., Xu, X. –L., & Zhou, G.H. (2014). Use of high-intensity ultrasound to improve functional properties of batter suspensions prepared from PSE-like chicken breast meat. *Food and Bioprocess Technology Journal*, 7 (12), 3466-3477

Litvinov, R.I., Faizullin, D.A., Zeuv, Y.F., Weisel, J.W. (2012). The α-Helix to β-Sheet Transition in Stretched and Compressed Hydrated Fibrin Clots. *Biophysical Journal*, 103, 1020-1027

Lui, J., Li,P., Jiang, Z., Yang, R., & Zhang, W. (2019). Characterization of peanut protein concentrates from industrial aqueous extraction processing prepared by spray and freeze drying methods. *International Journal of Food Science and Technology*,

Malomo, S.A., & Aluko, R.E. (2015). Conversion of a low protein hemp seed meal into a functional protein concentrates through enzymatic digestion of fibre coupled with membrane ultrafiltration. *Innovative Food Science and Emerging Technologies*, 31, 151-159

Mao, B., Tejero, R., Baker, D., & Montelione, G.T. (2014). Protein NMR structures refined with Rosetta have higher accuracy relative to corresponding X-ray crystal structures. *Journal of the American Chemical Society*, 136 (5), 1893-1906

Mao, X., & Hau, Y. (2014). Chemical composition, molecular weight distribution, secondary structure and effect on NaCl on functional properties of walnut (*Juglans regia L*.) protein isolates and concentrates. *Journal of Food Science Technologies*, 51 (8),1473-1482

Mao, X., & Hau, Y. (2012). Composition, structural and functional properties of protein concentrates and isolates produced from walnut (*Juglans regia L*.). *International Journal of Molecular Science*, 13, 1561-1581

Mune-mune, M.A., & Sogi, D.S. (2016). Emulsifying and foaming properties of protein concentrates prepared from cowpea and Bambara bean using different drying methods. *International Journal of Food Properties*, 19, 371-384

This article studied the relationships between functional properties of protein from cowpeas and Bambara nuts. The relationships observed in the article were compared to those in the present manuscript. This contextualized the results in the present research.

Nwachukwu, I.D., & Aluko, R.E. (2018). Physicochemical and emulsification properties of flaxseed (*Linum usitatissmun*) albumin and globulin fractions. *Food Chemistry Journal*, 255, 216-225

Orwa, C., Mutua, A., Kindt, R., Jamnadass, R., & Simons, A. (2009). Agroforestree Database: A tree reference and selection guide. Version 4

Pearce, K., & Kinsella, J. (1978). Emulsifying properties of proteins: evaluation of a turbidimetric technique. *Journal of Agriculture and Food Chemistry*, 26 (3), 716-723

Rao, A., Shallo, H.E., Ericson, A.P., & Thomas, R.L. (2002). Characterization of soy protein concentrates produced by membrane ultrafiltration. *Journal of Food Science*, 67 (4), 1412-1418

Riaz, T., Zeeshan, R., Zarif, F., Ilyas, K., Muhammad, N., Safi, S.Z., Rahim, A., Rizvi, S.A.A., and Rehman, I.U. (2018). FTIR analysis of natural and synthetic collagen. *Applied Spectroscopy Reviews Journal*, 53 (9), 703-746

Sathe, S., & Salunkhe, D. (1981). Functional properties of the great northern bean (*Phaseolus vulgaris L.*) proteins: emulsion, foaming, viscosity and gelation properties. *Journal of Food Science*, 46 (1), 71-81

Sreerama Y.N, Sashikala, V.B, Pratape V., & Singh V. (2012). Nutrients and antinutrients in cowpea and horse gram flour in comparison to chickpea flour: evaluation of their flour functionality. *Food Chemistry*, 131, 462-468

Shevkani, K., Singh, N., Kaur, A., and Rana, J.C. (2015). Structural and functional characterization of kidney bean and field pea protein isolates: A comparative study. *Food Hydrocolloids Journal*, 43: 679-689

Shevkani, K., Singh, N., Chen, Y., Kaur, A., & Yu, L. (2019). Pulse proteins: secondary structure, functionality and applications. *Journal of Food Science and Technology*, 56 (6), 2787-2798

Shivu, B., Seshadri, S., Li, J., Oberg, K.A., Uversky, V.N., & Fink, A.L. (2013). Distinct β-sheet structure in protein aggregates determined by ATR-FTIR spectroscopy. *Biochemistry Journal*, 52 (31), 5176-5183

Smith, L.J., Fiebig, K.M., Schwalbe, H., & Dobson, C.M. (1996). The concept of a random coil: Residual structure in peptides and denatured protein. *Folding and Design Journal*, 15 (1), R95-R108

Sun,H., Lu, G., Ren, S., Chen, J., & Wang, Y. (2011). Catabolism of branched-chain amino acids in heart failure: Insights from genetic models. *Pediatric Cardiology*, 32, 305-310

Tang, C.H., & Sun, X. (2011). A comparative study of physicochemical and conformational properties in three vicillins from Phaseolus legumes: Implications for the structure-function relationship. *Food Hydrocolloids Journal*, 25, 315-324

Teh, S.S., Bekhit, A.E.D., Came, A., & Birch, J. (2014). Effect of the defatting process, acid and alkaline extraction on the physicochemical and functional properties of hemp, flax and canola seed cake protein isolates. *Journal of Food Measurement and Charcterization*, 8, 92-104

Tsomele, G.F., Venter, E., Wokadala, O.C., Dlamini, B., Ngobese, N., & Siwela, M. (2021). Structural (gross and micro), Physical and Nutritional properties of *Trichilia emetica* and *Trichilia dregeana* seeds. *CyTA- Journal of Food*, 19 (1): 483-492

Van Wyk, B., Van Wyk, P., & Van Wyke, B. E. (2000). Photographic guide to trees of South Africa. Briza Publications, Pretoria, South Africa. ISBN-13:9781920217044

Vasconcelos, I.M., Maia, F.M.M., Farias, D., Campello, C.C., Carvalho, A.F.U., Moreira, R., & Oliveira, J.T.A. (2010). Protein fractions, amino acid composition, and antinutritional constituents of high-yielding cowpea cultivars. *Journal of Food Composition and Analysis*, 23 (1): 54-60

Wang, X-S., Tang, C-H., Yang, X-Q., & Gao, W-R. 2008. Characterization, amino acids composition and in vitro digestibility of hemp (Cannabis Sativa L) proteins. *Food Chemistry Journal*, 107, 11-18

Wang, K., Sun, D., Hongbin, P., & Wei, Q. (2017). Principles and applications of spectroscopic techniques for evaluating food protein conformational changes: A review. *Trends in Food Science and Technology Journal*, 63, 207-219

Wu, G. (2009). Amino acids: metabolism, functions and nutrition. *Amino acids*, 37, 1-17

Yu, J., Ahmedna, M., & Goktepe, I. (2007). Peanut protein concentrate: Production and functional properties as affected by processing. *Food Chemistry*, 103, 121-129

Zhang, Y., Zhau, W., Yang R., Ahmed, M.A., Hau, X., Zhang, W. (2013). Preparation and functional properties protein from heat-denatured soybean meal assisted by steam flash-explosion with dilute acids soaking. *Journal of Food Engineering*, 119 (1), 56- 64

Zhang, W., Waghmare, P. R., Chen, L., Xu, Z., & Mitra, S. K. (2015). Interfacial rheological and wetting properties of deamidated barley proteins. *Food Hydrocolloids*, 43, 400-409

Zhu, Z., Bassey, A. P., Cao, Y., Ma, Y., Huang, M. & Yang, H. (2022). Food protein aggregation and its application. *Food Research International*, 60, 111725

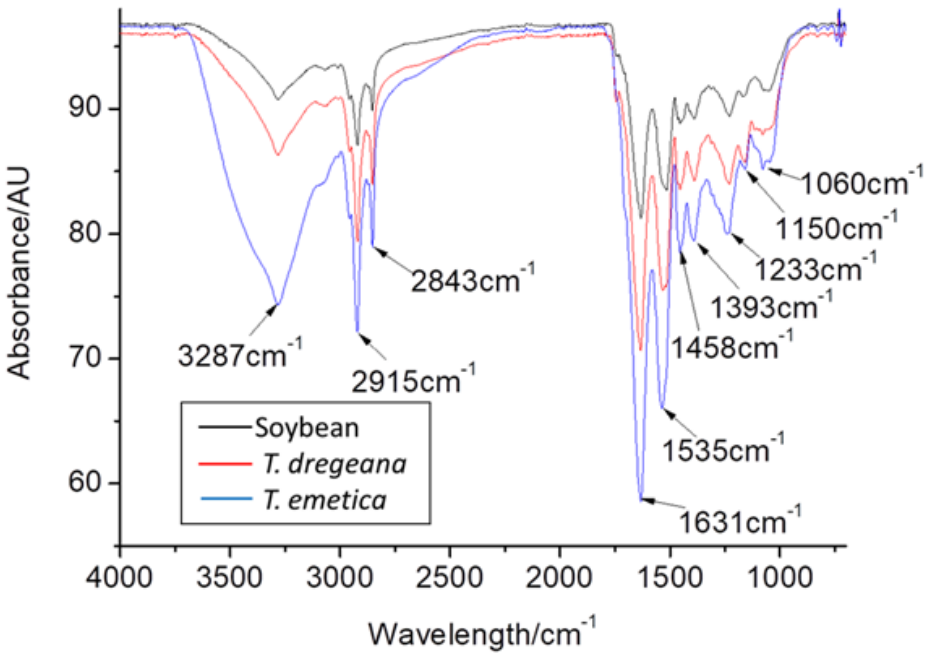

Fig 1: Illustration of Protein secondary structure of *T. emetica*, *T. dregeana* and Soybean using FTIR. The different peaks in the IR spectra represent the absorptions of different Amide bands within each protein concentrate

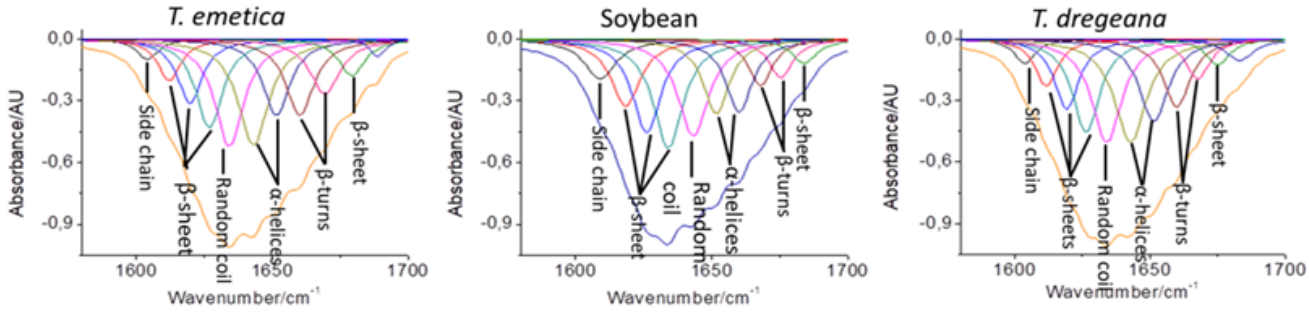

Fig 2: Deconvoluted *Trichilia* and soybean protein concentrates Amide I region represent peak-fitting of the secondary derived curves in the Amide I region from the IR spectra.

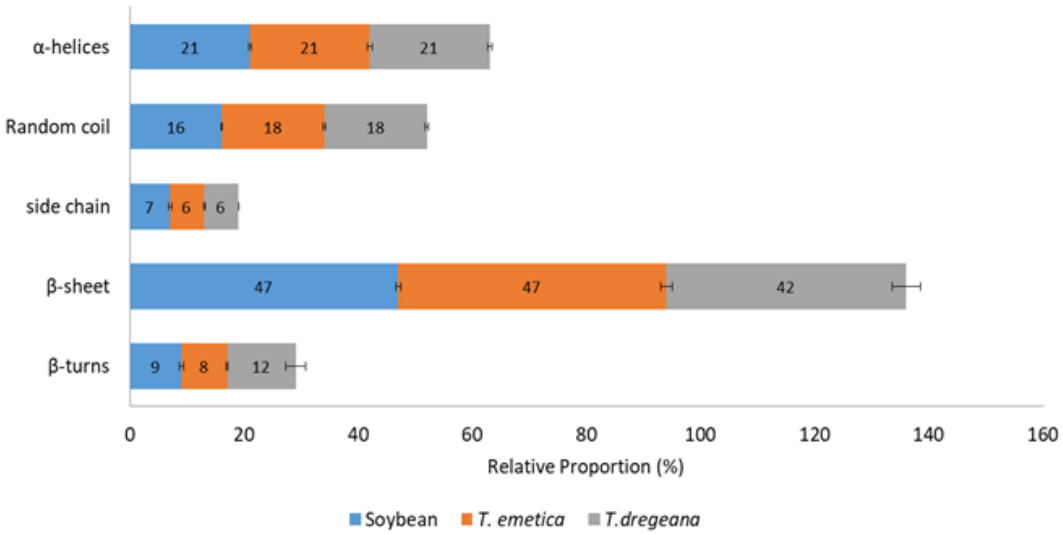

Fig 3: Relative proportion of different secondary structure in *T. emetica*, *T. dregeana* and Soybean protein concentrates from the Amide I IR Spectra. The combination/ addition of peak fitting derived areas for each secondary structure of FTIR Amide I (1600-1700 cm^{-1}) done according to Achouri et al., 2012 and Shevkani et al., 2019.

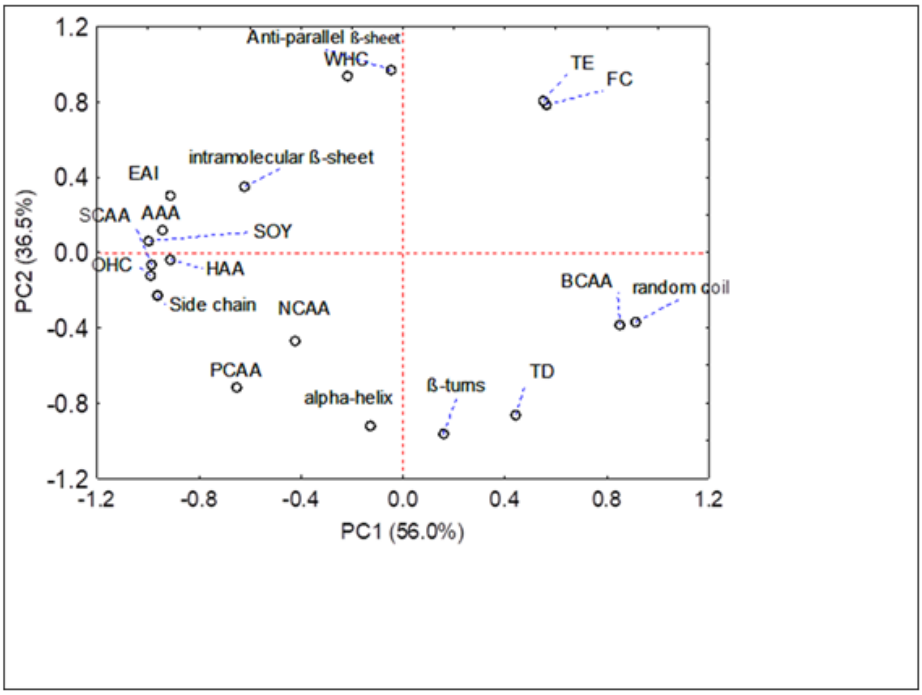

Fig 4: Principal component analysis plot describing the relationship between the amino acids classification composition, protein chemical structure and functional properties of *Trichilia emetica*, *Trichilia dregeana* and Soybean protein. AAA: aromatic amino acids; BCAA: branched amino acids; EAI: emulsion activity index; FC: foaming capacity; HAA: hydrophobic amino acids; NCAA: negatively charged amino acids; OHC: oil holding capacity; PCAA: positively charged amino acids; SCAA: Sulphur-containing amino acids; SOY: Soybean; TD: *Trichilia dregeana*; TE: *Trichilia emetica*; WHC: water holding capacity

Table 1: Amino acid profile (g/100g) of *T. dregeana* and *T. emetica* proteins in comparison with soybean proteins

Amino acid	*T. dregeana*	*T. emetica*	Soybean
Asp	15.24 ±0.16[a]	14.82 ±1.11[ab]	14.04 ±0.74[b]
Glu	17.13 ±1.02[b]	16.49 ±0.99[b]	18.94 ±0.45[a]
Ser	5.52 ±0.35[ab]	5.34 ±0.03[b]	6.02 ±0.17[a]
Gly	6.04 ±0.44[a]	5.85 ±0.11[a]	6.19 ±0.24[a]
His	3.62 ±0.23[b]	3.35 ±0.21[b]	4.35 ±0.01[a]
Arg	9.53 ±0.41[a]	4.46 ±0.21[b]	9.20 ±0.00[a]
Thr	2.93 ±0.22[b]	3.16 ±0.16[b]	3.94 ±0.04[a]
Ala	5.27 ±0.21[b]	5.72 ±0.64[ab]	6.18 ±0.21[a]
Pro	5.89 ±0.29[ab]	5.90 ±0.20[b]	6.32 ±0.18[a]
Val	9.91 ±0.41[a]	9.92 ±0.40[a]	8.21 ±0.14[b]
Tyr	3.17 ±0.16[a]	4.03 ±1.52[a]	3.59 ±0.03[a]
Met	3.11 ±0.08[b]	3.33 ±0.01[b]	4.84 ±0.18[a]
Cys	0.72 ±0.04[a]	0.74 ±0.07[a]	0.74 ±0.15[a]
Ile	9.16 ±0.13[a]	8.79 ±0.02[b]	7.22 ±0.16[c]
Leu	13.31 ±0.45[a]	13.13 ±0.42[a]	13.87 ±0.06[a]
Phe	11.68 ±0.47[b]	11.25 ±0.77[b]	15.81 ±0.01[a]
Lys	15.74 ±0.60[a]	14.40 ±0.98[a]	15.95 ±0.03[a]
Trp	1.79 ±0.15[b]	2.13 ±0.11[a]	2.12 ±0.01[a]

*The amino acids content and protein content as dry basis were expressed in g/100g protein

*Mean values of at least two replicates ± standard deviation (SD)

*Mean values within the same row with different superscript letters are significantly different ($p < 0.05$)

Table 2: The chemical and nutritional classification of *Trichilia* and soybean amino acids

Classification	*T. dregeana*	*T. emetica*	Soybean
HAA*	63.01 ± 2.11[b]	63.3 ±0.83[b]	68.49 ±1.09[a]
NCAA*	32.37 ±1.17[a]	31.31 ±2.10[a]	32.98 ±1.20[a]
PCAA*	28.88 ±1.24[a]	22.21 ±1.41[b]	30.22 ±0.01[a]
BCAA*	32.37 ±0.99[a]	31.83 ±0.83[a]	29.30 ±0.36[b]
AAA*	17.44 ±1.61[b]	17.40 ±0.64[b]	21.52 ±0.03[a]
SCAA*	3.83 ±0.03[c]	4.06 ±0.08[b]	5.57 ±0.33[a]
E/T (%)**	52.77 ±0.81[a]	53.42 ± 0.40[a]	52.99 ±0.44[a]

*The mean values of essential amino acid classification with standard deviation were expressed as mg/g crude protein. **E/T is the proportion of essential amino acid (E) to the total amino acids (T). HAA is the hydrophobic amino acids, BCAA: branched chains amino acids, SCAA: sulphur-containing amino acids, NCAA: Negatively charged amino acids, PCAA: Positively charged amino acids and AAA: Aromatic amino acids. Mean values within the same row with different superscript letters are significantly different ($p < 0.05$)

Table 3: Essential amino acid score of *Trichilia* protein concentrates compared with the scores for soybean protein concentrates

EAA	*T. dregeana* AAS		*T. emetica* AAS		Soybean AAS		Chia AAS		Sunflower AAS		Flax AAS		Dietary FAO Recommendation	
	Child	Adult	Child	Adult	Child	Adult	Child	Adult	Child	Adult	Child	Adult	Child	Adult
His	1.80±0.11[b]	2.40 ± 0.15[b]	1.68 ±0.11[b]	2.23±0.14[b]	2.18 ± 0.01[a]	2.90 ±0.01[a]	1.54-1.60	2.05-2.14	1.15	1.53	1.38	1.84	20	15
Ile	2.95±0.04[a]	3.05 ± 0.04[a]	2.83 ±0.01[a]	2.93±0.01[a]	2.33 ±0.05[b]	2.41 ±0.05[b]	1.08-1.09	1.11- 1.12	1.44	1.48	1.34	1.39	31	30
Leu	2.11±0.07[b]	2.26 ±0.07[a]	2.08 ±0.07[b]	2.22±0.07[a]	2.20 ±0.01[a]	2.35 ±0.01[a]	1.14-1.17	1.22- 1.25	1.35	1.44	0.97	1.04	63	59
Lys	3.03±0.11[a]	3.50 ±0.13[a]	2.77 ±0.19[a]	3.20±0.22[a]	3.07 ±0.01[a]	3.54 ±0.01[a]	1.15-1.16	1.57- 1.59	0.70	0.96	0.69	0.94	52	45
Met +Cys	1.37±0.01[b]	1.74 ±0.01[b]	1.45 ±0.03[b]	1.85±0.04[b]	1.99 ±0.12[a]	2.54 ±0.15[a]	1.49-1.66	1.90-2.11	1.34	1.70	0.85	1.08	28	22
Phe +Tyr	3.01±0.28[b]	4.11 ±0.38[b]	2.94 ±0.14[b]	4.02±0.20[b]	3.73 ±0.01[a]	5.11 ±0.01[a]	1.96-2.00	2.68-2.74	1.20	1.64	1.59	2.18	52	38
Thr***	1.08±0.08[b]	1.27 ±0.09[b]	1.17 ±0.01[b]	1.37±0.01[b]	1.46 ±0.06[a]	1.71±0.07[a]	1.32-1.34	1.55- 1.58	1.37	1.61	1.31	1.53	27	23
Val	2.36±0.10[a]	2.54 ±0.11[a]	2.36 ±0.10[a]	2.54 ±0.10[a]	1.95 ±0.03[b]	2.10±0.04[b]	1.04-1.08	1.12- 1.16	1.46	1.58	1.23	1.32	42	39
LEAA	THR		THR		THR		Val	Ile	Lys		Lys			

*The mean essential amino acid (EAA) values were expressed as % w/w of crude protein on dwb.

**AAS (amino acid score) =mg amino acid in 1 g protein of test sample／mg amino acid in requirement pattern (FAO/WHO, 2007)

***The AAS of Thr is the chemical score for the Trichilia proteins and Soybean proteins. Chemical score = AAS×100, truncated to 100% if >100.

*Mean values of at least two replicate with standard deviation

*Mean values within the same row with different superscript letters are significantly difference ($p < 0.05$)

Table 4: Non-essential amino acids of *Trichilia* and Soybean protein concentrates

NEAA*	*T. dregeana*	*T. emetica*	Soybean
Arg	9.53 ±0.41^{a}	4.46±0.21^{b}	9.20 ±0.00^{a}
Ser	5.52 ±0.35^{ab}	5.34 ±0.03^{b}	6.02 ±0.17^{a}
Gly	6.04 ±0.44^{a}	5.85 ±0.11^{a}	6.19 ±0.24^{a}
Asp	15.24 ±0.16^{a}	14.82 ±1.11^{ab}	14.04 ±0.74^{b}
Glu	17.13 ±1.02^{b}	16.49 ±0.99^{b}	18.94 ±0.45^{a}
Ala	5.27 ±0.21^{b}	5.72 ±0.64^{ab}	6.18 ±0.21^{a}
Pro	5.89 ±0.29^{ab}	5.90 ±0.20^{b}	6.32 ±0.18^{a}
Trp	1.79 ±0.15^{b}	2.13 ±0.11^{a}	2.12 ±0.01^{a}

*The mean values of the non-essential amino acids (NEAA) with standard deviation were expressed as g/100g crude protein. Mean values within the same row with different superscript letters are significantly different ($p < 0.05$)

Table 5: Functional properties of *T. emetica*, *T. dregeana* and soybean protein concentrates

Functional Properties	**Soybean**	***T. emetica***	***T. dregeana***
WHC (%)	800 ±50.00^{b}	921 ±48.02^{a}	450 ±0.00^{c}
OHC (%)	466 ±28.87^{a}	200 ±0.00^{c}	250 ±0.00^{b}
EAI (%)	54.54 ±0.65^{a}	52.32 ±0.45^{b}	51.5 ±0.71^{b}
FC (%)	23.5 ± 2.12^{c}	90 ± 5.66^{a}	26 ± 0.00^{b}
FS (%): 15 mins	19 ± 1.41^{c}	80 ± 2.83^{a}	21.5 ± 0.71^{b}
FS (%): 30 mins	15 ± 1.41^{c}	75 ± 7.07^{a}	20 ± 0.00^{b}
FS (%): 60 mins	10 ± 0.00^{c}	75 ± 7.07^{a}	20 ± 0.00^{b}

*Mean values of at least two replicate with standard deviation

*Mean values with different superscript letters within the same row are significantly different ($p < 0.05$)