Profiling proteins in nutraceutical formulations: Characterization of the constituents

Alessia Bellomaria\textsuperscript{a,h}, Ridvan Nepravishita\textsuperscript{a,1}, Mario Marchetti\textsuperscript{c}, Maurizio Paci\textsuperscript{a}

\textsuperscript{a} Department of Chemical Sciences and Technology, University of Rome Tor Vergata, via della Ricerca Scientifica 1, 00133 Rome, Italy
\textsuperscript{b} Cooperativa di Ricerca Finalizza, Via della Ricerca Scientifica 1, 00133 Rome, Italy
\textsuperscript{c} Istituto di Igiene, Via \textsuperscript{a} Tor Vergata, 8, Rome, Italy

\begin{abstract}
Several nutraceutical preparations containing proteins, amino acids and other small molecules are nowadays present on the market. In this work we propose NMR spectroscopy such as \textsuperscript{1}H NMR, \textsuperscript{1}H–\textsuperscript{1}H TOCSY and DOSY for their constituents characterization, identification and profiling, comparing these results with those obtained by electrophoretic technique such as SDS–PAGE. The \textsuperscript{1}H NMR spectroscopy was applied for measurements of the amino acids and other small compounds added from the manufacturer. Further the autocorrelation function obtained from the one dimensional spectrum was used without the complete assignment of the resonances of the NMR spectrum of proteins for the evaluation of the folding quality and stability. Finally the DOSY NMR technique was performed on the samples for the characterization of the mean molecular weight range of proteins. All these features considered together create an important set of data useful for the evaluation of the protein profiling and the characterization of such formulations.
\end{abstract}

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Proteins are an important source of amino acids essential for the correct growth and all other physiological functions. It is also well known that proteins are not conserved in reserve pools, as it usually happens with glucose or lipids, for this reason it is necessary that the protein intake is constant and based on good quality product. In fact an inadequate intake, especially during periods of growth and development, may generate serious problems that may affect all organs of the body, and particularly the brain, the heart and the immune system (Raman, Wijesinha-Bettoni, & Burlingame, 2016).

Among the most important sources of protein for the human diet, cow milk plays a central role. In fact, the milk provides about 32 g/l of protein. On average, the cow milk is composed of 87% water, 4–5% lactose, 3% protein, 3–4% fat and 0.1% of vitamins.

Proteins in cow’s milk can be divided into soluble and insoluble proteins. The soluble protein fraction contains proteins that are commonly named whey proteins, whereas insoluble proteins are primarily caseins (Jennings, 1979). Whey proteins, represent 20% of milk protein fraction, whereas caseins represent the remaining 80% part. Both fractions are classified as high-quality proteins considering human requirements for amino acids, digestibility, and bioavailability. In fact, milk proteins are frequently considered the best protein source taking into account the Protein Digestibility Corrected Amino Acid Score (PDCAAS) (Haug, Hesmark, & Harstad, 2007; Schafmeister, 2000, 2012).

Moreover, several bioactive peptides resulting from the enzymatic hydrolysis of these proteins have shown multiple biological roles that could exert a protective action in human health. These biological roles include antibacterial, antiviral, antifungal, antioxidant, antihypertensive, antimicrobial, antiinflammatory, opioid, and immune-modulatory effects. Further it was also demonstrated their role in absorption enhancement of other nutrients (Mills, Ross, Hill, Fitzgerald, & Stanton, 2011). In milk, the major two protein fractions, caseins and whey proteins, present a different composition of amino acids. Milk whey proteins, in particular, are composed mainly by branched chain amino acids whereas caseins are, principally, composed by aromatic and sulfur containing amino acids such as histidine, phenylalanine and methionine (Tang, Moore,
Kujbida, Tamposki, & Phillips, 2009). From the point of view of the constituents milk whey proteins are a heterogeneous class that differs in composition, structure and molecular weight (MW). Principal components are β-lactoglobulin, α-lactalbumin, serum albumin, immunoglobulins, and other minor fractions of milk whey proteins. Caseins are grouped in four major classes: αs1-, αs2-, β-, and κ-caseins; and are all phosphate-containing proteins that appear as micelles in the native form and, differently from whey proteins, precipitate at acid pH (pH ≈ 4.0). Furthermore, during digestion, the behavior of these caseins and milk whey proteins differs markedly. Milk whey proteins are rapidly evacuated from the stomach, whereas caseins clot or precipitate under the acidic gastric pH, which delays amino acid delivery to the intestine tract (Mahe, Roos, & Benamouzig, 1996). Whey proteins reveal a major role for their gastric emptying and absorption with respect to caseins, this is due to their composition and their solubility. For these reasons, some studies indicated that the milk whey protein ingestion mediate larger increases in post meal aminoacidemia than caseins (Boirie et al., 1997), although this is a point that needs to be better defined. In fact, other studies, reported that caseins slow-digestion have a major biological value with respect to whey proteins that are rapidly digested, demonstrating that milk whey proteins are unable to sustain the anabolic postprandial amino acid requirement (Lacroute et al., 2006; Meisel & FitzGerald, 2003).

Several effects of milk whey proteins on human health were already reported demonstrating an antimicrobial activity for milk whey components such as lysozyme, lactoferrin, lactoperoxidase (Jensen & Hancock, 2009; Min, Harris, & Kuncha, 2005) and suppressing action in tumor development for lactoferrin, β-lactoglobulin and α-lactalbumin (Parodi, 2007). It is also important to underline that β-lactoglobulin is a retinol carrier with antioxidant properties, while lactoferrin is a crucial element in iron absorption and exerts antioxidant and anticarcinogenic effects (Gonzalez-Chavez, Arevalo-Gallegos, & Arenas-Cruz, 2009).

In this study we have analyzed some nutraceutical formulations containing principally free amino acids and milk whey proteins. The use of 1H NMR spectroscopy as a fast and direct method to identify the amino acids and other small molecules permitted to measure their content in comparison with the composition declared by the manufacturer. Moreover the application of NMR spectroscopy allowed us to characterize the formulations in terms of proteins content and these data were compared to electrophoresis experiments performed on the samples. In particular inspection of dispersion of NMR resonance(s) in spectra was used to probe protein structural stability and folding (e.g., the presence of stable protein structures) in solution. This was performed using the statistical analysis such as the autocorrelation function also in the absence of the individual assignment the 1H chemical shifts of the proteins. In fact based on experimental data obtained on well-structured proteins and proteins that exist in a molten globule state or a partially-folded helical state, a well-defined threshold exists that can be used as a characteristic benchmark for protein structural stability (e.g., foldedness) in solution (Hoffmann, Eichmüller, Steinhauser, & Konrat, 2005).

Moreover the diffusion coefficients of proteins in the samples obtained by diffusion ordered spectroscopy (DOSY) made possible the further characterization of the integrity of proteins used in the formulations.

2. Materials and methods

2.1. Samples

Samples used in this study were gently offered from Italfarmacia Srl. Rome. All of them were composed by milk whey proteins with the addition from the producer of isoleucine, tryptophan, ornithine, taurine, citrulline and citrate. Pure isoleucine, tryptophan, ornithine, taurine, citrulline and citrate, useful for resonances assignments in the NMR spectra were purchased from Sigma Aldrich (Milan, Italy).

2.2. Determination of total protein concentration

Total protein concentration was determined colorimetrically with the Bradford assay (Bradford Reagent, Sigma–Aldrich, B-6916, USA). The standard curve was prepared using standard solution of bovine serum albumin (BSA) at 200 µg/ml (Pierce Chemical), was diluted in 50 mM sodium carbonate–sodium bicarbonate buffer (pH 10.0) according to the manufacturer's instructions. The concentrations of standards were 25, 125, 250, and 500 µg/ml. Absorbance readings at 595 nm were determined with a UV–Vis spectrophotometer Lamba Bio Jasco 600. All sample were diluted in the assay's buffer, preparing 0.5 mg/ml solutions and analyzed with Bradford kit for estimate the real concentration. The solutions using for the Bradford test were prepared by referring to the information on the packaging. Determinations were replicated at least four times for each preparation (Bradford, 1976).

2.3. Electrophoresis

The protein composition of samples was analyzed by SDS-PAGE, using 12% separating gel and 4% stacking gel. The samples were heated for 5 min at 100 °C in capped vials with 1% (w/v) SDS in the presence of β-mercaptoethanol. Electrophoresis was performed at a 125V in Tris–HCl buffer of pH 8.3. After electrophoresis, proteins in the separating gel were made visible by staining with Coomassie Brilliant Blue R-250. The standards with M.W. of 15, 25, 35, 50, 75, 100 and 150 kDa were used to make a plot of log molecular weight versus mobility. The standard curve plot was further used for the determination of the molecular weight of proteins in each sample.

2.4. Density analysis of SDS–PAGE gel

The density analysis of the protein bands separated by SDS-PAGE gel was performed with ImageJ (National Institutes of Health) software by measuring the Intensity (I) of each protein band.

2.5. NMR spectroscopy

2.5.1. Sample preparation

All the samples were prepared by directly dissolving 0.05 g/ml of each sample in H₂O present 10% D₂O for the lock signal and transferred on 5 mm NMR tubes for the measure using DSS as reference.

2.5.2. NMR experiments

Were performed on a 400.13 and 700.13 MHz Bruker Avance instruments. In order to suppress the water signal at 4.7 ppm the 1H NMR experiments were acquired with the zgpr pulse sequence of the Bruker library with 32Ktime domain accumulating 64 scans at 298 K.

2.5.3. NMR scalar correlation experiments

Scalar correlation experiments were performed as total correlated spectroscopy (TOCSY), using the Watergate pulse sequence to suppress the water signal (Braunschweiger & Ernst, 1983; Davies & Bax, 1985). A typical mixing time of 150 ms in order to observe either direct or both direct and remote connectivity was used. Two-dimensional NMR experiments were performed in the
phase sensitive mode typically using 2 K of memory for 512 increments. The number of scans was optimized to obtain a satisfactory signal-to-noise ratio.

2.5.4. Diffusion measurement by DOSY experiment

The DOSY spectra were performed by using the ledgbpapr2a pulse sequence of the Bruker library in order to suppress the water signal at 4.7 ppm. During the DOSY experiment 32 mono dimensional spectra were acquired with 64 scans in a linear increasing gradient varying from 5% to 95% with a Delta of 70 ms and a delta of 2 ms. The spectra were then analyzed using the DOSY module implemented in Bruker software TOPSPIN 3.1.

2.5.5. Autocorrelation

For the statistical analysis we used the autocorrelation function of protein $^1$H NMR spectra following the previously reported procedure (Hoffmann et al., 2005) acquired with the zgpr pulse sequence for water signal saturation. The $^1$H NMR spectra were processed with the binning module in Mestre Nova at 0.01 ppm (Mestrelab research, Spain) and the calculation of the autocorrelation function was performed with Matlab software (MathWorks, USA). Residual water signal was excluded from the calculation deleting the spectral region 4.50–4.55 ppm as previously reported (Hoffmann et al., 2005). The autocorrelation function was obtained by the Fourier transform of the product between the free induction decay and its complex conjugate and is thus related to the distribution function of the frequency and relaxation rate differences, respectively (Hoffmann et al., 2005). The obtained autocorrelation functions C(1) are normalized to the value at the smallest available frequency difference (0.01 ppm) and numerically smoothed. The values of the autocorrelation function at frequency 0.5 ppm, C (0.5) were used as a measure of protein structural folding (Hoffmann et al., 2005).

3. Results

Seven commercial nutraceutical formulations containing milk whey proteins, amino acids and other small molecules, currently present in the Italian market were examined during this study.

3.1. SDS–PAGE characterization

The first approach was the use of electrophoresis in, SDS–PAGE, as reported in Section 2. In order to normalize the amount of proteins in all samples, the solutions were checked by Bradford assay and then was calculated the right quantity of protein suitable to load for each line on the SDS–PAGE. For this experiment, the samples were treated with 6-mercaptopropanol and boiled before loading on the gel in order to ensure that the proteins have been separated in the absence of intermolecular disulfide bridges which may change their electrophoretic behavior. The experiments were repeated at different concentrations of samples and at different acrylamide percentage in the gel, in order to obtain the best separation of the components. As one can observe in Fig. 1, reporting the electrophoresis on SDS–PAGE of all the samples, seven main bands separated can clearly be identified, labeled with the letters from a to f. By comparison with the common protein standards, the bands observed were assigned to proteins of the following molecular weights: band a = 90 kDa; band b = 69–60 kDa; band d–d1 = 45 kDa; band e = 30 kDa; band f = 20 kDa. The protein profiling for each line respectively was used for comparison of these results. From the graphs in Fig. 1 an interesting observation about the profile of the protein fractions in these formulations can be done. With regard to the component a we note that this component, although with different concentrations, is present in all of the formulations with the exception of sample G. Components b and c, are present in all the formulations, but specifically in A, B and E one can note that component b presents a higher concentration than the component c, while in other samples they are nearly equivalent. The component d is present in all the samples while d1 was found only in E and G samples. The e component was not found in the F sample while the f component resulting from proteins with relatively low molecular weight, was found in all the samples with a major concentration in F and D samples. About the typology of proteins the bands of high molecular weight can be assigned to lactoferrin, lactalbumin; the bands around 20–30 kDa can be associated with the soluble fraction of caseins in their different isoforms that are present in the raw material used for the declared formulation of the nutraceutical samples. The bands at molecular weight less than 20 kDa are to be attributed to lactoglobulins. However, the low molecular weight fractions can be also attributed to fragments caused from proteolysis or fragmentation during the manufacturing of raw material. It is important to note that, in the samples analyzed, there is not a common profile of protein components. Probably this variability may be due to the source of protein used as raw material from manufacturer. In fact there are several procedures used to obtain milk whey proteins applied by the raw material producers, which may have a strong impact on the protein profile of milk whey proteins used for these preparations.

3.2. NMR spectroscopy

The electrophoretic technique demonstrated to be very informative for the profiling of protein components present in each formulation but no information on small molecules like amino acids and other small molecules added from each producer to the formulations was available. Moreover SDS–PAGE experiments reported above did not give information about the protein quality such as the status of folding, that may be another distinguishing characteristic between samples. In this study we applied NMR spectroscopy using the standard pulse sequences for $^1$H NMR, $^1$H–$^1$H TOCSY and $^1$H–$^1$H DOSY spectroscopy to obtain more information about this point. A careful analysis of the spectra were carried out in order to:

a. assign small molecules (amino acids) added to the formulations by the manufacturer and by checking the correspondence between the declared and the actual concentrations,

b. check the status of the folding of the proteins as indicator of a good production procedure,

c. evaluation of the mean molecular weight of proteins in the samples by DOSY technique in comparison with the results of SDS–PAGE.

This procedure now defined "profiling" it is far from a "quantitative" determination that requires complete assignment and a complete determination (as in Gowda Nagana & Raftery, 2014) while it is a fast procedure for the characterization during the production process of protein nutraceuticals.

3.2.1. Identification of small molecules and amino acids

In Fig. 2, the spectra acquired on sample A and B are reported as examples. The resonances of the component were assigned and by integration the concentration all of the amino acids and other substances were evaluated in order to compare their values with those declared by the manufacturers. The only exception was tryptophan due to its very low concentration. The assignments were performed adding to the samples small amounts of the amino acid and of the other components added to the protein preparation by factories as declared (Fig. 2). Then the concentration of these few amino acids and of other molecules added was measured by the
Fig. 1. SDS-PAGE analysis and their densitometric analysis of the seven bands present in the SDS-PAGE gel. The samples loaded were 0.5 mg/L concentrated for each commercial nutraceutical formulation.

![SDS-PAGE Analysis](image)

Fig. 2. (a) $^1$H NMR total spectrum and (b) the enlargement of the spectral region between 0 and 4 ppm in the spectrum of sample A (upper panels) and sample B (lower panels).

![$^1$H NMR Spectra](image)

use of an external reference (DSS at 0.1%, data not shown). At this stage the presence of protein resonances were not an interfering presence due to their low concentration, anyway appropriate NMR pulse were adopted for filtering the protein resonances. It must be said that resonances from proteins shows in general a rather dispersed chemical shifts as will be discussed later. The results of the measured concentrations of the previously cited molecules added to the preparations obtained by $^1$H NMR spectroscopy were in good agreement with the quantities declared from the manufacturer (data not to be disclosed). This confirms that NMR spectroscopy is a valuable tool to control in a fast and direct way the formulation of this kind of preparations.

3.2.2. Evaluation of protein folding

The characterization of the protein folding of the proteins in these samples was evaluated directly by NMR spectroscopy of samples. In fact the spectra obtained by $^1$H NMR and $^1$H- $^1$H TOCSY spectra give information for this purpose. In this case the close inspection of the regions of the TOCSY spectrum (Fig. 3), particularly in the fingerprint region, indicates a large spread of chemical shifts of the resonances with respect to the simple distribution expected for a simple mixture of the 20 single amino acids. This dispersion can be evaluated by direct peak-picking in the spectral 2D region. Moreover the dispersion of resonances in the NMR spectrum of protein due to the magnetic environment(s) produced by the partial or complete tertiary structure can be used as an indicator of a stable tertiary structure. In fact the origin of the $^1$H NMR chemical shifts is governed by the details of the 3D solution structures of proteins. The magnetic microenvironment due to the secondary or tertiary structure of the protein generates the dispersion of resonances visible in the high resolution NMR spectra. Several $^1$H NMR experimental data exists (Seavey, Farr, Westler, & Marikley, 1991), that are mainly used as an initial prerequisite for the structural studies of proteins. On the contrary a low resonances dispersion is assumed safely as an indicator of the poorly folded structure. In fact the resonances of the protons of the usual amino acids alone or in a completely unfolded proteins show a very simple spectrum and show a nearly null dispersion. Moreover
the presence of a number of ring current shifts in the methyl region, namely resonances with chemical shifts between 1 and, particularly in the negative region less than 0 ppm are also diagnostic indicators of globular character of the protein and good indicators of their stability (Perkins & Wuthrich, 1979). An approach of this observation was performed by evaluating on the resonances of the $^1$H NMR spectra by the autocorrelation function also in the absence of the individual assignment. This use of the statistics of the $^1$H chemical shift distribution to probe protein structural stability in solution has been proposed (Hoffmann et al., 2005). The method uses the autocorrelation function directly on the $^1$H spectra demonstrating that a significant correlation exists between the autocorrelation function and the topological complexity (expressed as the relative contact order), and the structural stability of the protein. The study on diverse sets of folded proteins with native structures or partially folded proteins indicated that the method can be used in a high-throughput manner to screen for protein structural stability. The evaluation of this function, was performed by taking into consideration the value of the autocorrelation function at 0.5 ppm C(0.5). The calculation of the autocorrelation function obtained as reported in Materials and methods gave the results reported in Fig. 4. The autocorrelation function on the chemical shift values indicated different results for the proteins of the sample. Also the spreading of the cross peaks in the TOCSY spectrum both of the CHz region and the NH region were considered as good indicators of the presence of a three dimensional structure. In Fig. 4 and Table 1 both the autocorrelation function calculated from $^1$H NMR and the qualitative evaluation of TOCSYS are reported in accordance between them. In fact, among native folded proteins display C(0.5) values >0.5, while partially folded or unfolded proteins have values <0.4. In general terms of folding we can observe from Table 1 that the Highest Score (HS) was achieved by sample E, followed by sample A and C with an Intermediate High Score (IHS). Samples B, F, G presented an intermediate low score (ILS) while D was classified with a Very Low Score (VLS).

3.2.3 Mean molecular weight evaluation of the protein fraction

NMR diffusion experiments (DOSY) using constant physical surrounding environment such as viscosity, temperature, for each chemical species in solution, provides a way to separate the different compounds in a mixture based on the difference translation in diffusion coefficients (D) and offers a tool to know the size and shape of the molecule. In our case we were interested on the estimate of the hydrodynamic radius which may be related to the mean molecular weight of the proteins and other molecules in the samples. It is known that in the case of proteins, it is very difficult or nearly impossible to obtain a general calibration plot of self-diffusion data on the hydrodynamic radius. In fact the very large differences among the protein structures and their differences in hydrophilicity makes that measure affected of large errors. In polymers composed by the same monomers with only different MW the diffusion is considered an extremely useful tool with a high reliable calibration curve. In our case only the D of some standard proteins like lysozyme and bovine serum albumin were used as samples of references (not shown). In Fig. 5 upper panels we can observe the enlargement of two DOSY experiments as example. The DOSY experiments are shown together in the lower panels, and the lines are for guidance of the eye as indication of the
diffusive front. The diffusion coefficient (D) is reported as logD and is related to the radius of the molecular species by the Stokes-Einstein equation (Einstein, 1956). Due to the fact that the radius of the species in the solution is only approximately related to their MW, the measure of D in our case can give information on the mean MW of the protein fraction of the sample under investigation (Auguin, Gottran, Deluc, & Roumestand, 2004). Without the knowledge in advance of the mean MW of the samples we can only classify them on the basis of their diffusion coefficient as samples with High Mean Molecular Weight (HMMW) which present a small D coefficient, corresponding to a low diffusivity, with Intermediate Mean MW (IMMW), and Low Mean Molecular Weight (LMMW), with a dynamic high diffusivity, presenting a high value of D. The results are summarized in Table 1.

4. Discussion

The results obtained indicate that NMR spectroscopy represents a useful, fast, direct and precise protocol for the characterization of the protein, amino acids and other small molecules in nutraceuticals preparations and, in general, of pharmaceutical compositions of this type. The use of the electrophoresis technique allowed us to crosscheck the results with those of the NMR spectroscopy in profiling of proteins used for the preparation of nutraceutical formulations. In this type of analysis resulted important to distinguish between whey proteins and hydrophobic proteins, principally caseins. In the electrophoretic examination, it was also highlighted the presence of a significant quantity of casein (bands at ≈20 kDa). These proteins showed a high tendency to coagulate into a micellar dispersion due to their strong hydrophobicity and to their minor content in branched chain amino acids. The casein clot is able to provide a slow and sustained release of amino acids into the bloodstream, which can last for several hours. This feature is considered very important to maintain an anabolic and anti-catabolic state on muscle tissue for a relatively long time, where the serum proteins promote greater protein synthesis, but are not able to limit protein catabolism. On the other hand the possibility to detect possible species that may be due to proteolysis of the initial protein pool is another powerful tool for their characterization in final products. These fragments may highlight change(s) in the protein pool during the initial phase of preparation of the nutraceutical formulation or during the storage also depending on the storage conditions and can be eventually used for the determination of the shelf life of the product. Moreover using $^1$H NMR spectroscopy was possible to identify the amino acids and other small molecule added to the formulation by the manufacture verifying the conformity of the results with the declared amounts. This is very important for evaluation of product conformity with regard to the quality control. In fact the NMR spectroscopy permits to analyze directly the intact product in a simple and fast manner. Another aspect of the formulations considered in this study was the evaluation of the protein folding and the evaluation of the mean MW used as a marker of both product quality and stability. The folding was obtained from both TOCSY and $^1$H NMR spectroscopy by qualitative evaluation respectively. The evaluation of the protein folding was performed by the autocorrelation function applied to $^1$H NMR spectra. On the other hand the use of the DOSY NMR permitted us to estimate the mean molecular weight of the single protein. These results taken together indicate that the NMR spectroscopy and the application of DOSY NMR are able to characterize the profile of the proteins present in nutraceuticals products in very good agreement with the electrophoretic technique. These characteristics are strictly related with the grade of quality of components of the final commercial product here examined. The proposed procedure can be very useful for the analysis of the compositions of the nutraceutical formulations preexisting proteins, amino acids and other small molecules and may give an important information about the quality of raw products used and/or the finished product and upon its particular storage conditions.

5. Conclusions

In this work we demonstrated the use of electrophoresis technique and NMR spectroscopy for conformity to the declared amounts of added components and for protein fraction profiling of some nutraceutical products. The electrophoresis technique was very useful for the profiling of the protein fraction as a simple and inexpensive tool for the analysis of these formulations. In fact from SDS–page analysis we were able to observe which were the principal differences and the similarities between samples. Moreover from the NMR experiments here proposed were obtained useful information on protein folding, mean MW present in these formulations. When it is needed the tests described may characterize in a quick and simple manner the overall state of the product and can be used as preliminary screening tests on raw or finished product.

Acknowledgments

We are grateful to Mr. Fabio Bertocchi for NMR technical support during this work.
References


