

Sequence determination of α_{s1} -casein isoforms from donkey by mass spectrometric methods

Vincenzo Cunsolo,^{a*} Elisa Cairone,^a Debora Fontanini,^b Andrea Criscione,^c Vera Muccilli,^a Rosaria Saletti^a and Salvatore Foti^a

Four co-eluting components, with experimentally measured M_r of 23 658, 23 786, 24 278 and 24 406 Da, were detected by reversed-phase high-performance liquid chromatography (RP-HPLC) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis in the dephosphorylated casein fraction of a milk sample collected at middle lactation stage from an individual donkey belonging to the Ragusano breed. By coupling RP-HPLC, two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), enzymatic digestions, MALDI-TOF MS and capillary RP-HPLC/nano-electrospray ionization tandem mass spectrometry (nESI-MS/MS) analyses, the four components were identified as donkey's α_{s1} -CNs and their sequences completely characterized, using the known mare's α_{s1} -CN (GenBank Acc. No. AAK83668; M_r 23750.7 Da) as reference. The proteins with M_r of 23 786 and 23 658 Da differ in the presence of a glutamine residue at position 83 in the full-length component and present the amino acid substitutions Q⁸ → H and H¹¹⁵ → Y with respect to the mare's α_{s1} -CN. The other two components with M_r 24 406 and 24 278 Da, which also differ in the presence of a glutamine residue at position 88 in the full-length component, show the insertion of the pentapeptide HTPRE between Leu³³ and the Glu³⁴. The two α_{s1} -CNs bearing the pentapeptide insertion were named variants A (202 amino acids; M_r 24 406) and A₁ (201 amino acids; M_r 24 278), whereas the two α_{s1} -CNs without the pentapeptide were named variants B (197 amino acids; M_r 23 786) and B₁ (196 amino acids; M_r 23 658). Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: donkey milk; α_{s1} -casein; milk allergy; RP-HPLC/nESI-MS/MS; 2D-PAGE

Introduction

Cow's milk allergy constitutes a problem of social relevance, especially because cow's milk is used in the diet of newborns as a substitute for breast milk when the mother's milk is not available. The efficacy of donkey's milk, which has been traditionally used to feed problematic children, has been recently confirmed by clinical investigations on children with severe cow's milk allergy. These studies have shown that in many cases milk from donkey represents a safe and alternative food in both immunoglobulin E (IgE)-mediated and non-IgE-mediated cow's milk protein allergy (CMPA), providing nutritional adequacy and good palatability.^[1–3] Analogous results were obtained with mare's milk.^[4] Although the mechanism of this tolerance has not yet been fully clarified, it is reasonable to hypothesize that the reduced allergenic properties of donkey's milk can be related to structural differences of its protein components with respect to bovine milk. The high lactose content, the characteristic protein profile with a lower casein content (~50%) with respect to cow's milk (~80%)^[5] and the relevant percentage of essential amino acids make this milk a potential new dietetic food and a promising breast milk substitute.^[6]

However, compared with bovine milk, the characterization at molecular level of *Equidae* (mare and donkey) milk proteins is at a relatively early stage and limited data are available for its genetic polymorphism. Information is mainly available on whey proteins,^[7–9] whereas the investigation of casein (CN) components is still at a relatively early stage of progress. In particular, the complete amino acid sequence of a κ -CN,^[10] two β -CNs (nrNCBI Acc. No. Q9GKK3)^[11] and small fragments identified as equine α_{s2} -like caseins^[8,12] have been characterized from

mare's milk. In addition, the complete cDNA deduced amino acid sequences of two mare's α_{s1} -CNs (GenBank Acc. Nos AAL05435 and AAK83668) and of a κ -CN (GenBank Acc. No. AAK83669) have been reported.^[13,14] The α_{s1} -CN primary structure displays polymorphic patterns, due to alternative splicing processes leading to casual exon skipping events involving exons 7 and 14.^[13]

Conversely, to date the knowledge of the donkey's CNs is scant and only the cDNA deduced amino acid sequences of a κ -CN (GenBank Acc. No. EU448385)^[15] and two α_{s2} -CNs (GenBank Acc. Nos FM946022 and CAX65660.1)^[16,17] and the recent characterization of the primary structure of two β -CNs have been reported (UniProtKB/Swiss-Prot Acc. No. P86273).^[18]

In the frame of our research line oriented to the characterization of the donkey's milk proteins, we report here the identification and direct sequence determination of the primary structure of two donkey's α_{s1} -CN variants and their respective isoforms by coupling reversed-phase high-performance liquid chromatography (RP-HPLC), 2D-PAGE, enzymatic reactions and mass spectrometric methods.

* Correspondence to: Vincenzo Cunsolo, Department of Chemical Science, University of Catania, Viale A. Doria 6, I-95125 Catania, Italy.
E-mail: vcunsolo@unict.it

a Department of Chemical Science, University of Catania, Viale A. Doria 6, I-95125 Catania, Italy

b Department of Biology, University of Pisa, Via L. Ghini 5, 56126 Pisa, Italy

c Department of Agronomical, Agrochemical and Animal Production Sciences (DACPA), University of Catania, Via Valdisavoia 5, 95125 Catania, Italy

Experimental

Materials and methods

Acetic acid, formic acid (FA), trifluoroacetic acid (TFA), sinapinic acid (SA), α -cyano-4-hydroxycinnamic acid (CHCA), cytochrome *c* from bovine heart, myoglobin from horse skeletal muscle, trypsinogen from bovine pancreas, endoproteinase Glu-C (*Staphylococcus aureus* strain V8) and β -lactoglobulin from bovine milk, were obtained from Aldrich (St Louis, Missouri, USA). Modified porcine trypsin was purchased from Promega (Madison, WI, USA). Ammonium bicarbonate, sodium acetate, HPLC grade H₂O and CH₃CN were provided by Carlo Erba (Milan, Italy). Alkaline phosphatase from bovine intestinal mucosa was purchased from Fluka (St Gallen, Switzerland). The General Protein/Mass Analysis for Windows software (<http://welcome.to/gpmaw>) was used for all sequence handling and storage. The ClustalW2 tool (<http://www.clustal.org/>) was used for sequence alignment.

Preparation of milk and enzymatic dephosphorylation of casein fraction

The individual milk sample was collected in Eastern Sicily from a donkey belonging to the Ragusano breed at middle lactation stage. After milking, the sample was immediately frozen and stored at -20°C until used. The milk was warmed to 30°C and defatted by centrifugation at 4000 rpm, 4°C for 60 min. The fat layer was solidified at -24°C and drawn up. The casein fraction was precipitated from the skimmed milk at pH 4.3 with sodium acetate/acid acetic buffer and separated from the supernatant by centrifugation at 4000 rpm. The sodium caseinate precipitate was washed twice with the buffer, dialyzed against water and finally freeze-dried and stored at -24°C .

An aliquot of the casein fraction was dephosphorylated using alkaline phosphatase from bovine intestinal mucosa.^[19] The casein fraction was dissolved in 0.4% NH₄HCO₃ aqueous solution, pH 8.5, at a concentration of 1 $\mu\text{g}/\mu\text{l}$ and a 20 U enzyme/mg protein ratio was used. The solution was incubated at 37°C for 2 h and the digestion was stopped by freezing.

Isolation of dephosphorylated α_{s1} -CNs by RP-HPLC

An enriched fraction of α_{s1} -CNs was obtained by separation of the dephosphorylated casein mixture by analytical RP-HPLC carried out on a Varian 9010 high-performance liquid chromatograph (Sunnyvale, CA, USA) equipped with a Varian 9050 UV detector. Data were acquired with a PC using the software system 'Peak Simple II'. The dephosphorylated casein fraction solution (in NH₄HCO₃) was filtered on micro-spin filters (Alltech, Milan, Italy) and loaded onto a reversed-phase Vydac C₄ column (4.6×250 mm, 300 Å, 5 μm). The column was eluted at 50°C with a linear gradient of solvent B (CH₃CN + 0.05% TFA) in A (H₂O + 0.05% TFA) from 35 to 55% in 60 min at a flow rate of 1 ml/min. Peaks were detected by their absorption at 224 nm, manually collected and freeze-dried.

Two-dimensional electrophoresis

The enriched α_{s1} -CNs fraction was subjected to a 2D-PAGE separation. The protein sample was dissolved in 25 mM Tris-HCl, pH 9.0, containing 2% 3-((3-Cholamidopropyl)dimethylammonium)-1-propanesulfonate (CHAPS), 6 M urea, 2 M thiourea, 0.5% Immobilized pH Gradient (IPG) buffer (pH 4–7; GE Healthcare, Milan, Italy) and traces of bromophenol blue (BPB), to give a final protein

concentration of 0.4 $\mu\text{g}/\mu\text{l}$ in 250 μl . The proteins were reduced with 43 mM Dithiothreitol (DTT) for 2 h and then alkylated for 1 h with 60 mM acrylamide.

The pH 4–7 linear IPG strips (13 cm; GE Healthcare, Milan, Italy) used for the first dimension separation, were rehydrated overnight in 250 μl of the protein samples. Focusing was performed with a multiphor apparatus (GE Healthcare, Milan, Italy) at 20°C with the following program: 1 min at 200 V followed by a linear voltage gradient that went to 3500 V in 90 min and by 3 h of constant voltage (3500 V).

The strips were equilibrated for the second dimension separation under continuous shaking in equilibration buffer [50 mM Tris-HCl, pH 8.8, containing 6 M urea, 30% glycerol, 2% sodium dodecyl sulfate (SDS) and traces of BPB] containing 25 mM DTT. After 20 min, the strips were alkylated in equilibration buffer containing 360 mM acrylamide for 20 min under continuous agitation, and applied on 12% SDS-polyacrylamide gel electrophoresis (PAGE) gels (14×16 cm). The second dimension separation was performed at 20 mA/gel to allow for the tracking dye to exit the IPG strip and penetrate the resolving gel, and then at 30 mA/gel until the tracking dye reached the gel bottom. The gels were stained with Coomassie Brilliant Blue (CBB) R-350 (Amersham Pharmacia Biotech AB, Uppsala, Sweden) in a mixture of CH₃OH, H₂O and acetic acid (3:6:1 v/v/v). The gel images were acquired using a Epson 1680 Pro scanner.

In-gel digestion of protein spots

In-gel digestion of the protein spots was carried out on selected gel pieces manually excised from the Coomassie blue-stained two-dimensional electrophoresis (2-DE) gels. The excised gel pieces were transferred to 1.5 ml microcentrifuge tubes, washed three times with H₂O/CH₃CN (1:1 v/v) at 37°C and subjected to in-gel digestion using modified porcine trypsin or endoproteinase Glu-C, according to the method of Shevchenko *et al.*^[20] After soaking trypsin (or endoproteinase Glu-C) into the gel pieces, the supernatant containing an excess of enzyme was discarded and the gel pieces were covered with 50 μl of 50 mM NH₄HCO₃. Digestion was allowed to proceed overnight at 37°C . The enzymatic reactions were stopped by cooling the gel pieces and supernatant solution at -24°C .

MALDI-TOF MS analysis

Matrix-assisted laser desorption/ionization (MALDI) mass spectra were acquired using a Voyager DE-PRO time-of-flight (TOF) mass spectrometer (Applied Biosystems, Foster City, CA, USA) equipped with a UV nitrogen laser (337 nm).

For MALDI-TOF analysis of the intact proteins, samples of the chromatographic fractions isolated from the RP-HPLC run were prepared according to the dried droplet method^[21] using SA (dissolved in CH₃CN/0.1% TFA, 40:60, v/v at a concentration of 10 $\mu\text{g}/\mu\text{l}$) as matrix. Spectra were obtained in linear positive ion mode over a *m/z* range 10 000–30 000 and were averaged from about 150 laser shots to improve the signal-to-noise (*S/N*) ratio. Mass calibration was made using a mixture (1:1.5 v/v/v) of cytochrome *c* (*m/z* 12 232), myoglobin (*m/z* 16 952) and trypsinogen (*m/z* 23 982) as external standards. The accuracy obtained by MALDI-TOF measurement of the molecular masses of the proteins was 0.04%.

The mixtures of enzymatic peptides were subjected to a desalting/concentration step^[22] prior to analysis by MALDI-TOF

MS. In detail, 10 μ l of the solution of the enzymatic mixture of peptides were loaded onto a homemade 5-mm nanocolumn packed with C₁₈ resin (POROS R2; PerSeptive Biosystems, Foster City, CA, USA) in a constricted GELoader tip (Eppendorf Scientific, Westbury, NY). A syringe was used to force liquid through the column by applying gentle air pressure. The column was washed with 10 μ l of 0.1% TFA and the bound peptides subsequently eluted directly onto the MALDI target with 0.6 μ l of matrix solution [10 μ g/ μ l CHCA in 70% (v/v) CH₃CN and 0.1% (v/v) TFA].

Spectra were obtained in reflectron positive ion mode over a m/z range 600–3000 and were averaged from about 150 laser shots to improve the signal-to-noise ratio. The m/z software (Proteometrics Ltd, New York, NY, USA) was used to analyze the MALDI-TOF spectra, which were internally calibrated using porcine trypsin autolysis products (m/z 842.51 and 2211.10). When trypsin autolysis peaks could not be detected, bovine β -lactoglobulin tryptic peptides (m/z 837.48, 2313.26 and 2707.38) were used for external mass calibration.

Capillary RP-HPLC/nESI-MS/MS analysis

Capillary RP-HPLC/nano-electrospray ionization tandem mass spectrometry (nESI-MS/MS) analysis of the in-gel enzymatic digests was performed using an Ultimate 3000 LC system combined with an autosampler and a flow splitter 1:100 (Dionex Corporation, Sunnyvale, CA, USA), coupled on-line with a linear ion trap nano-electrospray mass spectrometer (LTQ, Thermo Fischer Scientific, San Jose, CA, USA). Ionization was performed with liquid junction using an uncoated capillary probe (20 mm i.d.; New Objective, Woburn, MA, USA). After in-gel digestion, the digested solution was transferred into a clean 0.5-ml tube. The peptides were extracted from gel pieces with 40 μ l of 0.1% TFA and subsequently with an equal volume of CH₃CN. This extraction procedure was repeated three times. The total extracts were pooled, joined with the first supernatant, lyophilized and re-dissolved in 20 μ l of 0.1% TFA.

Then, 10 μ l of this solution were loaded onto a μ -pre-column C₁₈ (0.3 $\times$ 5 mm, 100 Å, 5 μ m, PepMap, Dionex) equilibrated with 0.5% FA aqueous solution at a flow rate of 20 μ l/min for 4 min. Then, the solution was switched onto a reversed-phase C₁₈ column (0.18 $\times$ 150 mm, 300 Å, 5 μ m, BioBasic, ThermoFisher Scientific) and peptides were separated by elution at room temperature with a linear gradient of solvent B (CH₃CN + 0.5% FA) in A (H₂O + 0.5% FA) from 10 to 50% in 35 min at a flow rate of 2 μ l/min. The nESI source operated under the following conditions: source temperature 220 °C, source voltage 1.9 kV and capillary voltage 42 V. Repetitive mass spectra were acquired in positive ion mode in the m/z range 350–2000.

Characterization of peptide ions was performed by the data-dependent method as follows: (1) full scan MS in the m/z range 350–2000; (2) zoom scan of the five most intense ions (isolation width: 2); (3) MS/MS analysis of the five most intense ions (normalized collision energy: 24 a.u., activation Q: 0.250). Mass calibration was made using a standard mixture of caffeine (M_r , 194.1 Da), MRFA peptide (M_r , 423.6 Da) and Ultramark (M_r , 1621 Da). Data acquisition and data analysis were performed using the Excalibur v. 1.4 and Bioworks Browser v. 3.2 software (ThermoFisher Scientific), respectively. *De novo* amino acid sequence of the peptides investigated was also performed by manual interpretation of the MS/MS spectra assisted by the use of the AminoCalc v. 1.13 software (Protana, Odense, Denmark).

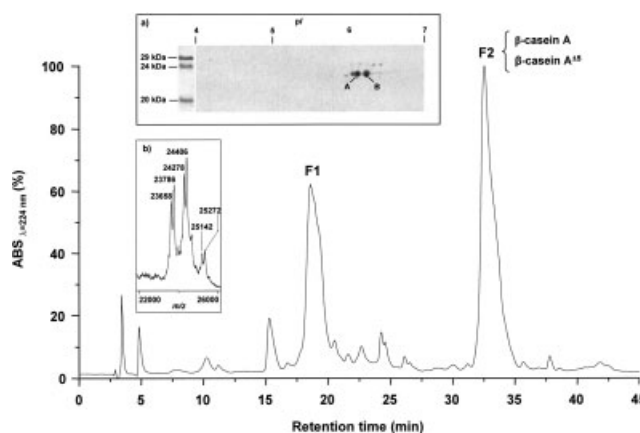


Figure 1. UV profile ($\lambda = 224$ nm) of the RP-HPLC run of the dephosphorylated casein fraction obtained from milk of an individual donkey belonging to the Ragusano breed. Insets show the 2D-PAGE analysis (a) and the MALDI-TOF mass spectrum (b) of the chromatographic fraction F1.

Results and Discussion

Characterization of the primary structure of the α_{s1} -caseins

The dephosphorylated casein mixture was separated by RP-HPLC with an analytical C₄ column into two main fractions named F1 and F2 (Fig. 1). The MALDI-TOF MS analysis of the collected fractions allowed to detect that the F2 fraction contained exclusively two β -CN variants (UniProtKB/Swiss-Prot Acc. No. P86273) previously characterized^[18] and showed the co-elution in the F1 fraction of four unidentified major components, having an experimentally measured M_r of 23 658, 23 786, 24 278 and 24 406 $\pm$ 10 Da, together with two minor ones with experimentally measured M_r of 25 142 and 25 272 $\pm$ 10 Da [Fig. 1, inset (b)].

In order to separate these unknown components, the F1 fraction was further resolved by 2D-PAGE run. The 2-DE map of this fraction [Fig. 1, inset (a)] revealed the presence of two main protein spots, labeled as spot A and spot B, having an apparent molecular mass of 23 kDa and pI 6–6.3. The high resolution obtained by using IPG strips of narrow pH range (pH 4–7) allowed the four most abundant proteins to separate as two couples into two distinct spots due to a slight difference in their pIs. On the other hand, the differences in the relative molecular masses of the proteins were not sufficient to obtain a separation based on migration rate along the SDS-PAGE direction. These two major spots were manually excised, subjected to in-gel digestion with modified porcine trypsin and the resulting peptide mixtures were analyzed by MALDI-TOF MS.

Figure 2 shows the MALDI-TOF mass spectra of the tryptic digests of the two spots investigated. Although most of the signals present in the two MALDI-TOF mass spectra were common, the signals at m/z 1712.8, 1840.9 and 1997.0 were exclusively present in the spectrum of spot A (Fig. 2(a)), whereas the signals at m/z 1172.6, 1179.6, 1300.7 and 2461.3 were only present in the spectrum of spot B (Fig. 2(b)). Almost all the signals present in the MALDI-TOF mass spectra can be interpreted as arising from the theoretical cleavage of the α_{s1} -CN of mare (GenBank Acc. No. AAK83668).

Specifically, the common signals at m/z 633.3, 1675.7, 1831.9 and 2042.0 are attributable to the MH⁺ of the T₂₂, T₁₇, T₍₁₇₊₁₈₎ and T₍₁₆₊₁₇₎, respectively (Table 1). The three signals at m/z 1712.8, 1840.9 and 1997.0 present only in the spectrum of the spot A (Fig. 2(a)) correspond to the theoretical tryptic fragments T₈, T₍₇₊₈₎

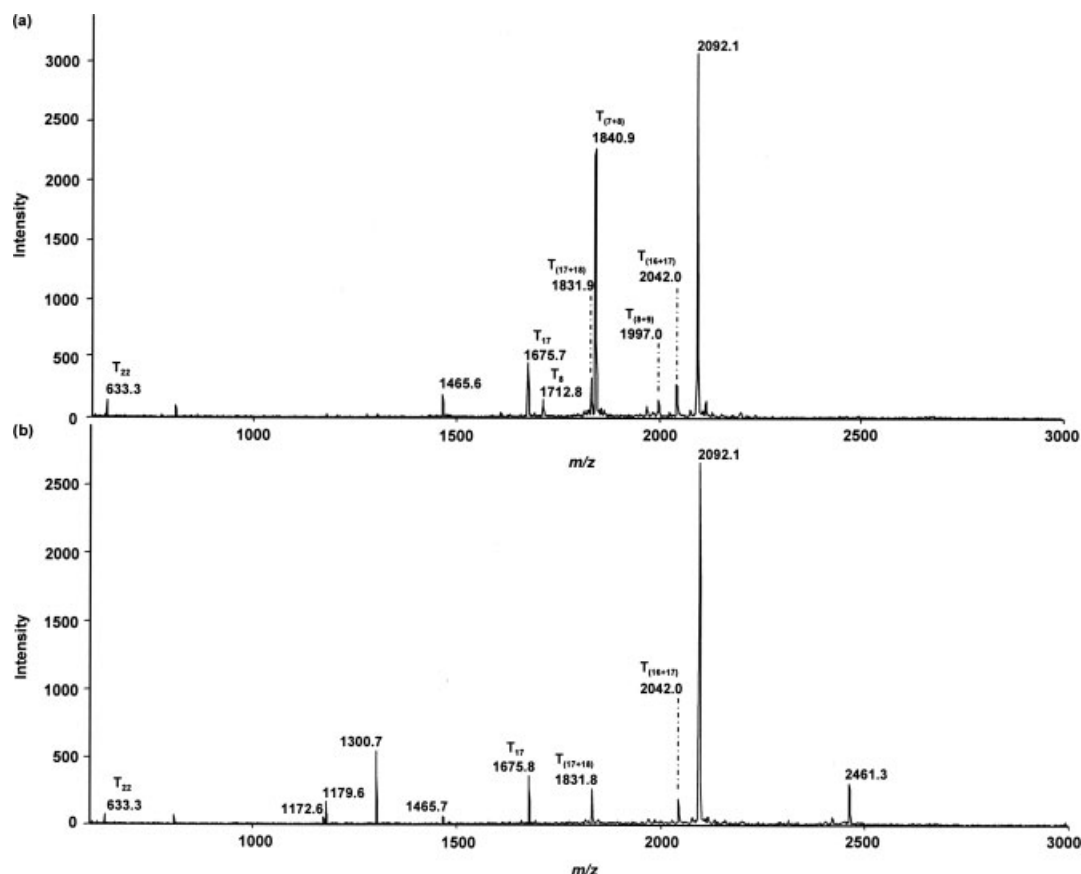


Figure 2. MALDI-TOF mass spectra of the in-gel tryptic digest of the spots A (a) and B (b) present in the 2D map [Fig. 1, inset (a)]. The spectra were acquired in the m/z range 600–3000, obtained operating in reflectron positive ion mode and using CHCA as matrix. The spectra were calibrated using a tryptic digest of β -Lactoglobulin (LG), as external standard. The nomenclature of the tryptic fragments is referred to the Table 1.

and $T_{(8+9)}$ (Table 1), respectively. This finding clearly indicates that the components of fraction F1 are donkey's α_{s1} -CNs.

In addition to the attributable peaks, the two MALDI-TOF mass spectra show several signals, including some of remarkable intensity, which remain unassigned because they are probably due to donkey's α_{s1} -CNs tryptic peptides containing amino acid substitutions, deletions or insertions with respect to the sequence of the reference mare's α_{s1} -CN; therefore, they cannot be identified by the peptide mass fingerprinting investigation.

In order to obtain additional informations about the amino acid sequence of the donkey's α_{s1} -CNs, the tryptic mixtures were analyzed by capillary RP-HPLC/nESI-MS/MS and the results interpreted using the sequence of the mare's α_{s1} -CN as reference.

Characterization of the primary structure of the proteins with M_r 23 658 and 23 786 (spot A)

Figure 3(a) shows the total ion current (TIC) chromatogram obtained by capillary RP-HPLC/nESI-MS of the tryptic digest of the spot A. By the MS/MS spectra of the doubly or triply charged molecular ions of the parent peptides, characterization of all the fragments already identified by MALDI-TOF MS was confirmed. The capillary RP-HPLC/nESI-MS/MS analysis also allowed the detection of peptides previously not detected by MALDI analysis. Particularly, the theoretical tryptic fragments T_{14} (MH^+ 3282.9), $T_{(12+13+14)}$ (MH^+ 4048.6) and $T_{(25+26)}$ (MH^+ 1845.2) were identified and characterized (Table 1). Additionally, three peptides with MH^+ 1260.7, 2022.0 and 2653.6 were identified as the sub-fragments

Val¹³⁴–Tyr¹⁴³, Tyr¹⁴³–Tyr¹⁵⁸ and Ala¹⁵⁹–Lys¹⁸¹ of the previously undetected T_{24} peptide, and arising from unspecific tryptic cleavages at the Tyr¹⁴³–Leu¹⁴⁴, Phe¹⁴²–Tyr¹⁴³ and Tyr¹⁵⁸–Ala¹⁵⁹ bonds. These fragments were probably produced by the activity of pseudotrypsin which is generated by autolysis of trypsin and cleaves at the C-terminal side of aromatic amino acids.^[23,24]

Furthermore, the MS/MS spectrum of the triply charged ion at m/z 1052.3 (MH^+ 3154.9, Fig. 4) allowed to deduce the sequence EYLIEDPEQQESSSTSSSEEVPIINTEK, which corresponds to the tryptic fragment T_{14} with the deletion of Gln residue at position 83 (Table 1). This result was confirmed by the presence of the triply charged molecular ion at m/z 1307.5 (MH^+ 3920.5), which corresponds to the theoretical tryptic fragment $T_{(12+13+14)}$ in which the Gln residue at position 83 is deleted (Table 1).

Finally, interpretation of the MS/MS spectra of the peptides responsible for the unassigned MALDI-MS peaks at m/z 1465.6 and 2092.1, allowed the identification of two amino acid substitutions in the donkey's α_{s1} -CN with respect to the mare's counterpart. In detail, the interpretation of the MS/MS spectra of the doubly charged molecular ions at m/z 733.4 [MH^+ 1465.8, Fig. 5(a)] and m/z 1046.7 [MH^+ 2092.4, Fig. 5(b)] allowed to deduce the sequences HPEIIQNEQDSR and YNQLQLQAIYAQEQLIR, respectively. The deduced sequence HPEIIQNEQDSR corresponds to the tryptic fragment T_3 (Gln⁸–Arg¹⁹) of the mare's α_{s1} -CN, carrying the amino acid substitution Gln⁸ $\rightarrow$ His (Table 1), whereas the deduced sequence YNQLQLQAIYAQEQLIR matches with the

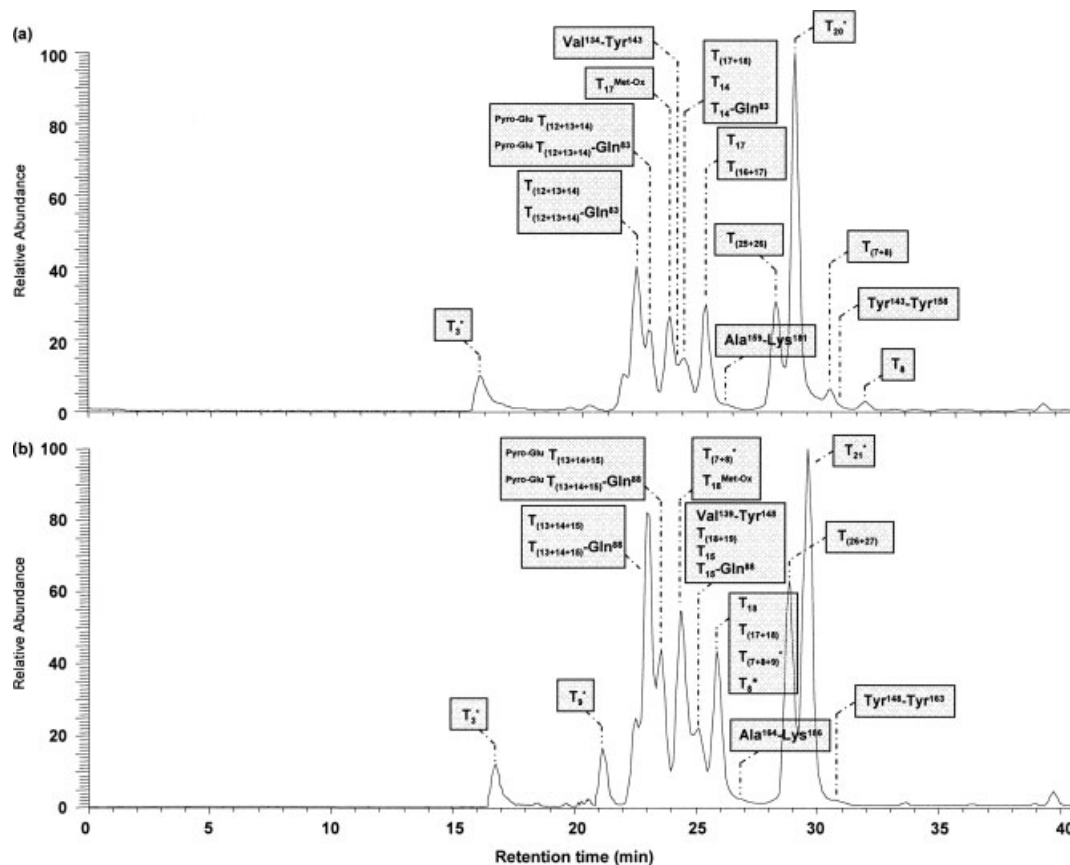


Figure 3. Total ion current (TIC) chromatogram obtained by capillary RP-HPLC/nESI-MS of the tryptic digests of the 2D protein spot A (a) and spot B (b). The nomenclature of the tryptic fragments reported in panels (a) and (b) is referred to the Tables 1 and 3, respectively.

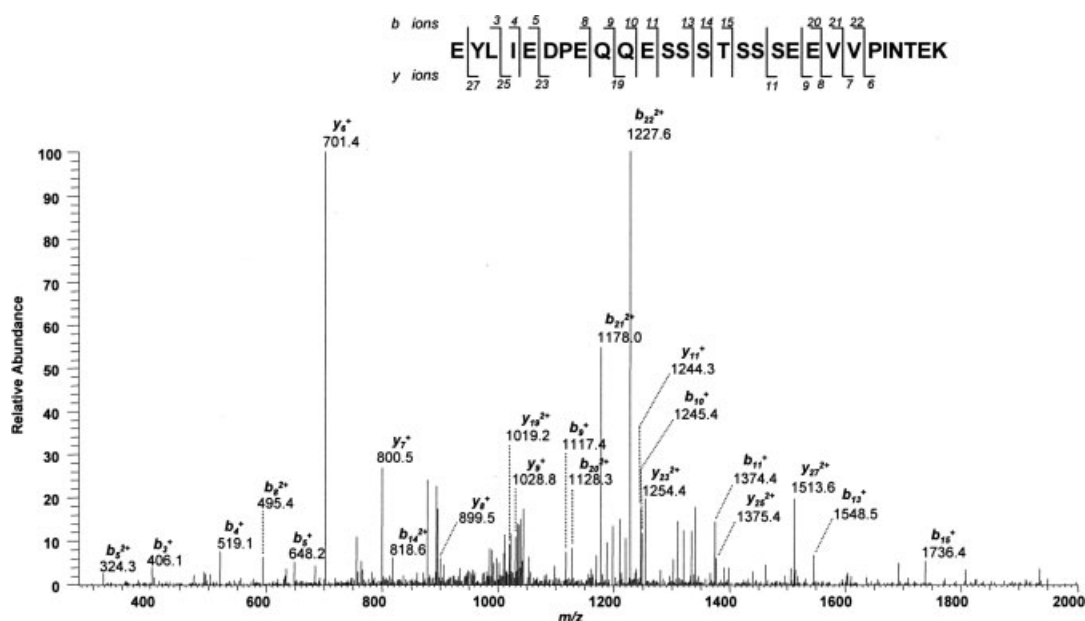


Figure 4. ESI-MS/MS spectrum of the triply charged molecular ion at m/z 1052.3 (MH^+ 3154.9). The sequence of the peptides is displayed with the fragment ions observed in the spectrum.

Table 1. Fragment nomenclature, position, calculated monoisotopic and experimentally measured MH^+ of tryptic fragments of α_{s1} -CNs present in the spot A [Fig. 1, inset a)]

Fragment	Position	Calculated MH^+	Measured MH^+		Reference sequence (α_{s1} -CN from mare; GenBank Acc. No. AAK83668)
			MALDI-MS	ESI-MS/MS (z)	
T ₁	1–3	400.3	–	–	RPK
T ₂	4–7	522.3	–	–	LPHR
T ₃	8–19	1456.7	–	–	QPEIIQNEQDSR
T ₃ ^{*a}	8–19	1465.7	1465.6	1465.8 (2)	<u>H</u> PEIIQNEQDSR ^a
T ₄	20–21	276.2	–	–	EK
T ₅	22–24	359.3	–	–	VLK
T ₆	25–26	304.2	–	–	ER
T ₇	27–27	147.1	–	–	K
T ₈	28–41	1712.9	1712.8	1713.0 (2)	FPSFALEYINELNR
T ₍₇₊₈₎	27–41	1841.0	1840.9	1841.2 (2)	
T ₉	42–43	303.2	–	–	QR
T ₍₈₊₉₎	28–43	1997.0	1997.0	1997.2 (2)	
T ₁₀	44–47	502.3	–	–	ELLK
T ₁₁	48–49	276.2	–	–	EK
T ₁₂	50–51	275.2	–	–	QK
T ₁₃	52–55	528.2	–	–	DEHK
T ₁₄	56–84	3282.5	–	3282.9 (3)	EYLIEDPEQQESSSTSSSEEVVPINTEKQ
T ₍₁₂₊₁₃₊₁₄₎	50–84	4047.9	–	4048.6 (3)	
T ₁₄ ^b	56–83	3154.4	–	3154.9 (3)	EYLIEDPEQQESSSTSSSEEVVPINTEK ^b
T ₍₁₂₊₁₃₊₁₄₎ ^b	50–83	3919.8	–	3920.5 (3)	QKDEHKEYLIEDPEQQESSSTSSSEEVVPINTEK ^b
T ₁₅	85–85	175.1	–	–	R
T ₁₆	86–88	385.3	–	–	IPR
T ₁₇	89–101	1675.8	1675.7	1676.0 (2)	EDMLYQHTLEQLR
T ₍₁₆₊₁₇₎	86–101	2042.0	2042.0	2042.2 (2)	
T ₁₈	102–102	175.1	–	–	R
T ₍₁₇₊₁₈₎	89–102	1831.9	1831.9	1832.0 (2)	
T ₁₉	103–105	347.2	–	–	LSK
T ₂₀	106–122	2066.1	–	–	YNQLQLQAIHAQEQLIR
T ₂₀ ^{*c}	106–122	2092.1	2092.1	2092.4 (2)	YNQLQLQAIYAQEQLIR ^c
T ₂₁	123–124	278.1	–	–	MK
T ₂₂	125–129	633.3	633.3	633.4 (1)	ENSQR
T ₂₃	130–133	531.3	–	–	KPMR
T ₂₄	134–181	5734.7	–	–	VVNQEQAIFYLEPFQPSYQLDVYPYAAWFHPAQIMQHVAYSFPFHDTAK
Val ¹³⁴ –Tyr ¹⁴³	134–143	1260.6	–	1260.7 (1)	
Tyr ¹⁴³ –Tyr ¹⁵⁸	143–158	2021.9	–	2022.0 (2)	
Ala ¹⁵⁹ –Lys ¹⁸¹	159–181	2653.3	–	2653.6 (2)	
T ₂₅	182–190	990.5	–	–	LIASENSEK
T ₂₆	191–197	873.4	–	–	TDIPEW
T ₍₂₅₊₂₆₎	182–197	1844.9	–	1845.2 (2)	

The cDNA derived sequence of α_{s1} -CN (GenBank Acc. No. AAK83668) from mare was taken as reference. The modified tryptic fragments of donkey's α_{s1} -CNs are marked with an asterisk, whereas the amino acid substitutions with respect to the mare's reference α_{s1} -CN are underlined. The glutamine residue at position 83 deleted in the donkey's α_{s1} -CN with M_r 23 658 is reported in italics.

^a Fragment carrying the substitution Gln⁸ → His with respect to the α_{s1} -CN from mare.

^b Specific tryptic fragments of the α_{s1} -CN having M_r 23 658, bearing the Gln⁸³ deletion.

^c Fragment carrying the substitution His¹¹⁵ → Tyr with respect to the α_{s1} -CN from mare.

tryptic fragment T₂₀ (Tyr¹⁰⁶–Arg¹²²) of the mare's α_{s1} -CN with the amino acid substitution His¹¹⁵ → Tyr (Table 1).

In summary, MS/MS analysis of the tryptic digest allowed the characterization of about 85% of the sequence of two donkey's α_{s1} -CN isoforms, giving clear evidence that the two proteins differ in the absence of the Gln⁸³ in the shorter form. The absence of this glutamine residue, which is the first amino acid encoded by exon 11, is probably related to an alternative splicing event occurring during primary transcript processing. On the other hand,

the existence of a cryptic splice site on the first codon (CAG) of exon 11, encoding a glutamine residue, has already been noted in ovine,^[25] caprine^[26] and bovine^[27] α_{s1} -CNs.

The results obtained by the analysis of the tryptic digest are summarized in Table 1, where the tryptic fragments carrying the substitutions are marked with an asterisk. Table 1 also shows that the uncharacterized peptides were mostly low molecular weight or very hydrophilic peptides, which were not detected because their molecular masses were below the scanned mass range or

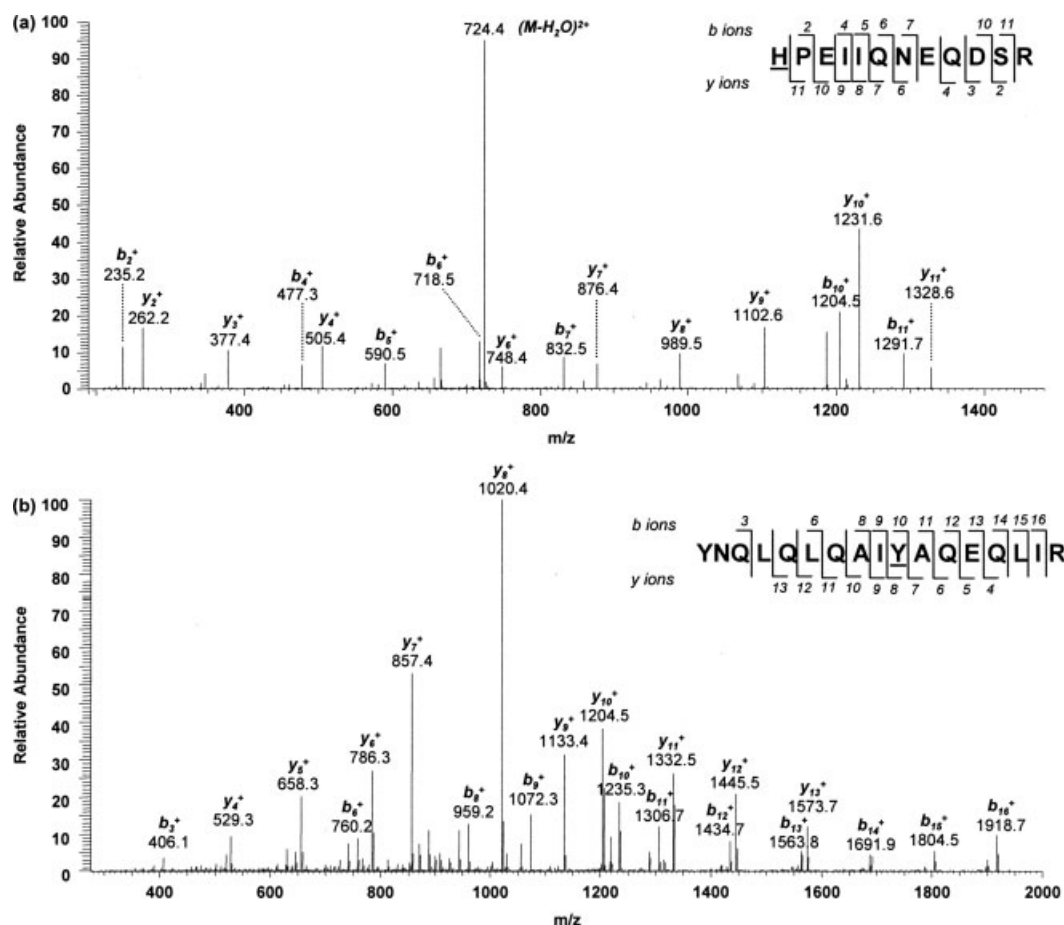


Figure 5. ESI-MS/MS spectra of the doubly charged molecular ions at m/z 733.4 (MH^+ 1465.8) (a) and m/z 1046.7 (MH^+ 2092.4) (b). The sequence of the peptides is displayed with the fragment ions observed in the spectra. The amino acid substitutions with respect to the mare's α_{51} -CN are underlined.

were not recovered because eluting in the void volume of the HPLC column, or both.

In order to obtain complementary data for the characterization of the amino acid sequence of the two donkey's α_{51} -CNs, spot A was digested with endoproteinase Glu-C and the resulting peptide mixture was analyzed by capillary RP-HPLC/nESI-MS/MS. The analysis of the Glu-C digest confirmed most of the results already obtained by the analysis of the tryptic mixture and, in addition, allowed to complete the coverage of the amino acid sequence. Selected fragments identified in the Glu-C digestion mixture of the two donkey's α_{51} -CNs and used to complete the sequence coverage are reported in Table 2.

In conclusion, the results obtained show that the two α_{51} -CNs present in the spot A correspond to the components with M_r 23 786 and 23 658 present in the chromatographic fraction F1. The two donkey's α_{51} -CNs differ in the absence of the Gln⁸³ residue in the shorter form (196 amino acids; M_r 23 658 Da), which explains the difference of 128 Da in their molecular masses. In addition, they show two amino acid substitutions (Gln⁸ $\rightarrow$ His and His¹¹⁵ $\rightarrow$ Tyr) with respect to the homologous mare's α_{51} -CN (Fig. 6).

Characterization of the primary structure of the proteins with M_r 24 278 and 24 406 (spot B)

Figure 3(b) shows the TIC chromatogram obtained by capillary RP-HPLC/nESI-MS of the tryptic digest of the spot B. Taking into account the amino acid sequences of the two characterized

donkey's α_{51} -CNs identified in spot A and considering the signals present in the MALDI mass spectrum of the spot B (Fig. 2(b)), the interpretation of the MS/MS spectra of the doubly or triply charged molecular ions of all the eluting peptides enabled to reconstruct almost all the primary structure of the donkey's α_{51} -CNs with M_r 24 278 and 24 406 Da. Specifically, the characterization of the peptides with MH^+ 3282.9 and 4048.3, corresponding to fragments T₁₅ and T₍₁₃₊₁₄₊₁₅₎, and of the peptides with MH^+ 3154.6 and 3920.3 corresponding to the same fragments T₁₅ and T₍₁₃₊₁₄₊₁₅₎ without the Glu⁸⁸ (Table 3), allows to assert that these two donkey's α_{51} -CNs differ in the absence of the glutamine residue at position 88 in the shorter form. Moreover, the identification and characterization of the doubly charged ions at m/z 733.4 (MH^+ 1465.8, Table 3) and m/z 1046.6 (MH^+ 2092.2, Table 3) show that they present, in comparison with the homologous mare's α_{51} -CNs, the same amino acid substitutions observed in the variants co-migrating in spot A. The MS/MS spectra of the doubly charged ions at m/z 650.9 [MH^+ 1300.8, Fig. 7(a)] and m/z 590.3 [MH^+ 1179.6, Fig. 7(b)] allowed to deduce the sequences KFPSFALHTPR and EEYINELNR, respectively, indicating that the two donkey's α_{51} -CNs co-migrating in spot B present the insertion of the pentapeptide HTPRE between the amino acids Leu³³ and Glu³⁴ (Table 3, Fig. 6). This result was confirmed by the MS/MS spectrum of the fragment with MH^+ 2461.2, from which the sequence KFPSFALHTPREYINELNR could be deduced (Table 3). This peptide therefore corresponds to the theoretical tryptic fragment T₍₇₊₈₎ (K^{27} FPSFALEYINELNR⁴¹) of the

Table 2. Position, calculated and measured monoisotopic MH^+ of selected fragments identified in the Glu-C digestion mixture of donkey's α_{s1} -CNs present in the spot A [Fig. 1, inset (a)], used to increase the coverage of the amino acid sequence

Position	Calculated MH^+	Measured MH^+ by ESI-MS/MS (z)	Reference sequence (α_{s1} -CN from mare; GenBank Acc. No. AAK83668)
1–20 ^a	2479.3	2479.6 (3)	RPKLPHRH PEIIQNEQDSRE ^a
21–34	1692.0	1692.2 (2)	KVLKERK FPSFALE
39–53	1927.0	1927.2 (2)	LN RQRELLKEK QKDE
84–98	1929.0	1929.2 (2)	K RIPRED MLYQHTLE
99–118 ^b	2463.3	2463.6 (3)	QL RRLSKYNQLQLQAIYAQ E ^b
119–125	917.5	917.6 (1)	QL IRMKE
126–138	1585.8	1586.0 (2)	NS QRKPM RVVNQE

Sequences not previously covered by tryptic cleavage are reported in bold. The sequence of α_{s1} -CNs (Acc. No. AAK83668) from mare was taken as reference. The amino acid substitutions with respect to the mare's reference α_{s1} -CN are underlined.

^a Fragment carrying the substitution Gln⁸ → His with respect to the mare counterpart.

^b Fragment carrying the substitution His¹¹⁵ → Tyr with respect to the mare counterpart.

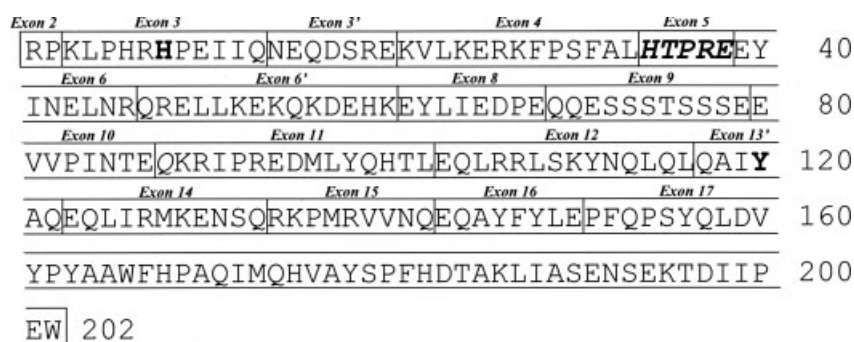


Figure 6. Exon modular structure of the primary structure of the α_{s1} -CNs identified in a milk sample collected from an individual donkey belonging to the Ragusano breed. The exon nomenclature is that used by Miranda *et al.*^[8] Amino acid sequences were characterized by coupling enzymatic digestions, MALDI-TOF MS and RP-HPLC/nESI-MS/MS analysis and using the known mare's α_{s1} -CN (GenBank Acc. No. AAK83668) as reference. The insertion of the pentapeptide HTPRE present in the variants A and A₁ and endocued by exon 5, is reported in bold italics. The glutamine residue absent in the variants A₁ and B₁ is reported in italics. The two amino acid differences with respect to the mare's α_{s1} -CN are reported in bold.

mare's α_{s1} -CN, carrying the insertion of the pentapeptide HTPRE between the amino acids Leu³³ and Glu³⁴.

In order to complete the determination of the amino acid sequence of the two donkey's α_{s1} -CNs, spot B was digested with endoproteinase Glu-C and the resulting peptide mixture was analyzed by capillary RP-HPLC/nESI-MS/MS. The analysis of the Glu-C digest allowed the complete characterization of the sequence, confirming the insertion of the pentapeptide in the N-terminal region of the two donkey's α_{s1} -CNs co-migrating in spot B. Selected fragments identified in the Glu-C digestion mixture of the two donkey's α_{s1} -CNs and used to complete the sequence coverage are reported in Table 4.

In summary, the results obtained revealed that the two α_{s1} -CNs present in the spot B correspond to the components with M_r 24406 and 24278 [Fig. 1, inset (b)]. These two isoforms differ in the absence of the Gln⁸⁸ residue in the shorter form (201 amino acids; M_r 24278 Da), which explains the difference of 128 Da in their molecular masses. In comparison with the two α_{s1} -CNs co-migrating in the spot A, they show the insertion of the pentapeptide HTPRE between the amino acids Leu³³ and Glu³⁴ (Fig. 6) that accounts for the increase of 620 Da in their M_r .

Primary structure comparison of the donkey's and cow's α_{s1} -caseins

Studies of large populations of patients allergic to cow's milk proteins show that α_{s1} -CN, the most abundant protein in the casein

fraction, represents one of the major cow's milk allergens.^[28] In bovine α_{s1} -CN disulphide bonds are absent and therefore it shows a reduced rigidity of the tertiary structure.^[29] As a consequence, the major allergenic epitopes of this protein are linear rather than conformational.

As results of several investigations, six major (AA 17–36, AA 83–102, AA 109–120, AA 139–154, AA 159–174 and AA 173–194) and three minor (AA 39–48, AA 69–78 and AA 123–132) sequential IgE-binding regions in bovine α_{s1} -CN were identified. Among these, two epitopes (AA 69–78 and AA 173–194) are considered to contribute to persistent cow's milk allergy.^[30] Therefore, the wealth of knowledge about the allergenic epitopes of bovine α_{s1} -CN, together with the determination of the primary structure of α_{s1} -CN from donkey reported here provide the basis for some comparison of their sequences. The best alignment of the bovine and donkey's α_{s1} -CN (Fig. 8) variants B reveals that these two proteins share a low sequence homology (41% of identity and 57% of similarity). Particularly, alignment reveals that the IgE-binding linear epitopes of cow's α_{s1} -CN and the corresponding domains present in donkey's α_{s1} -CN have remarkable differences in their amino acid sequences. It should also be noted that goat's and cow's α_{s1} -CNs show 87% of identity and 93% of similarity, and that anaphylactic reactions to goat's milk in children with CMPA have been reported.^[31] Moreover, individuals with CMPA may also be allergic to goat's milk^[32] and a cross-reactivity between cow's

Table 3. Fragment nomenclature, position, calculated monoisotopic and experimentally measured MH^+ of tryptic fragments of α_{s1} -CNs present in the spot B [Fig. 1, inset (a)]

Fragment	Position	Calculated MH^+	Measured MH^+		Reference sequence (α_{s1} -CN from mare; GenBank Acc. No. AAK83668)
			MALDI-MS	ESI-MS/MS (z)	
T ₁	1–3	400.3	–	–	RPK
T ₂	4–7	522.3	–	–	LPHR
T ₃	8–19	1456.7	–	–	QPEIIQNEQDSR
T ₃ [*]	8–19	1465.7	1465.6	1465.8 (2)	<u>H</u> PEIIQNEQDSR ^a
T ₄	20–21	276.2	–	–	EK
T ₅	22–24	359.3	–	–	VLK
T ₆	25–26	304.2	–	–	ER
T ₇	27–27	147.1	–	–	K
T ₈	28–41	1712.9	–	–	FPSFALEYINELNR
T ₈ ^{*b}	28–37	1172.6	1172.6	1172.6 (2)	FPSFAL HTPR ^b
T ₍₇₊₈₎ ^{*b}	27–37	1300.7	1300.7	1300.8 (2)	
T ₉ ^{*b}	38–46	1179.6	1179.6	1179.6 (2)	E EYINELNR ^b
T ₍₇₊₈₊₉₎ ^{*b}	27–46	2461.3	2461.3	2461.2 (2)	
T ₁₀	47–48	303.2	–	–	QR
T ₁₀	49–52	502.3	–	–	ELLK
T ₁₂	53–54	276.2	–	–	EK
T ₁₃	55–56	275.2	–	–	QK
T ₁₄	57–60	528.2	–	–	DEHK
T ₁₅	61–89	3282.5	–	3282.9 (3)	EYLIEDPEQQESSSTSSSEEVVPINTEQK
T ₍₁₃₊₁₄₊₁₅₎	55–89	4047.9	–	4048.3 (3)	
T ₁₅ ^c	61–88	3154.4	–	3154.6 (3)	EYLIEDPEQQESSSTSSSEEVVPINTEK ^c
T ₍₁₃₊₁₄₊₁₅₎ ^c	55–88	3919.8	–	3920.3 (3)	QKDEHKEYLIEDPEQQESSSTSSSEEVVPINTEK ^c
T ₁₆	90–90	175.1	–	–	R
T ₁₇	91–93	385.3	–	–	IPR
T ₁₈	94–106	1675.8	1675.8	1675.8 (2)	EDMLYQHTLEQLR
T ₍₁₇₊₁₈₎	91–106	2042.0	2042.0	2042.2 (2)	
T ₁₉	107–107	175.1	–	–	R
T ₍₁₈₊₁₉₎	94–107	1831.9	1831.8	1832.0 (2)	
T ₂₀	108–110	347.2	–	–	LSK
T ₂₁	111–127	2066.1	–	–	YNQLQLQAIHAQEQLIR
T ₂₁ ^{*d}	111–127	2092.1	2092.1	2092.2 (2)	YNQLQLQAIYAQEQLIR ^d
T ₂₂	128–129	278.1	–	–	MK
T ₂₃	130–134	633.3	633.3	633.3 (1)	ENSQR
T ₂₄	135–138	531.3	–	–	KPMR
T ₂₅	139–186	5734.7	–	–	VVNQEQAIFYLEPFQPSYQLDVYPYAAWFHPAQIMQHVAYSFPFHDATK
Val ¹³⁹ –Tyr ¹⁴⁸	139–148	1260.6	–	1260.7 (1)	
Tyr ¹⁴⁸ –Tyr ¹⁶³	148–163	2021.9	–	2022.2 (2)	
Ala ¹⁶⁴ –Lys ¹⁸⁶	164–186	2653.3	–	2653.4 (2)	
T ₂₆	187–195	990.5	–	–	LIASENSEK
T ₂₇	196–202	873.4	–	–	TDIPEW
T ₍₂₆₊₂₇₎	187–202	1844.9	–	1845.0 (2)	

The modified tryptic fragments of donkey's α_{s1} -CNs are marked with an asterisk, whereas the amino acid substitutions with respect to the mare's reference α_{s1} -CN are underlined. The insertion of the pentapeptide HTPRE is reported in bold italics. The glutamine residue at position 88 deleted in the donkey's α_{s1} -CN with M_r 24 278 is reported in italics.

^a Fragment carrying the substitution Gln⁸ → His with respect to the α_{s1} -CN from mare.

^b Fragments showing the pentapeptide insertion with respect to the α_{s1} -CN from mare.

^c Specific tryptic fragments of the α_{s1} -CN having M_r 24 278, bearing the Gln⁸⁸ deletion.

^d Fragment carrying the substitution His¹²⁰ → Tyr with respect to the α_{s1} -CN from mare.

and goat's milk in placebo-controlled food challenges has been documented.^[33]

Therefore, the considerable differences between the primary structure of donkey and bovine α_{s1} -CNs could be related to the already demonstrated low allergenic properties of donkey's milk and could contribute to explain the better tolerance of donkey's

milk with respect to the goat's one for children suffering from CMPA.^[34]

Although these considerations seem reasonable, further biochemical studies and clinical experiences are needed to establish precise relationship between donkey's milk protein structures and its low allergenic properties.

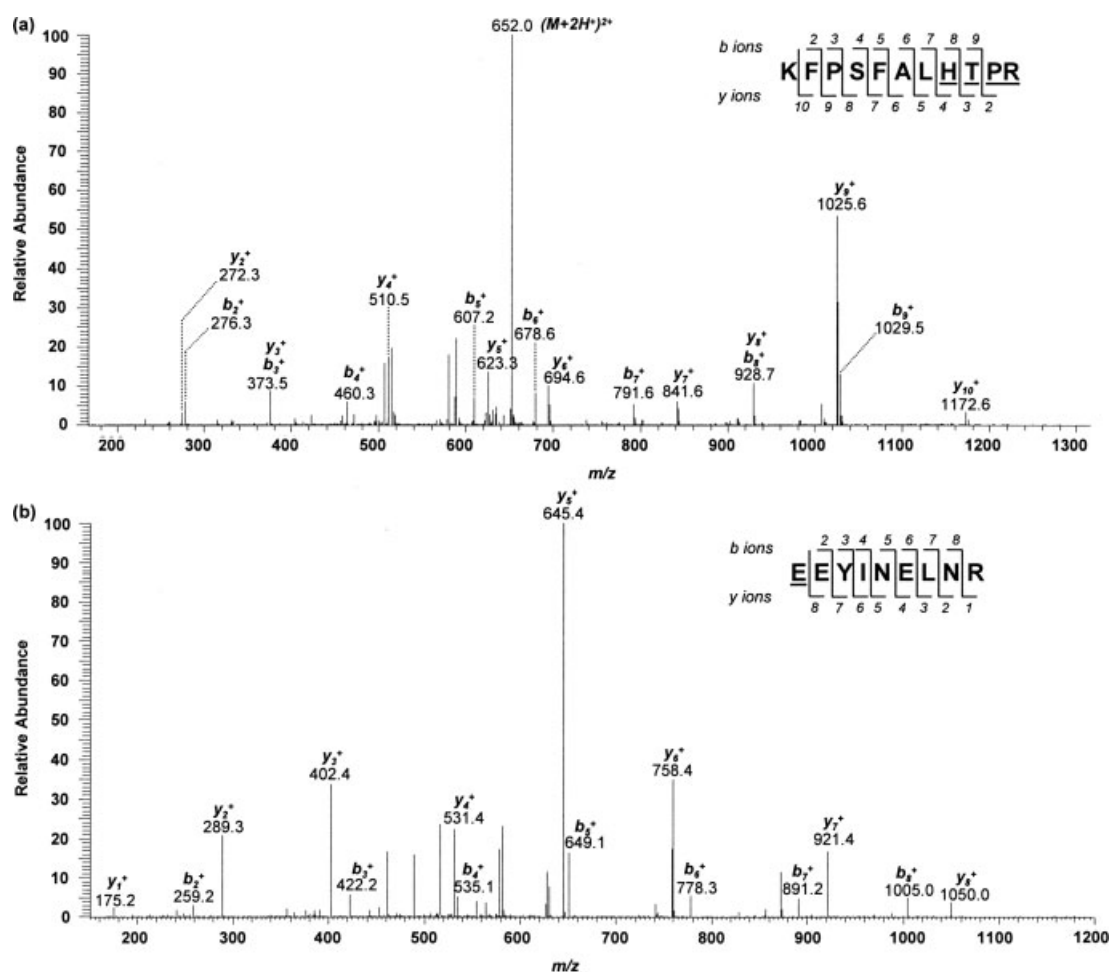


Figure 7. ESI-MS/MS spectra of the doubly charged molecular ions at m/z 650.9 (MH^+ 1300.8) (a) and m/z 590.3 (MH^+ 1179.6) (b). The sequence of the peptides is displayed with the fragment ions observed in the spectra. The amino acid differences with respect to the mare's α_{s1} -CN are underlined.

Table 4. Position, calculated and measured monoisotopic MH^+ of selected fragments identified in the Glu-C digestion mixture of donkey's α_{s1} -CNs present in the spot B [Fig. 1, inset (a)], used to increase the coverage of the amino acid sequence

Position	Calculated MH^+	Measured MH^+ by ESI-MS/MS (z)	Reference sequence (α_{s1} -CN from mare; GenBank Acc. No. AAK83668)
1–20 ^a	2479.3	2479.6 (3)	RPKLPHR <u>H</u> PEIIQNEQDSRE ^a
21–43	2831.5	2831.6 (3)	KVLKER KFPSPFALHTPREEYINE
50–80	3650.7	3650.9 (3)	LLKEKQKDEHKEYLIEDPEQQESSSTSSSEE
39–53	1927.0	1927.2 (2)	LNRR QELLKEKQKDE
84–98	1929.0	1929.2 (2)	KRIPREDMLYQHTLE
104–123 ^b	2463.3	2463.6 (3)	QLRR LSKYNQLQLQAIY AQE ^b
124–130	917.5	917.6 (1)	QLIR MKE
131–143	1585.8	1586.0 (2)	NSQR KPMRVVNQE

Sequences not previously covered by tryptic cleavage are reported in bold. The sequence of α_{s1} -CNs (Acc. No. AAK83668) from mare was taken as reference. The amino acid substitutions with respect to the mare's reference α_{s1} -CN are underlined. The insertion of the pentapeptide is reported in italics.

^a Fragment carrying the substitution Gln⁸ → His with respect to the mare counterpart.

^b Fragment carrying the substitution His¹²⁰ → Tyr with respect to the mare counterpart.

Conclusions

In this study, four major co-eluting proteins, having an experimentally measured M_r of 23 658, 23 786, 24 278 and 24 406 ± 10 Da, were detected, together with two minor ones with experimentally measured M_r 25 142 and 25 272 ± 10 Da, by RP-HPLC analysis

and MALDI-TOF mass spectrometry of the dephosphorylated casein fraction of a milk sample collected at middle lactation stage from an individual donkey belonging to the Ragusano breed.

By coupling RP-HPLC, 2D-PAGE, enzymatic digestions, MALDI-TOF MS and capillary RP-HPLC/nESI-MS/MS analyses, the four

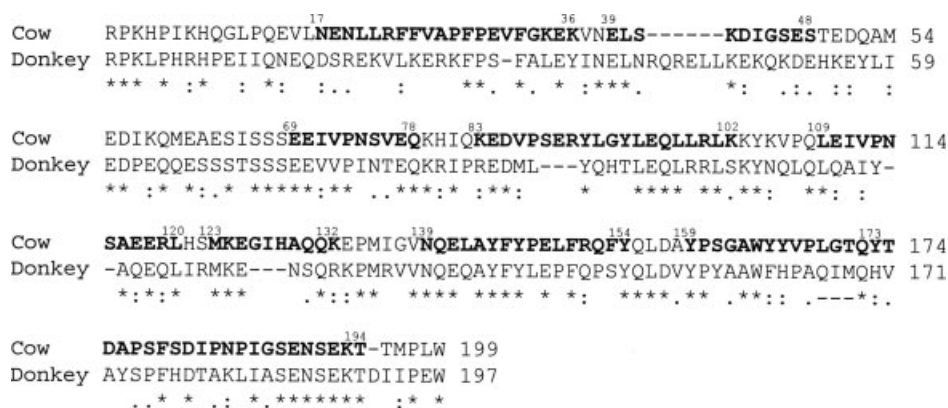


Figure 8. Alignment of the primary structure of the cow's (199 AA; GenBank Acc. No. CAB57792.1) and donkey's (197 AA) α_{s1} -CN variants B. The regions corresponding to the IgE-binding epitopes of cow's α_{s1} -CN are shown in bold. Alignment displays the consensus symbols, which denote the degree of conservation observed in each column. The symbol '*' means that the amino acids in that column are identical in both sequences. The symbol ':' means that conserved substitutions have been observed. The symbol '.' means that semi-conserved substitutions are observed.

major components were identified as donkey's α_{s1} -CNs and their sequences completely characterized, using the known mare's α_{s1} -CN as reference. The proteins with M_r of 23 786 and 23 658 Da differ in the presence of a glutamine residue at position 83 in the full-length component and present the amino acid substitutions Gln⁸ → His and His¹¹⁵ → Tyr with respect to the mare's α_{s1} -CN (Fig. 6). The other two components with M_r 24 406 and 24 278 Da, which also differ in the presence of a glutamine residue at position 88 in the full-length component, show the insertion of the pentapeptide HTPRE between Leu³³ and the Glu³⁴, which accounts for the increase of 620 Da in their observed molecular masses.

Taking into account the exon modular structure of the α_{s1} -CN in other species and in mare,^[8,14] it is reasonable to assume that the pentapeptide HTPRE between Leu³³ and the Glu³⁴ is encoded by the exon 5, which is instead constitutively spliced in mare's α_{s1} -CNs, where it should express the sequence SIPREVRK.^[14] On the other hand, the existence of the two isoforms lacking a glutamine residue is probably related to a cryptic splice site occurring at the first codon (CAG) of exon 11 (encoding this glutamine residue) as already reported for ewe,^[25] goat,^[26] cow and water buffalo.^[27] α_{s1} -CNs.

The two α_{s1} -CNs bearing the pentapeptide insertion were named variants A (202 amino acids; M_r 24 406) and A₁ (201 amino acids; M_r 24 278), whereas the two α_{s1} -CNs without the pentapeptide were named variants B (197 amino acids; M_r 23 786) and B₁ (196 amino acids; M_r 23 658).

The absence of the pentapeptide HTPRE confers to α_{s1} -CNs variants, B and B₁, a more acidic nature (theoretical pIs of 6.01) with respect to the variants A and A₁ (theoretical pIs of 6.12), thus explaining the co-migration of the first ones in the spot A and of the last ones in the spot B [Fig. 1, inset (a)]. The other two minor components originating the MALDI peaks at m/z 25 142 and 25 272 might correspond to two additional α_{s1} -CN isoforms (Fig. 1, inset (b)), according to the results on the characterization of α_{s1} -CN in mare's milk reported in a paper published during the revision of this work.^[35] However, these minor components were undetectable in the 2-DE gel analysis probably because to their very low abundance. The existence of a full-length mare's α_{s1} -CN and seven splicing variants^[35] also clearly indicates that the polymorphism we observed in the individual donkey's milk sample studied in the present work is not unique to this sample, but instead it is typical of equine milk and reflects the complex exon/intron

structure of α_{s1} -CN, which consists of many short exons and allows alternative splicing of shorter exons from pre-mRNA.^[14]

The primary structure of donkey's α_{s1} -CN shows remarkable differences with respect to that of bovine α_{s1} -CN, which represents one of the major cow's milk allergens. These differences could be related to the already demonstrated low allergenic properties of donkey's milk.

Finally, the methodological approach reported here, which involves coupling of RP-HPLC, 2-DE and mass spectrometric analyses, demonstrates to be successful in the characterization of protein isoforms with very similar sequences. Indeed, whereas the RP-HPLC or electrophoretic analyses alone present limitations in fully resolving proteins with very similar sequence, the coupling of these two techniques before the use of mass spectrometry proves to be very helpful for the detection and characterization of protein isoforms present in a complex mixture.

The protein sequence data reported in this paper will appear in the UniProt Knowledgebase under the accession number P86272 (direct submission 02-APR-2009).

Acknowledgements

This work was supported by MIUR PRIN 2006, project number 2006075417 and FIRB 'Italian Human ProteomeNet' RBRN07BMCT.

References

- [1] G. Iacono, A. Carroccio, F. Cavataio, G. Montalto, M. Soresi, V. Balsamo. Use of ass's milk in multiple food allergy. *Journal of Pediatric Gastroenterology and Nutrition* **1992**, 14, 177.
- [2] A. Carroccio, F. Cavataio, G. Montalto, D. D'Amico, L. Alabrese, G. Iacono. Intolerance to hydrolysed cow's milk proteins in infants: clinical characteristics and dietary treatment. *Clinical and Experimental Allergy* **2000**, 30, 1597.
- [3] G. Monti, E. Bertino, M. C. Muratore, A. Coscia, F. Cresi, L. Silvestro, C. Fabris, D. Fortunato, M. G. Giffurda, A. Conti. Efficacy of donkey's milk in treating highly problematic cow's milk allergic children: an *in vivo* and *in vitro* study. *Pediatric Allergy and Immunology* **2007**, 18, 258.
- [4] L. Businco, P. G. Giampietro, P. Lucenti, F. Lucaroni, C. Pini, G. Di Felice, P. Iacovacci, C. Curadi, M. Orlandi. Allergenicity of mare's milk in children with cow's milk allergy. *Journal of Allergy and Clinical Immunology* **2000**, 105, 1031.
- [5] S. C. Zicker, B. Lonnerdal. Protein and nitrogen composition of equine (*Equus caballus*) milk during early lactation. *Comparative Biochemistry and Physiology* **1994**, 108, 411.

- [6] H. Y. Guo, K. Pang, X. Y. Zhang, L. Zhao, S. W. Chen, M. L. Dong, F. Z. Ren. Composition, physicochemical properties, nitrogen fraction distribution, and amino acid profile of donkey milk. *Journal of Dairy Science* **2007**, *90*, 1635.
- [7] M. Herrouin, D. Molle, J. Fauquant, F. Ballestra, J. L. Maubois, J. Leonil. New genetic variants identified in donkey's milk whey proteins. *Journal of Protein Chemistry* **2000**, *19*, 105 and reference therein.
- [8] G. Miranda, M. F. Mahé, C. Leroux, P. Martin. Proteomic tools to characterize the protein fraction of Equidae milk. *Proteomics* **2004**, *4*, 2496 and reference therein.
- [9] V. Cunsolo, R. Saletti, V. Muccilli, S. Foti. Characterization of the protein profile of donkey's milk whey fraction. *Journal of Mass Spectrometry* **2007**, *42*, 1162 and reference therein.
- [10] S. Iametti, G. Tedeschi, E. Oungre, F. Bonomi. Primary structure of κ -casein isolated from mares' milk. *Journal of Dairy Research* **2001**, *68*, 53.
- [11] J. M. Girardet, L. Miclo, S. Florent, D. Mollé, J. L. Gaillard. Determination of the phosphorylation level and deamidation susceptibility of equine β -casein. *Proteomics* **2006**, *6*, 3707.
- [12] B. Ochirkhuyag, J. M. Chobert, M. Dalgalarondo, T. Haertlé. Characterization of mare caseins. Identification of α_{s1} - and α_{s2} -caseins. *Lait* **2000**, *80*, 223.
- [13] D. Milenkovic, P. Martin, G. Guerin, C. Leroux. A specific pattern of splicing for the horse α_{s1} -Casein mRNA and partial genomic characterization of the relevant locus. *Genetics Selection Evolution* **2002**, *34*, 509.
- [14] T. Lenasi, I. Rogelj, P. Dovc. Characterization of equine cDNA sequences for α_{s1} -, β - and κ -casein. *Journal of Dairy Research* **2003**, *70*, 29.
- [15] S. Hobor, T. Kunej, P. Dovc. Direct Submission in EMBL/GenBank/DDBJ Databases. February **2008**.
- [16] L. Ramunno. Direct Submission in EMBL/GenBank/DDBJ Databases. December **2008**.
- [17] L. Ramunno. Direct Submission in EMBL/GenBank/DDBJ Databases. April **2009**.
- [18] V. Cunsolo, E. Cairone, R. Saletti, V. Muccilli, S. Foti. Sequence and phosphorylation level determination of two donkey beta-caseins by mass spectrometry. *Rapid Communication in Mass Spectrometry* **2009**, *23*, 1907.
- [19] G. Mamone, S. Caira, G. Garro, A. Nicolai, P. Ferranti, G. Picariello, A. Malorni, L. Chianese, F. Addeo. Casein phosphoproteome: identification of phosphoproteins by combined mass spectrometry and two-dimensional gel electrophoresis. *Electrophoresis* **2003**, *24*, 2824.
- [20] A. Shevchenko, M. Wilm, O. Vorm, M. Mann. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Analytical Chemistry* **1996**, *68*, 850.
- [21] M. Kussmann, E. Nordhoff, H. Rahbek-Nielsen, S. Haebel, M. Rossel-Larsen, L. Jakobsen, J. Gobom, E. Mirgorodskaya, A. Kroll-Kristensen, L. Palm, P. Roepstorff. Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes. *Journal of Mass Spectrometry* **1997**, *32*, 593.
- [22] J. Gobom, E. Nordhoff, E. Mirgorodskaya, R. Ekman, P. Roepstorff. Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *Journal of Mass Spectrometry* **1999**, *34*, 105.
- [23] V. Keil-Dlouhá, N. Zylber, J. M. Imhoff, N. T. Tong, B. Keil. Proteolytic activity of pseudotrypsin. *FEBS Letters* **1971**, *16*, 291.
- [24] B. Keil. Enzymatic cleavage of proteins. In *Methods in Protein Sequence Analysis*, M. Elzinga (Ed). Humana Press: Clifton, New Jersey, **1981**, 291.
- [25] P. Ferranti, A. Malorni, G. Nitti, P. Laezza, R. Pizzano, L. Chianese, F. Addeo. Primary structure of ovine α_{s1} -caseins: localization of phosphorylation sites and characterization of genetic variants A, C and D. *Journal of Dairy Research* **1995**, *62*, 281.
- [26] P. Ferranti, F. Addeo, A. Malorni, L. Chianese, C. Leroux, P. Martin. Differential splicing of pre-messenger RNA produces multiple forms of mature caprine α_{s1} -casein. *European Journal of Biochemistry* **1997**, *249*, 1.
- [27] P. Ferranti, S. Lilla, L. Chianese, F. Addeo. Alternative nonallelic deletion is constitutive of ruminant α_{s1} -casein. *Journal Protein Chemistry* **1999**, *18*, 595.
- [28] L. Monaci, V. Tregoat, A. J. van Hengel, E. Anklam. Milk allergens, their characteristics and their detection in food: a review. *European Food Research Technology* **2006**, *223*, 149.
- [29] T. F. Kumosinski, E. M. Brown, H. M. Farrell Jr. Three-dimensional molecular modeling of bovine casein: α_{s1} -casein. *Journal of Dairy Science* **1991**, *74*, 2889.
- [30] P. Chatchatee, K. M. Jarvinen, L. Bardina, K. Beyer, H. A. Sampson. Identification of IgE- and IgG-binding epitopes on α_{s1} -casein: differences in patients with persistent and transient cow's milk allergy. *Journal of Allergy and Clinical Immunology* **2001**, *107*, 379.
- [31] F. Pessler, M. Nejat. Anaphylactic reaction to goat's milk in a cow's milk-allergic infant. *Pediatric Allergy and Immunology* **2004**, *15*, 183.
- [32] N. Clein. Cow's milk allergy in infants and children. *International Archives of Allergy and Immunology* **1958**, *13*, 245.
- [33] B. Belloni-Businco, R. Paganelli, P. Lucenti, P. G. Giampietro, H. Perborn, L. Businco. Allergenicity of goat's milk in children with cow's milk allergy. *Journal of Allergy and Clinical Immunology* **1999**, *103*, 1191.
- [34] D. Vita, G. Passalacqua, G. Di Pasquale, L. Caminiti, G. Crisafulli, I. Rulli, G. B. Pajno. Ass's milk in children with atopic dermatitis and cow's milk allergy: crossover comparison with goat's milk. *Pediatric Allergy and Immunology* **2007**, *18*, 594.
- [35] A. Matéos, L. Miclo, D. Mollé, A. Dary, J. M. Girardet, J. L. Gaillard. Equine α_{s1} -casein: characterization of alternative splicing isoforms and determination of phosphorylation levels. *Journal of Dairy Science* **2009**, *92*, 3604.