### Abstract
Caseins, which are contained in milk, play a significant role in the development of clinical symptoms of allergic reactions in adults and children. To date, the properties of caseins have been studied, their primary structure has been identified. However, despite available scientific reviews and original articles, an issue of the structural organization of milk caseins micelles and their biological functions is still very controversial. In this regard, the proposed review is relevant, since it most fully reflects current information about various types of caseins, their physicochemical and immunobiological properties, and analyzes in detail the existing theories on the structural organization of casein micelles. Of particular interest in our review are the data on the comparative analysis of the structure and properties of caseins in both human breast milk and cow’s milk, as well as the section dealing with the allergenic activity of cow’s milk caseins and their cross-reactivity with milk proteins of other mammalian species. The objective of this review is to summarize current data on the structure and properties of casein proteins, and to determine their role in the formation of sensitization to dairy products.

### Keywords
- cow’s milk, human breast milk, casein micelle, β-casein, κ-casein.

### Introduction

The nutrition of an infant in the first year of life has a huge impact on the maturation of the immune system, along with physical and cognitive development. Human breast milk (HBM) is recognized as the ‘gold standard’ of nutrition for infants [1, 2]. If breastfeeding is impossible, its artificial substitutes are used [2]. Most milk formulae are based on cow’s milk. In the future, cow’s milk and products made from it constitute a substantial share of the human diet [3]. Cow’s milk allergy is the most common form of food allergy at an early age [4].

It is reported that currently 0.6–3% of children under 6 years of age, 0.3% of older children and youths, and less than 0.5% of adults are allergic to cow’s milk. Remarkably, most children with milk allergies ‘outgrow’ their allergies by becoming able to consume milk and dairy products. However, in 15% of children, milk allergy persists into their adulthood [5].

HBM contains approximately 9 g of protein per liter, while cow’s milk contains 32 g of protein per liter [2, 6]. The casein fraction predominates in cow’s milk (76–86%), while whey protein fraction accounts for 14–24% of the total protein content in cow’s milk. A different ratio of casein and whey proteins was found in HBM: it is 40% to 60%, respectively [6, 7]. Besides, there are differences in the ratios and content of protein subclasses in the composition of cow’s milk and HBM [2]. The αS1 fraction of the α-casein family constitutes the largest proportion in cow’s milk. Unlike cow’s milk, β-caseins predominate in HBM [6, 7]. Casein is believed to play an important role in the formation of persistent allergies; hence, it is considered the main allergen in adults [7].

The objective of this review is to summarize current data on the structure and properties of casein proteins, and to identify their role in forming the sensitization to dairy products.

### Caseins

Casein (from Latin caseus meaning ‘cheese’) represents the main protein fraction of milk. Caseins constitute approximately 80% of all milk proteins. The information base of allergens, created in recent years, includes cow’s milk allergens. In the official list of the World Health Organization and the Subcommittee on Allergen Nomenclature of the International Union of Immunological Societies, cow’s milk allergens are designated by the Latin name of the allergen source — *Bos domesticus* (i.e., domestic cow) [8, 9]. According to the official list of allergens, caseins are classified under the general term Bos d 8 [9]. However, despite this common name, the individual components of caseins have received different identifying names. Caseins are encoded by various genes located on the same chromosome [5]. The most important are: αS1-casein (Bos d 9), αS2-casein (Bos d 10), β-casein (Bos d 11), and κ-casein (Bos d 12), accounting for 40%, 12.5%, 35% and 12.5% of casein fraction in milk, respectively [8]. All caseins have genetic polymorphisms resulting in several protein variants and contributing to their high heterogeneity [7]. These variants are characterized by a point mutation of amino acids, deletion of peptide fragments of different sizes, or post-translational modifications, such as glycosylation, phosphorylation, or partial hydrolysis, which may affect their properties and allergenic potential [10, 11, 12].
Caseins are phosphoproteins [7]: they contain 1-11 phosphoric acid anions (organic phosphorus), which form an ester bond, mainly with the hydroxyl group of serine. The content of serine phosphate residues in the polypeptide chains of a protein determines its sensitivity to calcium cations [6, 12, 13, 14]. By their function, αS1, αS2, and β-caseins are calcium-binding proteins, whereas κ-casein is a stabilizing protein [5].

All fractions of casein have been studied in terms of their physicochemical properties; their primary structure has been deciphered as well [11, 15]. It was shown that all caseins have a molecular weight of 19-25.2 kDa, an isoelectric point (pI) of 4.7. The primary structure of caseins does not have high homology [6, 15, 16]. The polypeptide chain contains a large amount of proline (17 Pro in αS1; 10 Pro in αS2; 35 Pro in β-casein, and 20 Pro in κ-casein) [12]. Secondary structures, such as α-helices and β-sheets, are rare, which makes caseins flexible, unfolded, or random helical peptides, capable of generating intermolecular interactions (such as electrostatic, hydrogen and hydrophobic). The tertiary structure of casein is a loose, indistinct globule [6, 12, 18]. The quaternary structure of casein is called a micelle [7, 12, 17, 18]. It is comprised of a central hydrophobic core (calcium-sensitive αS1-, αS2- and β-caseins) and a peripheral hydrophilic layer (κ-casein) [11, 17].

In milk, caseins have pronounced acidic properties. Free carboxyl groups of amino acids and hydroxyl groups of phosphoric acid easily interact with calcium ions, as well as with other salts of alkali metals and alkaline earth metals (Na+, K+, Mg2+), forming caseinates [14].

Table 1. Secondary structural elements of αS1- and αS2-caseins (Adapted from: Treweek T., 2012) [12]

<table>
<thead>
<tr>
<th>Protein</th>
<th>α-helix (%)</th>
<th>β-sheet (%)</th>
<th>β-turns (%)</th>
<th>Disordered structure (%)</th>
<th>Not detected (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>αS1-casein</td>
<td>14</td>
<td>28</td>
<td>22</td>
<td>35</td>
<td>1</td>
</tr>
<tr>
<td>αS2-casein</td>
<td>18</td>
<td>26</td>
<td>22</td>
<td>33</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 1. Schematic diagram of linear chain distribution of charged hydrophilic and hydrophobic regions for the most common genetic types of caseins at the pH of milk (6.6). (Sensu Tomasz Konrad et al, 2017) [19]. αs1, αS1-casein; αs2, αS2-casein; β, β-casein; κ, κ-casein.
αS2-Caseins (Bos d 10) have a molecular weight of 25.2-25.4 kDa and contain 207 amino acid residues. This is the most hydrophilic protein that has 10-13 serine phosphate residues [15, 19]. Its hydrophobic regions are located in the range of 90-120th and 160207th amino acid residues [19] (Figure 1). Its secondary structure has few α-helices and β-sheets (Table 1) [12].

β-Casein (Bos d 11) has a molecular weight of 24 kDa and contains 209 amino acid residues. It is a disordered 'random coil' with β-sheet and β-turn structures [12, 19]. β-Casein is the most hydrophobic of all caseins [19, 20]. It occurs in five genetic variants: A (A1, A2, and A3), B, C, D, and E [11, 12, 21]. Bovine β-casein is usually present as one form with five phosphates [14, 20] on serine residues: Ser15, Ser17, Ser18, Ser19, and Ser35. The first four of those form a phosphorylation center [20]. Just two genetic variants of this protein (C and D) seem to have altered phosphorylation profiles [15, 21]. β-Casein has a negatively charged hydrophilic N-terminal region (1-40th amino acid residues). The hydrophobic properties of this protein increase from the N to the C end from the 136th to 209th amino acid residues [15, 19, 20, 22]. The discussed properties of β-casein, in contrast to other unstructured and disordered proteins, under physiological conditions, and even in an acidic medium, determine its ability to self-assemble into micelles. In this case, the hydrophobic part of the β-casein molecule is located inside the micelle, and the hydrophilic part, in which the phosphorylation center is located, is located outside [19, 20, 23] (Figure 1). β-Casein, like αS1-casein, has chaperone activity [12].

Previously, γ-casein was considered as a separate fraction, which accounted for 3% of the total amount of casein. Later it was shown that γ-casein was identical to the C-terminal part of β-casein [22, 24]. Many researchers suggested that γ-casein was a degradation product of β-casein and could be formed as a result of
a trypsin-like proteolysis of the latter [5, 22]. It was demonstrated that β-casein was hydrolyzed by milk proteinases (plasmin) in three regions, resulting in forming six peptides: y1, y2 and y3-caseins, proteases and peptones (thermostable, acid-soluble phosphoproteins) [24].

κ-Casein (Bos d 12), unlike α- and β-caseins, is a glycoprotein and contains about 5% of carbohydrates. The κ-casein molecule has a molecular weight of 19 kDa and consists of 169 amino acid residues (including two cysteine residues) and a single phosphate group bound to serine [19]. The fragment of the 1–105th amino acid residues of this chain is hydrophobic, it is called para-κ-casein. The hydrophilic C-terminal region of the molecule (glycomacropeptide) from 106th to 169th residues has a negative charge at pH 6.6 and may contain tetrasaccharide chains [15, 25, 26]. Amino acid residues of threonine (Thr121, Thr131, Thr133, Thr135, Thr142, Thr165) and serine (Ser141) undergo glycosylation [27]. The oligosaccharide at the C-terminus has a negatively charged N-acetylleucaminic acid, which increases the negative charge of the C-terminus in this casein [25].

Several κ-casein isoforms can coexist, depending on the degree of glycosylation in milk. There are 11 variants of κ-casein due to differences in the number of attached oligosaccharides. The number of glycosylation sites may vary from 0 to 7; therefore, both non-glycosylated and glycosylated isoforms exist in milk [5]. Para-κ-casein occurs in two main variants, A and B (Figure 1 and 2) [11].

κ-Casein contains two cysteine residues and can form disulfide bonds (Cys11 and Cys88) [25, 28, 29]. It was shown that κ-casein, similar to other caseins, does not have a stable three-dimensional structure, which enables it to change its conformation at different pH values [30]. Compared with calcium-sensitive αs- and β-caseins, κ-casein is insensitive to relatively high calcium levels [26].

**Casein micelle structure**

Caseins form complex aggregates (micelles) [5, 12]. The actual internal structure of the casein micelle remains not fully understood, and various models have been built to describe it [5, 16, 17]. Since casein is a secret of epithelial cells [14], the micelle structure is a result of evolution and plays an important biological role. Calcium phosphate binds to casein molecules to prevent the formation of amyloid fibrils in breast tissue. Accordingly, mothers are able to feed their newborns without negative consequences [19].

**Sensu the classical theory, a micelle consists of submicelles [7, 31]. All electron microscopic research methods prove the submicellar structure of casein micelles [16, 27]. It is shown that the micelle body is comprised of discrete blocks and is not quite spherical. The formation of micelles proceeds stepwise in the Golgi apparatus [31, 32]. Depending on the type of casein, the resulting polypeptide chain undergoes glycosylation (i.e., carbohydrates are attached to the surface of the protein globule) or/and phosphorylation [26, 31, 32].

The diameter of the submicelles is 12-15 nm. Casein proteins in submicelles are oriented by their hydrophobic part to the center of the micelle, and by their hydrophilic part they face outside. The bonds between molecules in submicelles are hydrogen, hydrophobic, and electrostatic [6, 17]. Casein molecules containing a significant amount of phosphate esters possess a large negative charge. As the submicelles become saturated with Ca2+ and phosphate anions, the submicelles begin to coalesce, forming micelles [31, 32]. Submicelles are connected by colloidal calcium phosphate; the latter has a free bond and can form a calcium bridge between the carboxyl and serine phosphate groups of casein molecules [11]. When calcium is added to the hydroxyl group of the phosphoric acid, a casein-calcium phosphate complex is formed [14, 31, 32] (Figure 3, 4).

Hence, casein micelles are ordered protein structures, in the formation of which calcium cations and phosphoric acid anions take part. When calcium combines with phosphoric acid anions, calcium-phosphate bridges are formed that bind casein molecules together. The resulting casein–calcium phosphate complex may also contain some other polar compounds (citric and acetic acids, vitamins, etc.) and ions (Na+, K+, Mg2+, Cl–, etc.) [27, 32].

κ-Casein is found mainly on the outer side of micelles, which is confirmed by electron microscopy [31]. The hydrophilic, negatively charged C-terminal part of most κ-casein molecules sticks out of the micelle in the form of flexible filaments [12, 22, 27]. The hydrodynamic thickness of the filament layer is about 7 nm [16]. Micelles form a colloidal suspension in the presence of a stabilizing factor (κ-casein) [33].

The formation of casein micelles is explained by the nanocluster theory. According to it, a micelle is a homogeneous matrix of caseins [11]. The homogeneity of the mass distribution inside the casein micelle is indicated by studies using neutron and X-ray light scattering [12, 34]. Colloidal calcium phosphate nanoclusters are located inside the homogeneous matrix of caseins [12, 17, 31], dispersed into very small (about 2 nm) ‘cherry pits’, the distance between which is, on average, 18.6 nm. Phosphorylation centers (3-5 adjacent phosphorylated amino acid residues) of caseins are attached to the surface of nanoclusters. These centers cross-link proteins, which leads to the formation of a network structure of a micelle or matrix [34]. The micelle is stabilized by κ-casein [17, 31, 34]. It is known that the ratio of κ-casein to other components of casein is higher in small micelles and lower in large micelles [35]. The aggregation of micelles is prevented by electrostatic and steric repulsion due to the κ-casein layer [34].

Thus, despite the fact that the tertiary structure of a single casein molecule is a disordered coil, several protein molecules of casein during ionic and flexible intermolecular interactions form a rigid structure, which is a micelle, the equilibrium of which is maintained by the outer κ-casein layer. The κ-casein layer prevents further micelle growth, reaction with calcium cations and, accordingly, precipitation.

The functional features of caseins depend directly on their structural species specificity and quantitative ratios in milk. The main physiological function of casein micelles is to supply proteins, phosphates and calcium to newborns [16]. Besides, during the digestion of both bovine and human β-casein, β-casomorphins are formed, which can act as ligands to opioid receptors. Experiments on animals have shown that oral administration of β-casomorphins affects the motility of the digestive tract and generates an analgesic effect [5, 21, 23, 36].

**Human breast milk caseins**

Human κ-casein is similar in structure to bovine casein. Just like the latter, human κ-casein is located on the micelle surface and contributes to its stabilization [36].
κ-Casein of human breast milk is a polypeptide chain comprising 158 amino acid residues. The hydrophobic region (1-93rd amino acid residues) contains one cysteine (Cys4), which determines the formation of a dimer, a para-κ-casein. The presence of two cysteine residues in bovine κ-casein indicates its ability to form polymers [35, 37, 38]. In human κ-casein, the carbohydrate component constitutes approximately 55% of the molecule. The bovine counterpart contains only 5% of carbohydrates [35, 36, 39]. The hydrophilic fragment of human κ-casein (from amino acid residues 94-158th) is a glycomacropeptide with antipathogenic and bifidogenic activity [36, 40, 41]. Additionally, it was shown that κ-caseins are capable of inhibiting the adhesion of *H. pylori* to the cells of the gastric mucosa [36, 41].

Functional and structural features of different casein types (α-, β-, κ-), along with their proportions, determine main differences between the micelle of human breast milk and cow’s milk [25]. As in cow’s milk, caseins of human breast milk form a stable calcium caseinate complex with colloidal calcium phosphate. The aggregate properties of β-casein are responsible for formation of micelles [35]. The average size of micelles in cow’s milk is 150–180 nm, whereas in human breast milk, it is 60-80 nm [24, 30, 41, 42]. The conducted studies demonstrated that β-caseins in human milk, which have fewer phosphorylation sites, in most cases, form smaller polymers. With an increase in phosphorylation groups, β-caseins aggregate [43]. It was shown that large micelles contain the least amount of carbohydrates in κ-casein (~25%). Consequently, the level of κ-casein glycosylation is associated with the micelle size [25, 35].

Human breast milk also contains numerous specific minor proteins, associated with casein micelles. The method of liquid chromatography and mass spectrometry identified 82 proteins in a casein micelle, 18 of which are not present in their whey fraction. Thirty-two proteins, specifically associated with casein micelles, have not previously been identified in human breast milk or colostrum [44].

The obtained data provide a new insight into the proteomic profile of breast milk casein micelle and its physiological significance.

**Cow’s milk protein allergy**

Due to the fact that caseins have a flexible noncompact structure, they are classified as weak allergens [5]. In addition, they are effectively broken down in the digestive tract. Some peptide fragments of casein, present in milk and formed as a result of proteolysis, are conserved domain regions of the casein molecule. Such fragments determine the allergenicity of native proteins [6]. An absence of a clear tertiary structure in caseins suggests the presence of predominantly linear epitopes [11]. As of now, 35 IgE-binding epitopes have been identified in caseins: 6 in αS1-casein, 10 in αS2-casein, 9 in β-casein, and 8 in κ-casein [5].

Of eight major IgE-binding epitopes found in κ-casein, three are detected in 93% of serum samples from patients with cow’s milk allergy — specifically, IgE-binding regions between 9-26th, 21-44th, and 47-68th amino acid residues. Hence, the region between residues at positions 9-68th (at the N-terminal) may play an important role in the allergenicity of this protein [5].

Casein proteins, present in milk of various species of ruminant animals, have high homology (>80-90%) and similar structural, functional and biological properties [7, 11]. For example, the sequences of αS1-, αS2-, and β-caseins of cows, goats and sheep are homologous at 87-98% [5].

The conserved domain regions of caseins, responsible for IgE cross-reactivity, include regions with serine residues, at which phosphorylation occurs [10]. IgE-binding epitopes of α-caseins are not available, as they are localized in the hydrophobic region of the molecule. Therefore, the antigenic properties of α-caseins are manifested solely in cleaved or denatured molecules. Most children with cow’s milk allergy (93-98%) exhibit sensitization to the casein from sheep and goat [5, 6, 7].

β-Caseins of humans and cattle have approximately 50% homology [5, 11, 45]. Bernard H. et al. demonstrated IgE cross-reactivity between β-casein in human breast milk and cow’s milk [7, 45]. In the Han N study, two potential cross-reactive IgE-binding epitopes of human and bovine κ-caseins have been demonstrated [46]. These facts suggest the possibility of forming cross food allergies even to such ‘ideal’ food source as human breast milk, especially when using a mixed type of feeding or after starting the introduction of complementary foods.

According to our studies, sensitization to cow’s milk proteins was found in 48.2% of patients in Moscow and Moscow Oblast suffering from IgE-mediated food allergies [3]. We were able to denote some of the features inherent in casein sensitization. It was established that, in contrast to the majority of whey fractions of milk (β-lactoglobulin and bovine serum albumin), allergic sensitivity to which significantly and reliably declined with age, the occurrence of IgE to casein did not significantly change in the older age groups of patients. In children with sensitization to milk proteins, allergic sensitivity to caseins occurred much less frequently than to whey proteins, and amounted to 43% and 91%, respectively, and only 9.0% of the group, studied by us, had specific IgE exclusively to caseins. Analysis of specific IgE to various milk proteins implied that the incidence of sensitization to caseins was comparable to β-lactoglobulin, but still was higher than to bovine serum albumin, both in frequency and in sensitization level [3].

**Conclusion**

There is currently no treatment for milk allergy. Once the diagnosis is made, the prevention of an allergic reaction depends mainly on elimination measures. However, avoiding milk altogether can cause a decrease in nutritional status and affect the growth of infants and children. This problem could be dealt with via optimizing novel strategies for milk processing with the goal of destroying or modifying the structure of allergenic proteins and, consequently, reducing or eliminating their allergenicity, as well as via developing the methods for allergen-specific immunotherapy with these allergens. It is encouraging that some success has been achieved in oral immunotherapy with cow’s milk proteins [47, 48].

**Acknowledgments**

We acknowledge the state source of funding for the research on the topic, “Development of Native and Molecular Forms of Allergens Intended for The Diagnosis and Treatment of Allergic Diseases in Pediatric Practice.” The study was carried out with the Collective Usage Center, I.I. Mechnikov Research Institute of Vaccines and Sera, Moscow, Russia.

**Conflict of interest**

The authors declare that they have no conflicts of interest.


Authors:

Stanislava Yu. Petrova – MD, PhD, Senior Researcher, Laboratory of Allergens, I.I. Mechnikov Research Institute of Vaccines and Sera, Moscow, Russia. https://orcid.org/0000-0003-3034-0148.

Svetlana V. Khlgatian – DSc, Lead Researcher, Laboratory of Allergens, I.I. Mechnikov Research Institute of Vaccines and Sera, Moscow, Russia. https://orcid.org/0000-0001-8354-7682.

Olga Yu. Emelyanova – PhD, Senior Researcher, Laboratory of Allergens, I.I. Mechnikov Research Institute of Vaccines and Sera, Moscow, Russia. https://orcid.org/0000-0002-1228-523X.

Larisa A. Pishchulina – Junior Researcher, Laboratory of Allergens, I.I. Mechnikov Research Institute of Vaccines and Sera, Moscow, Russia. https://orcid.org/0000-0003-0450-9368.

Valentina M. Berzhets – DSc, Professor, Head of the Laboratory of Allergens, I.I. Mechnikov Research Institute of Vaccines and Sera, Moscow, Russia. https://orcid.org/0000-0001-5055-7593.