

Structure and function of milk allergens

Proteins (CMP) involved in milk allergy are numerous and heterogeneous, with very few structural or functional common features. This heterogeneity is complicated by their genetic polymorphism, resulting in several variants for each protein. These variants are characterized by point substitutions of amino acids or by deletions of peptide fragments of varying size or by post-translational modifications such as phosphorylation or glycosylation. All of these modifications may affect allergenicity. No common molecular structure can be associated with allergenicity, although some homologous regions such as casein phospho-peptides can explain an IgE cross-reactivity. Three-dimensional structure is an important feature in CMP allergenicity but denatured and linear epitopes are also involved. Epitopes are numerous and widely spread along the CMP molecule. They may be located in hydrophobic parts of the molecule where they are inaccessible for IgE antibodies in the native conformation of the protein but become bioavailable after digestive processes. Peptides as short as ca. 12–14 amino acid residues may account for a significant part of the allergenicity of the whole molecule, which justifies the need to be careful before proposing any CMP hydrolysate for highly allergic children.

J. M. Wal

Key words: caseins; whey proteins; peptides; IgE-epitopes; phosphorylation sites.

Dr J. M. Wal.
Laboratoire Associé INRA-CEA d'Immuno-Allergie
Alimentaire
DRM-SPI Bât. 136
CEA de Saclay
91191 Gif sur Yvette
France

Cow's milk proteins and IgE repertoire

Cow's milk contains about 30–35 g of proteins (CMPs) per liter. The action of chymosin (rennin), or the acidification of the milk to pH 4.6 enables two fractions to be obtained: lactoserum (whey), about 20% of the CMPs, and coagulum (curd), about 80% of the CMPs. Whey contains essentially globular proteins. The major ones, β -lactoglobulin (BLG) and α -lactalbumin (ALA), are synthesized in the mammary gland, while others, such as bovine serum albumin (BSA), come from the blood. In the coagulum, casein fraction (CAS) comprises four proteins coded by different genes carried on the same chromosome: α_{S1} -, α_{S2} -, β -, and κ -caseins.

Sensitivity to the various CMP has been proved to be widely distributed. Studies on large populations of allergic patients showed that most of the patients are sensitized to several proteins, including BLG, CAS, ALA, BSA, lactoferrin (LF), and immunoglobulins (Ig). A great variability is observed in IgE response. CAS and BLG but also ALA are major allergens. However, proteins present in very low quantities, such as BSA, Igs and especially LF, also appear to be of great importance, since 35–50% of patients are sensitized to those proteins and sometimes to those proteins only. It is noteworthy that in the last years, sensitivity to casein has increased in terms of both frequency and intensity of IgE response. Sensitization to CAS, BLG and ALA is closely linked. In contrast, sensitivity to BSA appears to be completely independent, 50% of the patients being allergic to BSA regardless of their sensitivity to other milk allergens (1).

The main characteristic is the multiplicity and diversity of proteins that are involved in cow's milk allergy (CMA). Polysensitization to several proteins is most often common and all milk proteins appear to be potential allergens. The question may thus arise whether those different CMPs share similar functions or whether their structures display common features that could be related to their allergenic potential.

Structure and function of CMPs

The primary structure and part of the secondary and tertiary structures of CMPs are known. However, caseins cannot be crystalized and their tertiary structure is approached by molecular modeling.

Whey proteins

The major allergens of lactoserum are BLG and ALA. BLG occurs naturally in the form of a 36 kDa dimer. It has no homologous counterpart in human milk. Each subunit corresponds to a 162-residue polypeptide. There are two main isoforms of BLG, i.e. genetic variants A and B, which differ by only two point mutations on residues 64 and 118, aspartic acid and valine in BLG A, and glycine and alanine in BLG B, respectively. The molecule possesses two disulfide bridges and one free cysteine (2). This structure is responsible for the main physicochemical properties and also for interaction with casein during heat treatments (3). The relative resistance of BLG to acid hydrolysis as well as to proteases allows some of the

protein to remain intact after digestion, to be absorbed through the intestinal mucosa, and then to be presented to the immunopotent cells.

The tertiary structure of BLG is known. It belongs to the lipocalin family and is considered a retinol-binding protein. Lipocalins bind and transfer hydrophobic ligands. This kind of molecule has a high allergenic potential and several allergens of animal origin belong to this family. They share well-conserved sequence homologies in their N-terminus moiety where tryptophan at position 19 is always present. Crystallography studies revealed very similar folding, called β barrel structure, with the same arrangements of 8 (or 10) antiparallel β sheets, in spite of the mutations that have occurred during evolution (4).

ALA is a monomeric globular protein of 123 amino acid residues with a 14.4 kDa molecular weight and four disulfide bridges. It is a regulatory component of the enzymatic system of galactosyl transferase responsible for the synthesis of lactose. It possesses a high affinity binding site for calcium, and this bond stabilizes its secondary structure. On the basis of their amino acid sequences and three-dimensional structures, ALA and hen's egg white lysozyme are closely homologous, having evolved by divergence from a common ancestral gene. The complete amino acid sequence of bovine ALA shows also extensive homology with human ALA since 74% of the residues are identical and another 6% chemically similar (2). However, despite this great similarity, bovine ALA has been identified as a major cow's milk allergen (1).

Caseins

Each casein, α S1-, α S2-, β -, and κ -casein, represents a well-defined chemical compound but they crosslink to form ordered aggregates, micelles, in suspension in lactoserum. The proportion of caseins in the micelles is relatively constant, ca. 37%, 37%, 13% and 13%, respectively. The distribution of caseins is not uniform within these micelles, which comprise a central hydrophobic part and a peripheral hydrophilic layer where the C-terminal polar fragment of the κ -casein molecules (e.g. casein macropeptide) is exposed. Polar domains of the other caseins (α S1-, α S2-, and β -) and notably major sites of phosphorylation containing phosphoserine residues are also present at the periphery of the micelles, in relation to the calcium binding and transfer properties of CAS (e.g. whole casein fraction) (2).

α S1-, α S2-, β -, and κ -casein have little primary structure homology. Their functional properties also differ since three of them, α S1-, α S2-, and β -casein, appear to be calcium-sensitive, while κ -casein is not. However, the four caseins display common features

that are quite unusual, differing greatly from other milk proteins. They are phosphorylated proteins, with a loose tertiary, highly hydrated structure. CAS is often considered poorly immunogenic because of this flexible, non-compact structure and because it is rapidly and extensively degraded by proteolytic enzyme during digestion. Caseins are not significantly affected by severe heat treatments but are very susceptible to all proteinases and exopeptidases (1, 2).

The four caseins constituting CAS are present in the milk of other ruminant species with high sequence homologies varying from 80% to more than 90%. This could lead to adverse reactions in patients allergic to cow's milk who consume ewe or goat milk as a substitute (5).

α S1- and β -casein are the major components of CAS. They share a dipolar-type structure with amphipathic properties. The molecules have a globular hydrophobic domain and a highly solvated and charged domain. A striking common characteristic is the homology in the acidic peptide sequence containing the cluster of phosphoserine residues: (13–21) in β -casein and (62–70) in α S1-casein. The phosphate groups seem to be located at the β turns, and the folding of the molecule brings them into proximity with each other to form a very hydrophilic domain. A similar sequence, (7–12) and (55–60), appears twice in α S2-casein variant A, which contains 11 phosphoserine residues. As for α S1-casein, a genetic variant of α S2-casein (variant D) results from a deletion in the amino acid sequence. This deletion removes a cluster of three phosphoserine residues so that the structural characteristics and, consequently, the allergenic potential is different in the two α S2-casein variants (6). β -casein possesses five main genetic variants with different levels of phosphorylation. There is a 50% homology in the amino acid sequences of human and bovine β -caseins, which corresponds to conserved regions widely spread all along the molecules. However, two domains appear more interesting. One contains the major sites of phosphorylation, i.e. fragments (5–14) and (14–23) in human and bovine β -casein, respectively. The other particular conserved region includes a predicted α -helical segment, NLHLP, at positions (123–127) and (132–136) in human and bovine β -casein, respectively (7).

Several forms of caseins have been described derived from partial hydrolysis.

In conclusion, CMPs are very heterogeneous, with very few structural or functional features in common. This heterogeneity is complicated by their genetic polymorphism, resulting in several variants for each protein. These variants are characterized by point substitutions of amino acids, by deletions of peptide fragments of varying size or by post-translational modifications such as phosphorylation, glycosylation. All of these modifications may affect allergenicity.

Molecular characterization of allergenic structures on milk proteins

To better understand the relationship between structure and immunoreactivity and to establish a rationale for developing hypoallergenic formulae, characterization of "allergenic epitopes", or at least identification of fragments with IgE binding capacity, within CMP molecules is necessary. However, few studies have been performed using sera from patients with CMA.

β -Lactoglobulin

It is generally accepted that the hydrolysis of milk proteins considerably reduces their allergenicity. However, Haddad et al. (8) have shown that specific IgE from 10 patients with CMA recognized enzymatic digestion products of BLG by pepsin or pepsin + trypsin (10 patients out of 10) and that the recognition of peptides was even better than that of the intact molecule in four patients out of 10. Indeed, cleavage of the intrachain disulfide bonds within the BLG molecule and, consequently, the loss of the conformation of the molecule had little if any effect on its immunoreactivity, suggesting that linear epitopes are implicated (9).

Using tryptic and synthetic peptides Selo et al. showed that sensitization involved many epitopes that are widely spread all along the BLG molecule. Some of them were identified as short linear sequences, while other immunoreactive structures corresponded to quite large fragments that might encompass conformational epitopes or parts of epitopes (9). The best recognized peptides, by more than 90% of the patients, were fragments (41–60), (102–124) and (149–162), each accounting for 10–15% of whole BLG immunoreactivity. The C-terminus (149–162) peptide forms a turn, it adopts a short structure of α -helix and appears to be quite mobile according to the crystal structure (4). Peptides (41–60) and (102–124) form loops stabilized by either hydrogen or disulfide bonds. They are present on the surface of the molecule (4) and hence are accessible to antibodies. Less recognized epitopes could be classified into two groups. The first one comprises peptides (1–8), (25–40), and (92–100), recognized by 52–65% of sera. The second group consists of peptides (9–14) and (84–91), recognized by 40% of sera, and fragments (78–83) and (125–135), which were very poorly recognized by human IgE. However, it is noteworthy that the hydrophilicity profile of BLG suggests a high antigenicity for sequence (124–134), which is on the surface of the protein and was indeed described as very immunoreactive (10). Peptide (84–91), which was recognized by approximately 40% of human sera and does not appear to be on the surface of the intact BLG molecule, was also described as a minor epitope by Ball et al. (11). The two other peptides, (9–14) and (92–100), are well protected against enzymatic attacks and are not exposed at the surface of the

molecule, which could provide an indirect explanation for their relative minor importance in terms of immunoreactivity (9).

Reducing the length of fragment (8–24) obtained by chemical cleavage using cyanogen bromide (CNBr) to the sequence (9–14) led to a drastic decrease of the specific IgE binding. This suggests a predominant role of sequence (15–24) in conformation and epitope presentation. Further hydrolysis of the long (25–107) CNBr fragment into smaller peptides, e.g. fragments (25–40), (41–60), (78–83), (84–91) and (92–100), had no effect on the total IgE binding capacity in 66% of sera. This suggests that at least five different linear epitopes were present on this sequence

α -Lactalbumin

The peptide comprising loop (60–80):S-S:(91–96) has been considered the most antigenic region using animal models (12). However, it was not identified as a major epitope for human IgE in a study on sera from 19 allergic patients using purified tryptic digest fragments of ALA (13). IgE binding to native ALA and to large peptides confirmed the importance of conformational epitope(s). However, in some patients, reduced peptides, such as (59–94), exhibited a similar or a higher IgE-binding capacity than the native corresponding fragment, suggesting the existence of sequential epitope(s) exposed through protein denaturation. Moreover, IgE-binding sequences are also located in strongly hydrophobic regions of the ALA molecule, i.e. (99–108), where antigenicity is very unlikely to be predicted, and/or within parts of the molecule, e.g. (17–58) and (109–123), having a very high sequence homology (81% and 87%, respectively) with human ALA (13).

Caseins

Most patients allergic to CAS are sensitized to each of the four caseins. The intensity of the IgE responses varied greatly between the sera and the caseins but seemed closely related to the proportion of the four caseins in milk (14). These results probably correspond to a co-sensitization to the different casein components after the disruption of the casein micelles during the digestive process. However, polysensitization also appeared to be due to cross-sensitization mechanisms since, in some sera, the sum of the IgE titers against the different caseins greatly exceeded the IgE titer to the whole casein, suggesting the presence of IgE antibodies cross-reacting with different caseins through common or closely related epitopes. As already noticed, the region containing the major site of phosphorylation is conserved in bovine α s1- α s2- and β -caseins, as well as in the caseins of other species. Hence, this sequence plays an important part in cross-reactivity, especially since it has already been

described as being immunoreactive and resistant to digestive degradation (15). Using overlapping synthetic decapeptides, Spuergin et al. (5) have described three sequences, peptides (19–30), (86–103) and (141–150), recognized by sera of 15 allergic patients, the most immunoreactive one being peptide (86–103). These three major epitopes are located in hydrophobic regions of the molecule, where they are not accessible to antibodies unless the casein is denatured or degraded, e.g. during digestion. Moreover, Bernard et al. (6) have recently shown that the IgE response to caseins is affected (i.e. much decreased) by modifying or eliminating the major phosphorylation sites using the naturally occurring common variants A1 and A from β - and α_{s2} -caseins, respectively, which were purified in the native form, and then dephosphorylated, a purified rare variant D of α_{s2} -casein which lacks one major phosphorylation site, and the native and dephosphorylated tryptic fragment (1–25) from β -casein.

Interestingly, β -casein, which induced a high IgE response, is also abundant in human milk. It has recently been shown that conserved regions shared by the two β -caseins and particularly the C-terminus moiety of the molecules constituting the major sites of phosphorylation are responsible for IgE cross-reactivity. A fraction of the polyclonal anti bovine β -casein IgE antibodies, which is generally small and variable

depending upon the allergic patient, thus reacts with human β -casein, possibly with a lower apparent affinity (7).

Conclusion

It appears that no single protein can account for a major part of milk proteins' allergenicity. No particular structure or functional feature shared by all milk proteins can be associated with allergenicity, although some homologous regions such as casein phosphopeptides can explain an IgE cross-reactivity. "Epitopes" are numerous and widely spread along the CMP molecule. They may be located in hydrophobic parts of the molecule, inaccessible for IgE antibodies in the native conformation of the protein but become bioavailable after digestive processes. Three-dimensional structure is an important feature in CMP allergenicity but besides the conformational epitopes, IgE binding studies also showed the presence of linear epitopes. Peptides as short as 12–14 amino acid residues have been characterized to account for a significant part of the allergenicity of the whole molecule in some patients. These observations justify the need to be careful using goat or ewe milk as a substitute for cow's milk, using peptides for hyposensitization or proposing any whey protein hydrolysate in the diet of highly allergic children.

References

1. WAL JM. Cow's milk allergens. *Allergy* 1998;**53**:1013–1022.
2. FOX PF. Developments in dairy chemistry. 1. Proteins. London: Applied Science Publishers, 1982.
3. REDDY IM, KELLA NKD, KINSELLA JE. Structural and conformational basis of the resistance of β -lactoglobulin to peptic and chymotryptic digestion. *J Agric Food Chem* 1988;**36**:737–741.
4. BROWNLOW S, MORAIS CABRAL JH, COOPER R, et al. Bovine β -lactoglobulin at 1.8 Å resolution – still an enigmatic lipocalin. *Structure* 1997;**5**:481–495.
5. SPURGIN P, WALTER M, SCHILTZ E, DEICHMANN K, FORSTER J, MUELLER H. Allergenicity of α -caseins from cow, sheep, and goat. *Allergy* 1997;**52**: 293–298.
6. BERNARD H, MEISEL H, CREMINON C, WAL JM. Phosphorylation is a post translational event which affects IgE binding capacity of caseins. *FEBS Lett* 2000;**467**:239–244.
7. BERNARD H, NEGRONI L, CHATEL JM, et al. Molecular basis of IgE cross-reactivity between human β -casein and bovine β -casein, a major allergen of milk. *Mol. Immunol.* 2000;**37**:161–167.
8. HADDAD ZH, KALRA V, VERMA S. IgE antibodies to peptic and peptic-tryptic digests of β lactoglobulin: significance in food hypersensitivity. *Ann Allergy* 1979;**42**:368–371.
9. SELO I, CLEMENT G, BERNARD H, et al. Allergy to bovine β -lactoglobulin: specificity of human IgE to tryptic peptides. *Clin Exp Allergy* 1999;**29**:1055–1063.
10. ADAMS SL, BARNETT D, WALSH BJ, PEARCE RJ, HILL DJ, HOWDEN ME. Human IgE-binding synthetic peptide of bovine β -lactoglobulin and α -lactalbumin. *In vitro* cross-reactivity of the allergens. *Immunol Cell Biol* 1991;**69**:191–197.
11. BALL G, SHELTON MJ, WALSH BJ, HILL DJ, HOSKING CS, HOWDEN MEH. A major continuous epitope of bovine β -lactoglobulin recognized by human IgE binding. *Clin Exp Allergy* 1994;**24**:758–764.
12. HOPP TP, WOODS KR. Immunochemical studies on β -lactalbumin. *Mol Immunol* 1982;**19**:1453–1463.
13. MAYNARD F, JOST R, WAL JM. Human IgE binding capacity of tryptic peptides from bovine α -lactalbumin. *Int Arch Allergy Immunol* 1997;**113**:478–488.
14. BERNARD H, CRÉMINON C, YVON M, WAL JM. Specificity of the human IgE response to the different purified caseins in allergy to cow's milk proteins. *Int Arch Allergy Immunol* 1998;**115**: 235–244.
15. MEISEL H, FRISTER H. Chemical characterization of bioactive peptides from *in vivo* digests of casein. *J Dairy Res* 1989;**56**:343–349.