

Molecular modeling and conformational IgG epitope mapping on bovine β -casein

Fahui Liu^{1,2} · Jinyan Gao³ · Xin Li^{1,3} · Hongbing Chen^{1,4}

Received: 6 January 2016 / Revised: 22 February 2016 / Accepted: 9 April 2016 / Published online: 20 April 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Characterizing conformational B-cell epitopes is a key step for understanding the immunological basis of allergy contributed by β -casein. There is no resolved conformational structure of β -casein in protein data bank, and most of the previous research on epitope identification of β -casein focused on linear epitopes. Hence, the aim of this study is to map conformational epitopes on β -casein using molecular modeling and phage display. Phage peptides bound by the rabbit polyclonal antibodies were selected. Ten mimotopes were obtained after three rounds of biopanning. Analysis on the structural and geometrical properties of the antibody–antigen complexes that were compiled from immune epitope database exhibited that the number of the amino acids comprised in epitope ranged from 11 to 25 with a median number of 17. Epitopes spanned an area range from 541.4 to 1131.9 Å² with a median value of 782.9 Å². Epitopes were enriched with asparagine, tryptophan, tyrosine, glycine and lysine, whereas alanine, cysteine, aspartate, asparagine and serine were prone to separate the epitopes into fragments. These characteristics of epitope were used to select epitope candidates based on mimotopes and conformational structure of β -casein

that was built by homology modeling. Consequently, four B-cell IgG epitopes were mapped on β -casein.

Keywords β -casein · Phage display · Epitope mapping · Food allergy · Homology modeling

Introduction

Epitope, originally defined in the 1960s [1], can be typically divided into two categories, linear and conformational. Linear epitope consists of 8–10 amino acids that are contiguous in the primary sequence of proteins or peptides, whereas conformational epitopes are often comprised of several segments scattered along the sequence, but spatially close as a result of protein folding. Although it is well acknowledged that conformational epitopes play a more important role, experimental epitope mapping has mainly focused on linear epitopes [2]. Several linear epitopes on β -casein (β -CN) have been reported [3–5]. Research on the conformational epitopes, however, is still largely insufficient. This results from the fact that the antigen–antibody interactions are conformation dependent, making the B-cell epitope prediction less tractable than T-cell epitope prediction [6]. The direct approach to identifying conformational epitopes on allergens is to co-crystallize antibody–antigen complex. However, this method is time-consuming and cannot be applied to β -CN due to the difficulty for β -CN to form co-crystal complexes.

Phage display random peptides library (Ph.D.-C7C Phage Display Peptide Library Kit) consists of a large number (1.2×10^9) of random peptides. The peptides that mimic linear and conformational epitopes can bound to antibody with high affinity and subsequently be selected and amplified by infecting fresh bacterial host. In this

✉ Hongbing Chen
chenhongbing@ncu.edu.cn

¹ State Key Laboratory of Food Science and Technology, Nanchang University, 235 Nanjing Dong Road, Nanchang 330047, Jiangxi, China

² Department of Agrotechnology and Food Sciences, Wageningen University, Wageningen, The Netherlands

³ College of Food Science, Nanchang University, Nanchang, Jiangxi, China

⁴ Sino-German Joint Research Institute, Nanchang University, Nanchang, Jiangxi, China

“bio-panning” process, phages that display peptide ligands that can bind antibody are captured on the surface. Non-binding and non-specifically bound phages are washed away. After 3–4 rounds of panning, the selected phages are sequenced. The selected and sequenced peptides are termed “mimotope” [7]. The consensus sequence of mimotopes cannot always be found in the sequence of antigen. In these cases, the mimotopes that partly represent the “true” epitopes may be mimics of conformational epitopes. Therefore, the task to identify epitopes on the base of mimotopes is not always straightforward. Since mimotopes may lack sequence homology or contain amino acids from disparate regions of the antigen structure, the work to forward the mimotope information into conformational epitopes more accurately remains to be a challenge in bioinformatics and computational biology. Nevertheless, other than crystallography, epitope mapping by phage display is still the most powerful technique in the identification of conformational epitopes and is more widely applicable than crystallography.

Many computational algorithms have been developed to map epitope on allergens [8–12]. However, the general consensus is that none of the predictive methods can give a high level of correct prediction [13]. Almost in all of the computational algorithms that are based on mimotope to identify epitopes, such as MIMOP [9], Persurf [8], Episearch [10], and MIMOX [14], the precise conformational structure of allergen plays an important role. Beta-CN constitutes up to 38 % of the casein in bovine milk [15]. It is the most hydrophobic CN because of its large hydrophobic C-terminal domain. It is also one of the calcium-sensitive caseins (α S1-CN, α S2-CN and β -CN). Since inducing crystal is difficult in the case of β -CN up to date, conformational structure of β -CN is not available in PDB database. However, a conformational model has been built up via energy minimization method [16]. The accuracy has been evaluated via intramolecular crosslinking in conjunction with mass spectrometry method [17]. As more and more conformational structures of proteins have been resolved in PDB, the work of predicting protein model could become more accurate. In addition, to predict epitope on allergens and evaluate the performance of the computational algorithms, characteristics of epitopes derived from the resolved co-crystal complexes of antigen–antibody were usually taken into account. With the increase of available resolved structure of antigen–antibody complex in PDB database, it is now possible to perform a large-scale analysis on the epitope characteristics, which could be used as criteria in the selection of epitope residues.

In our study, a novel algorithm was developed to localize conformational IgG epitopes on β -CN based on the tertiary structure and the properties of IgG epitopes. A large-scale analysis on the qualified antigen–antibody complexes from

IEDB [18, 19] web server was performed. The properties of epitopes, such as epitope size, amino acid preference, interval amino acids, number of amino acids and secondary structure preference, were extracted from the complexes dataset. In the step of identifying the epitopes on β -CN, these properties were taken into account to select epitope candidates. To obtain mimotopes, Japanese white male rabbits were immunized with β -CN to produce polyclonal antibody, followed by isolating specific IgG from the anti-serum by affinity chromatography. The purified antibody was incubated with random phage. Phages bound to antibodies with high affinity were amplified to ensure the pool enriched with binding phages, resulting in positive phages for mimotopes. Since there is no solved tertiary structure of β -CN, the models of β -CN were built up using three modeling approaches, followed by an optimization on the predicted models. Moreover, to test the accuracy of the modeling, the optimal model was chosen and compared with the experimental structure data previously reported by Livney [17]. Based on the mimotope and conformational model of β -CN, a series of filtering steps were applied to the mimotope for identifying the conformational epitopes from qualified epitope candidates.

Materials and methods

Purification of bovine β -CN

Bovine milk was freshly collected from Jiangxi Yangguang Dairy Co. Ltd., (Nanchang, China). After being filtered through sterilized gauze, the milk was defatted by centrifuging at $6000\times g$ at 4 °C for 20 min. After skimming the top layer, the pH of the defatted milk was adjusted to 4.6, followed by incubation at 4 °C for 30 min. Then, the milk was centrifuged at 4 °C for another 10 min. The precipitate was collected and freeze-dried. Dried casein powder was dissolved in Tris/HCl buffer containing 3.3 M urea (Sangon Biotech, China) and 0.1 % (v/v) β -mercaptoethanol (Bio Basic Inc., Canada) before being loaded to DEAE-Sepharose Fast Flow column (GE Healthcare, USA). Samples were loaded onto the pre-equilibrated column (1.6×25 cm) and eluted at a flow rate of 1.5 mL/min. UV detection of eluted peaks was carried out at 280 nm. The target protein was identified by SDS-PAGE and isoelectric focusing (IEF)-PAGE. Then, purified protein was concentrated using Amicon Ultra-15 Centrifugal Filter Unit with a molecular weight cutoff of 3 kDa.

Preparation of polyclonal antibody

The polyclonal antibodies were generated by immunizing Japanese white male rabbits according to a standard

procedure as previously described [21]. Briefly, 2 mg β -CN was dissolved in 1 mL 0.05 M PBS under stirring, followed by mixing with 1 mL complete and incomplete Freund's adjuvant (Sigma, USA) for the first immunization and the booster injections, respectively. 1 mL of the mixed emulsion was injected to rabbits. Three booster injections were carried out at 14-day interval. The reactivity of antiserum against the purified β -CN was measured by indirect ELISA after each injection. 1 week later after the last injection, the antiserum was collected and stored at -70°C until further use. The performance of animal experiments was authorized by Institute of Occupational Medicine of Jiangxi (permission number SYXK (Gan) 2008-003).

The antiserum was then loaded onto Sepharose 4B affinity chromatography column (GE Healthcare, USA). All the purification procedures were performed according to the manufacturer's instructions. The specificity and purity of the collected antibody were defined by indirect ELISA and SDS-PAGE [21], respectively.

Selection of mimotopes

Phage display bio-panning technique was used to identify rabbit IgG-binding epitopes on β -CN. The purified specific polyclonal antibody was coated on 96-well microplate. The Ph.D.-C7C Phage Display Peptide Library (New England Biolabs, Inc., USA) was applied to perform bio-panning against purified antibody according to the manufacturer's instructions. Briefly, purified rabbit IgG was diluted to 100 $\mu\text{g/mL}$ with 0.1 M NaHCO_3 pH 8.6. Microplate was coated with the diluted IgG and incubated at 4°C overnight. The wells were then blocked with 3 % bovine serum albumin at 37°C for 1 h, followed by three times washing with TBST (50 mM Tris-HCl, pH 7.5, 150 mM NaCl and 0.1 % Tween 20). Phage with concentration of 2×10^{11} in 100 μL TBST was added to the wells and incubated at 37°C for 1 h. After that, unbound phages were washed away. The bound phages were eluted with glycine/HCl buffer (0.2 M, pH 2.2), followed by amplification in *E. coli* ER2738. After three rounds of panning and amplification, the individual clones were identified and DNA sequenced. The amplified culture was centrifuged at 10,000 rpm for 1 min. Then the supernatant was transferred to a fresh tube containing 200 μL PEG/NaCl, followed by an incubation at room temperature for 10 min. The solution was centrifuged for 10 min at 10,000 rpm. The resulting pellet was resolved in 100 μL iodide buffer. Then, 250 μL ethanol was added and incubated on ice for 10 min. After centrifugation, the pellet was collected and washed using 70 % ethanol. Finally, the pellet was suspended in 30 μL TE buffer and stored at -20°C until DNA sequencing by Sangon Biotech (Shanghai, China).

Conformational structure modeling of β -CN

Homology modeling, fold recognition and ab initio method were used to predict conformational structure of β -CN. For homology modeling, the Blast tool of PDB (<http://www.rcsb.org/pdb/home/home.do>) was used to search for an appropriate template. Fold recognition and ab initio method were carried out based on Zhang's I-TASSER and QUARK technique (<http://zhanglab.ccmb.med.umich.edu/research/>). The steric conflicts and geometry of loop regions were corrected according to the reported method [22]. All of the models were applied to an optimizing step of energy minimization and molecular dynamics simulations. To evaluate the quality of the resulting models, Ramachandran plot (<http://dicsoft1.physics.iisc.ernet.in/rp/>), WHAT_CHECK [23], ERRAT and VERIFY_3D of the NIH server (<http://nihserver.mbi.ucla.edu/>) were used to analyze the percentage of amino acids in allowed region, check the stereochemical parameters of the residues in the model, analyze the statistics of non-bonded interactions and determine the compatibility of an atomic model, respectively. Then, the differences between the selected optimal model and the conformational data of β -CN reported by Livney et al. [17] were analyzed. Conformational models of β -CN were displayed with PyMOL v0.99 (<http://www.pymol.org/>).

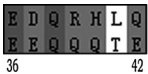
Epitope data construction and analysis

All the antigen-antibody complexes containing B-cell epitope information and solved 3-D structure of the antigen were retrieved from IEDB server (<http://www.immuneepitope.org/>). The complexes were then subjected to a filtering step according to the previously reported criteria [24]. The 3-D structures of qualified complexes were downloaded from PDB.

Epitopes of every complex were obtained from IEDB web server. Then, the number of amino acids in epitope and surface area that every epitope covers were calculated. Amino acid frequency in epitope was calculated to evaluate amino acid preference in epitope. Different amino acids show different frequencies in protein, and therefore the epitope may not prefer the amino acids even if these amino acids may show higher frequencies in epitopes. To address this problem, frequency of amino acid in epitopes was compared with that in proteins.

Epitope fragments that are separated by 1–3 amino acids can be incorporated to one segment. We termed the amino acids located between every two epitope fragments as the "interrupt amino acid," and the frequencies of them were calculated and compared with those of the corresponding amino acids in protein. In addition, the accessible surface area of epitope residues and secondary structure preference were calculated according to the methods of Rubinstein [24].

Table 1 Alignment of phage peptide (mimotopes) with the primary sequence of β -CN

Phage	Peptide sequence	Alignment
1	YYKRHLQ	 Minotope β-CN
2	EDQRHLQ	 Minotope β-CN
3	YYPRHLQ	 Minotope β-CN
4	ATKRHLQ	 Minotope β-CN
5	LCPRHLQ	 Minotope β-CN
6	AQQRHLQ	 Minotope β-CN

Epitope mapping on β -CN

The sequenced mimotopes were aligned to the primary sequence of bovine β -CN. The amino acids located around the geometrical center of the sequence that corresponds to mimotope were defined as central amino acid, which also should show high frequency in epitope and low frequency in “interrupt amino acid.” Taking the central amino acid as central point, scout for amino acids included in an area extends $1130 \text{ \AA}^2$ on the surface of conformational structure of protein. The selected amino acids were then subjected to a filtering step. Firstly, the amino acids with accessible surface area (ASA) in the 3-D structure of protein less than 5 % of their maximal ASA [25] were eliminated. ASA was calculated using Surface Racer Program [26] with a probe radius of $1.4 \text{ \AA}$. In addition, we defined Q as the relative amino acid frequency which was calculated as $Q = (e - n)/n$, where e is amino acid frequency in epitope, n denotes the amino acid frequency in protein. As the Q value increased from -100% to 0 , the number of the amino acids comprised in the epitopes decreased. Once the number of amino acids in all epitopes decreased to less than 17, the Q value was set as a threshold to additionally eliminate amino acids with Q below this value from epitope residue candidates.

Results

Preparation of antibodies and selection of mimotopes

Beta-CN with 90 % of purity was used for white rabbit immunization. After three times of booster immunizations,

the titer of serum that was detected by indirect ELISA increased to 1:240,000. Thirty phage clones specifically reactive to antiserum against the purified β -CN were selected for DNA sequencing, which resulted in 10 different mimotopes after three rounds of bio-panning. Six regions were recognized when these mimotopes aligned with the primary sequence of β -CN (Table 1).

Molecular modeling

Conformational models of β -CN predicted using different approaches were presented in Fig. 1. The structure of α -helix can be found in all of the models, whereas there were no β -sheets and some other secondary structures. The accuracy of the models was evaluated by the NIH structure analysis and verification server (<http://nihserver.mbi.ucla.edu/SAVES/>). After that, the homology model (Fig. 1a) was proved to be the optimal model in this study.

The optimal model was energy-minimized and applied to molecular dynamics simulations under different force fields. After that, intramolecular distances of some residue pairs were measured and compared with the length of crosslinks [17] that link the corresponding residue pairs (Table 2). Length (L) of crosslinks that connect some amino acid pairs has been analyzed previously using mass spectrometry [17]. The distances between corresponding amino acid pairs were also calculated in the optimal model that was built in this study as well as in the model of Kumosinski [16], denoted as D and D' , respectively. The less differences between D and L (d) or D' and L (d'), the more similarities between the models and the conformational data reported by Livney et al. As shown in Table 2, D was in good agreement with L in some amino acid pairs, especially for K97-K107, K48-K113 and R1-K113. Except for the residue pairs R1-K176 and K32-K113, the differences between the distances calculated by our model and lengths (L) of crosslink were less than that between the distances calculated by the model of Kumosinski [16].

Epitope mapping on β -CN

A dataset of 259 antibody–antigen co-crystal complexes was obtained from IEDB. Then, it was subjected to a filtering process, resulting in seventy-five qualified complexes. All of the conformational structures of the qualified complexes were downloaded manually from PDB. The number of amino acids, size of the epitope area, amino acid preference and interrupt amino acid preference of the complexes were analyzed and are shown in Fig. 2. These analyses revealed that the number of the amino acids (AAs) comprised in epitopes ranged from 11 to 25 with a median number of 17. Epitopes were enriched with asparagine

Fig. 1 Predicted models of β -CN using homology modeling (a), fold recognition (b) and ab initio method (c). The accuracy of these models was evaluated by the NIH web server. The optimal model was applied to further energy minimization and MD stimulations. Distances (shown in blue dash line) between some residue pairs of the final model (d) were calculated by PyMOL v0.99. All the models were illustrated with PyMOL v0.99

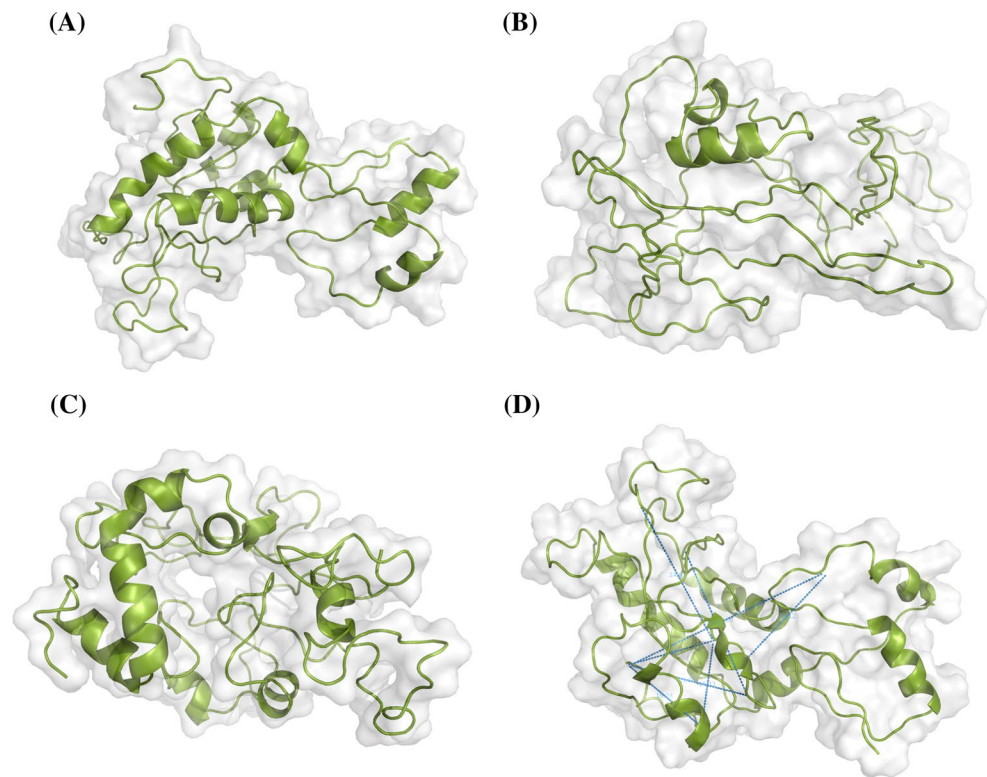


Table 2 Model tested by the length of crosslinks between some residue pairs of β -CN

Cross-linked residues	Distance (D) by our model (Å)	Distance (D') by reported model (Å)*	Length (L) of crosslink (Å) [#]	Difference (d) ($d = L - D$) (Å)	Difference (d') ($d' = L - D'$) (Å)
R1-K32	19.2	19.7	15.1	−4.1	−4.6
K105-K176	18.1	40.6	23.9	5.8	−16.7
K32-K113	15.1	21.4	21.4	6.3	0
R1-K113	14.9	4.9	15.1	0.2	10.2
R1-K107	23.0	13.4	18.8	−4.2	5.4
R1-K176	34.6	25.7	21.4	−13.2	−4.3
K97-K107	23.5	9.6	21.4	−2.1	11.8
K48-K113	24.6	22.2	26.4	1.8	4.2

* Reported by Kumosinski et al. [1]

[#] Reported by Livney et al. [2]

(N), tryptophan (W), tyrosine (Y), glycine (G) and lysine (K), whereas leucine (L), phenylalanine (F), cysteine (C), methionine (M) and alanine (A) were not significantly overrepresented in epitopes. Although aspartate (D), arginine (R) and threonine (T) showed high frequencies in epitope, they also highly enriched in protein. There were no significant differences between their frequencies in epitope and protein. Moreover, epitopes span an area ranging from 541.4 Å² to 1131.9 Å² with a median value of 782.9 Å². Alanine (A), cysteine (C), aspartate (D), asparagine (N) and serine (S) tend to separate the epitopes into several fragments. In addition, further analyses on the secondary

structure preference of epitope have been performed, and no general conclusion can be drawn from the results (data not shown).

According to the results of alignment and statistical analyses of dataset, four central amino acids, Q38, K107, N132 and R183, representing four epitope regions were located. Based on the central amino acids, four clusters of epitope candidates were obtained after searching for the epitope candidates on the surface of β -CN (Table 3). Then all these epitope candidates were subjected to a filtering process. The amino acids with relative ASA < 5 % were eliminated. When Q increased from −100 % to 0, the number of epitope

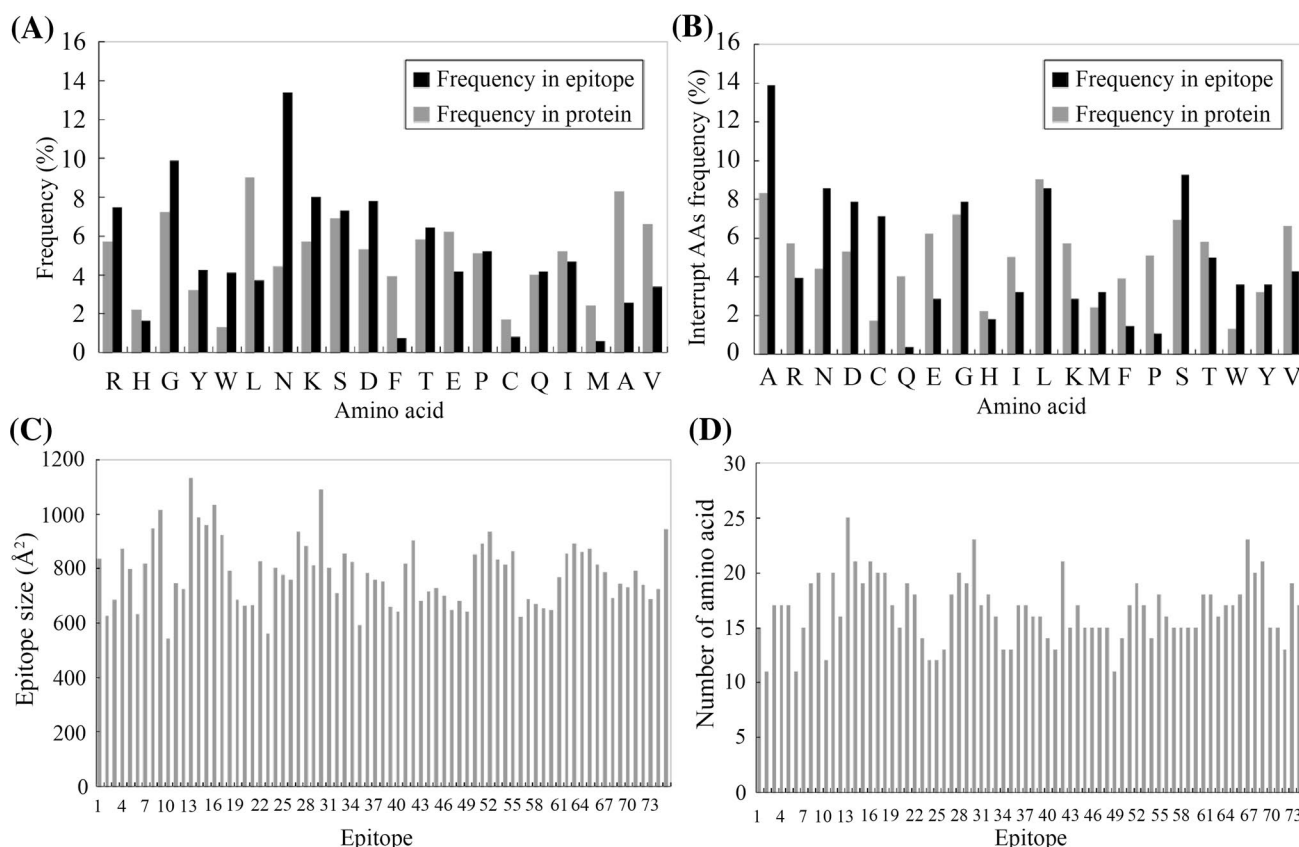


Fig. 2 Statistical analysis of B-cell epitopes on protein. Seventy-five qualified antigen–antibody complexes were used for the analyses of amino acids preference in epitope and protein (a), interrupt amino

acids frequency (b), epitope size (c) and the number of amino acids that are included in one epitope (d)

Table 3 Central amino acids of each IgG-binding region and epitope candidates selected by searching on the surface of antigen

Epitope number	Central AA	Epitope candidates
1	Q38	E4,E5,L6,N7,E11,V13,I30,K32,F33,Q34,S35,E36,E37,Q38,P138,L139, S142,W143,H145,Q146,Q149,P150,L151,T154,V155,F157,P158
2	K107	E11,V13,I23,R25,I26,K29,I30,F33,Q34,S35,E36,E37,S69,P71,K99,K105,H106,K107, E108,M109,P110,F157
3	N132	F87,L88,Q89,P90,E121,L127,T128,D129,V130,E131,N132,L133,L135,P136,L137,P138,L139
4	R183	L70,P71,I74,P75,K105,M156,F157,P158,P159,Q160,S161,V162,L163,S164,Q167,S168,K168,V170,L171, R183,L191,E195,P196,V197,L198,G199,I208

amino acids decreased. All the four epitopes comprise less than 17 amino acids when Q increased to a value between -50 and -25 % (Fig. 3). Therefore, the threshold of Q was set as -50 %. The amino acids with $Q < -50$ % were additionally eliminated. Finally, four B-cell IgG epitopes were localized on the 3-D structure of β -CN (Fig. 4).

As shown in Fig. 4, the four B-cell IgG-binding regions of β -CN consist of 15, 17, 18 and 18 residues, respectively. Each of them was also composed of more than one continuous segment. Most of these segments were separated by one or two amino acids residues.

Discussion

Prediction and identification of epitopes on allergens are still an uncertain and time-consuming task, especially for conformational epitopes. A predicted structure of β -CN was used in this study to map conformational epitopes. Beta-CN is the most hydrophobic one among the caseins. It is difficult for β -CN to form crystals, and this may be due to the fact that it is dependent on environment surrounding and its propensity for self-association [27]. Currently, most of the methods used for predicting conformational epitopes

Fig. 3 Change of the number of amino acids that are included in epitopes as a function of Q . Q was defined as the relative amino acid frequency and can be calculated as $Q = (e-n)/n$, where e is the amino acid frequency in epitope and n denotes the amino acid frequency in protein. The amino acids with a Q value lower than a certain Q threshold were eliminated from the epitope candidates' cluster

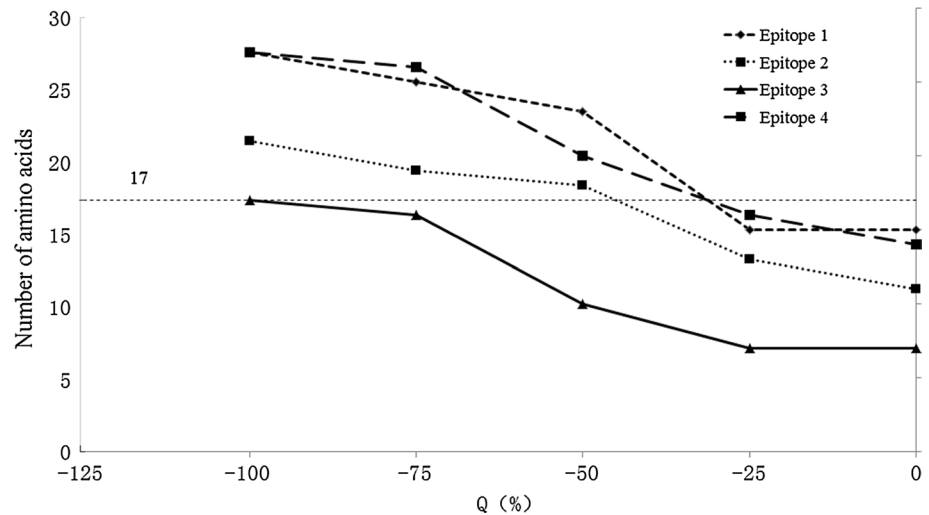
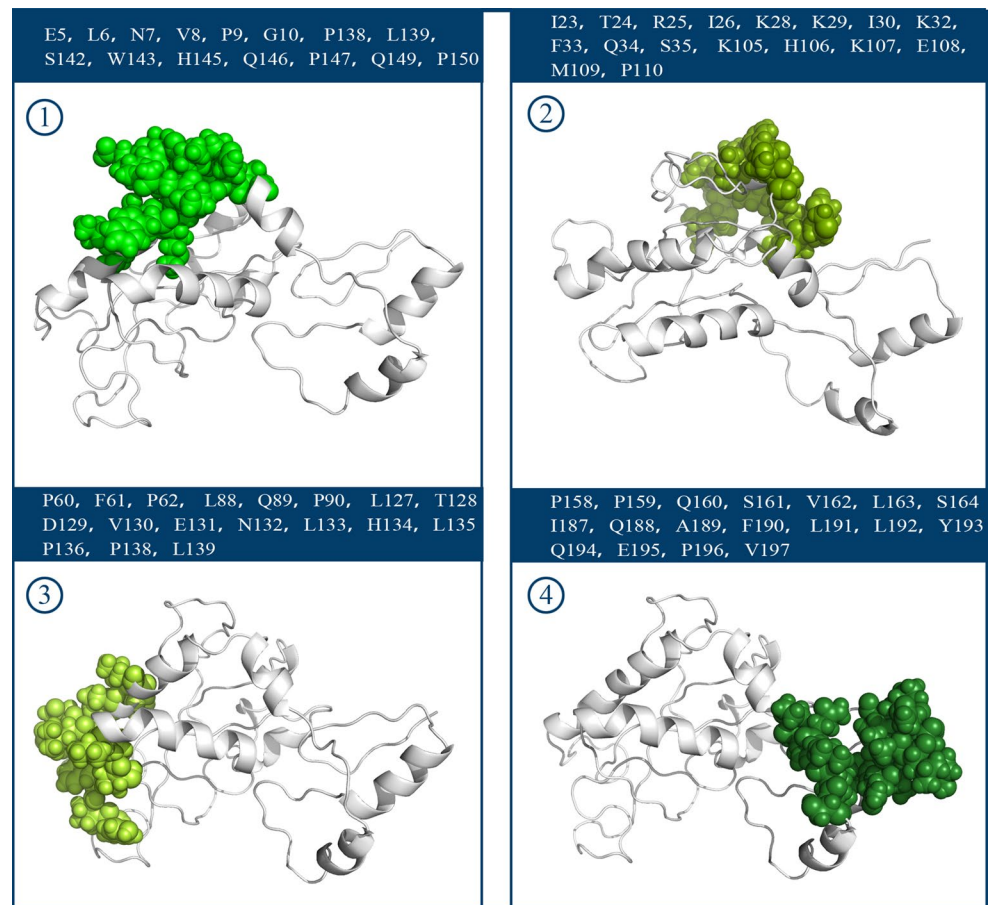


Fig. 4 Predicted conformational B-cell epitopes on β -casein. Antigen is indicated by white ribbons. Epitopes are colored green and shown with space-filling model. Residues that are included in each epitope are shown above the 3-D structures. All the models were illustrated with PyMOL v0.99



on allergens have been based on the tertiary structures. Therefore, an accurate conformational structure is essential to get more reasonable results of identifying epitope on allergen. Kumosinski et al. [16] predicted a model for β -CN based on ab initio method and the secondary structure information resulting from Raman spectroscopy. The author reported that this model was in good agreement with

biochemical cleavage results for plasmin and chymosin action on β -CN. Moreover, Livney et al. studied the tertiary structure of β -CN using crosslinking method and compared their results with the predicted model of Kumosinski. They found that their results were mostly inconsistent with this model. In this study, we built the homology model for β -CN and performed the energy minimization and

molecular dynamics simulations on it. As shown in Table 2, most of our results are in better agreement with the length of the crosslinks than the model of Kumosinski. However, the distance between the residue R1 and K176 calculated by our model is much longer than the length of crosslink, and this may be due to the use of urea with high concentration that may alter the conformational structure of proteins during the preparation of β -CN. The concentrated urea decreases the hydrophobic effect and directly binds to the amide units via hydrogen bonds, resulting in the changes in the conformational structure of proteins [28].

Characteristics of the 75 epitopes were comprehensively analyzed in our study. Asparagine (N), tryptophan (W), tyrosine (Y), glycine (G) and lysine (K) showed high propensities in epitope. Among them, N, W and Y are polar residues. The similar results have been found by other researchers [29]. Tryptophan (W) and tyrosine (Y) were also previously reported to with preference in epitopes [24, 30]. In addition, alanine (A), aspartate (D) and asparagine (N) were found to be liable to absent from epitopes [31]. The properties of residues in epitopes in our work may be different with those reported by other groups. This may be due to the fact that some studies compared the epitope residues with residues in protein, while other studies may just take residues located on the surface of allergens into account. Another factor is the differences in data sets. As the increase in the amount of the resolved co-crystal structure of antigen–antibody complex, the future data sets used for analysis of epitopes would become more complete. In addition, as the complex is usually different from the structure of respective binding sites in the free antibody and antigen molecules, the conformational structure information that is obtained from resolved co-crystal structure of antigen–antibody complex is not suitable to be directly used as criteria in the selection of epitope candidates [13]. We found that the epitopes consist of 11–25 residues whose side chains expose to a higher extent, and they cover an average area around 800 Å². Moreover, epitopes are enriched with polar and charged residues and located in a relatively protruding area. All these characteristics of the epitopes distinguish them from the remaining antigen surface. Neighbors of the B-cell epitope carry useful information that can help discriminate better between epitopes and non-epitopes [6]. This makes the recognition and binding between the epitope region and the complementarity-determining regions (CDR) of the antibody much easier.

Epitope candidates have been obtained from bio-panning and the followed “select and filter” steps. Phage display random peptides library was used to map conformational epitopes on β -CN. These cyclic peptides displayed on the phage are structurally constrained, and thus they are capable to bind CDR of the antibody. Unbound peptides were washed away, and bound phages were eluted for

amplification and the subsequent bio-panning. This bio-panning method has been widely used in selecting peptides with a high affinity to a given target. The results differ with the use of different peptide library. The peptide library used in our study consists of peptides with only seven residues. However, some epitopes consist of more residues, and not enough information can be obtained from the alignment of linear sequences of the selected peptides to the sequence of the antigen, especially for conformational epitopes. In this work, the selected peptides represent discontinuous fragments of amino acids that may not have a high sequence homology to the antigen, and thus the peptides can partly represent “true” epitopes. In addition, some of the mimotopes may reflect noisy biological observations and should be filtered out in the analysis. Hence, the combination of phage display method and computational method is necessary to map the panel of affinity-selected peptides on the surface of a target allergen.

Four conformational epitopes on β -CN were originally found in our study. Each epitope consists of several fragments. These fragments may “survive” from gastrointestinal digestion and then get access to immune system. This process obscures the boundary between the linear and conformational epitope. Previously, Senocq et al. [32] have found six antigenic determinants at sequence 4–8, 14–24, 33–48, 49–91, 178–183 and 184–209 of β -CN. These determinants consist of 1, 4, 1, 2, 1 and 5 epitopes, respectively. In respect of this work, the peptides 5–10, 60–62, 88–90 and 187–197 of β -CN were in good agreement with the results of Senocq [32], while peptides 138–150, 23–35, 105–110, 127–139 and 158–164 were not. Twenty-one murine monoclonal antibodies were used in their study, whereas we have mapped the conformational epitopes on β -CN that are recognized by rabbit polyclonal antibodies. This may partly account for the discrepancy between the two studies. The IgE-binding regions, AA 107–120 and AA 55–70, have been identified in the study of Lissou et al. [33]. These two epitopes overlap with the B-cell binding regions identified in this study, indicating that some peptides could be identified by both of IgE and IgG antibodies. Similar conclusion can be drawn from some previous reports [34]. In another work, determinants at sequence 1–25, 26–60, 61–93, 94–102, 103–109, 110–144 and 157–185 of β -CN were identified by rabbit polyclonal antibodies [35, 36]. Peptides 23–35 and 138–150 identified in our study are out of the determinant regions derived from the cited works. It may be caused by the dramatic difference between individual immune responses of rabbits. Moreover, parts of these two peptides were located at the α -helix region of our model, whereas there is no particular structure in these two regions according to the model of Kumosinski [16]. According to the casein micelle model created by Horne [37], phosphoserine-rich region of β -CN

was buried inside the micelle structure, and it is difficult for the fragments in this region to be identified by the immune system. In the study of Johansson [38], monoclonal antibodies specific to casein micelle did not identify the peptides in the phosphoserine-rich region (sequence 1–25). However, sequence 5–10 of β -CN was found as part of the conformational epitope in our study, and this discrepancy may originate from the use of β -CN instead of casein micelle.

Conclusions

Four B-cell IgG-binding regions on β -CN can be recognized by rabbit polyclonal IgG antibody. All of the IgG epitopes were composed of more than one continuous segment which were separated by one or two amino acids residues. The epitopes are enriched with asparagine (N), tryptophan (W), tyrosine (Y), glycine (G) and lysine (K), whereas the amino acids alanine (A), cysteine (C), aspartate (D), asparagine (N) and serine (S) are prone to separate the IgG epitopes into several fragments.

Acknowledgments The work was supported by National High Technology Research and Development Program of China (863 Program, No. 2013AA102205), the International Science & Technology Cooperation Program of China (No. 2013DFG31380), National Natural Science Foundation of China (No. 31171716) and the Research Program of State Key Laboratory of Food Science and Technology (SKLF-ZZB-201302).

Compliance with ethical standards

Conflict of interest None.

Compliance with ethics requirements The performance of animal experiments was authorized by Institute of Occupational Medicine of Jiangxi (permission number SYXK (Gan) 2008-003).

References

1. Jerne NK (1960) Immunological speculations. *Ann Rev Microbiol* 14(1):341–358
2. Flower DR (2007) *Immunoinformatics: predicting immunogenicity in silico*. Springer, New York
3. Cerecedo I, Zamora J, Shreffler WG, Lin J, Bardina L, Dieguez MC, Wang J, Muriel A, de la Hoz B, Sampson HA (2008) Mapping of the IgE and IgG4 sequential epitopes of milk allergens with a peptide microarray-based immunoassay. *J Allergy Clin Immunol* 122(3):589–594
4. Matsumoto N, Okochi M, Matsushima M, Ogawa A, Takase T, Yoshida Y, Kawase M, Isobe K, Kawabe T, Honda H (2009) Development of peptide arrays for detection of IgE-binding epitopes in cow's milk allergens. *J Biosci Bioeng* 107(3):324–330
5. Hayakawa K, Linko Y-Y, Linko P (1999) Mechanism and control of food allergy. *LWT Food Sci Technol* 32(1):1–11
6. EL-Manzalawy Y, Dobbs D, Honavar V (2008) Predicting linear B-cell epitopes using string kernels. *J Mol Recognit* 21(4):243–255
7. Geysen HM, Rodda SJ, Mason TJ (1986) A priori delineation of a peptide which mimics a discontinuous antigenic determinant. *Mol Immunol* 23(7):709–715
8. Mayrose I, Shlomi T, Rubinstein ND, Gershoni JM, Ruppin E, Sharan R, Pupko T (2006) Epitope mapping using combinatorial phage-display libraries: a graph-based algorithm. *Nucleic Acids Res* 35(1):69
9. Moreau V, Granier C, Villard S, Laune D, Molina F (2006) Discontinuous epitope prediction based on mimotope analysis. *Bioinformatics* 22(9):1088
10. Negi SS, Braun W (2009) Automated detection of conformational epitopes using phage display peptide sequences. *Bioinform Biol Insights* 3:71
11. Schreiber A, Humbert M, Benz A, Dietrich U (2005) 3D@Epitope@Explorer (3DEX): localization of conformational epitopes within three-dimensional structures of proteins. *J Comput Chem* 26(9):879–887
12. Benedé S, López-Expósito I, Giménez G, Grishina G, Bardina L, Sampson HA, López-Fandiño R, Molina E (2014) Mapping of IgE epitopes in in vitro gastroduodenal digests of β -lactoglobulin produced with human and simulated fluids. *Food Res Int* 62:1127–1133
13. Van Regenmortel MH (2006) Immunoinformatics may lead to a reappraisal of the nature of B cell epitopes and of the feasibility of synthetic peptide vaccines. *J Mol Recognit* 19(3):183–187
14. Huang J, Gutteridge A, Honda W, Kanehisa M (2006) MIMOX: a web tool for phage display based epitope mapping. *BMC Bioinform* 7(1):451
15. Swaisgood H, Brunner J (1962) Characterization of k-casein obtained by fractionation with trichloroacetic acid in a concentrated urea solution 1, 2. *J Dairy Sci* 45(1):1–11
16. Kumosinski T, Brown E, Farrell H Jr (1993) Three-dimensional molecular modeling of bovine caseins: an energy-minimized β -casein structure. *J Dairy Sci* 76(4):931–945
17. Livney Y, Schwan A, Dalgleish D (2004) A study of [beta]-casein tertiary structure by intramolecular crosslinking and mass spectrometry. *J Dairy Sci* 87(11):3638–3647
18. Zhang Q, Wang P, Kim Y, Haste-Andersen P, Beaver J, Bourne PE, Bui HH, Buus S, Frankild S, Greenbaum J (2008) Immune epitope database analysis resource (IEDB-AR). *Nucleic Acids Res* 36(suppl 2):W513
19. Kim Y, Sette A, Peters B (2011) Applications for T-cell epitope queries and tools in the immune epitope database and analysis resource. *J Immunol Methods* 374(1):62–69
20. Li X, Luo ZL, Chen HB, Cao YS (2008) Isolation and antigenicity evaluation of β -lactoglobulin from buffalo milk. *Afr J Biotechnol* 7(13):2258–2264
21. Diaz-Amigo C (2010) Towards a comprehensive validation of ELISA kits for food allergens. Case 2—milk. *Food Anal Methods* 3(4):351–356
22. Barre A, Jacquet G, Sordet C, Culerrier R, Rougé P (2007) Homology modelling and conformational analysis of IgE-binding epitopes of Ara h 3 and other legumin allergens with a cupin fold from tree nuts. *Mol Immunol* 44(12):3243–3255
23. Hooft R, Vriend G, Sander C, Abola EE (1996) Errors in protein structures. *Nature* 381(6580):272
24. Rubinstein ND, Mayrose I, Halperin D, Yekutieli D, Gershoni JM, Pupko T (2008) Computational characterization of B-cell epitopes. *Mol Immunol* 45(12):3477–3489
25. Miller S, Janin J, Lesk AM, Chothia C (1987) Interior and surface of monomeric proteins. *J Mol Biol* 196(3):641–656

26. Tsodikov OV, Record MT Jr, Sergeev YV (2002) Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. *J Comput Chem* 23(6):600–609
27. Farrell H, Jimenez-Flores R, Bleck G, Brown E, Butler J, Creamer L, Hicks C, Hollar C, Ng-Kwai-Hang K, Swaisgood H (2004) Nomenclature of the proteins of cows' milk—sixth revision. *J Dairy Sci* 87(6):1641–1674
28. Zou Q, Habermann Rottinghaus SM, Murphy KP (1998) Urea effects on protein stability: hydrogen bonding and the hydrophobic effect. *Proteins Struct Funct Bioinform* 31(2):107–115
29. Jackson RM (1999) Comparison of protein–protein interactions in serine protease inhibitor and antibody antigen complexes: implications for the protein docking problem. *Protein Sci* 8(3):603–613
30. Bogan AA, Thorn KS (1998) Anatomy of hot spots in protein interfaces. *J Mol Biol* 280(1):1–9
31. Soga S, Kuroda D, Shirai H, Kobori M, Hirayama N (2010) Use of amino acid composition to predict epitope residues of individual antibodies. *Protein Eng Des Sel* 23(6):441–448
32. Senocq D, Dupont D, Rolet-Repecaud O, Levieux D (2001) Monoclonal antibodies against bovine β -casein: production and epitope characterization. *Food Agric Immunol* 13(4):213–224
33. Lisson M, Lochnit G, Erhardt G (2013) Genetic variants of bovine β - and κ -casein result in different immunoglobulin E-binding epitopes after in vitro gastrointestinal digestion. *J Dairy Sci* 96(9):5532–5543
34. Garcia Alonso M, Caballero M, Umpierrez A, Lluch Bernal M, Knaute T, Rodríguez Pérez R (2015) Relationships between T cell and IgE/IgG4 epitopes of the *Anisakis simplex* major allergen Ani s 1. *Clin Exp Allergy* 45(5):994–1005
35. Otani H, Mine Y, Hosono A (1987) Studies on the antigenic structure of bovine β -casein. V. Antigenic activities of some peptides derived from residues 1–93. *Milchwissenschaft* 42(8):505–508
36. Otani H, Dong XY, Hosono A (1990) Antigen specificity of antibodies raised in rabbits injected with a chymotryptic casein digest with MW less than 1000. *Jpn J Dairy Food Sci* 39(2):A31–A37
37. Horne DS (2006) Casein micelle structure: models and muddles. *Curr Opin Colloid Interface Sci* 11(2):148–153
38. Johansson A, Lugand D, Rolet-Répécaud O, Mollé D, Delage MM, Peltre G, Marchesseau S, Léonil J, Dupont D (2009) Epitope characterization of a supramolecular protein assembly with a collection of monoclonal antibodies: the case of casein micelle. *Mol Immunol* 46(6):1058–1066