Polish Journal of Food and Nutrition Sciences

Pol. J. Food Nutr. Sci., 2011, Vol. 61, No. 2, pp. 153-158
http://journal.pan.olsztyn.pl

Short Report
Section: Food Chemistry

DOI:10-2478/v10222-011-0016-6

Update of the List of Allergenic Proteins From Milk, Based on Local Amino Acid Sequence Identity with Known Epitopes From Bovine Milk Proteins – a Short Report

Piotr Minkiewicz[1], Jerzy Dziuba[1], Izabela Gładkowska–Balewicz[1,2]*

[1]Chair of Food Biochemistry, University of Warmia and Mazury in Olsztyn, Plac Cieszyński 1, 10–726 Olsztyn, Poland

Key words: milk proteins, allergens, epitopes, cross-reactivity, bioinformatics, local sequence identity

The study involved the screening of protein sequence database nrdb 95 for sequences containing fragments identical with experimentally proven sequential epitopes of bovine milk proteins. Such fragments were found in proteins from milk of the buffalo (*Bubalus bubalis*), yak (*Bos grunniens*), goat (*Capra hircus*) and ewe (*Ovis aries*). Some proteins, such as α-lactalbumins (from the yak, buffalo and goat) and κ-caseins (from the goat and ewe), have not been previously considered as allergens. They were entered into a new database of allergenic proteins from foods and their epitopes, which is part of the BIOPEP database http://www.uwm.edu.pl/biochemia.

The proteins containing fragments identical with linear epitopes from known allergens should also be classified as allergens, based on local sequence identity. The absence of common linear epitopes with known allergens cannot be treated as the evidence that a given protein is not allergenic.

INTRODUCTION

Food allergy is becoming an increasingly serious problem, in particular among children. Among food allergens there are a number of food proteins that are known to trigger allergic reactions in sensitive patients [Skripak & Sampson, 2008; Ramesh, 2008; Jędrychowski *et al*., 2008]. The number of known allergens is growing rapidly due to, among others, the application of proteomic methods for the experimental detection of new allergens [for a review see González-Buitrago *et al*., 2007]. Proteomic experiments are supported and enhanced by a rapid growth of the number of databases of protein sequences and structures [for a review see Jelić *et al*., 2003; Akalin, 2006].

The application of bioinformatics methods in immunology is a "hot topic" [Korber *et al*., 2006; Tong *et al.*, 2007; Tong & Ren, 2009]. The bioinformatics strategies and methods for cross-reactivity prediction have been reviewed by Goodman [2006], Schein *et al*. [2007] and Ivanciuc *et al*. [2009c].

Searching for the motifs characteristic of the families of allergenic proteins, an analysis of distance between the physicochemical properties of amino acid residues in peptides or a Support Vector Machine algorithm for evaluating the possibility of cross-reactions [Ivanciuc *et al*., 2009a, b, c; Muh *et al*., 2009] may serve as recent examples of the use of bioinformatics methods in immunology and allergology. The official bioinformatics criteria recommended by WHO for the identification of proteins as allergens able to cross-react with previously known allergens are as follows: the presence of a common fragment containing at least 6–8 amino acid residues or the presence of a fragment containing at least 80 amino acid residues with an identity of at least 35% [for a review see Goodman, 2006; Ivanciuc *et al*., 2009c].

Milk proteins are considered potential food allergens being the most common cause of food allergy especially in the early childhood [for a review see Monaci *et al*., 2006; Restani *et al*., 2009; Järvinen & Chatchatee, 2009]. In some countries they are the most common cause of allergy. Due to this fact their epitopes can be identified using the epitope mapping procedure [for a review see Monaci *et al*., 2006]. Bovine milk protein epitopes have been extensively investigated and may serve as templates for computer-aided research on other proteins. Milk proteins originating from other species, such as ewe (*Ovis aries*) and goat (*Capra hircus*), reveal cross-reactivity to bovine milk [for a review see Järvinen & Chatchatee, 2009].

The aim of this study was to investigate about the occurrence of sequential epitopes shared between major bovine milk proteins and proteins from other species.

METHODS

The dataset of query sequences comes from a review article by Monaci and co-workers [2006]. The MS-BLAST program, available at the website: http://dove.embl-heidelberg.

* Corresponding author: Tel. +4889 523 3715
E-mail: *jerzy.dziuba@uwm.edu.pl* (Prof. Dr. Jerzy Dziuba)
[2] Present address: Nuffield Health, The Manor Hospital, Headington, OX3 7RP Oxford, Oxfordshire, UK

de/Blast2/msblast.html [Shevchenko *et al*., 2001] was used in the study. Screening was performed against the nrdb95 database [Holm & Sander, 1998], using the PAM 30 MS matrix and default values of all parameters attributed to searching. The application of MS-BLAST has been discussed by Darewicz *et al.* [2007] and Shevchenko *et al*. [2009].

RESULTS AND DISCUSSION

An example of the output of the MS-BLAST program is presented in Figure 1. Such an output format is typical of programs screening protein sequence databases. The output file displays information about the proteins found (in this case α-lactalbumin from the goat, *Capra hircus*), sequence alignment between the query sequence (the sequence of the epitope from bovine α-lactalbumin or sequence of entire protein precursor) and the subject (protein from the screened database), the location of a fragment in the subject sequence, the length and identity as well as the score [Smith & Waterman, 1981] estimating the statistical significance of alignment. In this study we searched for fragments whose length is identical to the length of the query sequence (12 amino acid residues in the presented example) and showing 100% identity, as shown in Figure 1a. The parameters describing the statistical significance of alignment, designed for homology searching, such as Smith-Waterman score, are not applicable in this case. Sequences presented in Figure 1b contain 142 amino acid residues and reveal 95% identity. This length and identity significantly exceeds official WHO criterion requiring presence of fragment containing at least 80 amino acid residues and at least 35 % identity [for a review see Goodman, 2006].

The proteins containing fragments identical with experimentally proven epitopes from major proteins of bovine milk are listed in Table 1. The sequences of common epitopes (occurring in both bovine milk proteins and their homologs from other species) are available in the BIOPEP database of allergenic proteins from foods (http://www.uwm.edu.pl/biochemia, click the icon of an allergen and then the "Homology" icon). Since some of the proteins listed in Table 1 are not included in the Allergen Atlas database [Tong *et al*., 2009] – the most up-to-date database of allergenic proteins, available at the website: http://tiger.dbs.nus.edu.sg/ATLAS/, we decided to assign them names according to the allergen nomenclature system. In view of the rules summarized by King and co-workers [1995], the name of an allergenic protein should include a species abbreviation and an allergen number. In the Allergen Atlas database, published in 2009, the number is often replaced by an abbreviated protein name. Such a modification of the rules is used to describe milk allergens. In our study we followed the principles contained in the Allergen Atlas, except for buffalo and goat caseins in whose case our proposed names are more detailed than those used previously.

As compared with the classical approach which involves searching for any identical fragments containing 6–8 amino acid residues [for a review see Goodman, 2006] or fragments containing 5 amino acid residues from epitopes [Lucchese *et al*., 2009], we proposed the following modifications:

– using experimentally proven epitopes as query sequences which, compared with using protein sequences as queries, facilitates the interpretation of results;
– searching for common fragments longer than 6–8 amino acid residues as synthetic peptides used for epitope mapping usually contain 10–15 amino acid residues and shorter fragments responsible for immunogenicity are not always well defined. The above applies also to fragments originating from proteolysis.

Until recently, a protein could be classified as an allergen if it contained a fragment consisting of 6–8 amino acid residues, identical to that in a known allergenic protein [Bindslev-Jensen *et al*., 2003; for a review see Goodman, 2006; Ivanciuc *et al*., 2009c]. However, the above criterion does not define any rules for selecting such a fragment, implying that it may be any fragment of an allergenic protein, whereas the peptides used for epitope mapping usually contain 10–15 amino acid residues. Using epitope sequences as queries, compared with using the entire protein sequences, makes it easier to interpret the results (even for someone not familiar with bioinformatics).

Figure 2 presents the possible place of searching for local sequence identity with experimental epitopes at the decision tree used for classifying proteins as allergenic or non-allergenic. Searching for fragments identical with experimental linear epitopes may be recommended as the first step in the bioinformatics-assisted searching process.

The use of local sequence identity of proteins with experimentally proven linear epitopes of known allergens as a criterion of allergenicity and cross-reactivity does not produce false positive results. The presence of such epitopes is the evidence that there exist people suffering from allergy to this protein or its fragment, *i.e.* the protein should be considered as an allergen. This approach is free from the limitations of other bioinformatics methods applied to predict immunogenicity of peptides and proteins, pointed out by Gowthaman & Agrewala [2008, 2009]. On the other hand, a protein which does not contain fragments identical with experimentally proven epitopes cannot be considered safe. The proposed criterion does not allow to find allergens that do not contain sequential epitopes identical with the epitopes of previously known allergens, even if the official WHO criteria (for instance 35% similarity to the fragment containing 80 amino acid residues) show those proteins should be allergenic. The application of the official WHO criteria to allergens summarized in the SDAP database revealed that one third of the allergens were missed, *i.e.* one third of the tested allergenic proteins were predicted to be non-allergenic [Ivanciuc *et al*., 2009c]. Experimental results show that the epitopes identical with those of bovine milk are not a single case of allergy against the milk of other animal species. Allergy to goat and sheep milk proteins is possible also without allergy to bovine milk proteins [Ah-Leung *et al*., 2006]. Thus, proteins which do not contain fragments identical with known epitopes should be subjected to further investigations to confirm or exclude allergenicity.

There are numerous bioinformatics tools sufficient to search for local sequence identity. BIOPEP [Dziuba & Iwaniak, 2006; Minkiewicz *et al*., 2008] and ALGPRED [Saha & Raghava, 2006] may serve as examples of programs us-

a

```
^ = swissnew|P00712|LALBA_CAPHI Alpha-lactalbumin precursor (Lactose synthase B
           protein).//:swiss|P00712|LCA_CAPHI Alpha-lactalbumin precursor
           (Lactose synthase B protein).//:trembl|X05149|CHLACTAR_1 product:
           "alpha-lactalbumin";  Goat mRNA for prealpha-lactalbumin
           //:gp|X05149|980 alpha-lactalbumin [Capra hircus]
           Length = 142

Total Score: 90

                       0         30         60         90        120
                       |          |          |          |         |       |142
swissnew|P00712|LALBA  | ____________________________________________________
Local hits (HSPs)      |                              ___

 Score = 90 (45.0 bits)
 Identities = 12/12 (100%), Positives = 12/12 (100%)

Query:     1 STEYGLFQINNK 12
             STEYGLFQINNK
Sbjct:    66 STEYGLFQINNK 77
```

b

```
^ = swissnew|P00712|LALBA_CAPHI Alpha-lactalbumin precursor (Lactose synthase B
           protein).//:swiss|P00712|LCA_CAPHI Alpha-lactalbumin precursor
           (Lactose synthase B protein).//:trembl|X05149|CHLACTAR_1 product:
           "alpha-lactalbumin";  Goat mRNA for prealpha-lactalbumin
           //:gp|X05149|980 alpha-lactalbumin [Capra hircus]
           Length = 142

Total Score: 1009

                       0         30         60         90        120
                       |          |          |          |         |       |142
swissnew|P00712|LALBA  | ____________________________________________________
Local hits (HSPs)      | ____________________________________________________

 Score = 1009 (494.3 bits)
 Identities = 135/142 (95%), Positives = 135/142 (95%)

Query:     1 MMSFVSLLLVGILFHATQAEQLTKCEVFRELKDLKGYGGVSLPEWVCTTFHTSGYDTQAI 60
             MMSFVSLLLVGILFHATQAEQLTKCEVF  LKDLK YGGVSLPEWVCT FHTSGYDTQAI
Sbjct:     1 MMSFVSLLLVGILFHATQAEQLTKCEVFQKLKDLKDYGGVSLPEWVCTAFHTSGYDTQAI 60

Query:    61 VQNNDSTEYGLFQINNKIWCKDDQNPHSSNICNISCDKFLDDDLTDDIMCVKKILDKVGI 120
             VQNNDSTEYGLFQINNKIWCKDDQNPHS NICNISCDKFLDDDLTDDI C KKILDKVGI
Sbjct:    61 VQNNDSTEYGLFQINNKIWCKDDQNPHSRNICNISCDKFLDDDLTDDIVCAKKILDKVGI 120

Query:   121 NYWLAHKALCSEKLDQWLCEKL 142
             NYWLAHKALCSEKLDQWLCEKL
Sbjct:   121 NYWLAHKALCSEKLDQWLCEKL 142
```

FIGURE. 1. Example of results generated by the MS-BLAST program. Goat α-lactalbumin is a target sequence. a) Result obtained using epitope sequence as a query; b) Result obtained using entire bovine α-lactalbumin as a query. Common epitopes are underlined in query, target and consensus sequence at Figure b.

TABLE 1. Milk allergens listed based on the presence of sequential epitopes identical with those of bovine milk proteins.

Precursor of query sequences	Proteins containing epitopes identical with query sequences	Sequences of common epitopes (identical with query sequences)[a]
α-lactalbumin, cattle, *Bos taurus* or *Bos domesticus*, Allergen Bos d 4[b], UniProt[c] Access. No P00711, BIOPEP[a] ID 13	α-lactalbumin, water buffalo, *Bubalus bubalis*, Allergen Bub b ALA[d], UniProt[c] Access. No Q645J6, BIOPEP[c] ID 67	EQLTKCEVFRELKDLK (20–35[e; f])
	α-lactalbumin, yak, *Bos grunniens*, Allergen Bos g ALA[d], UniProt[c] Access. No Q9TSR4, BIOPEP[3] ID 68	EQLTKCEVFRELKDLK (20–35[e; f]) STEYGLFQINNK (66–77[e; f]) KKILDKVGIN (112–121[e; f])
	α-lactalbumin, goat, *Capra hircus*, Allergen Cap h ALA[d], UniProt[c] Access. No P00712, BIOPEP[a] ID 69	STEYGLFQINNK (66–77[e; f]) KKILDKVGIN (112–121[e; f])
β-lactoglobulin, cattle, *Bos taurus* or *Bos domesticus*,, Allergen Bos d 5[b]; UniProt[c] Access. No P02754; BIOPEP[a] ID 14	β-lactoglobulin, ewe, *Ovis aries*, Allergen Ovi a BLG[b], UniProt[c] Access. No P67976; BIOPEP[3] ID 77	LLDAQSAPLRVYVEELKP (49–66[e; f]) AQKKIIAEKTKI (85–96[e; f]) KTKIPAVFKIDA (93–104[e; f]) KIPAVFKIDALNENKVLVLDTDYKKYLLFCM (95–115[e; f]) TDYKKYLLFCMENSAEPEQSL (113–133[e; f]) CQCLVRTPEV (135–144[e; f])
α^{s1}-casein, cattle, *Bos taurus* or *Bos domesticus*, Allergen Bos d 8 alpha s1[b]; UniProt[c] Access No P02663; BIOPEP[a] ID 9	α^{s1}-casein, goat, *Capra hircus*, Allergen Cap h casein alpha s1[g]; UniProt[c] Access No P18626; BIOPEP[c] ID 78	ELSKDIGSES (54–63[e; f]) KEDVPSERYLGYLEQLLRLK (98–117[e; f]) `
α^{s2}-casein, cattle, *Bos taurus* or *Bos domesticus*, Allergen Bos d 8 alpha s2[b]; UniProt[c] Access No P02662; BIOPEP[a] ID 10	α^{s2}-casein, water buffalo, *Bubalus bubalis*, Allergen Bub b casein alpha s2[h]; UniProt[c] Access No O62825; BIOPEP[c] ID 79	NEINQFYQKFPQYLQYLY (98–115[e]; 86–103[f]) PQYLQYLYQGPIVL (108–121[e]; 96–109[f]) VLNPWDQVKR (120–129[e]; 108–117[f]) VPITPTLNREQL (132–143[e]; 120–131[f]) STEVFTKKTKLTEEEK (158–173[e]; 125–144[f])
	α^{s2}-casein, goat, *Capra hircus*, Allergen Cap h casein alpha s2[d]; UniProt[c] Access No P33049; BIOPEP[a] ID 80	VLNPWDQVKR (120–129[e]; 121–130[f]) STEVFTKKTKLTEEEK (158–173[e]; 159–174[f])
β-casein, cattle, *Bos taurus* or *Bos domesticus*, Allergen Bos d 8 beta[b]; UniProt[c] Access No P02666; BIOPEP[a] ID 11	β-casein, goat, *Capra hircus*, Allergen Cap h casein beta[d]; UniProt[c] Access No Q95L76; BIOPEP[a] ID 81	LQDKIHPFAQ (60–69[e; f]) KEMPFPKYPVEPFT (122–135[e; f])
κ-casein, cattle, *Bos taurus* or *Bos domesticus*, Allergen Bos d 8 kappa[b]; UniProt[c] Access No P02668; BIOPEP[a] ID 12	κ-casein, water buffalo, *Bubalus bubalis*, allergen Bub b casein kappa[h]; UniProt[c] Access No P11840; BIOPEP[a] ID 82	KIAKYIPIQYVLSRYPSYGLNYYQ (42–65[e; f]) PVALINNQFLPYPYYAKPAAVR (68–89[e; f]) VRSPAQILQWQV (96–109[e; f])
	κ-casein, goat, *Capra hircus*, Allergen Cap h casein kappa[d]; UniProt[c] Access No P02670; BIOPEP[a] ID 83	KIAKYIPIQYVLSRYPSYGLNYYQ (42–65[e; f])
	κ-casein, ewe, *Ovis aries*, Allergen Ovi a casein kappa[d]; UniProt[c] Access No P02669; BIOPEP[a] ID 84	KIAKYIPIQYVLSRYPSYGLNYYQ (42–65[e; f])

[a] Data presented in the novel database of allergenic proteins from foods and their epitopes, available at the BIOPEP website (accessed 2010.07.06; The list of common epitopes will be enriched in the future.).
[b] Present in the Allergen Atlas database [Tong *et al.*, 2009] (accessed 2010.07.07).
[c] UniProt database of protein sequences is available at the website http://www.expasy.org [Jain *et al.,* 2009, UniProt 2010].
[d] Our proposed name. Protein not available in the Allergen Atlas database (accessed 2010.07.07).
[e] Location of the epitope in sequence of allergen from bovine milk. Precursor sequence is taken into attention.
[f] Location of the epitope in target sequence. Precursor sequence is taken into attention.
[g] Present in the Allergen Atlas database (accessed 2010.07.07) as allergen Cap h casein.
[h] Present in the Allergen Atlas database (accessed 2010.07.07) as allergen Bub b casein.

ing search engines enabling to find epitopes using a protein sequence as a query. The SDAP database [Ivanciuc *et al*., 2009c] may be used to search for proteins containing fragments with at least 6–8 amino acid residues identical with the corresponding fragment of a query sequence, according to the official WHO criteria.

The tools for sequence alignments such as BLAST [Altschul *et al*., 1997] and FASTA [Pearson, 2000] use the sequences of epitopes as queries. The Immune Epitope Database [Vita *et al*., 2010] is the most extensive source of linear epitope sequences. The UniProt protein sequence database [Jain *et al*., 2009; UniProt Consortium, 2010] may serve as a source of protein sequences. The Allergen Atlas database [Tong *et al*., 2009] contains over 1500 sequences of allergenic proteins and enables screening for global and local sequence similarity using the BLAST program. The SDAP database offers a similar option [Ivanciuc *et al*., 2009c]. All of the above databases are available *via* the website of the BIOPEP database (*via* the "Useful links" icon followed by "Protein Resources" or "Immunology of Proteins and Peptides"). The list of milk allergens will be updated in the future due to the rapid increase in the number of protein and epitope sequences

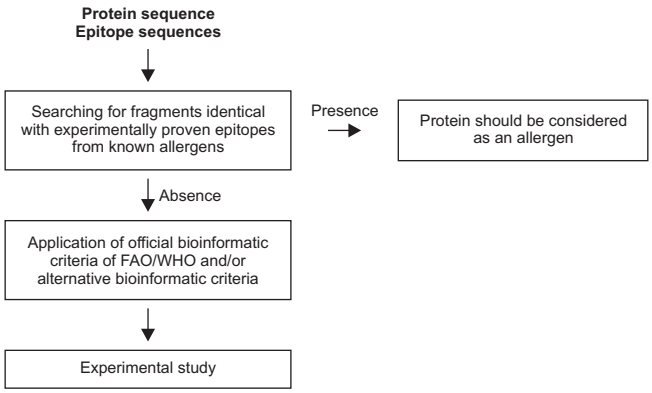

FIGURE. 2. Place of searching for local sequence identity with experimentally proven epitopes at the decision tree used for classifying proteins as allergenic or non-allergenic.

in databases available *via* the World Wide Web. The limitation of the proposed strategy is that it is restricted only to allergens containing experimentally found linear epitopes.

CONCLUSIONS

Experimentally found linear sequence epitopes from bovine milk proteins occur in proteins from milk of other animal species, such as the goat (*Capra hircus*), ewe (*Ovis aries*), buffalo (*Bubalus bubalis*) and yak (*Bos grunniens*). Therefore, proteins from the milk of those species should be classified as allergens revealing cross-reactivity with bovine milk proteins, based on local sequence identity. The search for local sequence identity using experimental epitopes as query sequences is the simplest possible bioinformatics strategy for finding new allergens. However, the absence of common linear epitopes with known allergens cannot be treated as the evidence that a given protein is not allergenic.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Science and Higher Education within project no. NN 312 286 333.

REFERENCES

1. Ah-Leung S., Bernard H., Bidat E., Paty E., Rancé F., Scheinmann P., Wal J.M., Allergy to goat and sheep milk without allergy to cow's milk. Allergy, 2006, 61, 1358–1365.
2. Akalin P.K., Introduction to bioinformatics. Mol. Nutr. Food Res., 2006, 50, 610–619.
3. Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res., 1997, 25, 3389–3402.
4. Bindslev-Jensen C., Sten E., Earl L.K., Crevel R.W.R., Bindslev-Jensen U., Hansen T.K., Skov P.S., Poulsen L.K., Assessment of the potential allergenicity of ice structuring protein type III HPLC 12 using the FAO/WHO 2001 decision tree for novel foods. Food Chem. Toxicol., 2003, 41, 81–87.
5. Darewicz M., Dziuba J., Minkiewicz P., Computational characterisation and identification of peptides for *in silico* detection of potentially celiac-toxic proteins. Food Sci. Technol. Int., 2007, 13, 125–133.
6. Dziuba J., Iwaniak A., Database of protein and bioactive peptide sequences. 2006, *in:* Nutraceutical Proteins and Peptides in Health and Disease (eds. Y. Mine, F. Shahidi). CRC – Taylor & Francis, Boca Raton – London – New York, pp. 543–563.
7. González-Buitrago J.M., Ferreira L., Isidoro-García M., Sanz C., Lorente F., Avila I., Proteomic approaches for identifying new allergens and diagnosing allergic diseases. Clin. Chim. Acta, 2007, 385, 21–27.
8. Goodman R.E., Practical and predictive bioinformatic methods for the identification of potentially cross-reactive protein matches. Mol. Nutr. Food Res., 2006, 50, 655–660.
9. Gowthaman U., Agrewala J.N., *In Silico* tools for predicting peptides binding to HLA-class II molecules: more confusion than conclusion. J. Proteome Res., 2008, 7, 154–163.
10. Gowthaman U., Agrewala J.N., *In silico* methods for predicting T-cell epitopes: Dr Jekyll or Mr Hyde? Expert Rev. Proteom., 2009, 6, 527–537.
11. Holm L., Sander C., Removing near-neighbour redundancy from a large protein sequence collection. Bioinformatics, 1998, 14, 423–429.
12. Ivanciuc O., Garcia T., Torres M., Schein C.H., Braun W., Characteristic motifs for families of allergenic proteins. Mol. Immunol., 2009a, 46, 559–568.
13. Ivanciuc O., Midoro-Horiuti T., Schein C.H., Xie L., Hillman G.R., Goldblum R.M., Braun W., The property distance index PD predicts peptides that cross-react with IgE antibodies. Mol. Immunol., 2009b, 46, 873–883.
14. Ivanciuc O., Schein C.H., Garcia, T. Oezguen N., Negi S.S., Braun W., Structural analysis of linear and conformational epitopes of allergens. Regulatory Toxicol. Pharmacol., 2009c, 54 (3 Suppl.), S11-S19.
15. Jain E., Bairoch A., Duvaud S., Phan I., Redaschi N., Suzek B.E., Martin M.J., McGarvey P., Gasteiger E., Infrastructure for the life sciences: design and implementation of the UniProt website. BMC Bioinform., 2009, 10, Article No 136.
16. Järvinen K.M., Chatchatee P., Mammalian milk allergy: Clinical suspicion, cross-reactivities and diagnosis. Curr. Opin. Allergy Clin. Immunol., 2009, 9, 251–258.
17. Jędrychowski L., Wróblewska B., Szymkiewicz A., State of the art on food allergens – a review. Pol. J. Food Nutr. Sci., 2008, 58, 165–175.
18. Jelić D., Toth D., Verbanac D., Macromolecular databases – a background of bioinformatics. Food Technol. Biotechnol., 2003, 41, 269–286.
19. King T. P., Hoffman D., Løwenstein H., Marsh D.G., Platt-Mills T.A.E., Thomas W., Allergen nomenclature. J. Allergy Clin. Immunol., 1995, 96, 5–14.
20. Korber B., LaBute M., Yusim K., Immunoinformatics comes of age. PLoS Comput. Biol., 2006, 2, 484–492.
21. Lucchese G., Stuffano A., Kanduc D., Proteome-guided search for influenza A B-cell epitopes. FEMS Immunol. Med. Microbiol., 2009, 57, 88–92.
22. Minkiewicz P., Dziuba J., Iwaniak A., Dziuba M., Darewicz M., BIOPEP database and other programs for processing bioactive peptide sequences. J. AOAC Int., 2008, 91, 965–980.
23. Monaci L., Tregoat V., van Hengel A.J., Anklam E., Milk allergens; their characteristics and their detection in food: a review. Eur. Food Res. Technol., 2006, 223, 149–179.

24. Muh H.C., Tong J.C., Tammi M.T., AllerHunter: a SVM-pairwise system for assessment of allergenicity and allergic cross-reactivity in proteins. PLoS ONE, 2009, 4, Article e5861.
25. Pearson W.R., Flexible sequence similarity searching with the FASTA3 program package. Methods Mol. Biol., 2000, 132, 185–219.
26. Ramesh S., Food allergy overview in children. Clin. Rev. Allergy Immunol., 2008, 34, 217–230.
27. Restani P., Ballabio C., Di Lorenzo C., Tripodi S., Fiocchi A., Molecular aspects of milk allergens and their role in clinical events. Anal. Bioanal. Chem., 2009, 395, 47–56.
28. Saha S., Raghava G.P.S., AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res., 2006, 34, W202-W209.
29. Schein C.H., Ivanciuc O., Braun W., Bioinformatic approaches to classifying allergens and predicting cross-reactivity. Immunol. Allergy Clin. North America, 2007, 27, 1–27.
30. Shevchenko A., Sunyaev S., Loboda A., Shevchenko A., Bork P., Ens W., Standing K.G., Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time of flight mass spectrometry and BLAST homology searching. Anal. Chem., 2001, 73, 1917–1926.
31. Shevchenko A., Valcu C.M., Jungueira M., Tools for exploiting proteomosphere. J. Proteom., 2009, 72, 137–144.
32. Skripak J.M., Sampson H.A., Towards a cure for food allergy. Curr. Opin. Immunol., 2008, 20, 690–696.
33. Smith T.F., Waterman M.S., Identification of common molecular subsequences. J. Mol. Biol., 1981, 147, 195–197.
34. Tong J.C., Lim S.J., Muh H.C., Chew F.T., Tammi M.T., Allergen atlas: A comprehensive knowledge center and analysis resource for allergen information. Bioinformatics, 2009, 25, 979–980.
35. Tong J.C., Ren E.C., Immunoinformatics: current trends and future directions. Drug Discov. Today, 2009, 14, 684–689.
36. Tong J.C., Tan T.W., Ranganathan S., Methods and protocols for prediction of immunogenic epitopes. Brief. Bioinform., 2007, 8, 96–108.
37. UniProt Consortium, The Universal Protein Resource (UniProt) in 2010. Nucleic Acids Res., 2010, 38, D142-D148.
38. Vita R., Zarebski L., Greenbaum J.A., Emami H., Hoof I., Salimi N., Damle R., Sette A., Peters B., The Immune Epitope Database 2.0. Nucleic Acids Res., 2010, 38, D854-D862.

Received March 2010. Revision received July and accepted August 2010.