



Journal of Biomolecular Structure and Dynamics

Publication details, including instructions for authors and subscription information:
<http://www.tandfonline.com/loi/tbsd20>

Accuracy of protein allergenicity prediction can be improved by taking into account data on allergenic protein discontinuous peptides

Anatoly O. Bragin^a, Pavel S. Demenkov^a, Nickolay A. Kolchanov^a & Vladimir A. Ivanisenko^a

^a Institute of Cytology and Genetics, Lavrentiev ave.10, Novosibirsk, 630090, Russia

Version of record first published: 18 Jul 2012

To cite this article: Anatoly O. Bragin, Pavel S. Demenkov, Nickolay A. Kolchanov & Vladimir A. Ivanisenko (2012): Accuracy of protein allergenicity prediction can be improved by taking into account data on allergenic protein discontinuous peptides, Journal of Biomolecular Structure and Dynamics, DOI:10.1080/07391102.2012.691362

To link to this article: <http://dx.doi.org/10.1080/07391102.2012.691362>



PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Accuracy of protein allergenicity prediction can be improved by taking into account data on allergenic protein discontinuous peptides

Anatoly O. Bragin*, Pavel S. Demenkov, Nickolay A. Kolchanov and Vladimir A. Ivanisenko

Institute of Cytology and Genetics, Lavrentiev ave.10, Novosibirsk 630090, Russia

Communicated by Vsevolod J. Makeev

(Received 29 April 2011; final version received 20 January 2012)

Allergy poses major health problems in industrialized countries, affecting over 20% of the population. Proteins from transgenic foods, cosmetics, animal hair, and other ubiquitous sources can be allergens. For this reason, development of improved methods for the prediction of potential allergenicity of proteins is timely. The currently available approaches to allergenicity prediction are numerous. Some approaches relied heavily on information on protein three-dimensional (3D) structure for allergenicity prediction. They required knowledge about 3D structure of query protein, thereby considerably restricting analysis to only those proteins whose 3D structure was known. As a consequence, many proteins with unknown structure could be overlooked. We developed a new method for allergenicity prediction, using information on protein 3D structure only for training. Three-dimensional structures of known allergenic proteins were used for representing protein surface as patches designated as discontinuous peptides. Allergenicity was predicted through search of such peptides in query protein sequences. It was demonstrated that the information on the discontinuous peptides made feasible better prediction of allergenic proteins. The allergenicity prediction method is available at <http://www-bionet.ssc.ru/psd/cgi-bin/programs/Allergen/allergen.cgi>.

Keywords: discontinuous peptide; allergenicity prediction; allergenic protein

Introduction

Allergy is today's health concern with a prevalence of more than 20% in the industrialized countries (Casolaro, Georas, Song, & Ono, 1996). Search of potential allergens is obviously needed. In addition to experimental methods designed to check proteins for food allergenicity, the World Health Organization (WHO) and the Food and Agriculture Organization (FAO) have suggested to compare sequences of the query proteins with those of known allergenic proteins. They introduced two *in silico* decision criteria. A protein was regarded as allergen, if its sequence contained at least six contiguous amino acids or showed more than 35% similarity over a window of 80 amino acids when compared with a known allergenic protein (FAO/WHO, 2003).

An approach has been developed for allergenicity prediction relying on alignment of a query sequence with known allergenic sequences using the FASTA3 algorithm and then the k-Nearest-Neighbor classification algorithm (Zorzet, Gustafsson, & Hammerling, 2002). An improved method was derived from this approach. Classifiers, such

as Bayesian linear Gaussian and the Bayesian quadratic Gaussian, were introduced (Soeria-Atmadja, Zorzet, Gustafsson, & Hammerling, 2004). A step forward was the motif-based allergenicity prediction, which assessed protein allergenicity by scanning a query sequence with motifs from known allergens (Stadler & Stadler, 2003).

Wavelet transform was proposed for identification of conserved motifs in allergenic proteins. Search of these motifs in analyzed sequences for allergenicity prediction was carried out using Hidden Markov Models (Li, Issac, & Krishnan, 2004).

Further improvement of allergenicity prediction was provided by increasing the number of identified motifs (Truffer, Stadler, Vogel, Mari, & Stadler 2006). An important finding was that search of multiple motifs in a protein of interest is more specific than the traditional single motif search (Kong, Tan, Tham, & Choo, 2007).

Credit is due to Saha and Raghava for development of methods on the basis of support vector machine (SVM) and for their hybrid approaches (Saha & Raghava, 2006a). Credit is also due to Muh et al. for presenting Allerhunter,

*Corresponding author. Email: ibragim@bionet.nsc.ru

a computational system for assessment of potential allergenicity through detection of motifs- or domain-sized similarities with known allergens (Muh, Tong, & Tammi, 2009). The SVM pairwise method was implemented for remote protein homology detection.

Most methods for prediction of protein allergenicity explicitly or implicitly had recourse to information on antigenic epitopes. For example, antigenic epitopes can be parts of motifs and of conserved regions of protein sequences used for prediction of allergenicity by the methods described above.

It is known that type I hypersensitive reaction is induced by interaction of allergen with IgE antibodies on the mast cells or basophils (Sutton & Gould, 1993). Antigenic epitopes with which IgE interacts may be both linear and discontinuous (Schramm et al., 2001; Takagi, Teshima, & Sawada, 2005). Discontinuous epitopes consist of amino acid residues spatially proximate in protein three-dimensional (3D) structure, yet remote from each other in primary protein structure.

Methods for prediction of continuous and discontinuous antigenic epitopes have become available. Continuous epitopes have been well predicted through analysis of primary protein structure (Odorico & Pellequer, 2003). To predict discontinuous epitopes, information on protein 3D structure is required (Kulkarni-Kale, Bhosle, & Kolas-kar, 2005; Liang, Zheng, Zhang, & Zacharias, 2009; Ponomarenko et al., 2008; Sun et al., 2009). Three-dimensional structure is not known for a wide range of proteins, nor can it be predicted by homology for membrane proteins, in particular. Prediction of discontinuous epitopes with the use of only protein primary structure is, so far, a challenge. This issue may be partially resolved under the assumptions that discontinuous epitopes from a protein may be continuous in another protein and that they may bind to the same monoclonal antibodies. In such a case, the idea is to identify continuous epitopes in query protein sequences, which mimic discontinuous epitopes from known allergenic proteins. The demonstration of functional mimicry of discontinuous epitopes by linear peptides (LPs) goes back to phage display (Smith, 1985). It proved that many monoclonal antibodies that bind to discontinuous epitopes are able to bind to LPs from phage library. There was frequently a good match between the amino acid composition of such peptides and their discontinuous epitopes; sequences of such peptides were observed as stretches of discontinuous epitopes in protein 3D structure. Furthermore, based on analysis of discontinuous protein functional sites, synthetic peptides were obtained and their capacity to mimic site functions was demonstrated. Such peptides were called discontinuous (Howie et al., 1998; Seelig, Prosis, & Taremi, 1994).

We have suggested an approach to prediction of discontinuous epitopes in proteins by taking advantage of similarity between discontinuous peptides, calculated

from protein 3D structure and those obtained by phage display capable of specifically binding to monoclonal antibodies. This approach assisted our identification of discontinuous epitopes in viruses (Loktev et al., 2002; Tumanova et al., 2002).

We addressed the question of whether information on similarity between the amino acids sequences of query proteins and discontinuous peptides on allergenic protein surface would improve allergenicity prediction accuracy. The results showed that the proposed method outperforms the one relying on only similarity between linear sequence fragments from known allergenic proteins and query proteins. The matches between discontinuous peptides in different allergenic proteins may supposedly evidence that the considered discontinuous peptides may be parts of potential discontinuous epitopes important to the expression of the allergenic properties of proteins. This approach may be useful in increasing the accuracy of methods for protein allergenicity prediction also for furthering innovative strategies for prediction of discontinuous epitopes.

The allergenicity prediction method was integrated into the Protein Structure Discovery System (Ivanisenko, Demenkov, Ivanisenko, & Kolchanov, 2011). The program is available at <http://www-bionet.ssc.ru/psd/cgi-bin/programs/Allergen/allergen.cgi>.

Methods

We proceeded on the assumption that accuracy of methods for predicting of protein allergenicity would be increased by taking into consideration information on discontinuous peptides on the surface of allergenic protein. To obtain assurance, we developed a simple method for allergenicity prediction based on calculation of similarity between known allergenic proteins (training sample) and the query protein sequences. The similarity was defined by comparing linear sequence fragments from a training sample with those from query proteins. To take into account discontinuous peptides, we additionally compared discontinuous peptides calculated for proteins of the training sample with linear sequence fragments of query proteins.

Creation of linear peptide set

The sequences of known allergenic proteins required for setting up the training sample were extracted from the Structural Database of Allergenic Proteins (Ivaniciu, Schein, & Braun, 2003). The PISCES program was applied to remove the highly homologous proteins (Wang & Dunbrack, 2003). The similarity threshold was set at 90%. In this way, a nonredundant set comprising 586 allergenic protein sequences was obtained.

In these sequences, LPs present in not less than two allergenic proteins were searched using sliding window eight amino acids in length. This length is frequently used

Table 1. Grouping of amino acids by similarity of their physicochemical properties.

Hydrophobic	V,M,I,L,F,Y,W
Positively charged	R,K,H
Negatively charge	E,D
Polar	S,T,N,Q
Small	A,G
Cysteine	C
Proline	P

in methods for predicting antigenic determinants and allergenicity prediction (Herman, Song, & Thirumalaiswamysekhar, 2009; Saha, & Raghava, 2006b; Silvanovich et al., 2006).

Comparisons of peptides from a sequence with peptides from another sequence were done using physicochemical properties of amino acids. To do so, peptide sequences were defined on a degenerate alphabet according to grouping of amino acids by similarity of their physicochemical properties (see Table 1).

The amino acids belonging to a group were replaced in peptide sequence by a corresponding group identifier. Two peptides were regarded as identical, if their sequence matched, when alphabet degeneracy was taken into account. In this way, a training sample set PepL consisting of over 44,000 LP sequences from allergenic proteins was provided.

Expansion of linear peptide set by discontinuous peptides

Discontinuous peptides were designed following the rules: (1) Two amino acids were accepted as bound in a discontinuous peptide, if the distance between their C-alpha atoms in protein 3D structure was not greater than 5Å. (2) The average residue solvent accessibility for amino acids for a discontinuous peptide was not less than 50%. (3) The length of the discontinuous peptides and the LPs were the same (eight amino acid residues). Figure 1 presents a scheme for protein surface carrying a set of discontinuous peptides.

To assess improved prediction accuracy for allergenic proteins, when discontinuous peptides were taken into account, the PepL set was expanded to PepLD set by including discontinuous peptides, besides linear. To cal-

culate discontinuous peptides, spatial structures of allergenic proteins were examined. The same allergenic proteins as those used for creating PepL were utilized for this task. Sixteen spatial structures were extracted from the Protein Data Bank. The 3D models of the other proteins were retrieved from a repository of predicted spatial structures (Kiefer, Arnold, Künzli, Bordoli, & Schwede, 2009). The total number of the obtained spatial structures of allergenic proteins was 345. A set of discontinuous peptides was established for each allergenic protein with known or predicted 3D structure.

A PepLD set consisting of 99,000 peptides, which included both linear from PepL (44,000) and discontinuous peptides (55,000), was created. Discontinuous peptides were deposited in PepLD, like PepL, under the condition that they repeatedly occurred in not less than two different allergenic proteins (bearing alphabet degeneracy in mind). PepLD was further used for predicting allergenic proteins with discontinuous peptides given due account.

Prediction of protein allergenicity

Prediction of protein allergenicity was based on search of similarity between linear fragments from a sequence of a query protein and peptides from the PepL and PepLD sets. A sequence of query protein was subdivided into linear fragments of eight amino acid residues applying a sliding window. Similarity between fragments and peptides was calculated taking into account degenerate alphabet (Table 1). In this way, two methods were compared. The LP method was based on analysis of LP only using the PepL set, the other linear and discontinuous peptides (LDP) method involved LDPs from the PepLD set. For each peptide from the PepL and the PepLD sets, a characteristic calculated for the occurrence frequencies of peptide in allergenic proteins was introduced [1]. This score was expressed as the ratio of the occurrence frequency of a peptide in allergenic proteins to its occurrence frequency in human proteins:

$$SA = \ln \left(\frac{Va}{Vh} \right) \quad (1)$$

where Va is the occurrence frequency of a peptide in allergenic proteins and Vh is the occurrence frequency of

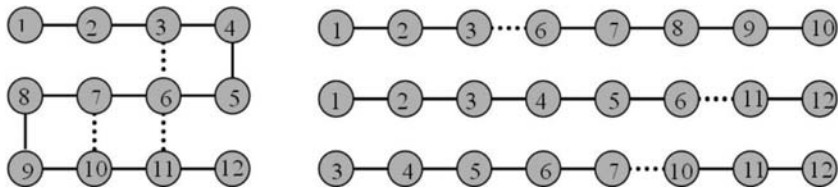


Figure 1. A schematic representation of protein surface (A) carrying a set of discontinuous peptides (B) Amino acids are represented with gray circles; black lines connecting them show peptide bonds. Amino acids with imaginary bonds in discontinuous peptides are connected by dotted lines (the distance between the C-alpha atoms of such amino acids is smaller than 5Å). Number denotes the order number of amino acids in a sequence.

Table 2. Examples of matches between LPs from allergenic protein Tri r 4 and discontinuous peptides from LDP sample.

Comparison of linear with discontinuous peptides	Positions of amino acids of LP in Tri r 4 protein sequence.	Amino acids positions of discontinuous peptide in allergenic protein sequence	Allergenic protein from which discontinuous peptide was taken
FHVSKYNF* FRVNRVSM**	45–52	68, 69, 79, 78, 77, 71, 72, 30	Lol p 2
VGIWISDA* LALVINEG**	105–112	359, 360, 297, 298, 299, 300, 301, 302	Gly m conglycinin
DGTFVGSS* DGSIWAQS**	583–590	31, 32, 33, 34, 35, 36, 37, 26	Mer a 1

*LPs from sequence Tri r 4; **discontinuous peptide from the PepLD set; identical amino acids from LDPs according to the standard alphabet are in bold, the other amino acids are identical according to the degenerate alphabet.

the peptide in human proteins. It was implicitly assumed that human proteins are not allergens, because allergens seldom occur among human proteins. For example, 22 allergens are known for *Arachis hypogaea*, with 781 *Arachis hypogaea* protein sequences stored (UniProt release 2011_04 – 5 April 2011). As for human, five proteins of over 96,000 protein sequences are allergens (Apweiler et al., 2004). The degenerate alphabet was also used in calculation of occurrence frequencies. The set of human proteins comprised about 3500 sequences chosen at random from SwissProt. To predict allergenicity of the analyzed protein, the sum of all the SA values for peptides from the PepL or PepLD sets, which matched with peptides of the query protein, was divided by the number of amino acids in the query protein [2]

$$DF = \frac{\sum_{i=1}^N SA_i}{L} \quad (2)$$

where SA_i is the characteristic for the occurrence frequencies of the i th peptide from the PepL or the PepLD sets in allergenic proteins matching with the peptide of the query protein; N is the number of matched peptides; and L is the length of the query protein (the total number of amino acids in the query protein). The negative SA values were taken as 0. A protein was regarded as allergenic, if the DF value was greater than a given threshold.

Results

Two methods LP and LDP for allergenicity prediction relying on LPs and on linear together with discontinuous peptides (LDP) were compared. Test sets of allergenic and nonallergenic proteins taken from (Muh et al., 2009) were used to estimate the accuracy of the prediction methods. A “testing.putative_non_allergen.fa” set composed of 497 nonallergens and a “testing.allergen.fa” sample composed of 141 allergens available on <http://tiger.dbs.nus.edu.sg/AllerHunter/> were used. We removed one protein from allergens with length shorter than eight amino acid residues. So we used 140 proteins in the test

set of allergenic proteins. These test sets were compared with our training sample. In accuracy estimates, we omitted those proteins from the training sample that occurred in the test sets. As a result, there remained 545 of allergenic proteins in the training sample. This means that 41 proteins were discarded from the training sample for the reason that they were presented in the test.

The number of matches between discontinuous peptides from LDP and LPs from the allergenic proteins in the test sample was rather high, when degenerate alphabet was taken into account. The proteins with numerous matches between the discontinuous and the LPs were possibly cross-reactive. It would be of interest to check this cross-reactivity experimentally. Examples of matches between LPs from Tri r 4 allergen and discontinuous peptides from PepLD are given in Table 2. The Tri r 4 protein was taken from the test sample.

Underprediction ($E1$) [3] and overprediction ($E2$) [4] errors for protein allergenicity prediction were calculated from

$$E1 = \frac{FN}{(TP + FN)} \quad (3)$$

$$E2 = \frac{FP}{(TN + FP)} \quad (4)$$

where TP are allergens predicted as allergens, FN are allergens predicted as nonallergens, TN are nonallergens predicted as nonallergens, and FP are nonallergens predicted as allergens.

The comparative results showed that the method combining linear with discontinuous peptides is superior to that using LPs alone (Table 3).

Table 3. Accuracy comparison of LP and LDP methods.

Method for allergenicity prediction	Underprediction error ($E1$)	Overprediction error ($E2$)
LP	0.171	0.068
LDP	0.143	0.068
AllerHunter*	0.163	0.068

*AllerHunter method (Muh et al., 2009).

The underprediction error was reduced by 2.8%, when using discontinuous peptides, as compared with the method of analysis using LPs alone (at the same overprediction error). Thus, accuracy of the method for prediction of protein allergenicity was increased by using discontinuous peptides. The accuracy of the LDP is quite comparable with that of the sophisticated AllerHunter method based on SVM technique (Muh et al., 2009). Our idea was not to exceed the accuracy achieved by others, much rather to clarify the contribution of the information on discontinuous peptides to the accuracy of protein allergenicity prediction. The question is: Would data on discontinuous peptides make better the available prediction methods? If so, the 3D structure of a protein should be known for discontinuous peptides to be calculated. The suggested method based on analysis of discontinuous peptides calculated for only proteins of known allergens does not require knowledge about the spatial structure of the query protein. This is an advantage enabling to use this approach in analysis of bulky sets of proteins, even whole proteomes. It is hoped that our approach will improve other methods available for prediction of protein allergenicity.

Acknowledgements

The work was supported in part by the EU-FP7 PATHOSYS project No. 260429, the Russian Ministry of Science & Education 14.740.11.0001, the Interdisciplinary Integrative Projects of the SB RAS No. 119 and 26, the Program of Russian Academy of Sciences No. 19. We thank Anna Fadeeva for translating the manuscript from Russian into English.

References

- Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., ... Yeh, L.S. (2004). UniProt: The universal protein knowledgebase. *Nucleic Acids Research*, 32, D115–119.
- Casolaro, V., Georas, S.N., Song, Z., & Ono, S.J. (1996). Biology and genetics of atopic disease. *Current Opinion in Immunology*, 8, 796–803.
- FAO/WHO. (2003). Codex principles and guidelines on foods derived from biotechnology. Rome, Italy Joint FAO/WHO Food Standards Programme.
- Herman, R.A., Song, P., & Thirumalaiswamysekhar, A. (2009). Value of eight-amino-acid matches in predicting the allergenicity status of proteins: An empirical bioinformatic investigation. *Clinical and Molecular Allergy*, 7, 9.
- Howie, S.E., Cotton, G.J., Heslop, I., Martin, N.J., Harrison, D.J., & Ramage, R. (1998). Synthetic peptides representing discontinuous CD4 binding epitopes of HIV-1 gp120 that induce T cell apoptosis and block cell death induced by gp120. *Federation of American Societies for Experimental Biology*, 12, 991–998.
- Ivanciu, O., Schein, C.H., & Braun, W. (2003). SDAP: Database and computational tools for allergenic proteins. *Nucleic Acids Research*, 31, 359–362.
- Ivanisenko, V.A., Demenkov, P.S., Ivanisenko, T.V., & Kolchanov, N.A. (2011). Protein structure discovery: A software package to computer proteomics tasks (Review). *Russian Journal of Bioorganic Chemistry*, 37, 17–29.
- Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., & Schwede, T. (2009). The SWISS-MODEL repository and associated resources. *Nucleic Acids Research*, 37, D387–392.
- Kong, W., Tan, T.S., Tham, L., & Choo, K.W. (2007). Improved prediction of allergenicity by combination of multiple sequence motifs. *In Silico Biology*, 7, 77–86.
- Kulkarni-Kale, U., Bhosle, S., & Kolaskar, A.S. (2005). CEP: A conformational epitope prediction server. *Nucleic Acids Research*, 33, 168–171.
- Li, K.B., Issac, P., & Krishnan, A. (2004). Predicting allergenic proteins using wavelet transform. *Bioinformatics*, 20, 2572–2578.
- Liang, S., Zheng, D., Zhang, C., & Zacharias, M. (2009). Prediction of antigenic epitopes on protein surfaces by consensus scoring. *BMC Bioinformatics*, 10, 302.
- Loktev, A.V., Kuvshinov, V.N., Melamed, N.V., Ivanisenko, V.A., Mishin, V.P., & Il'ichev, A.A. (2002). Localization of tick-borne encephalitis virus protein E antigenic determinant recognized by antihemagglutinating monoclonal antibodies using a phage-display peptide library. *Voprosy Virusologii*, 47, 31–34.
- Muh, H.C., Tong, J.C., & Tammi, M.T. (2009). AllerHunter: A SVM-pairwise system for assessment of allergenicity and allergic cross-reactivity in proteins. *PLoS ONE*, 4, e5861.
- Odorico, M., & Pellequer, J.L. (2003). BEPITOPE: Predicting the location of continuous epitopes and patterns in proteins. *Journal of Molecular Recognition*, 16, 20–22.
- Ponomarenko, J., Bui, H.H., Li, W., Fusseder, N., Bourne, P.E., Sette, A., & Peters, B. (2008). ElliPro: A new structure-based tool for the prediction of antibody epitopes. *Bioinformatics*, 9, 514.
- Saha, S., & Raghava, G.P. (2006a). AlgPred: Prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Research*, 34, W202–209.
- Saha, S., & Raghava, G.P. (2006b). Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins*, 65, 40–48.
- Schramm, G., Bufer, A., Petersen, A., Haas, H., Merget, R., Schlaak, M., & Becker, W.M. (2001). Discontinuous IgE-binding epitopes contain multiple continuous epitope regions: Results of an epitope mapping on recombinant Hol 1 5, a major allergen from velvet grass pollen. *Clinical and Experimental Allergy: Journal of the British Society for Allergy and Clinical Immunology*, 31, 331–341.
- Seelig, G.F., Prossise, W.W., & Taremi, S.S. (1994). Synthetic mimics of juxtaposed amino- and carboxyl-terminal peptide domains of human gamma interferon block ligand binding to human gamma interferon receptor. *The Journal of Biological Chemistry*, 269, 358–363.
- Silvanovich, A., Nemeth, M.A., Song, P., Herman, R., Tagliani, L., & Bannon, G.A. (2006). The value of short amino acid sequence matches for prediction of protein allergenicity. *Toxicological Sciences: An Official Journal of the Society of Toxicology*, 90, 252–258.
- Smith, G.P. (1985). Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science*, 228, 1315–1317.
- Soeria-Atmadja, D., Zorzet, A., Gustafsson, M.G., & Hammerling, U. (2004). Statistical evaluation of local alignment features predicting allergenicity using supervised classification algorithms. *International Archives of Allergy and Immunology*, 133, 101–112.
- Stadler, M., & Stadler, B. (2003). Allergenicity prediction by protein sequence. *Federation of American Societies for Experimental Biology*, 17, 1141–1143.

- Sun, J., Wu, D., Xu, T., Wang, X., Xu, X., Tao, L., & Cao, Z.W. (2009). SEPPA: A computational server for spatial epitope prediction of protein antigens. *Nucleic Acids Research*, 37, W612–616.
- Sutton, B.J., & Gould, H.J. (1993). The human IgE network. *Nature*, 366, 421–428.
- Takagi, K., Teshima, R., & Sawada, J. (2005). Determination of human linear IgE epitopes of Japanese cedar allergen Cry j 1. *Biological & Pharmaceutical Bulletin*, 28, 1496–1499.
- Truffer, R., Stadler, M.B., Vogel, M., Mari, A., & Stadler, B.M. (2006). *Computational resources: A regulatory need, a tool for research*. Frankfurt: Arbeiten aus dem Paul-Ehrlich-Institut (Bundesamt für Sera und Impfstoffe) zu, 11–15.
- Tumanova, O.Iu., Kuvshinov, V.N., Il'ichev, A.A., Nekrasov, B.G., Ivanisenko, V.A., Kozlov, A.P., & Sandakhchiev, L.S. (2002). Localization of the HIV-1 gp120 conformational epitope recognized by virus neutralizing monoclonal antibodies 2G12. *Molekuliarnaia Biologiya*, 36, 657–663.
- Wang, G., & Dunbrack, R.L.Jr (2003). PISCES: A protein sequence culling server. *Bioinformatics*, 19, 1589–1591.
- Zorzet, A., Gustafsson, M., & Hammerling, U. (2002). Prediction of food protein allergenicity: A bioinformatic learning systems approach. *In Silico Biology*, 2, 525–534.