Journal of Innovations in Pharmaceutical and Biological Sciences (JIPBS)
ISSN: 2349-2759
Available online at www.jipbs.com

Research article

Epitope-based molecular docking studies of allergenic proteins with immunoglobulin protein during type I hypersensitivity reaction

Anamika Basu[1*], Anasua Sarkar[2], Piyali Basak[3]

[1]*Department of Biochemistry, Gurudas College, Kolkata, India.*

[2]*Computer Science and Engineering, Jadavpur University, Kolkata, India.*

[3]*School of Bioscience and Engineering, Jadavpur University, Kolkata, India.*

Key words: Pollen allergens, Zea m 1, Phl p 1, B cell epitopes, molecular docking, cross reactivity prediction.

***Corresponding Author: Anamika Basu**, Department of Biochemistry, Gurudas College, Kolkata, India.

Abstract

Objectives: The identification of B-cell epitopes is a challenging approach to explore the antigen-antibody interactions for diagnosis and therapy of hypersensitivity reaction. In our present study, an *in-silico* approach is used to investigate the interaction of pollen allergen EXPB1 (Zea m 1), pollen allergen from maize with IgE molecules of human. **Material and Methods:** Paratope of human immunoglobulin E is identified using site-specific proABC predictor method. Phylogenetic analysis of Zea m 1 reveals that 13 pollen allergens from different grasses, maize, timothy grass, velvet grass, Bermuda grass, canary grass, rice and perennial rye grass are close homologs to our query allergen EXPB1. Among them Phl p 1 pollen allergen from *Phleum pratense* is identified with 60% identity with Zea m 1. Experimental B cell epitopes of Phl p 1 are known and we have verified those epitopes with PIPER, molecular docking software. Thus, interacting amino acids present both in epitopes and paratopes are visualized and confirmed with predicted paratopes. For all homologous allergens, the interacting amino acids i.e. epitopes and paratopes have been identified using the two docking programs, DOT and ZDOCK. **Results and Conclusions:** Negative binding energies of all pollen allergens with immunoglobulin E confirm their allergenicity. Thus, all allergens become cross reactive with maize allergen. The multiple sequence alignment for all homologous sequences reveals that the positions of antigenic peptide of Zea m 1 sequence are well conserved in its homologs and responsible for cross-reactivity. This cross-reactivity identification will help us to identify the immunotheraputics e.g. vaccine designing for these β expansin family protein allergens during pollinosis.

Introduction

Cosgrove [1] proposes the role of protein *expansins*, cell wall glycoproteins (250-275 amino acids long) in cell wall expansion by slippage or rearrangement of matrix polymers. In our earlier work, we found that the orthologous genes of expansin genes are first present in eudicots [2]. So expansin gene is included in developmental toolkit for plants considering an important developmental role inplant embryogenesis in presence of turgor pressure [2]. But among the four classes of expansin protein family, EXPB class of proteins are known in the immunological literature as group-1 grass pollen allergens [3]. These EXPBs proteins cause hay fever and seasonal asthma in humans [4]. Allergy symptoms are caused by these proteins when these allergens come into contact with the moist surface of the human respiratory tract [5]. As much as 20% of the general population may be affected by grass pollen as a major cause of allergic disease.

Identification of epitopes gives valuable information for designing an effective vaccine for leptospirosis [6]. Kumar *et al.*, 2015[6]in their in-silico work, have mapped sequential and conformational B-cell epitopes from the crystal structure of LipL32, the most abundant surface-associated protein of Leptospira and identified the order of antigenicity of four B cell epitopes. Radauer *et al.* [7] identify Zea m 1, Cyn d 1, Ory s 1, Pha a 1, Phl p 1, Lol p 1, Dac g 1, Poa p 1 and Hol l 1, major pollen allergens from maize, Bermuda grass, rice, canary grass, timothy grass, rye grass, orchard grass, Kentucky blue grass and velvet grass respectively. They reveal that these allergens are phylogenetically related pollen allergens belong to beta-expansin family proteins. These proteins cause allergic reaction in human, promoting Type I hypersensitivity reaction. In type I hypersensitivity, B-cells are stimulated to produce IgE antibodies specific to an antigen. During allergen – antibody interaction, the IgE antibodies bind to Fcε receptors on the surface of tissue mast cells and blood basophils, resulting in

degranulation of mast cells and the secretion of pharmacologically active mediators such as histamine, leukotriene (LTC4 and LTD4), and prostaglandin that act on the surrounding tissues. The principal effects of these products are vasodilation and smooth-muscle contraction. More than 25% of the world's population is affected by type I hypersensitivity reaction which causes mild to fatal effects on human beings.

The present study focuses on B cell epitope (BCE) analysis of EXPB1 (also called Zea m 1) a beta-expansin protein and group-1 pollen allergen from maize (PDB ID: 2HCZ) and its homologs. Though Allergome - *A Database of Allergenic Molecules* [8] accepts Zea m 1 as allergen, but detailed immune-mechanism of this allergen is not available. Therefore, in our present work, we have to identify the epitopes of Zea m 1 and their paratopes of human immunoglobulin E to understand the hypersensitivity reactions due to this allergen on human. At first, we have collected IgE specific BCEs from Immune Epitope Database (IEDB) [9, 10]. Subsequently, we have also analysed the binding parameters of other allergens, namely Cyn d 1, Ory s 1, Pha a 1, Phl p 1, Lol p 1, Dac g 1, Poa p 1 and Hol l 1 in this paper.

Role of allergens in hypersensitivity reaction

In general allergens are foreign proteins that when come in contact of part(s) of human body, stimulate the production of immunoglobulin types of proteins (antibodies). These allergens react with antibodies (immunoglobulin type E or IgE) and produces allergic reactions, also known as immediate-type hypersensitivity reactions.

Antibodies, expressed on the surface of B cell of human immune system, recognize antigenic determinants, also called epitopes on their antigen. The interacting part of the antibody involved in the anigen- antibody interaction, is called the paratope. Paratope is formed in combination of different amino acids in the Complementarity Determining Regions (CDRs) of antibody or immunoglobulin.

Sometimes the specific Ig E molecules for a specific allergen can identify other allergens from various sources, a cross-reactivity allergic reaction can occur. These types of cross-reactions are frequently seen between pollen allergens. Soheila *et al.* [11] discuss about cross-reactivity for computationally predicted IgE epitopes of walnut allergens with that of peanuts.

Zea m 1 is a world-wide well-known respiratory allergen, which causes polinogenesis, among people who are in contact with pollens of Zea maize. Oldenburg *et al.* [12] describes maize pollen as an important allergen in occupationally exposed workers. They also suggest cross-reactive reactions exist between Zea m 1 - Phl p 1, where Phl p 1 is known as a Major Timothy Grass Pollen Allergen (PDB ID: 1N10). The major pollen allergen Phl p 1of timothy grass is considered as one of the most potent and frequently recognized environmental allergen with crystal structure and known experimental B cell epitopes. But the molecular mechanism of cross reactions between Zea m 1 and Ph l p 1 allergens is not clear till now. Therefore, we are concentrating on the in-sillico analysis of these cross-reactions among these two pollen allergens.

B cell epitopes

Allergen-immunoglobulin interaction is based on the three-dimensional structural complementarity between the epitope of allergens and the paratope of immunoglobulins. There are two types of epitopes e.g. B cell epitopes and T cell epitopes which are related with B and T types of immunological cells respectively. Among them B-cell epitopes are very important for allergy reaction and in determining the cross-reactivity various types of allergens [13-15]. Not only that B cell epitopes play an important role vaccine design and also in the diagnosis of diseases [16-18]. Epitopes of an allergen are composed of 5-7 amino acids, may be continuous or discontinuous. Due to lack of information about experimental B cell epitopes computational tools are used to predict those. Several computer programs considering various physiochemical properties are available for prediction of B cell epitopes [19-22].

Cross-reactivity among pollen allergens

According to The World Health Organization guidelines, an allergen can be considered to cross-react with another allergen, if more than 35% identity exists between the amino acid sequences of two allergens. This identification test can be performed either using a window of 80 amino acids or by finding a peptide of 6-8 contiguous amino acids among their complete amino acids sequences. Subsequently, those two allergens must share similar three-dimensional structures with overall folding of protein molecules, which is essential for their same complementary binding with Ig E molecules. Thus, the probability of cross reaction between two allergens, belong to similar protein family, would be higher, when they are homologous proteins from phylogenetically close species. Three-dimensional modelling of these homologous proteins is required. Ara h 2, Bla g 4, and Aed a 2, three allergens are modelled and assessed using software tool, QMEAN (http://swissmodel.expasy.org /qmean/cgi/index.cgi) [23] by Power *et al,* in 2013 [24]. For prediction of cross-reactivity between homologous allergens, molecular docking study between allergens and human immunoglobin molecule is essential. During the discussion about cross reactivity among pollen allergens, Schein *et al.,* 2010 [25] conclude that the common epitopes of Zea m 1 and other grass pollen allergens Phl p 1 and Phl p 2 from timothy grass are present on these expansins. But they have also shown that these epitopes of group 1 grass allergens are not conserved among all

members of the expansin superfamily [25]. Weber *et al.* [26, 27] highlight the knowledge on cross - reactive pollen allergens, which is crucial for formulation of inhalant allergen immunotherapy.

Experimental

Materials and Method

Sequence retrieval

The primary sequence of Zea m 1from Zea maize in FASTA format has been retrieved from the RCSB PDB *(http://www.rcsb.org)*, a publicly available database [28]. In the present experiments, we use the structure of EXPB1 from x-ray crystallography of 2.75 Å resolution obtained from PDB ID: 2HCZ [29], as the 3D structure of allergen of our interest.

Phylogenetic Analysis

We perform the homolog search using NCBI BLINK Precomputed Blast (BLASTP 2.2.30) alignments for protein Zea m 1 sequence in the Entrez databases by relation to the 3D structure of protein with conserved protein domain [30]. Thus, several homologous beta-expansin family proteins have been identified.

Allergenicity prediction

We further search for the cross-reactive allergen of Zea m 1 according to FAO/ WHO allergenicity rules in the SDAP allergen database [31, 32]. SDAP integrates a database of allergenic proteins and provides structural knowledge related to allergens and characterization of their epitopes. For a specific B cell epitope sequence, the PD (Property distance) tool of SDAP allergen database [33] identifies the most similar epitopes from all the allergenic proteins stored among their database. The lower the PD between two peptides, the more similar they are (0 for identical). The results from this analysis have been added in Results and Discussions section.

Sequence retrieval for Zea m 1 and template selection

Dali server [34] is used to identify the most similar sequence of Zea m 1 having PDB structures of pollen allergens. The crystal structures of PDB IDs 1N10 and 4J4P for Phl p 1 and human IgE are retrieved from RCSB PDB database [35, 36].

Prediction of B cell epitopes

Experimental B cell linear epitopes for Phl p 1[accession number Q40967] for host organism Homo sapiens (human) for allergic diseases, have been identified from IEDB database [9, 10] for B cell positive assays.

CDRs identification

At the same time for human Ig E molecule, Complementarity Determining Regions (CDRs) are predicted using proABC method, based on the random forest automatic learning techniques. For every amino acid in the antibody sequence, proABC calculates the interaction probability of antibody with a specific antigen [37, 38]. proABC (Prediction of AntiBody Contacts) provides a prediction of non-bonded contacts, hydrogen bonds and hydrophobic interactions of a chosen amino acid sequence as an antibody with its probable antigens.

Molecular docking for allergens and immunoglobulin E molecule

Multi-stage approach to protein-protein docking

Molecular docking technique is used to determine the most stable antigen-antibody conformation for both Zea m 1 and Phl p 1 with human IgE using PIPER tool of Cluspro docking tool [39]. The protein–protein docking program, PIPER [40] is used for docking in our present work. Docked conformations have been generated using two Fast-Fourier Transform (FFT) based docking methods, DOT [41] and ZDOCK [42]. DOT runs retain 20,000 docked conformations, while ZDOCK runs retained 2000 structures. Only shape complementarity is used to sample approximately 1010 putative conformations. Among them, the top scoring 20,000 conformations are retained for filtering by desolvation and electrostatics parameters, based on the correlation approach with a 10◦ Euler angle increment, and default values of 1 Å grid-step and 4 Å surface layer.

Antigen–antibody docking (for allergens and immunoglobulin E molecule)

In the case of antibody–antigen complexes, which are mostly flat and less hydrophobic [43], the polar interactions are more sensitive to atomic positions than the hydrophobic interactions in docking. So, the total energy for this type of complexes is calculated considering the sum of terms representing shape complementarity, electrostatic and desolvation contributions and a pairwise interaction potential called Decoys as the Reference State (DARS) [44].

Assessment of the results

Among the top 10 predictions, which include at least one near-native complex, with an average RMSD (Root-Mean-Square Deviation) of 5 Å from the native structure, we identify the most stable docking structures for both allergens with human IgE molecule.

From predicted docking structures, specific experimental B cell epitopes and their interacting amino acids in immunoglobulin H and L chains of human IgE, have been identified by visualizing antigen bound antibody structure by using UCSB Chimera software [45]**.** Similar docking

method is used to predict the antigen binding sites for Zea m 1 allergen and its homologous allergens. All the results as obtained by these experiments are analysed in the Results section of this paper.

Molecular modelling for other similar allergens

Pre-computed BLAST search from NCBI BLINK for beta-expansin 1 [Zea m 1] has been done to identify structurally similar pollen allergens from different species. With homologous allergenic sequences for Zea m 1 from Allergen, Allerdata and Allfam databases [46, 47, 48], the structures of eight pollen allergens are modelled based on protein modelling using SWISS-MODEL Version 8.05.

SWISS-MODEL tool is an environment for protein structure homology modelling (swissmodel.expasy.org /workspace/) [49]. In this environment, the models are built by searching templates of allergens with BLAST [50] and HHBlits [51] against the SWISS-MODEL template library (SMTL). Consequently, for each identified template, the target-template alignment has been predicted using Promod-II program [52]. The coordinates which are conserved between the target and the template, are copied from the template to the built model. Furthermore, the insertions and deletions are remodelled using a fragment library. Side chains are then also rebuilt. Finally, the geometry of the resulting model is regularized by using a force field. The global and per-residue model quality for each model has been assessed using the QMEAN scoring function [53]. Finally, the locations of those specific epitopes are identified in the built model.

Results and Discussions

Secondary structure prediction

Dali server [34] identifies major timothy grass pollen allergen Phl p 1 (allergen from *Phleum pratense)* having crystal structure (PDB ID: 1N10) as the most similar PDB structure of allergen Zea m 1 (beta-expansin protein of Zea mays) and at the same time, sequence similarity percentage between them is calculated (as shown in Table 1).

Table 1. Result from Dali server

Sr. No.	Chain	rmsd	% id	Description
1	2HCZ_X	0.0	100	MOLECULE: BETA-EXPANSIN 1A
2	1N10_A	1.8	60	MOLECULE: POLLEN ALLERGEN PHL P 1
3	1N10_B	1.8	59	MOLECULE: POLLEN ALLERGEN PHL P 1

Sequence Analysis

Pre-computed BLAST results from NCBI BLINK for beta-expansin 1 [Zea maize] shows that the structure analysis of N terminal domain of some pollen allergens e.g. Phl p 1 and Zea m 1, contents a six stranded β barrel flanked by short loops and α-helices, termed as double-psi-β-barrel (DPBB). Not only that, among 13 amino acids sequences, 10 are amino acids sequences of pollen allergens from 7 different species, are obtained from BLASTP 2.2.30 search, which have score <810 (as shown in Table S1 in supplementary files).

Allergenicity prediction

Search result in the SDAP allergens database for allergen Zea m 1shows that, the following FASTA alignments between the query sequence 2HCZ and all SDAP major allergens (omitting isoallergens) have an E score higher than 0.010000 as shown in Table 2. These 13 pollen allergens are present in 8 plant species, as shown in Table 2.

Among the above-mentioned allergens, only Phl p 1, which has sequence homology with Zea m 1 allergen, has crystal structure with PDB ID: 1N10 with A and B polypeptide chains.

In-sillico B-cell epitope prediction

In IEDB database [9, 10], for a given input ph1 p 1, we obtain the results for experimental B cell epitopes as shown in Table S2.

Identification of antigen-binding regions (ABRs) (Paratopes)

From results of proABC predictor method [37], [38] shown in Table 3, for human IgE we can identify Paratope or Complementarity Determining Regions (CDRs) of human immunoglobulin E.

Result from Paratome server in Table 3, clearly shows that the paratope is formed in combination of different amino acids in the Complementarity Determining Regions (CDRs) of human Ig E by amino acids sequences from 46-56, 68-82 and 120- 135 for heavy chain and 45-54, 66-76 and 109-118 for light chains. So, it can be concluded that probably these regions of human immunoglobulin molecule are involved in antigen-antibody interactions, when Ig E of B cell comes in contact with allergens.

Molecular Ab-Ag docking for allergens and human Ig E

The molecular docking results for Phl p 1 and Zea m 1 with human IgE confirm the interacting amino acids of both allergens with immunoglobulin. For molecular docking study, the following structure of human Ig E, (heavy chain in green and light chain in yellow colour, as shown in Figure 1) has been used.

Table S1. Search results from Pre-computed BLAST results from NCBI BLINK for beta-expansin 1 [Zea mays]

No. of sequence	Score	Accession	Length	Protein Description
1	1469	P58738	269	RecName:Full=Expansin-B1; AltName: Full=Allergen Zea m 1d; AltName: Full=Beta-expansin-1a; AltName: Full=Pollen allergen Zea m 1; AltName: Full=ZmEXPB1; AltName: Allergen=Zea m 1; Flags: Precursor [*Zea mays*]
2	1451	Q07154	269	RecName: Full=Expansin-B9; AltName: Full=Beta-expansin-1b; AltName: Full=Pollen allergen Zea m 1; AltName: Full=ZmEXPB9; AltName: Allergen=Zea m 1; Flags: Precursor [*Zea mays*]
3	1130	Q7XCG7	269	RecName: Full=Expansin-B9; AltName: Full=Beta-expansin-9; AltName: Full=OsEXPB9; AltName: Full=OsaEXPb1.6; Flags: Precursor [*Oryza sativa* Japonica Group]
4	951	Q40638	267	RecName: Full=Expansin-B1; AltName: Full=Beta-expansin-1; AltName: Full=Major pollen allergen Ory s 1; AltName: Full=Ory s I; AltName: Full=OsEXPB1; AltName: Full=OsaEXPb1.2; AltName: Full=OsaEXPb1.3; AltName: Allergen=Ory s 1; Flags: Precursor [*Oryza sativa* Japonica Group]
5	932	Q8H7T4	267	RecName: Full=Expansin-B10; AltName: Full=Beta-expansin-10; AltName: Full=OsEXPB10; AltName: Full=OsaEXPb1.5; Flags: Precursor [*Oryza sativa Japonica* Group]
6	917	P0C1Y5	269	RecName: Full=Expansin-B11; AltName: Full=Beta-expansin-11; AltName: Full=Pollen allergen Zea m 1a; AltName: Full=Pollen allergen Zea m 1b; AltName: Full=ZmEXPB11; AltName: Allergen=Zea m 1; Flags: Precursor [*Zea mays*]
7	886	Q1ZYQ8	270	RecName: Full=Expansin-B10; AltName: Full=Beta-expansin-10; AltName: Full=Pollen allergen Zea m 1c; AltName: Full=ZmEXPB10; AltName: Allergen=Zea m 1; Flags: Precursor [*Zea mays*]
8	869	P43213	263	RecName: Full=Pollen allergen Phl p 1; AltName: Full=Allergen Phl p I; AltName: Allergen=Phl p 1; Flags: Precursor [*Phleum pratense*]
9	866	P43216	265	RecName: Full=Major pollen allergen Hol l 1; AltName: Full=Allergen Hol l 1.0101/1.0102; AltName: Full=Allergen Hol l I; AltName: Allergen=Hol l 1; Flags: Precursor [*Holcus lanatus*]
10	856	O04701	246	RecName: Full=Major pollen allergen Cyn d 1; AltName: Allergen=Cyn d 1 [Cynodondactylon]
11	847	Q41260	269	RecName: Full=Major pollen allergen Pha a 1; AltName: Full=Allergen Pha a I; AltName: Allergen=Pha a 1; Flags: Precursor [*Phalaris aquatica*]
12	835	Q7XCA7	275	RecName: Full=Expansin-B6; AltName: Full=Beta-expansin-6; AltName: Full=OsEXPB6; AltName: Full=OsaEXPb1.8; Flags: Precursor [*Oryza sativa* Japonica Group]
13	822	P14946	263	RecName: Full=Pollen allergen Lol p 1; AltName: Full=Allergen Lol p I; AltName: Full=Allergen R7; AltName: Allergen=Lol p 1; Flags: Precursor [*Lolium perenne*]

Table 2. Search results in the SDAP allergens database for allergen Zea m 1

No	Allergen	Sequence Link in SwissProt/ NCBI/PIR	Sequence Length	bit score	E score
1	Zea m 1	P58738	269	373.0	4.9e-105
2	Ory s 1	AAF72990	269	291.2	2.0e-80
3	Ory s 1	AAF72983	267	247.9	2.3e-67
4	Ory s 1	AAF72991	267	244.8	1.9e-66
5	Phl p 1	P43213	263	229.6	7.0e-62
6	Hol l 1	P43216	265	227.5	3.0e-61
7	Cyn d 1	AAL14078	262	224.4	2.5e-60
8	Pha a 1	Q41260	269	223.9	3.8e-60
9	Cyn d 1	O04701	246	223.4	4.9e-60
10	Hol l 1	CAA10140	263	223.5	4.9e-60
11	Ory s 1	AAF72987	275	223.1	6.5e-60
12	Poa p a	CAA10520	263	221.4	2.1e-59
13	Lol p 1	P14946	263	218.1	2.0e-58

Table S2. Results from IEDB for Phl p 1 for searching for B cell epitope for human- total 37 epitopes, 36 continuous epitopes and one non- continuous epitope

- Result for discontinuous epitope

Details	Epitope	Antigen	Organism
190361	K176, N179, D223	Phl p 1	*Phleum pratense* (timothy grass)

- Result for continuous epitope

Details	Epitope	Antigen	Organism
3912	APYHFDLSGHAFGAM	Phl p 1	*Phleum pratense* (timothy grass)
24100	HITDDNEEPIAPYHFDLSGHA	Phl p 1	*Phleum pratense* (timothy grass)
43602	NEEPIAPYHFDLSGHAFG	Phl p 1	*Phleum pratense* (timothy grass)
6223	CFEIKCTKPEACSGEPVVV	Phl p 1	*Phleum pratense* (timothy grass)
7168	CTKPEACSGEPVVVHITDDNEEPI APYHFDLSGH	Phl p 1	*Phleum pratense* (timothy grass)
9446	DNEEPIAPYHF	Phl p 1	*Phleum pratense* (timothy grass)
13661	EPIAPYHFDLSGH	Phl p 1	*Phleum pratense* (timothy grass)
13856	EQKLRSAGELELQFRRVKC	Phl p 1	*Phleum pratense* (timothy grass)
19398	GEPVVVHITDDNEEPIAPYHFDLS GHAFGAMAKKG	Phl p 1	*Phleum pratense* (timothy grass)
21535	GNTPIFKSGRGCGSCFEIKCTKPE ACSGEPVVVHITDDNEEPIAPYHF DL	Phl p 1	*Phleum pratense* (timothy grass)
21623	GPFTVRYTTEGGTKTE	Phl p 1	*Phleum pratense* (timothy grass)
22733	GTKTEAEDVIPEGWKADTSYESK	Phl p 1	*Phleum pratense* (timothy grass)
23380	GYKDVDKPPFSGMTGCGNTPIFK SGRGCGSCFEIKCTKPEACS	Phl p 1	*Phleum pratense* (timothy grass)
24099	HITDDNEEPIAPYHF	Phl p 1	*Phleum pratense* (timothy grass)
24101	HITDDNEEPIAPYHFDLSGHAFGA	Phl p 1	*Phleum pratense* (timothy grass)
25033	HVEKGSNPNYLALLVKYVNGDG DVVAV	Phl p 1	*Phleum pratense* (timothy grass)
27910	IPKVPPGPNITA	Phl p 1	*Phleum pratense* (timothy grass)
30091	KCTKPEACSGEPVVVHITDDNEEP IAPYHFDLS	Phl p 1	*Phleum pratense* (timothy grass)
32818	KPPFSGMTGCGNT	Phl p 1	*Phleum pratense* (timothy grass)
33641	KTEAEDVIPEGWKADTSYESK	Phl p 1	*Phleum pratense* (timothy grass)
39928	LTGPFTVRYTTEGGTKTEAEDVIP EGWKADTSYESK	Phl p 1	*Phleum pratense* (timothy grass)
47868	PIAPYHFD	Phl p 1	Phleum pratense (timothy grass)
47869	PIAPYHFDLSGHAFG	Phl p 1	*Phleum pratense* (timothy grass)
47897	PIFKSGRGCGSCFEI	Phl p 1	*Phleum pratense* (timothy grass)
50011	PVVVHITDDNE	Phl p 1	*Phleum pratense* (timothy grass)
50012	PVVVHITDDNEEPIAPYHFDLSGH AFG	Phl p 1	*Phleum pratense* (timothy grass)
56667	RYTTEGGTKTEAE	Phl p 1	*Phleum pratense* (timothy grass)
56668	RYTTEGGTKTEAEDVIPEGWKAD TSYESK	Phl p 1	*Phleum pratense* (timothy grass)
63135	TDDNEEPIAPYHFDLSG	Phl p 1	*Phleum pratense* (timothy grass)
63136	TDDNEEPIAPYHFDLSGH	Phl p 1	*Phleum pratense* (timothy grass)
63137	TDDNEEPIAPYHFDLSGHAFGAM A	Phl p 1	*Phleum pratense* (timothy grass)
63287	TEAEDVIPEGWKADTSYESK	Phl p 1	*Phleum pratense* (timothy grass)
64631	TKPEACSGEPVVVHITDDNEEPIA PYHFDLSGHAFGA	Phl p 1	*Phleum pratense* (timothy grass)
66583	TTEGGTKTEAEDV	Phl p 1	*Phleum pratense* (timothy grass)
70861	VRYTTEGGTKTEAEDVIPEGWKA DTSYESK	Phl p 1	*Phleum pratense* (timothy grass)
114414	IPKVPPGPNITA + HYL (P5, P8)	Phl p 1	*Phleum pratense* (timothy grass)

Table 3. Results of Paratome server for human IgE

CHAIN NAME	CDR REGIONS	AMINO ACID SEQUENCE	NUMBER OF AMINO ACIDS
4J4P_H (heavy chain)	ABR H1:	DSVSSNSAAWN	(46,47,48,49,50,51,52,53,54,55,56)
	ABR H2:	WLGRTYYRSKWYNDY	(68,69,70,71,72,73,74,75,76,77,78,79,80,81,82)
	ABR H3:	RDGEISYDYYYYGMDV	(120,121,122,123,124,125,126,127,128,129,130,131,132, 133,134,135)
4J4P_L (light chain)	ABR L1:	SSNIGNNGVN	(45,46,47,48,49,50,51,52,53,54)
	ABR L2:	LLIYYDDLLPS	(66,67,68,69,70,71,72,73,74,75,76)
	ABR L3:	EAWDDSLDGV	(109,110,111,112,113,114,115,116,117,118)

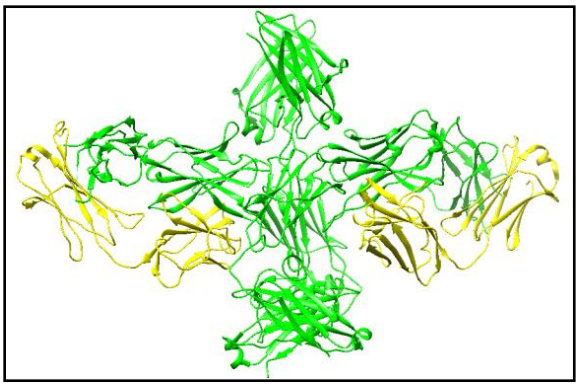


Figure 1. Structure of human Ig E

Molecular docking results for Phl p 1 and Zea m 1 using ClusPro algorithm

Antigen-Antibody docking result of Phl p 1 and IgE

ClusPro is an automated, fast rigid-body docking and discrimination algorithm. Binding of Phl p 1 with chain A and B with Ig E for cluster 0 using ClusPro algorithm [39] is shown in Figure 2.

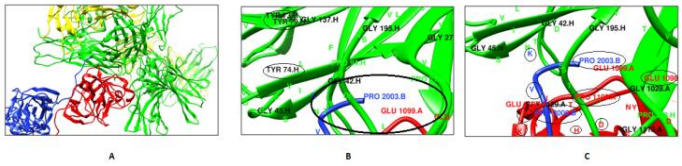


Figure 2. Heavy chain in green and light chain in yellow colour are shown for IgE. A chain and B chain of Phl p 1 are shown in red and blue colour respectively. A shows Phl p 1 and IgE interaction, B shows two epitopes PKVPPGPNIT and EPIAPYHFDLSG are interacting with paratope (containing Tyr 73, Tyr 74, Tyr 79 of Ig E (encircled in black), C shows interacting amino acids of epitope PKVPPGPNIT and EPIAPYHFDLSG

From the docking results from ClusPro algorithm [39], cluster 0 with lowest energy Weighted Score -323.7containing 71 members, is selected as model 0, as the most stable structure for antigen Phl p 1 and Ig E antibody interaction.

Using UCSB Chimera software [45] the most stable structure is visualized (Figure 2) and two epitopes PKVPPGPNIT and EPIAPYHFDLSG are identified among 36 continuous epitopes, which were obtained from IEDB database [9, 10]. As shown in Figure 3C, Pro 3, Lys 4, Val 5, Pro 6 of epitope PKVPPGPNIT and Glu 99, Pro 100, Ile 101, Ala 102, Pro 103, His 105 of epitope EPIAPYHFDLSG, interact with Ig E molecule. This docking result coincides with earlier experimental results from IEDB database [9,10]. Along with other non-covalent bonds, H bonds have been formed between following atoms of paratopes and epitopes, shown in Table 4.

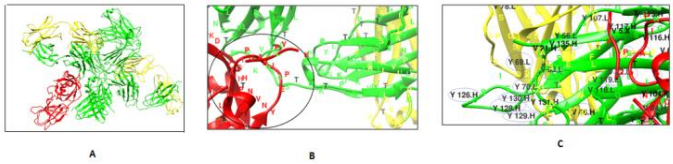


Figure 3. A Heavy chain in green and light chain in yellow colour are shown for IgE. X chain of Zea m 1 is shown in red colour. In B Zea m 1 and IgE interaction shows two epitopes PKVPPGPNIT and EPIAPYHFDLSG are interacting with paratope of Ig E (encircled in black), C shows interacting amino acids of paratope and epitopes

Table 4. Hydrogen boding between Phl p 1 and immunoglobulin E

Atoms of heavy chain	Atoms of Phl p 1	Bond length in angstrom
Arg 138.H HH21	Glu 98.AO	2.146
Asn 232.H HD21	Glu 98.A OE1	2.061
Arg 138.H HE	Glu 99.A OE1	2.528
Gln 25.H H	Glu 99.A OE2	2.526
Gln 24.H HE21	Pro 100.A O	2.108
Ser 44.H HG	Ile 101.A O	1.899

Antigen-antibody docking result of Zea m 1 and IgE

Similar docking method is followed using ClusPro algorithm [39] to identify cluster 0 with lowest energy Weighted Score -296.3 containing 43 members as model 0. This model is selected as the most stable structure for antigen Zea m 1 and Ig E antibody interaction. This structure is shown in Figure 3 using UCSB Chimera

software [45]. Pro 3, Lys 4, Val 5, Pro 6 of epitope PKVPPGPNIT and Glu 99, Pro 100, Ile 101, Ala 102, Pro 103, His 105 of epitope EPIAPYHFDLSG, interact with Val 48, Tyr 73, Tyr 126, Tyr 128, Tyr129, Tyr 130, Tyr 131, Val 135 for heavy chain and Val 53, Tyr 56, Tyr 69, Tyr 70, Val 118 for light chain of Ig E molecule as shown in Figure 3C. During antigen-antibody interaction, different non-covalent interactions e.g. electrostatic interaction, van der Waals interaction and H bonding play important roles. Following H bonds are formed between immunoglobulin E and Zea m 1 in paratope and epitope regions, as shown in **Table 5** with their bond lengths.
Thus, from molecular docking studies for both Phl p 1 and Zea m 1, two B cell epitopes PKVPPGPNIT and EPIAPYHFDLSG are identified.

Table 5. Hydrogen bonding between Zea m 1 and immunoglobulin E

Atoms of heavy chain	Atoms of Zea m 1	Bond length in angstrom
Arg 98.H HE	Tyr 104.X O	6.021
Asn 98.H HH 22	Tyr 104.X O	6.049

Homology mapping of linear epitopes

The amino acid sequences (26-35) and (124-133) of these two common epitopes in Zea m 1 and Phl p 1are shown using NCBI Multiple Alignment Viewer [54] in Figure 4. The positions of amino acids present in paratope of Ig E, which are obtained, using proABC predictor method [37], are also confirmed from the results of docking studies.

Figure 4. Amino acid sequences of two common two B cell epitopes PKVPPGPNIT and EPIAPYHFDLSG in Zea m 1 and Phl p 1 are shown in red boxes

Identification of potentially cross-reactive allergens with conserved epitopes

The search results for cross-reactivity among SDAP allergens with PD score 0 for epitopes EPIAPYHFDLSG and PKVPPGPNIT shows that for thirteen SDAP allergens (excluding isoallergenic proteins) the positions of amino acids of those epitopes are well conserved as shown in Figure 5. The detailed results for cross reactivity for two epitopes are shown in Table S3.

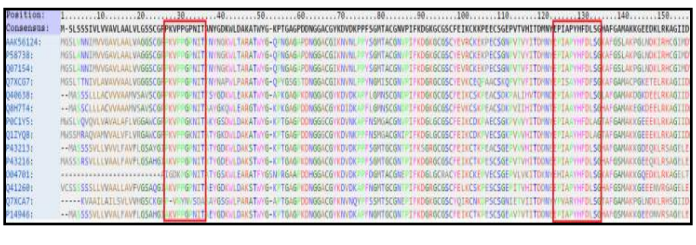


Figure 5. Positions of amino acids of epitopes PKVPPGPNIT and EPIAPYHFDLSG in thirteen SDAP allergens in red boxes.

Generation of models for pollen allergens containing conserved epitopes

Modelled structures for P0C1Y5 (ZmEXPB11), Q1ZYQ8 (ZmEXPB10) from Zea maize, P43216 (Holl 1) from velvet grass, O04701 (Cyn d 1) from bermuda grass, Q4160 (Pha a 1) from canary grass, Q7XCA7 (OsEXPB6) and (OsEXPB10) from rice and P14946 (Lol p 1) from perennial ryegrass have been generated with QMEAN scores as shown in Figure 6 and Table 6.

Table 6. Sequence identity and QMEN scores of modelled pollen allergens

Name of allergens	Template	Sequence identity	QMEAN scores
ZmEXPB11	1n10.2.A	73.42	-3.16
ZmEXPB10	1n10.2.A	73.84	-3.30
Hol l 1	1n10.2.A	92.92	-1.93
Cyn d 1	1n10.2.A	69.79	-2.67
Pha a 1	1n10.2.A	90.00	-2.32
OsEXPB6	1n10.1.A	71.13	-2.97
Lol p 1	1n10.2.A	89.58	-2.29

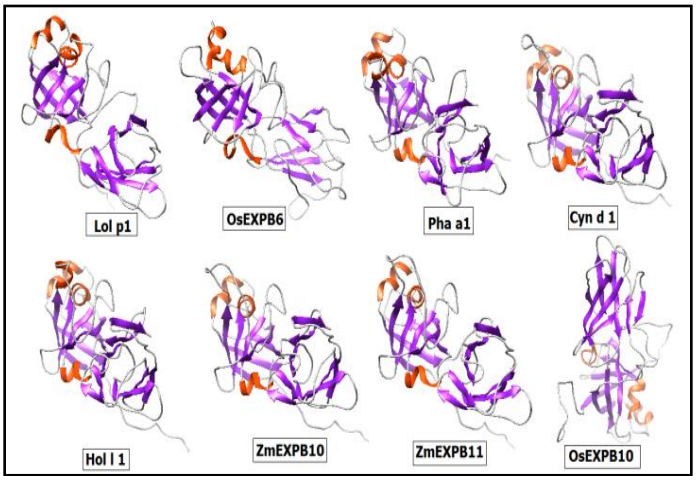


Figure 6. Modelled structures for different allergens containing epitopes PKVPPGPNIT and EPIAPYHFDLSG

Using these modelled structures, the molecular docking studies with human immunoglobin E molecules using two Fast-Fourier Transform (FFT) based docking methods; DOT [41] and ZDOCK [42] are performed. The results show that above mentioned all allergens bind with IgE with lowest energies -375.7, -345.1, -363.4, -372.0, -381.5, -351.1 and -378.7 respectively. More over for those seven allergens, the presence of two B cell epitopes EPIAPYHFDLSG and PKVPPGPNIT in binding structures with IgE, has been identified (images are shown in Table 7 along with lowest energy of bound structures).

Table S3. Search result for cross-reactivity among SDAP allergens with PD score 0 for epitope EPIAPYHFDLSG

Search Results: Peptide Similarity

Query peptide: EPIAPYHFDLSG

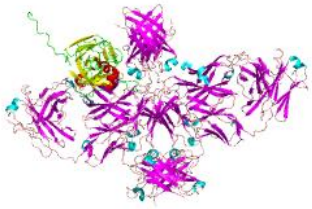
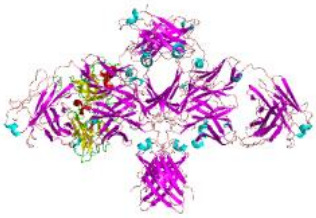
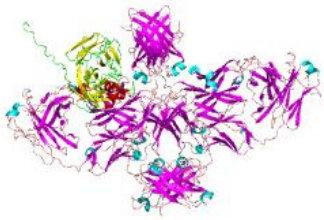
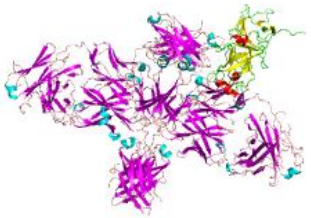
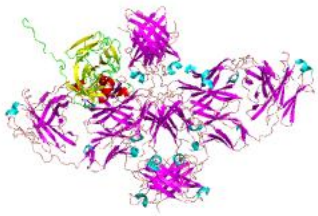
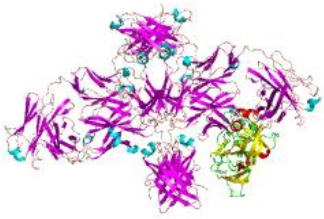
No	Allergen	Link to NCBI/ PIR/SwissProt	*PD* Sequence Similarity Index	*z* (*PD*, min)	*z*(*PD*, all)	Start Residue	Matching region	End Residue
1	Phl p 1	P43213	0.00	6.1277	8.7321	121	EPIAPYHFDLSG	132
2	Lol p 1	P14946	0.00	6.1277	8.7321	121	EPIAPYHFDLSG	132
3	Hol l 1.0102	1167836	0.00	6.1277	8.7321	106	EPIAPYHFDLSG	117
4	Lol p 1.0101	168316	0.00	6.1277	8.7321	121	EPIAPYHFDLSG	132
5	Pas n 1.0101	168419914	0.00	6.1277	8.7321	121	EPIAPYHFDLSG	132
6	Pha a 1	Q41260	0.00	6.1277	8.7321	127	EPIAPYHFDLSG	138
7	Hol l 1	P43216	0.00	6.1277	8.7321	123	EPIAPYHFDLSG	134
8	Lol p 1.0102	168314	0.00	6.1277	8.7321	110	EPIAPYHFDLSG	121
9	Zea m 1	P58738	0.00	6.1277	8.7321	123	EPIAPYHFDLSG	134
10	Zea m 1	Q07154	0.00	6.1277	8.7321	45	EPIAPYHFDLSG	56
11	Lol p 1.0103	6599300	0.00	6.1277	8.7321	121	EPIAPYHFDLSG	132

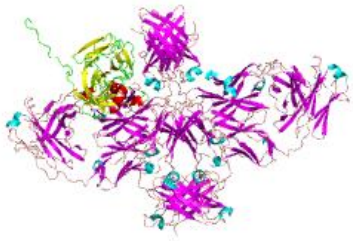
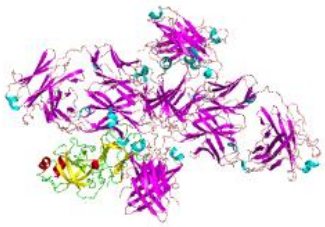
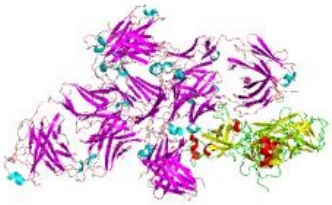
Search Results: Peptide Similarity

Query peptide: PKVPPGPNIT

No	Allergen	Link to NCBI/ PIR/SwissProt	*PD* Sequence Similarity Index	*z*(*PD*,min)	*z*(*PD*,all)	Start Residue	Matching region	End Residue
1	Zea m 1	P58738	0.00	6.3304	7.8804	27	PKVPPGPNIT	36
2	Phl p 1.0101	3901094	0.00	6.3304	7.8804	25	PKVPPGPNIT	34
3	Ory s 1	Q40638	0.00	6.3304	7.8804	25	PKVPPGPNIT	34
4	Phl p 1	P43213	0.00	6.3304	7.8804	25	PKVPPGPNIT	34
5	Pas n 1.0101	168419914	0.00	6.3304	7.8804	26	PKVPPGPNIT	35
6	Ory s 1	8118439	0.00	6.3304	7.8804	25	PKVPPGPNIT	34
7	Ory s 1	8118421	0.00	6.3304	7.8804	25	PKVPPGPNIT	34
8	Dac g 1.0101	Q7M1X8	0.00	6.3304	7.8804	2	PKVPPGPNIT	11
9	Ory s 1	8118437	0.00	6.3304	7.8804	27	PKVPPGPNIT	36
10	Hol l 1	3860384	2.26	5.1301	6.9278	25	AKVPPGPNIT	34
11	Lol p 1.0102	168314	2.26	5.1301	6.9278	14	AKVPPGPNIT	23
12	Pha a 1	Q41260	2.26	5.1301	6.9278	31	AKVPPGPNIT	40
13	Lol p 1.0101	168316	2.26	5.1301	6.9278	25	AKVPPGPNIT	34
14	Lol p 1.0103	6599300	2.26	5.1301	6.9278	25	AKVPPGPNIT	34
15	Hol l 1	P43216	2.26	5.1301	6.9278	27	AKVPPGPNIT	36
16	Hol l 1.0102	1167836	2.26	5.1301	6.9278	10	AKVPPGPNIT	19
17	Lol p 1	P14946	2.26	5.1301	6.9278	25	AKVPPGPNIT	34
18	Ant o 1.0101	Q7M1X6	2.26	5.1301	6.9278	2	AKVPPGPNIT	11
19	Poa p a	4090265	2.26	5.1301	6.9278	25	AKVPPGPNIT	34

Table 7. Bound structures of pollen allergens with human IgE

Name of allergen	Cluster no. with lowest energy	Bound structures
ZmEXPB11	Cluster 0, -346.8	
ZmEXPB10	Cluster 7, -350.0	
Hol l 1	Cluster 0,-455.9	
Cyn d 1	Cluster 0,-360.0	
Pha a 1	Cluster 0, -450.3	
OsEXPB6	Cluster 1 , -385.8	

Lol p 1	**Cluster 0, -452.8**	
Zea m 1	**Cluster 2,-316.0**	
Phl p 1	**Cluster 0, -356.4**	

Conclusions

Ball *et al.*, 1999 identify five continuous B cell epitopes of Phl p 1 by gene fragmentation techniques. They use one of them for immunotherapy in grass pollen allergic patients. This fragment is prepared to use as allergy vaccine for those patients. We are using this concept for proposing design of vaccine against Zea m 1 allergen, when the primary sequence of that allergen is the only available data. From our present work, it can be concluded that two B cell epitopes EPIAPYHFDLSG and PKVPPGPNIT of Zea m 1 are responsible for causing an allergic reaction in human, promoting Type I hypersensitivity reaction.First epitope is present on a double-psi beta barrel domain of this expansin protein and second one present on the same domain just after signal peptide region of that protein. Domain analysis of expansin proteins and pollen allergens reveals that N-terminal domain of these proteins that has a Barwin-like double psi beta-barrel structure (DPBB) (IPR007112). Homology modelling study of these major pollen allergens from maize, rice, timothy grass, canary grass, rye grass, velvet grass and Bermuda grass has been performed in this work. This study confirms that, not only the positions of these epitopes are same in their primary protein structures of these allergens, but also these well conserved linear B cell epitopes for human immunoglobulin E are present on the surface-exposed part of these allergens with minor variation in amino acid sequence. Our present in-silico analysis shows that all these pollen allergens (Phl p 1, Lol p 1,Pha a 1, Hol l 1, Ory s 1(OsEXPB9),Ory s 1(OsEXB1),Ory s 1(OsEXB10),Ory s 1(OsEXB6),Cyn d 1, Zea m 1 (ZmEXB1), Zea m 1 (ZmEXB9), Zea m 1 (ZmEXB11), Zea m 1 (ZmEXB10)) can bind spontaneously with IgE molecule of human. All these above-mentioned B cell epitopes EPIAPYHFDLSG and PKVPPGPNIT, which are present in thirteen pollen allergens can cause cross-reactivity during pollinosis. This prediction is helpful for avoiding other pollen allergens for patients having allergy for a specific pollen allergen. These two predicted epitopes could be used for designing epitope based vaccines or immunotherapy after considering the general therapeutic properties of epitopes like half-life, cytotoxicity and immune toxicity.

Acknowledgement for Funding: No supporting source of funding agency is involved in this work.

References

1. Cosgrove DJ: Loosening of plant cell walls by expansins. Nature 2000; 407(21):321–326.
2. Basu A, Sarkar A: Emphasizing the role of proteins in construction of the developmental genetic Toolkit in plants. Algorithmic and Artificial Intelligence Methods for Protein Bioinformatics. First Edition. Edited by Yi Pan, Jianxin Wang, Min Li.2014 John Wiley & Sons, Inc. Chapter 1.
3. Andersson K, Lidholm J: Characteristics and immunobiology of grass pollen allergens. Int Arch Allergy Immunol. 2003; 130(2):87-107.
4. Ball T, Edstrom W, Mauch L, Schmitt J, Leistler B, Fiebig H, Sperr WR, Hauswirth AW, Valent P, Kraft D, Almo SC, ValentaR, Gain of structure and IgE epitopes by eukaryotic expression of the major Timothy grass pollen allergen, Phl p 1.FEBS J. 2005; 272(1):217-27.

5. Knox RB, Suphioglu C: Pollen allergens: development and function. Sex Plant Reprod 1996; 9: 318–323.
6. Kumar, V., Damodharan, S., Pandaranayaka, E. P., Madathiparambil, M. G., & Tennyson, J: Molecular modelling and in-silico engineering of Cardamom mosaic virus coat protein for the presentation of immunogenic epitopes of Leptospira LipL32. Journal of Biomolecular Structure and Dynamics 2015 (ahead-of-print), 1-15.
7. Radauer C, Breiteneder H: Pollen allergens are restricted to few protein families and show distinct patterns of species distribution. J Allergy Clin Immunol. 2006;117(1):141-7.
8. (http://www.allergome.org) Mari A, Scala E, Palazzo P, Ridolfi S, Zennaro D, Carabella G: Bioinformatics applied to allergy: Allergen databases, from collecting sequence information to data integration. The Allergome platform as a model. Cell Immunol 2007; 244:97-100.
9. www.iedb.org
10. Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, Wheeler DK, Gabbard JL, Hix D, Sette A, Peters B: The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 2014 Oct 9. pii: gku938. [Epub ahead of print] PubMed PMID: 25300482.
11. Soheila J. M, Suzanne S. T, Hsiaopo C, Deliang C, Sarah S. C, Sanbao R: Computationally predicted IgE epitopes of walnut allergens contribute to cross-reactivity with peanuts. Allergy 2011; 66(12): 1522–1529.
12. Oldenburg M, Petersen A, Baur X: Maize pollen is an important allergen in occupationally exposed workers. Journal of Occupational Medicine and Toxicology 2011, 6:32.
13. Negroni L, Bernard H, Clement G, Chatel JM, Brune P, Frobert Y, Wal JM, Grassi J: Two-site enzyme immunometric assays for determination of native and denatured b-lactoglobulin. J Immunol Methods 1998; 220:25–37.
14. Selo I, Clement G, Bernard H, Chatel J, Creminon C, Peltre G, Wal J: Allergy to bovine b-lactoglobulin: specificity of human IgE to tryptic peptides. Clin Exp Allergy 1999; 29:1055–1063.
15. Clement G, Boquet D, Frobert Y, Bernard H, Negroni L, Chatel JM, Adel-Patient K, Creminon C, Wal JM, Grassi J: Epitopic characterization of native bovine b-lactoglobulin. J Immunol Methods 2002; 266:67–78.
16. Wiesmuller KH, Fleckenstein B, Jung G: Peptide vaccines and peptide libraries. Biol Chem 2001; 382:571–579.
17. Zauner W, Lingnau K, Mattner F, Von Gabain A, Buschle M: Defined synthetic vaccines. Biol Chem 2001; 382:581–595.
18. Van Regenmortel MH: Pitfalls of reductionism in the design of peptide-cased vaccines. Vaccine 2001; 19:2369–2374.
19. Pellequer JL, Westhof E. PREDITOP: A program for antigenicity prediction. J Mol Graphics 1993; 11:204–210.
20. Alix AJ: Predictive estimation of protein linear epitopes by using the program PEOPLE. Vaccine 1999; 18:311–314.
21. Odorico M, Pellequer JL: BEPITOPE: predicting the location of continuous epitope and patterns in proteins. J Mol Recognit 2003; 16: 20–22.
22. Saha, S., & Raghava, G. P. S: Prediction of continuous B- cell epitopes in an antigen using recurrent neural network. Proteins: Structure, Function, and Bioinformatics 2006; 65(1): 40-48.
23. Benkert P, Biasini M, Schwede T: Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 2011; 27:343–350.
24. Power, T. D., Ivanciuc, O., Schein, C. H., & Braun, W: Assessment of 3D models for allergen research. Proteins: Structure, Function, and Bioinformatics 2013; 81(4): 545-554.
25. Schein, C. H., Ivanciuc, O., Midoro-Horiuti, T., Goldblum, R. M., & Braun, W: An allergen portrait gallery: representative structures and an overview of IgE binding surfaces. Bioinformatics and biology insights 2010; 4: 113.
26. Weber, R. W: Guidelines for using pollen cross-reactivity in formulating allergen immunotherapy. Journal of Allergy and Clinical Immunology 2008; 122(1): 219-221.
27. Weber, R. W: Cross-reactivity of pollen allergens: impact on allergen immunotherapy. Annals of Allergy, Asthma & Immunology 2007; 99(3), 203-212.
28. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E: The Protein Data Bank, Nucleic Acids Research 2000; 28: 235-242.
29. Yennawar, N.H., Li, L.C., Dudzinski, D.M., Tabuchi, A., Cosgrove, D.J., Crystal structure and activities of EXPB1 (Zea m 1), a beta-expansin and group-1 pollen allergen from maize. Proc. Natl. Acad. Sci. Usa 2006; 103: 14664-14671.
30. Database resources of the National Center for Biotechnology Information. Nucleic acids research, 2003; 28-33.
31. Ivanciuc, O., Schein. C. H., and Braun, W: SDAP: Database and Computational Tools for Allergenic Proteins. Nucleic Acids Res. 2003; 31(1):359-362.
32. Ivanciuc, O., Schein. C. H., and Braun, W: Data Mining of Sequences and 3D Structures of Allergenic Proteins. Bioinformatics 2002, 18(10): 1358-1364.
33. Ivanciuc O, Midoro-Horiuti T, Schein CH, Xie L, Hillman GR, Goldblum RM, Braun W: The property distance index PD predicts peptides that cross-react with IgE antibodies. Mol Immunol. 2009; 46(5):873-83.
34. Holm L, Rosenström P: Dali server: conservation mapping in 3D. Nucl. Acids Res. 2010; 38: W545-549.
35. Fedorov, A.A., Ball, T., Leistler, B., Valenta, R., Almo, S.C. X-ray Crystal Structure of Phl p 1, a Major Timothy Grass Pollen Allergen (To be Published).
36. Drinkwater, N., Cossins, B.P., Keeble, A.H., Wright, M., Cain, K., Hailu, H., Oxbrow, A., Delgado, J., Shuttleworth, L.K., Kao, M.W., McDonnell, J.M., Beavil, A. J., Henry, A.J., Sutton, B.J.: Human immunoglobulin E flexes between acutely bent and extended conf ormations. Nat. Struct. Mol. Biol. 2014; 21: 397-404.
37. Kunik V, Peters B, Ofran Y: Structural Consensus among Antibodies Defines the Antigen Binding Site. PLoS Comput Biol 2012; 8(2): e1002388. doi: 10.1371/journal.pcbi.1002388.
38. Kunik V, Ashkenazi S, Ofran Y. Paratome: An online tool for systematic identification of antigen binding regions in antibodies based on sequence or structure. Nucleic Acids Res. 2012; 40(Web Server issue): W521-4.
39. Brenke R, Hall DR, Chuang G-Y, Comeau SR, Bohnuud T, Beglov D, Schueler-Furman O, Vajda S, Kozakov D: Application of asymmetric statistical potentials to antibody-protein docking. Bioinformatics 2012; 28(20): 2608-2614.
40. Kozakov D, Brenke R, Comeau SR, Vajda S. PIPER: An FFT-based protein docking program with pair wise potentials. Proteins. 2006 Aug 24
41. Mandell, J.G., Roberts, V.A., Pique, M.E., Kotlovyi, V., Mitchell, J.C., Nelson, E., Tsigelny, I. and Ten Eyck, L.F: Protein docking using continuum electrostatics and geometric fit. Protein Eng. 2001; 14: 105–113
42. Chen, R., Li, L. and Weng, Z.: ZDOCK: an initial-stage protein docking algorithm. Proteins 2003; 52: 82–87.
43. Lo Conte, L. *et al.:* The atomic structure of protein-protein recognition sites. J. Mol. Biol.1999; 285: 2177–2198.
44. Chuang, G. *et al.:* DARS (Decoys as the Reference State) potentials for protein–protein docking. Biophys. J. 2008; 95: 4217–4227.
45. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE: UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004; 25(13):1605-12.
46. http://www.allergen.org.
47. http://www.allerdata.com/
48. http://www.meduniwien.ac.at/allergens/allfam/
49. Arnold, K., Bordoli, L., Kopp, J. and Schwede, T: The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 2006; 22, 195-201.
50. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997; 25: 3389-3402.
51. Remmert, M., Biegert, A., Hauser, A. and Soding, J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods 2012; 9: 173-175.
52. Guex, N. and Peitsch, M.C: SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modelling. Electrophoresis 1997; 18: 2714-2723.
53. Benkert, P., Biasini, M. and Schwede, T: Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 2011; 27: 343-350.
54. www.ncbi.nlm.nih.gov
55. Ball, T; Fuchs, T; Sperr, WR; Valent, P; Vangelista, L; Kraft, D; Valenta, R: B cell epitopes of the major timothy grass pollen allergen, phl p 1, revealed by gene fragmentation as candidates for immunotherapy. FASEB J. 1999; 13(11):1277-1290.