CONFORMATIONAL EPITOPE PREDICTION OF BIRCH BETV 1 AND HAZEL COR A1 TOWARDS B-CELLS

ABSTRACT

Background: White birch and hazel allergens, namely Bet v1 and Cor a1 are known allergens, but their allergen specificity is not yet characterized. **Objective:** To map the antigenic determinants responsible for IgE binding utilizing *in silico* modelling and docking of the peptides against IgE antibody. **Methods:** The antigen sequences were cut into peptides are docked against the IgE antibody and those with the highest docking scores are further studied for the bond interactions. The overlapping sequences of the high score peptides are observed in the whole antigen model to predict their position. The residues at bond interactions also been reported for these overlapping peptide sequences. **Results:** The validation is done by antigen-antibody docking studies to confirm the predicted epitope. 25% of the world population suffers from allergic rhinitis and 15% of them develop asthma. **Conclusion:** Negative binding energies of the studied pollen allergens with IgE confirm their allergenicity. Based on the results of overlapping peptides PF 3,4 and PF 16,17 to play a key role in the allergenic response of white birch and Common hazel.

Keywords: Bet v1, Cor a1, Birch, Hazel, IgE, B-cell

INTRODUCTION

In the adaptive immune system, B cells play an essential role in protecting the human body against various pathogenic molecules. Specifically, B cells belong to humoral immunity that is mediated through antibodies. In response to exposure to pathogens, B cells develop antibodies that bind to and neutralize the target. However, pathogens are not identified by B-cells as a whole, but through molecular components known as antigens. The component of an antigen that is detected by the immune system, primarily by B cells or T cells, is the epitope, also known as the antigenic determinant. The majority of the B cell epitopes are conformational (discontinuous), while the remaining (merely 10 percent) are linear B-cell epitopes (continuous**)**. For immunologists, accurate detection of conformational B-cell epitopes is still a major hurdle **[1]**. The major plant allergens of white birch (*Betula verrucosa*) and common hazel (*Corylus avellana*) are the antigens chosen for this analysis. Allergic hypersensitivity or allergy is a reaction that happens in an individual when the same allergen is introduced into that person who has developed IgE antibodies in response to that antigen or allergen priorly **[2]**. The most common causative agents of allergic reactions are

pollen grains from plants, foods, bee stings, dust mites, molds, fungal spores, animal epithelia, fur and feathers, animal dander, latex. Studies have indicated that allergic disorders such as anaphylaxis, hay fever, atopic dermatitis, eczema, asthma, and many other respiratory and pulmonary diseases affected about 25% of the world population which are primarily caused by aero-allergens. The major hazel allergen Cor a1 and major birch pollen allergen Bet v1 are homologous to each other **[3]**. According to researchers, about 53% of people who are allergic to birch pollen also have cross-reactivity to Cor a1 as well. People who are allergic to birch pollen (70%) also exhibit hypersensitive responses towards different seeds, fruits, nuts, and roots than those without allergy to birch pollen (19%). Hazel allergy is very widespread across European countries and has also been shown to be the most prevalent source of food allergy mediated by IgE **[4].** The present work focused on the *in silico* molecular characterization of white birch and common hazel Bet v1 and Cor a1 allergen-derived peptides, respectively.

METHODS AND MATERIALS

Retrieval of allergen sequences and antibody structure

Bet v1 (P15494) and cor a1 (Q08407) allergens were retrieved for the FASTA sequences of the major allergens from uniprot belonging to a white birch and hazelnut. In general, immunoglobulin E (IgE) is an instinctive response shown by the human immune system to any antigen that has breached into the body. The IgE antibody structure with zero mutations and having a Fab macromolecule also was retrieved from the RSCB PDB database **[5-6]**.

Overlapping and cutting of peptides

To determine the epitope present in the allergen, the individual peptides are docked against the antibody. For this, the entire allergen sequences that were retrieved from UniProt are cut into short overlapping peptide fragments using Sigma-Aldrich tools [Peptide Library design and calculator tool]. Sigma Aldrich's "Overlapping Peptide Fragment Library" tool is used to chop the peptides with a convenient amino acid gap into the appropriate lengths and also hydropathy index. This tool of Sigma Aldrich is used to cut the amino acid sequence into short peptides of length 10 and with five overlapping amino acids **[7-8]**.

Protein-Peptide Docking

HPEPDOCK is used to conduct a docking check for the binding site attributable to a protein receptor structure and a peptide sequence, allowing the peptide to be completely flexible and

predicting the protein-peptide complex structure, beginning from random peptide conformations and locations. Computational docking techniques are used to scan for rotational space between a protein receptor and its peptide-binding partner in all possible binding modes. Antibody is docked against each overlapping peptide fragment of the pollen sequence. All potential models for each peptide fragment along with their docking scores are estimated by the HPEPDOCK docking. Peptide model with the highest docking score is identified and picked for every overlapping peptide fragment and docked against the antibody, among all potential models that have been predicted by the server **[9].**

Docking analysis

The antibody and peptide interactions are studied through Schrodinger's software, for the presence of non-covalent bonds and pi-pi interactions between antibody and antigen peptides and are analyzed through the Maestro Viewer (data not provided). The 3D structures of both the antibody and the antigen peptide model with the highest docking score according to HPEPDOCK are viewed in the Maestro workspace to identify the bond interactions between them. The type of bond that is formed between the atoms, the name & number of the atom, residues, and chains at where the bonds are formed for both antibody and peptide are noted. **[10]**

Homology Modelling

Template selection is done through the BLAST tool for the prediction of the secondary structure of the antigen sequence that is retrieved from UniProt. Areas of similarities amongst various biological sequences are detected through BLAST, which helps in the comparison of protein or nucleotide sequences against database sequences and measures their statistical significance. Due to the very high similarity between both the antigen sequences, the protein data structure (ID:4a86, Birch major allergen) with 72.33% of similarity was selected as a template for building secondary structure for Hazel major allergen (Cor a1). The model is built or generated using SWISS-MODEL **[11]**.

Antigen-Antibody Docking Studies

The complete antigen 3D models of both Birch (Bet v1) and Hazel (Cor a1) are docked against the IgE antibody for validation. The structure of the Birch allergen is retrieved from Protein Data Bank (PDB) (ID: 4A86). Whereas the structure of Hazel allergen is modelled through SWISS-MODEL. Both these structures are docked against IgE antibody whose

structure is retrieved from PDB (ID: 2vxq). Docking is carried out through the software ClusPro which is a web-based server that is useful for direct protein-protein docking. From all the models that are predicted by the ClusPro server, the models with high scores for both birch and hazel are considered for validation **[12]**.

Results:

Overlapping peptide fragments:

The birch pollen (Bet v1) and hazel pollen (Cor a1) sequences are cut into a length of 10-mers with a gap of five amino acids. Both the allergen sequences of chosen plant species are cut into 31 peptides each with the help of the Sigma Aldrich tool (Table 1). Peptide fragments with highest docking score acquired through HPEPDOCK (Table 2). The bond interactions is seen between overlapping peptide fragments of Bet v1, Cor a1 and IgE antibody (Table 3 and 4). Overlapping peptide fragments of Bet v1 and IgE antibody is shown in figure 1 and overlapping peptide fragments of Cor a1 and IgE antibody is shown in figure 2 respectively.

Birch peptide (Bet v1)			**Hazel peptide (Cor a1)**		
S.NO	**Overlapping Peptide**	**HIGH SCORE**	**S.NO**	**Overlapping Peptide**	**HIGH SCORE**
1	MGVFNYETET	-165.059	1	MGVFNYEVET	-182.768
2	YETETTSVIP	-152.866	2	YEVETPSVIP	-145.724
3	TSVIPAARLF	-182.945	3	PSVIPAARLF	-178.316
4	AARLFKAFIL	-183.929	4	AARLFKSYVL	-184.417
5	KAFILDGDNL	-142.398	5	KSYVLDGDKL	-145.843
6	DGDNLFPKVA	-142.191	6	DGDKLIPKVA	-120.536
7	FPKVAPQAIS	-168.564	7	IPKVAPQAIT	-163.24
8	PQAISSVENI	-152.31	8	PQAITSVENV	-164.734
9	SVENIEGNGG	-111.494	9	SVENVEGNGG	-124.516
10	EGNGGPGTIK	-134.82	10	EGNGGPGTIK	-134.802
11	PGTIKKISFP	-173.31	11	PGTIKNITFG	-176.457

12	KISFPEGFPF	-194.601	12	NITFGEGSRY	-177.941
13	EGFPFKYVKD	-166.444	13	EGSRYKYVKE	-172.549
14	KYVKDRVDEV	-144.435	14	KYVKERVDEV	-145.131
15	RVDEVDHTNF	-156.022	15	RVDEVDNTNF	-131.66
16	DHTNFKYNYS	-186.802	16	DNTNFTYSYT	-189.485
17	KYNYSVIEGG	-174.386	17	TYSYTVIEGD	-172.279
18	VIEGGPIGDT	-118.157	18	VIEGDVLGDK	-108.643
19	PIGDTLEKIS	-162.986	19	VLGDKLEKVC	-122.362
20	LEKISNEIKI	-131.581	20	LEKVCHELKI	-141.496
21	NEIKIVATPD	-134.962	21	HELKIVAAPG	-176.674
22	VATPDGGSIL	-126.187	22	VAAPGGGSIL	-144.764
23	GGSILKISNK	-154.886	23	GGSILKISSK	-155.694
24	KISNKYHTKG	-181.217	24	KISSKFHAKG	-182.18
25	YHTKGDHEVK	-162.461	25	FHAKGDHEIN	-155.299
26	DHEVKAEQVK	-120.969	26	DHEINAEEMK	-128.345
27	AEQVKASKEM	-124.307	27	AEEMKGAKEM	-109.097
28	ASKEMGETLL	-126.662	28	GAKEMAEKLL	-108.428
29	GETLLRAVES	-144.573	29	AEKLLRAVET	-140.016
30	RAVESYLLAH	-169.747	30	RAVETYLLAH	-185.11
31	YLLAHSDAYN	-172.705	31	YLLAHSAEYN	-193.964

Table 1: Overlapping peptide fragments of Bet v1 and Cor a1

Birch (Bet v1)				
S.NO	**Peptide**		**Peptide number**	**High score**
1	3	beta 1-alpha 1	TSVIPAARLF	-182.945
	4	alpha 1	AARLFKAFIL	-183.929

2	11	beta 6	PGTIKKISFP	-173.310
	12	beta 6	KISFPEGFPF	-194.601
3	16	beta 5-beta 4	DHTNFKYNYS	-186.802
	17	beta 4	KYNYSVIEGG	-174.386
Hazel (Cor a1)				
S.NO	**Peptide**		**Peptide number**	**High score**
1	3	beta 1-alpha 1	PSVIPAARLF	-178.316
	4	alpha 1	AARLFKSYVL	-184.417
2	16	beta 5	DNTNFTYSYT	-189.485
	17	beta 4	TYSYTVIEGD	-172.279
3	30	alpha 2	RAVETYLLAH	-185.110
	31	alpha 2	YLLAHSAEYN	-193.964

Table 2: Birch vs Hazel peptide mapping

Pep	Type of bond	Pep/Ab	Atom	Atom No.	Residue	Chain
1	Pi-pi stacking	peptide	----	----	TYR6	B
		Antibody	----	----	HID36	A
	Pi-pi stacking	peptide	----	----	TYR6	B
		Antibody	----	----	TYR59	H
	Salt bridge	peptide	N_N^{+1}	3968	MET1	B
		Antibody	O_{OE2}^{1-}	2335	GLU1	L
2	Salt bridge	peptide	O_{OXT}^{1-}	80	PRO10	B
		Antibody	N_{NZ}^{1+}	3227	LYS107	L
3	Salt bridge	peptide	O_{OXT}^{1-}	76	PHE10	B
		Antibody	N_{NZ}^{1+}	2733	LYS42	L
	Salt bridge	peptide	N_N^{1+}	1	THR1	B
		Antibody	O_{OE2}^{1-}	3670	GLU165	L
4	Pi-cation	peptide	----	----	PHE8	B
		Antibody	N_{NZ}^{1+}	2678	LYS45	L
	Salt bridge	peptide	O_{OXT}^{1-}	4049	LEU10	B
		Antibody	N_{NZ}^{1+}	2678	LYS45	L
5	Salt bridge	peptide	N_{NZ}^{1+}	9	LYS1	B
		Antibody	O_{OD2}^{1-}	3689	ASP167	L
	Salt bridge	peptide	N_N^{1+}	1	LYS1	B
		Antibody	O_{OD2}^{1-}	3689	ASP167	L
6	Pi-pi stacking	peptide	----	----	PHE6	B
		Antibody	----	----	HID38	L
	Salt bridge	Peptide	C_{CB}	5	ASP1	B
		Antibody	O_{OE2}^{1-}	3501	GLU143	L
7	----	----	----	----	----	----
8	----	----	----	----	----	----

9	Salt bridge	Peptide	O_{OXT}^{1-}	68	GLY10	B
		Antibody	N_{NZ}^{1+}	3694	LYS169	L
	Salt bridge	Peptide	O_{OE2}^{1-}	22	GLU3	B
		Antibody	N_{N}^{1+}	799	VAL1	H
10	----	----	----	----	----	----
11	----	----	----	----	----	----
12	----	----	----	----	----	----
13	----	----	----	----	----	----
14	Salt bridge	Peptide	O_{OO2}^{1-}	45	ASP5	B
		Antibody	N_{NZ}^{1+}	3714	LYS169	L
	Salt bridge	Peptide	N_{NH2}^{1+}	56	ARG6	B
		Antibody	O_{OO2}^{1-}	3722	ASP170	L
15	Salt bridge	Peptide	N_{N}^{1+}	1	ARG1	B
		Antibody	O_{OE2}^{1-}	2422	GLU1	L
	Salt bridge	peptide	O_{OE2}^{1-}	35	GLU4	B
		Antibody	N_{NZ}^{1+}	1146	LYS44	H
16	Salt bridge	peptide	N_{N}^{1+}	3968	ASP1	B
		Antibody	O_{OE2}^{1-}	3130	GLU105	L
17	Salt bridge	peptide	N_{NZ}^{1+}	9	LYS1	B
		Antibody	O_{OE2}^{1-}	3674	GLU165	L
18	Salt bridge	peptide	O_{OD2}^{1-}	59	ASP9	B
		Antibody	N_{NH2}^{1+}	3483	ARG142	L
19	Salt bridge	peptide	N_{NZ}^{1+}	60	LYS8	B
		Antibody	O_{OE2}^{1-}	3205	GLU105	L
20	----	----	----	----	----	----
21	Salt bridge	peptide	O_{OXT}^{1-}	77	ASP10	B
		Antibody	N_{NZ}^{1+}	2714	LYS39	L
22	----	----	----	----	----	----
23	Salt bridge	peptide	N_{NZ}^{1+}	39	LYS6	B
		Antibody	O_{OE2}^{1-}	1917	GLU151	L
	Salt bridge	peptide	O_{OXT}^{1-}	71	LYS10	B
		Antibody	N_{NZ}^{1+}	2302	LYS204	H
24	Salt bridge	peptide	N_{NZ}^{1+}	78	LYS9	B
		Antibody	O_{OE2}^{1-}	3677	GLU165	L
25	Salt bridge	peptide	O_{OE2}^{1-}	69	GLU8	B
		Antibody	N_{NZ}^{1+}	1145	LYS44	H
	Salt bridge	peptide	O_{OD2}^{1-}	50	ASP6	B
		Antibody	N_{NZ}^{1+}	3200	LYS103	L
26	Pi-pi stacking	peptide	----	----	HID2	B
		Antibody	----	----	TYR32	L
	Salt bridge	peptide	N_{N}^{1+}	1	ASP1	B
		Antibody	O_{OE2}^{1-}	382	GLU40	A
	Salt bridge	peptide	N_{NZ}^{1+}	43	LYS5	B
		Antibody	O_{OE2}^{1-}	301	GLU30	A
	Salt bridge	peptide	N_{NZ}^{1+}	43	LYS5	B
		Antibody	O_{OE2}^{1-}	317	GLU32	A
27	Salt bridge	peptide	N_{NZ}^{1+}	4026	LYS8	B
		Antibody	O_{OE2}^{1-}	3594	GLU165	L

	Salt bridge	peptide	O_{OE2}^{1-}	4035	GLU9	B
		Antibody	N_{NZ}^{1+}	2657	LYS42	L
28	Salt bridge	peptide	O_{OE2}^{1-}	3996	GLU4	B
		Antibody	N_{NZ}^{1+}	3439	LYS145	L
29	Salt bridge	peptide	N_{NH2}^{1+}	47	ARG6	B
		Antibody	O_{OE2}^{1-}	3669	GLU165	L
30	Pi-cation	peptide	Ring		HID10	B
		Antibody	N_{NZ}^{+1}	2739	LYS42	L
31	----	----	----	----	----	----

Table 3: The bond interactions between overlapping peptide fragments of Bet v1 and IgE antibody.

Pep	Type of bond	Pep/Ab	Atom	Atom No.	Residue	Chain
1	----	----	----	----	----	----
2	----	----	----	----	----	----
3	Salt bridge	Peptide	O_{OXT}^{1-}	76	PHE10	B
		Antibody	N_{NZ}^{1+}	3190	LYS103	L
4	Salt bridge	peptide	N_{NHZ}^{1+}	3988	ARG3	B
		Antibody	O_{OE2}^{1-}	3594	GLU165	L
5	Salt bridge	Peptide	N_{NZ}^{1+}	71	LY9	B
		Antibody	O_{OE2}^{1-}	3505	GLU143	L
	Salt bridge	peptide	O_{OD2}^{1-}	62	ASP8	B
		Antibody	N_{NZ}^{1+}	3227	LYS107	L
6	Salt bridge	Peptide	O_{OXT}^{1-}	74	ALA1O	B
		Antibody	N_{NZ}^{1+}	2752	LYS45	L
	Salt bridge	Peptide	O_{OD1}	7	ASP1	B
		Antibody	N_{NZ}^{1+}	2711	LYS39	L
7	Salt bridge	Peptide	N_{N}^{1+}	3968	ILE1	B
		Antibody	O_{OE2}^{1-}	3594	GLU165	L
8	----	----	----	----	----	----
9	Salt bridge	Peptide	O_{OE2}^{1-}	46	GLS44	B
		Antibody	N_{NZ}^{1+}	1126	LYS44	H
	Salt bridge	Peptide	N_{N}^{1+}	1	SERI	B
		Antibody	O_{OE2}^{1-}	2402	GLUI	L
10	----	----	----	----	----	----
11	Salt bridge	Peptide	N_{N}^{1+}	1	PRO1	B
		Antibody	O_{OE2}^{1-}	3204	GLU105	L
	Pi –Pi stacking	Peptide	----	----	PHE9	B
		Antibody	----	----	PHE149	H
	Salt bridge	Peptide	N_{N}^{1+}	35	LYS5	H
		Antibody	O_{OE2}^{1-}	3668	GLU165	B
12	Salt bridge	Peptide	O_{OE2}^{1-}	47	GLU6	B

		Antibody	N_{NZ}^{1+}	1140	LYS44	H
	Pi-pi stacking	Peptide	----	----	RHE4	B
		Antibody	----	----	HID38	L
13	Salt bridge	Peptide	$OOE2^{1-}$	9	GLU1	B
		Antibody	NN^{1+}	3505	ARG142	L
	Salt bridge	Peptide	NN^{1+}	1	GLU1	B
		Antibody	$OOE2^{1-}$	3219	GLU105	L
	Salt bridge	Peptide	$NNH2^{1+}$	30	ARG4	B
		Antibody	$OOE2^{1-}$	3683	GLU165	L
	Pi-pi stacking	Peptide	----	----	TYR7	B
		Antibody	----	----	PHE149	H
14	Salt bridge	Peptide	O_{OXT}^{1-}	89	VAL10	B
		Antibody	N_{NZ}^{1+}	2357	LYS209	H
	Salt bridge	Peptide	O_{E2T}^{1-}	81	GLU9	B
		Antibody	N_{NZ}^{1+}	2320	LYS204	H
	Salt bridge	Peptide	$OOD2^{1-}$	72	ASP8	B
		Antibody	N_{NZ}^{1+}	2320	LYS204	H
15	Salt bridge	Peptide	$OOE2^{1-}$	35	GLU4	B
		Antibody	N_{H2}^{1+}	2879	ARG61	L
	Salt bridge	Peptide	$OOE2^{1-}$	35	GLU4	B
		Antibody	N_{NZ}^{1+}	2722	LYS39	L
16	----	----	----	----	----	----
17	Salt bridge	Peptide	NN^{1+}	1	THR1	B
		Antibody	$OOD2^{1-}$	3692	ASP 167	L
18	Salt bridge	Peptide	$OOD2^{1-}$	4030	ASP9	B
		Antibody	$NNH2^{1+}$	2501	ARG24	L
19	Salt bridge	Peptide	$OOE2^{1-}$	53	GLU7	B
		Antibody	N_{NZ}^{1+}	2307	LYS204	H
	Salt bridge	Peptide	$OOD2^{1-}$	27	ASP4	B
		Antibody	N_{NZ}^{1+}	2733	LYS42	L
	Salt bridge	Peptide	N_{NZ}^{1+}	36	LYS5	
		Antibody	$OOE2^{1-}$	3670	GLU165	L
20	----	----	----	----	----	----
21	Salt bridge	Peptide	NN^{1+}	1	HID1	B
		Antibody	$OOe2^{1-}$	3667	GLU165	L
	Salt bridge	Peptide	$OOE2^{1-}$	19	GLU2	B
		Antibody	$NNh2^{1+}$	3489	ARG142	L
	Pi - stacking	Peptide	----	----	HID1	B
		Antibody	N_{NZ}^{1+}	3187	LYS103	L
22	Pi - cat	peptide	N_{N}^{1+}	1	VAL1	B
		Antibody	----	----	HID41	H
23	----	----	----	----	----	----
24	Salt bridge	peptide	N_{NZ}^{1+}	4005	LYS5	B
		Antibody	O_{E2}^{1-}	3594	GLU165	L
	Salt bridge	peptide	N_{NH2}^{1+}	3976	LYS1	B
		Antibody	O_{OE2}^{1-}	3611	ASP167	L
	Salt bridge	peptide	N_{NZ}^{1+}	3976	LYS44	B

		Antibody	O_{OD2}^{1-}	3634	ASP170	L
25	Salt bridge	peptide	O_{OXT}^{1-}	83	ASN10	B
		Antibody	N_{NZ}^{1+}	1142	LYS44	H
	Salt bridge	peptide	O_{OE2}^{1-}	66	GLU8	B
		Antibody	N_{NH2}^{1+}	1103	ARG39	H
	Salt bridge	peptide	O_{OE2}^{1-}	66	GLU8	B
		Antibody	N_{NH2}^{1+}	1335	ARG67	H
26	Salt bridge	peptide	N_{H}^{N}	1	Asp1	B
		Antibody	O_{OE2}^{1}	3676	Glu165	L
27	Salt bridge	peptide	O_{OE2}^{1-}	14	GLU2	B
		Antibody	N_{NZ}^{1-}	1875	LYSI46	H
	PI -cat	peptide	N_{NZ}^{1+}	58	LYS8	B
		Antibody	----	----	H1D189	L
28	----	----	----	----	----	----
29	Salt bridge	peptide	N_{NH2}^{1+}	4017	ARG6	B
		Antibody	O_{OE2}^{1-}	218	GLU3O	A
	Salt bridge	peptide	N_{N}^{1+}	3968	ALA1	B
		Antibody	O_{OE2}^{1--}	299	GLU40	A
30	Salt bridge	peptide	O_{OXT}^{1--}	4050	HID10	B
		Antibody	N_{NZ}^{1+}	2678	LYS45	L
	Salt bridge	peptide	O_{OE2}^{1--}	3999	GLU4	B
		Antibody	N_{NZ}^{1+}	2657	LYS42	L
31	Salt bridge	peptide	O_{OE2}^{1--}	63	GLU8	B
		Antibody	N_{NZ}^{1+}	2741	LYS42	L
	Salt bridge	peptide	O_{OXT}^{1--}	84	ASN10	B
		Antibody	N_{NZ}^{1+}	3710	LYS169	L

Table 4: The bond interactions between overlapping peptide fragments of Cor a1 and IgE antibody.

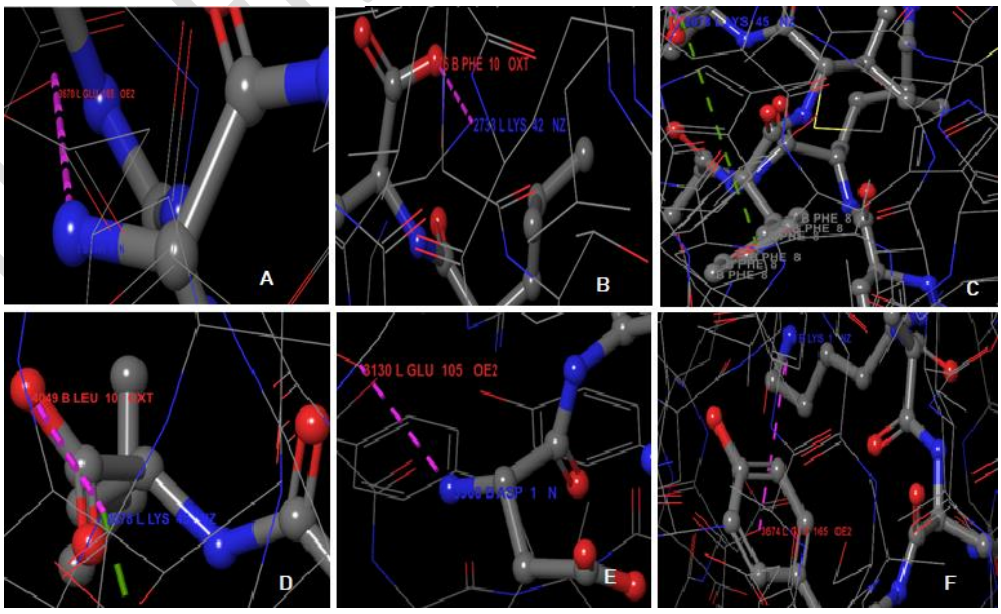

Figure 1: A) Salt Bridge in 3rd peptide of Bet v1 at residue 'T'. **B)** Salt Bridge in 3rd peptide of Bet v1 at residue 'F'. **C)** Pi-cation bond in 4th peptide of Bet v1 at residue 'F'. **D)** Salt Bridge in 4th peptide of Bet v1 at residue 'L'. **E)** Salt Bridge in 16th peptide of Bet v1 at residue 'D'. **F)** Salt Bridge in 17th peptide of Bet v1 at residue 'K'

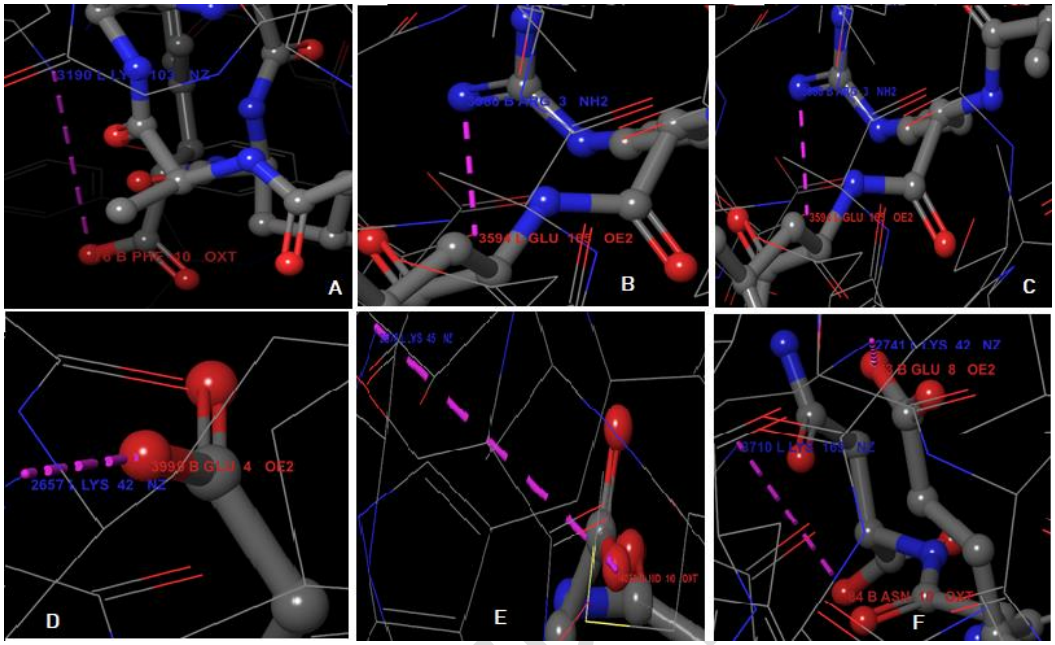

Figure 2: A) Salt Bridge in 3rd peptide of Cor a1 at residue 'P'. **B)** Salt Bridge in 4th peptide of Cor a1 at residue 'R'. **C)** Salt Bridge in 17th peptide of Cor a1 at residue 'T'. **D)** Salt Bridge in 30th peptide of Cor a1 at residue 'E'. **E)** Salt Bridge in 30th peptide of Cor a1 at residue 'H'. **F)** Salt Bridges in 31st peptide of Cor a1 at residue 'E' and 'N'

Homology modelling:

Template selection: The target sequence was searched with BLAST against the primary amino acid sequence contained in the SMTL. A total of 108 templates were found. Among them, the template with the highest sequence identity, MAJOR POLLEN ALLERGEN BET V1-A (PDB ID:4a86) is selected as a template with 72.33% similarity as shown in Table 5. Three-dimensional structures of Birch pollen (Bet v1) were available in the PDB (4A86). Hazel pollen (Cor a1) structure was determined using the homology modelling application of the SWISS-MODEL as shown in figure 3.

Template ---- 4a86.1.A			
Seq. Identity	72.33	**Resolution**	1.59Å
Oligo-state	monomer	**Seq. Similarity**	0.52

QSQE	0.00	**Range**	2 - 160
Found by	HHblits	**Coverage**	0.99
Method	X-ray	**Description**	Major pollen allergen BET V1-A

Table 5: Template with the highest sequence identity, BET V1-A

Top hit from the Blastp analysis of Hazel pollen (Cor a1) with PDB ID 4A86 as a template and energy-based model was developed. The structural alignment of the model as evaluated by Ramachandran plot indicated that most of the (93.4%) amino acids fit into the most favored regions, 5.8% of the modelled Cor a1 residues fall into the additional allowed regions and the remaining were found in generously allowed regions. The ERRAT overall quality factor was 93.617, specifying that the model predicted was good. To identify allergen-IgE interacting sites, an IgE-allergen (protein-protein) docking study was undertaken. IgE antibody (PDB ID: 2VXQ) was retrieved and prepared by using Schrödinger's protein preparation wizard. Concurrently, all the simulated trajectory frames of the modelled allergen of Cor a1 were clustered based on the energy and deviations. The cluster center frame showing minimal energy, deviations, and fluctuations was chosen for docking studies. Tail-end sequences of the allergen were found intact with the paratope region of the antibody by the end of docking studies. To validate the importance of other amino acids in the allergen, the sequence was divided into overlapping peptides. The allergen sequences of Bet v1, as well as Cor a1, were processed using overlapping peptide fragment library software, and 31 different 10-mer peptides, were designed (Bet v1-31 and Cor a1-31) with an overlap of five amino acids.

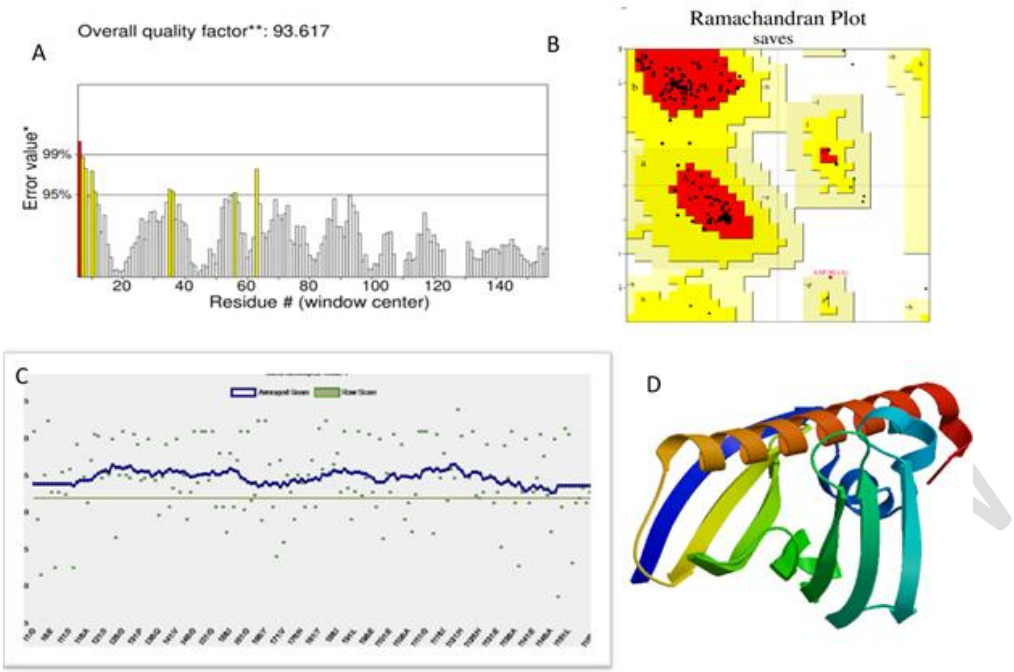

Figure 3: A) ERRAT overall quality factor was 3.617, specifying that the model predicted was good. **B)** Structural alignment of the model as evaluated by Ramachandran plot **C)** VERIFY 3D analysis indicated that 100% of the residues have averaged 3D-ID score >=0.2. **D)** 3-D structure of Hazel allergen (Cor a1) modelled through SWISS-MODEL.

Initially, the amino acids covering the paratope region of the IgE antibody with respect to bonding interactions with Bet v1 were identified using the Sitemap module: 5 residues from the heavy chain and 28 residues from the light chain were reported. Similarly, 14 residues from the heavy chain and 31 residues from the light chain of IgE were found to interact with overlapping peptide fragments of Cor a1. Among the various overlapping peptide fragments studied for their interactions with IgE, the model showed good stereo-chemical property in terms of overall G-factor value for overlapping peptide fragments 3,4; 11,12; and 16,17 of Bet v1. On the other hand, overlapping peptide fragments 3, 4; 16, 17; and 30, 31 of Cor a1 exhibited the highest G-scores. Since both Bet v1 and Cor a1 sequences exhibited 74% similarity, it was logical that overlapping peptide fragments 3, 4, and 16, 17 were commonly found to exhibit the highest G-scores for both sequences. Based on the in silico analysis, in Bet v1 and Cor a1 overlapping PF 3,4 and 16,17 were identified to have specific IgE paratope interactions and their binding poses are represented in Figure 4.

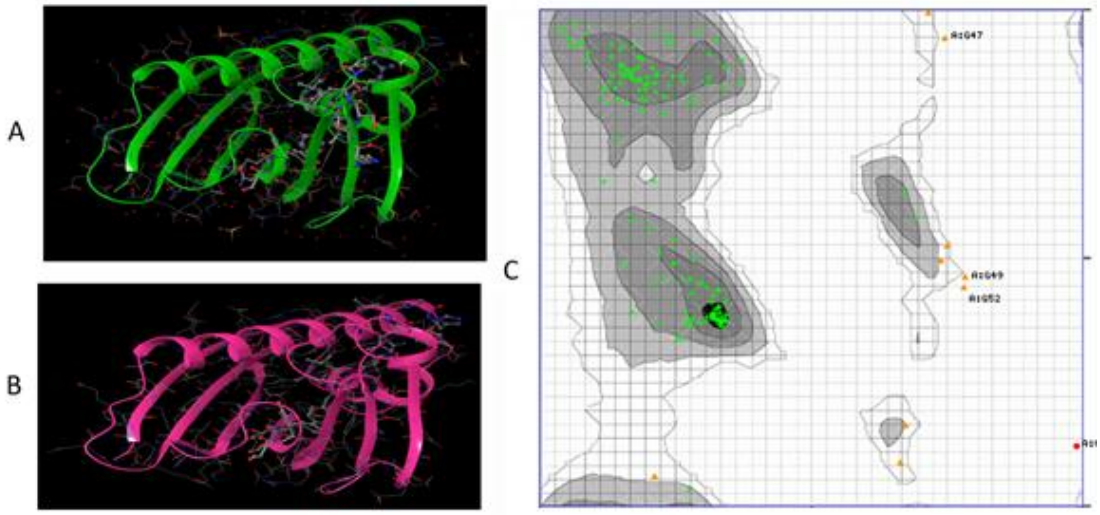

Figure 4: A) Positions of overlapping peptide fragments in Birch. **B)** Positions of overlapping peptide fragments in Hazel allergen. **C)** Density plot showing the value of Z-score**.**

DISCUSSION

Bet v1 is responsible for 60% of allergies with birch (*Betula verrucosa*) pollen released into the air affecting millions of people in spring **[13]**. The birch pollen allergen has different isoforms, all of which exhibit identical conformations, but different allergenic potentials **[14].** IgE and IgG antibodies of patients with allergy to birch pollen serve as tools to define the allergen **[15]**. Up to 90% of the Bet v1-exposed individuals do exhibit IgE-mediated allergic cross-reactions (oral allergy syndrome) to Bet v1-homologous food allergens, such as hazel nut **[16]**. The three-dimensional structure of Bet v1 and related pollen and food allergens including Cor a1 from hazelnut belong to the family of class 10 pathogenesis-related proteins (PR-10) within the Bet v1 superfamily. PR-10 proteins comprise about 160 amino acid residues with a molecular weight of 17.5 kDa. These proteins exhibit a canonical fold consisting of a seven-stranded antiparallel β-sheet (β1-β7) and three α helices (α1-α3). The two short, consecutive helices α1 and α2 interrupt the β-sheet between strands β1 and β2 while the long C-terminal helix α3 is located above the β-sheet, creating a large and fairly hydrophobic cavity in the protein interior **[17].** Cor a1 shares 67% sequence identity with Bet v1 and shared similar tertiary structures based on the homology modelling. As with Cor a11, structural flexibility in Bet v1 is distributed across the entire PR-10 scaffold, including secondary structure elements and loops **[18]**. Whether an allergen induces strong immediate-type hypersensitivity reactions in sensitized allergic patients is largely determined by its

ability to induce IgE-mediated degranulation of mast cells and basophils **[19]**. The process of degranulation is dependent on cross-linking of cell-bound IgE antibodies and hence requires the presence of at least two IgE epitopes on the allergen **[20].** The IgE antibodies appear to recognize primary conformational epitopes on allergens **[21].** Conformational epitope mapping using conventional strategies such as testing for IgE reactivity to recombinant or synthetic allergen fragments is not easy because fragmentation of proteins often leads to the loss of the three-dimensional structure of the protein and hence to loss of IgE reactivity **[22].** The onset of birch pollen-related food allergy is believed to be induced by primary sensitization to pollen allergens and subsequent development of secondary food allergy caused by IgE cross-reactivity between homologous pollen and food allergens **[23]**. Bet v1-specific IgE antibodies were shown to cross-react at the T-cell level with Cor a1. **[24].** In order to characterize IgE binding as a measure for allergenicity, we characterized the antibody-binding behavior of the Bet v1 **[25].** IgE recognition of Bet v1 is not influenced by the bound ligands such as flavonoids **[26]**. We also sought to map the IgE epitopes on the three-dimensional structure of Cor a1 **[27]**. Due to the lack of crystal structure of Cor a1, a homology 3D-model was employed for characterizing the epitopes on the surface of Cor a1 **[28]**. For each of the allergens, namely, Bet v1 and Cor a1, and their interaction profiles with Ig E antibodies, the antigen sequence was fragmented into a series of overlapping peptides and their binding modes against IgE were studied. RMSD and RMSF from the simulation results were found to be in the acceptable range of 1-3 A°. The ERRAT score indicates the overall stability of the modelled Cor a1 protein. Sequential IgE epitope analysis was performed to study IgE epitopes that recognize birch pollen and hazelnut allergens at the level of peptides **[29]**. Our results confirmed a few sequential IgE epitopes, which were found in similar locations and the homology of the amino acid composition of the epitopes of the two allergens was relatively high **[30].** The identified sequential epitopes mapped to the Bet v1 three-dimensional structures indicate that these residues are exposed on the protein surface and are spread over the β1-α1 regions, β6, β5 and β4 regions in case of Bet v1 **[31]**. On other hand, in the case of Cor a1, it involved the β1-α1 regions, β5, β4, and α2 regions. Amino acids 2-11 constitute β1, 113-123 constitute β2, 96-106 constitute β3, 79-87 constitute β4, 68-75 constitute β5, 51-57 constitute β6, 40-45 constitute β7 **[32]**.Similarly, amino acids 15-33 constitute α1, 131-154 constitute α2. The knowledge of the IgE epitopes on the Bet v1 and Cor a1 allergens should contribute to the design of effective active and passive immunotherapy strategies for birch pollen and related allergies.

CONCLUSION

The generated model could be supportive to understand the functional characteristics of Cor a1 and Bet v1 against IgE. The *in silico* molecular modeling and validation studies is helpful to understand the structure, function and mechanism of proteins action. We here display the usefulness of allergen-specific IgE antibody as a tool in studies of the crucial molecular interaction taking place at the initiation of an allergic response. Such studies may aid us in development of better diagnostic tools and guide us in the development of new therapeutic compounds.

REFERENCES

1. Potocnakova L, Bhide M, Pulzova LB. An introduction to B-cell epitope mapping and in silico epitope prediction. Journal of immunology research. 2016 Oct;2016.
2. Scadding GK, Scadding GW. Innate and adaptive immunity: ILC2 and Th2 cells in upper and lower airway allergic diseases. The Journal of Allergy and Clinical Immunology: In Practice. 2021 Feb 20.
3. Valenta R, Breiteneder H, Pettenburger K, Breitenbach M, Rumpold H, Kraft D, Scheiner O. Homology of the major birch-pollen allergen, Bet v I, with the major pollen allergens of alder, hazel, and hornbeam at the nucleic acid level as determined by cross-hybridization. Journal of allergy and clinical immunology. 1991 Mar 1;87(3):677-82.
4. Schocker F, Lüttkopf D, Müller U, Thomas P, Viehs S, Becker WM. IgE binding to unique hazelnut allergens: identification of non pollen-related and heat-stable hazelnut allergens eliciting severe allergic reactions. European journal of nutrition. 2000 Sep 1;39(4):172-80.
5. Valcour A, Lidholm J, Borres MP, Hamilton RG. Sensitization profiles to hazelnut allergens across the United States. Annals of Allergy, Asthma & Immunology. 2019 Jan 1;122(1):111-6.
6. Dimitrov I, Flower DR, Doytchinova I. AllerTOP-a server for in silico prediction of allergens. InBMC bioinformatics 2013 Apr (Vol. 14, No. 6, pp. 1-9). BioMed Central.
7. Dang HX, Lawrence CB. Allerdictor: fast allergen prediction using text classification techniques. Bioinformatics. 2014 Apr 15;30(8):1120-8.
8. Bateman A, Martin MJ, Orchard S, Magrane M, Agivetova R, Ahmad S, Alpi E, Bowler-Barnett EH, Britto R, Bursteinas B, Bye-A-Jee H. UniProt: the universal protein

knowledgebase in 2021. Nucleic Acids Research. 2020 Nov 25.
9. UniProt Consortium. Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic acids research. 2012 Jan 1;40(D1):D71-5.
10. Wang J, Alekseenko A, Kozakov D, Miao Y. Improved modeling of peptide-protein binding through global docking and accelerated molecular dynamics simulations. Frontiers in molecular biosciences. 2019 Oct 30;6:112.
11. Fiser A. Template-based protein structure modeling. Computational biology. 2010:73-94.
12. Ambrosetti F, Jiménez-García B, Roel-Touris J, Bonvin AMJJ. Modeling Antibody-Antigen Complexes by Information-Driven Docking. Structure. 2020 Jan 7;28(1):119-129.e2.
13. Asam C, Batista AL, Moraes AH, de Paula VS, Almeida FC, Aglas L, Kitzmüller C, Bohle B, Ebner C, Ferreira F, Wallner M. Bet v1–a Trojan horse for small ligands boosting allergic sensitization?. Clinical & Experimental Allergy. 2014 Aug;44(8):1083-93.
14. Vrtala S, Hirtenlehner K, Susani M, Akdis M, Kussebi F, Akdis CA, Blaser K, Hufnagl P, Binder BR, Politou A, Pastore A. Genetic engineering of a hypoallergenic trimer of the major birch pollen allergen, Bet v1. The FASEB Journal. 2001 Sep;15(11):2045-7.
15. Jarolim E, Rumpold H, Endler AT, Ebner H, Breitenbach M, Scheiner O, Kraft D. IgE and IgG antibodies of patients with allergy to birch pollen as tools to define the allergen profile of Betulaverrucosa. Allergy. 1989 Aug;44(6):385-95.
16. Herkenrath C, Gottmann-Lückerath I, Steigleder GK. Combined allergy against hazel pollen and hazel nuts. Zeitschrift fur Hautkrankheiten. 1982 Oct 1;57(19):1399-405
17. Fernandes H, Michalska K, Sikorski M, Jaskolski M. Structural and functional aspects of PR-10 proteins. The FEBS journal. 2013 Mar 1;280(5):1169-99.
18. Fernández-Quintero ML, Loeffler JR, Waibl F, Kamenik AS, Hofer F, Liedl KR. Conformational selection of allergen-antibody complexes—surface plasticity of paratopes and epitopes. Protein Engineering, Design and Selection. 2019 Nov;32(11):513-23.
19. Valenta R, Kraft D. Recombinant allergen molecules: tools to study effector cell activation. Immunological reviews. 2001 Feb;179(1):119-27.
20. Segal DM, Taurog JD, Metzger H. Dimeric immunoglobulin E serves as a unit signal for mast cell degranulation. Proceedings of the National Academy of Sciences. 1977 Jul 1;74(7):2993-7.
21. Gadermaier E, Flicker S, Aberer W, Egger C, Reider N, Focke M, Vrtala S, Kundi M,

Valenta R. Analysis of the antibody responses induced by subcutaneous injection immunotherapy with birch and Fagales pollen extracts adsorbed onto aluminum hydroxide. International archives of allergy and immunology. 2010;151(1):17-27.

22. Vrtala S, Hirtenlehner K, Vangelista L, Pastore A, Eichler HG, Sperr WR, Valent P, Ebner C, Kraft D, Valenta R. Conversion of the major birch pollen allergen, Bet v1, into two nonanaphylactic T cell epitope-containing fragments: candidates for a novel form of specific immunotherapy. The Journal of clinical investigation. 1997 Apr 1;99(7):1673-81.
23. Irañeta SG, Seoane MA, Laucella SA, Apicella C, Alonso A, Duschak VG. Antigenicity and immunocrossreactivity of orange tree pollen and orange fruit allergenic extracts. International archives of allergy and immunology. 2005;137(4):265-72.
24. Jahn-Schmid B, Radakovics A, Lüttkopf D, Scheurer S, Vieths S, Ebner C, Bohle B. Bet v1142-156 is the dominant T-cell epitope of the major birch pollen allergen and important for cross-reactivity with Bet v1–related food allergens. Journal of allergy and clinical immunology. 2005 Jul 1;116(1):213-9.
25. Spangfort MD, Mirza O, Ipsen H, Van Neerven RJ, Gajhede M, Larsen JN. Dominating IgE-binding epitope of Bet v1, the major allergen of birch pollen, characterized by X-ray crystallography and site-directed mutagenesis. The Journal of Immunology. 2003 Sep 15;171(6):3084-90.
26. Kofler S, Asam C, Eckhard U, Wallner M, Ferreira F, Brandstetter H. Crystallographically mapped ligand binding differs in high and low IgE binding isoforms of birch pollen allergen bet v1. Journal of molecular biology. 2012 Sep 7;422(1):109-23.
27. Negi SS, Braun W. Cross-React: a new structural bioinformatics method for predicting allergen cross-reactivity. Bioinformatics. 2017 Apr 1;33(7):1014-20.
28. Vyas VK, Ukawala RD, Ghate M, Chintha C. Homology modeling a fast tool for drug discovery: current perspectives. Indian journal of pharmaceutical sciences. 2012 Jan;74(1):1.
29. Dall'Antonia F, Pavkov-Keller T, Zangger K, Keller W. Structure of allergens and structure based epitope predictions. Methods. 2014 Mar 1;66(1):3-21.
30. McClain S. Bioinformatic screening and detection of allergen cross-reactive IgE-binding epitopes. Molecular nutrition & food research. 2017 Aug;61(8):1600676.
31. Ahammer L, Grutsch S, Kamenik AS, Liedl KR, Tollinger M. Structure of the major apple allergen Mal d 1. Journal of agricultural and food chemistry. 2017 Mar 1;65(8):1606-12.

32. Lotta LA, Scott RA, Sharp SJ, Burgess S, Luan JA, Tillin T, Schmidt AF, Imamura F, Stewart ID, Perry JR, Marney L. Genetic predisposition to an impaired metabolism of the branched-chain amino acids and risk of type 2 diabetes: a Mendelian randomisation analysis. PLoS medicine. 2016 Nov 29;13(11):e1002179.