In Silico Characterization and Comparative Analysis of Allergenicity of Allergic Proteins from Different Food Sources

^{1,2}Sourav Chakraborty, ^{1,2}Shakhawat Hossain Foysal, ^{1,2}Md. Nazmul Hasan and ^{1,2}Nahian Khan

¹Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh ²CANSi Research Institute, SUST, Sylhet, Bangladesh

Article history Received: 14-08-2014 Revised: 11-09-2014 Accepted: 18-11-2014

Corresponding Author: Sourav Chakraborty Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh Email: souravgeb91@gmail.com Abstract: Allergy is a steadily increasing health problem for all age groups. In general, it's recommended that 10-35% of our daily calories come from protein. A complete protein source is one that provides all of the essential amino acids called high quality proteins. Although animalbased foods; for example, meat, poultry, fish, milk, eggs, are considered complete protein sources these are the common sources of causing allergenicity. In case of meat Bos taurus, prawn Penaeus monodon and egg Gallus gallus are found to be the most responsible for triggering allergenicity. The current study has disclosed the best alternative sources for meat, prawn and egg through in silico characterization and comparative analysis of allergic proteins (Myoglobin Ovomucoid, Lysozyme, Ovalbumin, Ovotransferrin, Tropomyosin) with other common sources of meat (Capra hircus, Ovis aries, Gallus gallus, Sus scrofa), prawn (Fenneropenaeus merguiensis, Macrobrachium rosenbergii, Metapenaeus ensis, Pandalus borealis) and egg (Anas platyrhynchos). Analyzing the results we found that Gallus gallus, Macrobrachium rosenbergii, Anas platyrhynchos would be the safe source for meat, prawn and egg respectively.

Keywords: Allergenicity, High Quality Proteins, Allergy, Allergic Proteins

Introduction

The term food allergy is used to describe an adverse immune response to foods (Johansson et al., 2004). Allergy is a steadily increasing health problem for all age groups in the United States. Food allergies, mostly against milk, eggs, peanuts, soy, or wheat, affect up to 8% of infants and young children (Sampson, 1999a; 2005). A 2008 Centers for Disease Control and Prevention report indicated an 18% increase in childhood food allergy from 1997 to 2007, with an estimated 3.9% of children currently affected. Branum and Lukacs, (2008). One hypothesis is that this late onset may be the result of individuals being sensitized by long-term exposure to environmental factors that contain proteins similar to those in the known triggers of allergenic response (Sampson, 1999b; Vanek-Krebitz et al., 1995; Scheurer et al., 1999; Rabjohn et al., 1999). Recent studies have identified common molecular features of proteins from different sources, which could account for clinically important cross-reactivity (Breiteneder and Ebner, 2000; Jenkins et al., 2005) and sensitivity (Ferreira et al., 2004; Mari, 2001). Some common animal proteins from meat, egg, shrimp, cow's milk (Das et al., 2005) have been identified and characterized as major allergens. Myoglobin protein of mammalian meats causes allergy may be based on subtle changes of amino acids. Other cause may be a heat-resistant nature of the protein of 17 kDa that could be implicated in those patients that do not tolerate well-cooked meat (Fuentes et al., 2004). The only major allergen (Pen a 1) identified in shrimp is the muscle protein, tropomyosin (Daul et al., 1994). Surprisingly, there is no report of *M. rosenbergii* allergy in any medical literature. From the anaphylaxis study at the Siriraj hospital, Thailand, there were subpopulations of shrimp allergic patients who developed anaphylaxis to freshwater shrimp but could tolerate seawater shrimp or vice versa



© 2015 Sourav Chakraborty, Shakhawat Hossain Foysal, Md. Nazmul Hasan and Nahian Khan. This open access article is distributed under a Creative Commons Attribution (CC-BY) 3.0 license

(Jirapongsananuruk et al., 2007). The specific role of egg ovalbumin has been found in patients allergic to cow milk, casein, along with other two milk proteins immunoreacted with IgE antibody (Szabo and Eigenmann, 2000). Although infants can, in theory, be allergic to any food, one of the major food allergies is hen's egg (Du Toit et al., 2009). Most people who are allergic to hen's eggs have antibodies which react to one of four proteins in the egg white: Ovomucoid, ovalbumin, ovotransferrin and lysozyme (Platts-Mills and Ring, 2005). Fish allergy is one of the most common food allergies mediated by IgE antibody. Consumption of fish products could lead to symptoms like skin rash, dermatitis, urticarial, angioedema, gastrointestinal problems, diarrhoea, respiratory distress and even fatal systemic anaphylactic reactions (Pascual et al., 1992; O'Neil et al., 1993). Allergenic proteins that have been isolated from primary food sources, such as egg (Mine and Rupa, 2004) (ovomucoid (Mizumachi and Kurisaki, 2003; Mine and Zhang, 2002; Mine et al., 2003) and lysozyme (Fremont et al., 1997), shrimp and related species (tropomyosins (Ayuso et al., 2002; Reese et al., 2002; Samson et al., 2004)). The present study is an attempt to analyze the componential and structural similarity among different allergic proteins present in different foods to compare allergenicity among different allergic protein and identify common motifs. The newly discovered sequence motif along with the analyses of the structures of these allergens will not only help in the understanding of structure-function relationship of these allergens but also in the identification of new allergic protein.

Materials and Methods

The amino sequences of 18 different allergic were retrieved from NCBI proteins the (http://www.ncbi. nlm.nih. gov/genomes/FLU/). The Protparam tool at ExPASy (http://www. expasy.org/tools/) was used to analyze physic-chemical properties of two proteins i.e., amino acid compositions in all the species under consideration. The SOPMA tool at ExPASy server was exploited for comparative secondary structure analysis. The protein sequences of meat allergens are aligned using the Clustal W2 (http://www.ebi.ac.uk/tools/clustalw2). program Phylogenetic trees were constructed by the neighborjoining method. The computer software of the Molecular Evolution Genetic Analysis (MEGA), version 5.2 was utilized in this study for phylogenetic analysis of selected sequences. The MEME (http://meme.nbcr.net/meme/) software was used to elect the motifs from different protein sequences.

Results

The 18 protein sequences of allergens were retrieved from NCBI. The accession number of retrieved sequences along with species names is listed in Table 1. Multiple Sequence Alignment using ClustalW2 are shown in three separate figures (Fig. 1 to 3). The phylogenetic tree using neighbor joining mode revealed three major clusters of protein sequences (Fig. 4). Biochemical features for 18 allergic protein were obtained by using ProtParam are listed in Table 2. Secondary structure analysis were done by using SOPMA software and shown in Table 3. MEME analysis resulted in frequently observed 3 motifs (Table 4).

Discussion

The sequences were characterized for homology search, multiple sequences alignment, biochemical features, phylogenetic tree construction and motifs search using various bioinformatics tools. Multiple Sequence Alignment using ClustalW2 provided that the myoglobin protein of Bos taurus shows similarity score of 97.4 with Capra hircus and Ovis aries. The myoglobin protein of Gallus gallus and Sus scrofa shown similarity score of 72.73 and 88.31 respectively with *Bos taurus* (Fig. 1). The tropomyosin protein of Penaeus monodon showed similarity score of 100 with Fenneropenaeus merguiensis and with Macrobrachium rosenbergii, Metapenaeus ensis, Pandalus borealis it was ranges from 96.83-98.24 (Fig. 2). The egg allergic proteinsovomucoid, lysozyme, ovalbumin, ovotransferrin of Anas platyrhynchos and Gallus gallus shown similarity score of 73.68, 80.27, 76.68, 79.45 respectively (Fig. 3).

The phylogenetic tree based on protein sequences revealed three major clusters. Cluster 1, a cluster containing 5 sequences under study, includes meat allergens (Fig. 4).

Biochemical features for this cluster are listed in Table 2. The total number of amino acid residues was 154 with variable molecular weights. pI values of this cluster ranged from 6.63 to 7.96. Variations among various allergens in this group in terms of other physiochemical parameters like positively charged and negatively charged residues, hydropathicity (GRAVY) are given in Table 2. Aliphatic index analysis reveals uniformity in this group of allergens within the range of 79.87 to 86.88. Aliphatic index of protein measures the relative volume occupied by aliphatic side chains of the amino acids: Alanine, valine, leucine and isoleucine. Globular proteins with high aliphatic index have high thermo stability and an increase in aliphatic index increases protein thermo stability (Ikai, 1980; Rawlings et al., 2006).

gil378405239lspl87U9B5.3IMYG_C MCLSDCEWTLVLNAW gil164414449ireflNP_001072126. MCLSDCEWQLVLNAW gil27806939ireflNP_776306.11 MCLSDCEWQLVLNAWCI gil47523546ireflNP_999401.11 MCLSDCEWQLVLNAWCI gil268607704ireflNP_001161224. MCLSDQEWQQVLTIWI	VGKVEADVAGHGQEVLIRLFTGHPETLEKFDKFKHL 50 VGKVEAGVAGHGQEVLIRLFTGHPETLEKFDKFKHL 50 KVEADVAGHGQEVLIRLFTGHPETLEKFDKFKHL 50 KVEADVAGHGQEVLIRLFKGHPETLEKFDKFKHL 50 GKVEADIAGHGHEVLMRLFHDHPETLDRFDKFKGL 50
gil378405239lspl87U985.3IMYG_C gil164414449lreflNP_001072126: gil27806939lreflNP_776306.1I gil47523546lreflNP_999401.1I gil268607704lreflNP_001161224. KTPDQMKGSEDLKKHGN	CNTVLTALGGILKKKGHHEAEVKHLAESHANKHKI 100 NTVLTALGGILKKKGHHEAEVKHLAESHANKHKI 100 IVLTALGGILKKKGHHEAEVKHLAESHANKHKI 100 IVLTALGGILKKKGHHEAELTPLAQSHATKHKI 100 GATVLTQLGKILKQKGNHESELKPLAQTHATKHKI 100
gil378405239lspl87U985.3IMYG_C gil164414449lreflNP_001072126. gil27806939lreflNP_776306.1 gil27806939lreflNP_999401.1 gil268607704lreflNP_001161224. PVKYLEFISDAIIHVLHAKH PVKYLEFISDAIIHVLHAKH PVKYLEFISEAIIQVLQSKH PVKYLEFISEVIIKVIAEKI	AKHPSDFGADAQGAMSKALELFRNDMAAQYKVL 150 KHPSDFGADAQGAMSKALELFRNDMAAQYKVL 150 IPSDFGADAQAAMSKALELFRNDMAAQYKVL 150 PGDFGADAQCAMSKALELFRNDMAAKYKEL 150 HAADFGADSQAAMKKALELFRNDMASKYKEF 150
gil378405239lspl87U985.3IMYG_C gil164414449irefINP_001072126. CFQC 154 gil27806939irefINP_776306.1 CFHC 154 gil47523546irefINP_999401.1 CFQC 154 gil268607704irefINP_001161224. CFQC 154	



gi 335347355 gb ADC55381.4	MKLEKDNAMDRADTLEQQNKEANNRAEKSEEEVHNLQKRM 40
gi 308191589 sp A1KYZ2.1 TPM_P	MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANNRAEKSEEEVHNLQKRM 50
gi 6094504 sp Q25456.1 TPM_MET	MKLEKDNAMDRADTLEQQNKEANNRAEKSEEEVHNLQKRM 40
gi 288819271 gb ADC55380.1	MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANNRAEKSEEEVHNLQKRM 50
gi 312831088 emb CBY17558.1	MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANNRAEKSEEEVFGLQKKL 50
gil335347355igblADC55381.4i	QQLENDLDQVQESLLKANIQLVEKDKALSNAEGEVAALNRRIQLLEEDLE 90
gil308191589isplA1KY22.1ITPM_P	QQLENDLDQVQESLLKANIQLVEKDKALSNAEGEVAALNRRIQLLEEDLE 100
gil6094504isplQ25456.1ITPM_MET	QQLENDLDQVQESLLKANIQLVEKDKALSNAEGEVAALNRRIQLLEEDLE 90
gil288819271igblADC55380.1i	QQLENDLDSVQEALLKANQHLEEKDKALSNAEGEVAALNRRIQLLEEDLE 100
gil312831088iemblCBY17558.1i	QQLENDLDSVQEALLKANQHLEEKDKALSNAEGEVAALNRRIQLLEEDLE 100
gil335347355lgblADC55381.4	RSEERLNTATTKLAEASQAADESERMRKVLENRSLSDEERMDALENQLKE 140
gil308191589lsplA1KYZ2.1lTPM_P	RSEERLNTATTKLAEASQAADESERMRKVLENRSLSDEERMDALENQLKE 150
gil6094504lsplQ25456.1lTPM_MET	RSEERLNTATTKLAEASQAADESERMRKVLENRSLSDEERMDALENQLKE 140
gil288819271lgblADC55380.1	RSEERLNTATTKLAEASQAADESERMRKVLENRSLSDEERMDALENQLKE 150
gil312831088lemblCBY17558.1	RSEERLNTATTKLAEASQAADESERMRKVLENRSLSDEERMDALENQLKE 150
gi 335347355 gb ADC55381.4	ARFLAEEADRKYDEVARKLAMVEADLERAEERAETCESKIVELEEELRVV 190
gi 308191589 sp A1KYZ2.1 TPM_P	ARFLAEEADRKYDEVARKLAMVEADLERAEERAETGESKIVELEEELRVV 200
gi 6094504 sp Q25456.1 TPM_MET	ARFLAEEADRKYDEVARKLAMVEADLERAEERAETCESKIVELEEELRVV 190
gi 288819271 gb ADC55380.1	ARFLAEEADRKYDEVARKLAMVEADLERAEERAETCESKIVELEEELRVV 200
gi 312831088 emb CBY17558.1	ARFLAEEADRKYDEVARKLAMVEADLERAEERAETCESKIVELEEELRVV 200
gi 335347355 gb ADC55381.4	CNNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARAEFAERSVQKLQ 240
gi 308191589 sp A1KYZ2.1 TPM_P	CNNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARAEFAERSVQKLQ 250
gi 6094504 sp Q25456.1 TPM_MET	CNNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARAEFAERSVQKLQ 240
gi 288819271 gb ADC55380.1	CNNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARAEFAERSVQKLQ 250
gi 312831088 emb CBY17558.1	GNNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARAEFAERSVQKLQ 250
gi 335347355igb ADC55381.4	KEVDRLEDELVNEKEKYKSITDELDQTFSELSGY 274
gi 308191589 sp A1KYZ2.1 TPM_P	KEVDRLEDELVNEKEKYKSITDELDQTFSELSGY 284
gi 6094504 sp Q25456.1 TPM_MET	KEVDRLEDELVNEKEKYKSITDELDQTFSELSGY 274
gi 288819271 gb ADC55380.1	KEVDRLEDELVNEKEKYKSITDELDQTFSELSGY 284
gi 312831088 emb CBY17558.1	KEVDRLEDELVNEKEKYKSITDELDQTFSELSGY 284



Sourav Chakraborty *et al.* / American Journal of Biochemistry and Biotechnology 2015, 11 (1): 17.24 DOI: 10.3844/ajbbsp.2015.17.24







Fig. 4. Phylogenetic tree constructed by NJ method based on allergic protein sequences

Sourav Chakraborty *et al.* / American Journal of Biochemistry and Biotechnology 2015, 11 (1): 17.24 DOI: 10.3844/ajbbsp.2015.17.24

Table 1. List of retrieved protein sequences from NCBI their accession number				
Source organism	Protein	Accession No		
Bos taurus	Myoglobin	NP 776306.1		
Capra hircus	Myoglobin	B7U9B5.3		
Ovis aries	Myoglobin			
	NP 001072126.1			
Gallus gallus	Myoglobin			
-	NP_001161224.1			
Sus scrofa	Myoglobin	NP_999401.1		
Gallus gallus	Ovomucoid	ACJ04729.1		
Gallus gallus	Lysozyme	AAA48943.1		
Gallus gallus	Ovotransferrin	CAA26040.1		
Gallus gallus	Ovalbumin	AAB59956.1		
Anas platyrhynchos	Lysozyme	P00705.1		
Anas platyrhynchos	Ovotransferrin	P56410.1		
Anas platyrhynchos	Ovalbumin	EOA97657.1		
Anas platyrhynchos	Ovomucoid	EOA95437.1		
Fenneropenaeus merguiensis	Tropomyosin	ADC55381.4		
Macrobrachium rosenbergii	Tropomyosin	ADC55380.1		
Metapenaeus ensis	Tropomyosin	Q25456.1		
Penaeus monodon	Tropomyosin	A1KYZ2.1		
Pandalus borealis	Tropomyosin	CBY17558.1		

Table 2. Biochemical characteristics of allergic proteins

					Total number of negatively	Total number of positively			
		Source	Number of		charged residues	charged residues	Instability	Aliphatic	
Protein name	organisms	amino acids	MW	Theoretical pI	(Asp + Glu)	(Arg + Lys)	index	index	Gravy
Myoglobin	Bos taurus	154	17077.5	6.90	21	20	15.13	86.88	-0.364
Myoglobin	Capra hircus	154	16955.4	7.23	20	20	11.93	86.23	-0.342
Myoglobin	Ovis aries	154	16997.4	6.63	21	19	16.36	86.23	-0.357
Myoglobin	Sus scrofa	154	17084.6	6.76	22	21	20.50	84.94	-0.444
Myoglobin	Gallus gallus	154	17422.0	7.96	22	23	10.76	79.87	-0.542
Ovomucoid	Gallus gallus	210	22608.4	4.67	30	18	28.99	64.00	-0.170
Lysozyme	Gallus gallus	147	16238.6	9.36	9	18	19.86	81.70	-0.150
Ovalbumin	Gallus gallus	386	42881.2	5.19	47	35	37.11	89.95	-0.001
Ovotransferrin precursor	Gallus gallus	705	77893.6	6.50	91	88	35.68	74.28	-0.354
Lyozyme DL1	Anas platyrhynchos	147	16362.7	9.41	11	21	26.44	78.37	-0.276
Ovotransferrin	Anas platyrhynchos	686	75632.8	6.19	92	86	37.44	74.78	-0.386
Ovalbumin	Anas platyrhynchos	420	47749.5	4.92	55	38	48.01	77.52	-0.158
Ovomucoid	Anas platyrhynchos	171	18629.0	4.69	24	15	56.96	52.92	-0.436
Tropomyosin	Fenneropenaeusmerguiensis	274	31704.0	4.66	71	46	39.46	79.85	-1.109
Tropomyosin	Macrobrachium rosenbergii	284	32846.4	4.73	73	49	38.61	77.08	-1.129
Tropomyosin	Metapenaeus ensis	274	31704.9	4.66	71	46	37.58	78.43	-1.138
Tropomyosin	Penaeus monodon	284	32849.4	4.72	72	49	39.00	79.12	-1.094
Tropomyosin	Pandalus borealis	284	32753.3	4.70	73	49	37.74	78.45	-1.088

Table 3. Secondary structure of allergic proteins

Serial no.	Accession No	Alpha helix (Hh) (%)	Extended strand (Ee) (%)	Beta turn (Tt) (%)	Random coil (Cc) (%)
1	NP 776306.1	69.48	1.30	9.74	19.48
2	B7U9B5.3	69.48	1.30	9.74	19.48
3	NP 001072126.1	67.53	1.30	9.74	21.43
4	NP_001161224.1	70.13	2.60	7.14	20.13
5	NP 999401.1	66.23	0.65	11.04	22.08
6	ACJ04729.1	20.00	12.86	2.86	64.29
7	AAA48943.1	43.54	14.97	12.24	29.25
8	AAB59956.1	45.85	14.51	5.96	33.68
9	CAA26040.1	31.35	19.15	7.52	41.99
10	P00705.1	32.65	15.65	8.16	43.54
11	P56410.1	29.74	19.39	7.00	43.88
12	EOA97657.1	42.86	16.67	6.19	34.29
13	EOA95437.1	11.70	9.36	2.92	76.02
14	ADC55381.4	100.00	0.00	0.00	0.00
15	ADC55380.1	99.30	0.00	0.00	0.70
16	Q25456.1	99.27	0.00	0.00	0.73
17	A1KYZ2.1	99.30	0.00	0.00	0.70
18	CBY17558.1	100.00	0.00	0.00	0.00

Sourav Chakraborty et al. / American Journal of Biochemistry and Biotechnology 2015, 11 (1): 17.24 DOI: 10.3844/ajbbsp.2015.17.24

Motifs number	Motif present in number of sequence	Sequence
1	5	"HKIPVKYLEFIS[DE][AV]II[HKQ]V[LI] [HAQ] [AES]KH[PA][SAG] DFGAD[AS]Q[GA]AM[SK] KALELFRNDMA[AS][QK]Y"
2	5	"AEARAEFAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDQTFSE LSGY"
3	6	[GQ][KG][LV][CE][RS][IQR]C[EFK][GLQ][ADT]A[AK]][DMT][KQ] [ACR][LR][EGR][LN][DI][NP][YS][RSW][GV][YE]S[GLQ] [AGT] [N F][GW]

Table 4 Distribution of commonl	a observed motifs in different all	largic protain saguancas
rabic 4. Distribution of common	y observed mouns in unicient an	leight protein sequences

Cluster 2 includes 5 protein sequences and represents shrimp allergen sequences. The total number of amino acid residues was in the range of 274 to 284 and the pI values range from 4.66 to 4.73. It has less variation in its pI as compared to cluster 1 sequences. Aliphatic index of this cluster sequences was uniform in the range of 77.08 to 79.85. Cluster 3 includes 8 protein sequences and represents egg allergen sequences. Various biophysical parameters for this group of sequences reveal amino acid residues ranging from 147 to 705, while pI value of the majority of sequences was in range of 4.67 to 6.50 except for lysozyme [Gallus gallus] (9.36). Aliphatic index of this group of sequences reveals in the range of 64.0 to 81.70 and ovalbumin was the highest thermo stable allergen (90.18) among all three clusters. Secondary structure analysis exhibits that the instability index is used to measure in vivo half-life of a protein (Guruprasad et al., 1990). The proteins which have been reported as in vivo half-life of less than 5 h showed instability index greater than 40, whereas those having more than 16 h half-life (Rogers et al., 1986) have an instability index of less than 40. Instability index of allergic protein sequences under the study was found less than 40 (Table 2).

Secondary structure analysis demonstrated that the myoglobin protein of beef was more similar with goat and sheep meats. In case of shrimp Alpha helix ranges from 99.27-100.00% and Random coil is 0.70-0.73%, except CBY17558.1, ADC55381.4. Extended strand and Beta turn are absent in all the sequences of the cluster2. In egg allergen Alpha helix ranges from 11.70-45.85%, Extended strand 9.36-19.39%, Beta turn 2.86-12.24%, Random coil 76.02-29.25% (Table 3).

MEME analysis provided that five amino acid residues of Cluster1 representing motif 1.Cluster 2, representing motif 2 in its sequences, it contains a 50 amino acid residues long unique motif. Motif 3 was present in 6 protein sequences representing cluster 3.

Conclusion

Comparing variation among biochemical features of myoglobin from different organisms we get that goat, sheep and pig meat are most similar with beef than chicken meat in causing allergenicity. Heat stability of an allergic protein is one of the most important reasons of becoming more allergic and it depends on the value of aliphatic index. The aliphatic index of myoglobin protein of Gallus gallus is lower than other four organisms. So, chicken meat can be a convenient alternative to beef allergic peoples. In case of shrimp the biochemical features of tropomyosin of Penaeus monodon shows more with Fenneropenaeus similarity merguiensis, Metapenaeus and Pandalus borealis ensis than Macrobrachium rosenbergii. However, the aliphatic index of tropomyosin of Macrobrachium rosenbergii is lower than the other. As for that Macrobrachium rosenbergii can be a suitable substitute to Penaeus monodon allergic peoples. The allergic proteins of hen's egg white are more thermo stable because of having comparatively high aliphatic index than duck. So, duck egg can be an appropriate alternative to hen's egg allergic infants. Beside these phylogenetic clustering and variation among biochemical features of different allergens might contribute in further classification of highly diverse allergens and their selection for various application purposes. Conserved sequences in motifs may be utilized for designing specific degenerate primers for identification and isolation of type and class of allergens as numerous allergens are being isolated to assure food security. Variation in biochemical features may be a key source of information for the screening of novel allergens and comparison with other classes of allergens. Functional attributes are needed to verify experimentally for conserved motifs found. This in silico analysis might be used for future genetic engineering of assuring food security.

Acknowledgment

We are thankful to Auditi Purkaystha, Jahed Ahmed and Tanjia Afrin Chowdhury (MS student of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet, Bangladesh) for their valuable support and discussions related to this research.

Funding Information

This research work had been conducted with the selfcontribution of the four authors. We are ever grateful to American Journal of Biochemistry and Biotechnology for publishing our manuscript without taking necessary charges as because of the inhabitant of Bangladesh.

Author's Contributions

Sourav Chakraborty: Manuscript writing and result verification.

Shakhawat Hossain Foysal: Secondary structure analysis by using SOPMA software.

Nazmul Hasan: Biochemical features data preparation for 18 proteins using ProtParamsoftware.

Nahian Khan: Phylogenetic tree construction and common motif identificationusing MEGA5.2 and MEMEsoftware respectively.

Ethics

This article is original and contains unpublished materials. I'm Sourav Chakraborty-the corresponding author confirms that all of the other authors have read and approved the manuscript.

References

- Ayuso, R., S.B. Lehrer and G. Reese, 2002. Identification of continuous, allergenic regions of the major shrimp allergen Pen a 1 (tropomyosin). Int. Arch Allergy Immunol., 127: 27-37. DOI: 10.1159/000048166
- Branum, A.M. and S.L. Lukacs, 2008. Food allergy among U.S. children: Trends in prevalence and hospitalizations. NCHS Data Brief, 10: 1-8. DOI: 10.1542/peds.2009-1210
- Breiteneder, H. and C. Ebner, 2000. Molecular and biochemical classification of plant-derived food allergens. J. Allergy Clin. Immunol., 106: 27-36. DOI: 10.1067/mai.2000.106929
- Das, A., P. Chakraborti, U. Chatterjee, G. Mondal and B.P. Chatterjee, 2005. Identification of allergens in Indian fishes: Hilsa and Pomfret exemplified by ELISA and immunoblotting. Ind. J. Exp. Biol., 43: 1170-1175. PMID: 16359130
- Daul, C.B., M. Slattery, G. Reese and S.B. Lehrer, 1994.
 Identification of the major brown shrimp (Penaeus aztecus) allergen as the muscle protein tropomyosin. Int. Arch Allergy Clin. Immunol., 105: 49-55. PMID: 7916224
- Du Toit, G., A. Santos, G. Roberts, A.T. Fox and P. Smith *et al.*, 2009. The diagnosis of IgE-mediated food allergy in childhood. Pediatr. Allergy Immunol., 20: 309-319.

DOI: 10.1111/j.1399-3038.2009.00887.x

Ferreira, F., T. Hawranek and P. Gruber, 2004. Allergic cross-reactivity: From gene to the clinic. Allergy, 59: 243-267.

- Fremont, S., G. Kanny, J.P. Nicolas and D.A. Moneret-Vautrin, 1997. Prevalence of lysozyme sensitization in an egg-allergic population. Allergy, 52: 224-228. DOI: 10.1111/j.1398-9995.1997.tb00980.x
- Ikai, A., 1980. Thermostability and aliphatic index of globular proteins. J. Biochem., 88: 1895-1898. PMID: 7462208
- Jenkins, J.A., S. Griffiths-Jones, P.R. Shewry, H. Breiteneder and E.N. Mills *et al.*, 2005. Structural relatedness of plant food allergens with specific reference to cross-reactive allergens: An in silico analysis. J. Allergy Clin. Immunol., 115: 163-170. PMID: 15637564
- Jirapongsananuruk, O., W. Bunsawansong, N. Piyaphanee, N. Visitsunthorn and T. Thongngarm *et al.*, 2007. Features of patients with anaphylaxis admitted to a university hospital. Ann. Allergy Asthma Immunol., 98: 157-162. DOI: 10.1016/S1081-1206(10)60689-8
- Johansson, S.G., T. Bieber, R. Dahl, P.S. Friedmann and B.Q. Lanier *et al.*, 2004. Revised nomenclature for allergy for global use: Report of the nomenclature review committee of the world allergy organization, October 2003. J. Allergy Clin. Immunol., 113: 832-836. DOI: 10.1016/j.jaci.2003.12.591
- Guruprasad, K., B.V.B. Reddy and M.W. Pandit, 1990. Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Eng., 4: 155-161. DOI: 10.1093/protein/4.2.155
- Mari, A., 2001. Multiple pollen sensitization: A molecular approach to the diagnosis. Int. Arch Allergy Immunol., 125: 57-65. DOI: 10.1159/000053797
- Fuentes, M.M., R. Palacios, M.M. Garcés, M.L. Caballero and I. Moneo, 2004. Isolation and characterization of a heat-resistant beef allergen: Myoglobin. Allergy, 59: 327-331. PMID: 14982516
- Mine, Y. and P. Rupa, 2004 Immunological and biochemical properties of egg allergens. Worlds Poult. Sci. J., 60: 321-330. DOI: 10.1079/WPS200420
- Mine, Y., E. Sasaki and J.W. Zhang, 2003. Reduction of antigenicity and allergenicity of genetically modified egg white allergen, ovomucoid third domain. Biochem. Biophys Res. Commun., 302: 133-137. PMID: 12593859.
- DOI: 10.1016/S0006-291X(03)00117-7 Mine, Y. and J.W. Zhang, 2002. Identification and fine mapping of IgG and IgE epitopes in ovomucoid.
- Biochem. Biophys Res. Commun., 292: 1070-1074. PMID: 11944924. DOI: 10.1006/bbrc.2002.6725
- Mizumachi, K. and J. Kurisaki, 2003. Localization of T cell epitope regions of chicken ovomucoid recognized by mice. Biosci Biotechnol Biochem., 67: 712-719. PMID: 12784609. DOI: 10.1271/bbb.67.712

DOI: 10.1046/j.1398-9995.2003.00407.x

- Rawlings, N.D., F.R. Morton and A.J. Barrett, 2006. MEROPS: The peptidase database. Nucleic Acids Res., 34: 270-272. PMID: 16381862
- O'Neil, C., A.A. Helbling and S.B. Lehrer, 1993. Allergic reactions to fish. Clin. Rev. Allergy, 11: 183-183. DOI: 10.1007/BF02914470
- Pascual, C., E. M. Martin and J.F. Crespo, 1992. Fish allergy: Evaluation of the importance of cross-reactivity, J Pediatr, 121: 29-34.DOI: 10.1016/S0022-3476(05)81403-9
- Rabjohn, P., E.M. Helm, J.S. Stanley, C.M. West and G.A. Bannon *et al.*, 1999. Molecular cloning and epitope analysis of the peanut allergen Ara h 3. J. Clin. Invest, 103: 535-42. DOI: 10.1172/JCI5349
- Reese, G., R. Ayuso, S.M. Leong-Kee, M. Plante and S.B. Lehrer *et al.*, 2002. Epitope mapping and mutational substitution analysis of the major shrimp allergen Pen a 1 (tropomyosin). J. Allergy Clin. Immunol., 109: 307-307.

DOI: 10.1016/S0091-6749(02)82078-4

- Samson, K.T., F.H. Chen, K. Miura, Y. Odajima and M. Adachi *et al.*, 2004. IgE binding to raw and boiled shrimp proteins in atopic and nonatopic patients with adverse reactions to shrimp. Int. Arch Allergy Immunol., 133: 225-232. PMID: 14976390. DOI: 10.1159/000076828
- Rogers, S., R. Wells and M. Rechsteiner, 1986. Amino acid sequences common to rapidly degraded proteins: The PEST hypothesis. Science, 234: 364-368. DOI: 10.1126/science.2876518

- Sampson, H.A., 1999a. Food allergy. Part 1: Immunopathogenesis and clinical disorders. J. Allergy Clin. Immunol., 103: 717-728. DOI: 10.1016/S0091-6749(99)70411-2
- Sampson, H.A., 1999b. Food allergy. Part 2: Diagnosis and management. J. Allergy Clin. Immunol., 103: 981-989. DOI: 10.1016/S0091-6749(99)70167-3
- Sampson, H.A., 2005. Food allergy: When mucosal immunity goes wrong. J. Allergy Clin. Immunol., 115: 139-141. DOI: 10.1016/j.jaci.2004.11.003
- Scheurer, S., D.Y. Son, M. Boehm, F. Karamloo and S. Vieths *et al.*, 1999. Cross-reactivity and epitope analysis of Pru a 1, the major cherry allergen. Mol. Immunol., 36: 155-67.

DOI: 10.1016/S0161-5890(99)00033-4

- Szabo, I. and P.A. Eigenmann, 2000. Allergenicity of major cow's milk and peanut proteins determined by IgE and IgG immunoblotting. Allergy, 55: 42-49. DOI: 10.1034/j.1398 9995.2000.00256.x
- Vanek-Krebitz, M., K. Hoffmann-Sommergruber, M.L.D. Machado, M. Susani and H. Breiteneder *et al.*, 1995. Cloning and sequencing of Mal d 1, the major allergen from apple (Malus domestica) and its immunological relationship to bet v 1, the major birch pollen allergen. Biochem. Biophys. Res. Commun., 214: 538-51. DOI: 10.1006/bbrc.1995.2320
- Platts-Mills, T. and J. Ring, 2005. Allergy in Practice. 1st Edn., Springer Science and Business Media, Berlin, ISBN-10: 3540265848, pp: 294.