Evolutionary distance from human homologs reflects allergenicity of animal food proteins

John A. Jenkins, PhD,[a] Heimo Breiteneder, PhD,[b] and E. N. Clare Mills, PhD[a]

Norwich, United Kingdom, and Vienna, Austria

Background: ***In silico*** **analysis of allergens can identify putative relationships among protein sequence, structure, and allergenic properties. Such systematic analysis reveals that most plant food allergens belong to a restricted number of protein superfamilies, with pollen allergens behaving similarly.**

Objective: We have investigated the structural relationships of animal food allergens and their evolutionary relatedness to human homologs to define how closely a protein must resemble a human counterpart to lose its allergenic potential.

Methods: Profile-based sequence homology methods were used to classify animal food allergens into Pfam families, and *in silico* analyses of their evolutionary and structural relationships were performed.

Results: Animal food allergens could be classified into 3 main families—tropomyosins, EF-hand proteins, and caseins—along with 14 minor families each composed of 1 to 3 allergens. The evolutionary relationships of each of these allergen superfamilies showed that in general, proteins with a sequence identity to a human homolog above approximately 62% were rarely allergenic. Single substitutions in otherwise highly conserved regions containing IgE epitopes in EF-hand parvalbumins may modulate allergenicity.

Conclusion: These data support the premise that certain protein structures are more allergenic than others. Contrasting with plant food allergens, animal allergens, such as the highly conserved tropomyosins, challenge the capability of the human immune system to discriminate between foreign and self-proteins. Such immune responses run close to becoming autoimmune responses.

Clinical implications: Exploiting the closeness between animal allergens and their human homologs in the development of recombinant allergens for immunotherapy will need to consider the potential for developing unanticipated autoimmune responses. (J Allergy Clin Immunol 2007;120:1399-405.)

Key words: *Allergen bioinformatics, protein families, animal food allergen structures, tropomyosin, parvalbumin, casein, evolutionary relatedness, human homology*

Food allergy is a worldwide problem that is on the increase. Recent studies have estimated that almost 4% of the US population, and as many as 6% of infants less than 3 years of age, are afflicted with IgE-mediated food allergy.[1] A limited number of foods are responsible for the majority of reactions, with seafood and tree nut allergies predominating in adults, whereas cow's milk and hen's egg are important in infants.[1] To relate allergenic properties of proteins to sequence and structure, we have systematically classified allergens by using profile-based sequence homology methods. Previous analysis of plant food allergens revealed that most belong to a highly restricted number of protein superfamilies.[2] A similar situation has emerged for pollen allergens.[3]

Unlike plant food or pollen allergens, almost all the animal food allergens have homologs in the human proteome that may affect the way in which they are recognized by the human immune system. Others have alluded to the fact that allergenic mammalian proteins lie at the limits of the capability of the human immune system to discriminate between foreign and self-proteins.[4] Thus, we have considered the relatedness of animal food allergens to human homologs to answer the question of how closely a foreign protein has to resemble a human homolog before it loses its ability to act as an allergen.

METHODS

Animal food allergen databases

Lists of known animal food allergens were obtained from version 2.1 of the InformAll (http://www.foodallergens.ifr.ac.uk) and version 5.0 of the Food Allergy Research and Resource Programme (FARRP) (http://www.allergenonline.com) databases. The InformAll set (http://foodallergens.ifr.ac.uk/allergenlist.html) was limited to proteins that have been shown to bind IgE in sera from at least 3 patients with clinical allergy to the food from which the allergen originated. This was done to ensure that only allergens involved in the development of clinical allergy were included and that they were not included simply by virtue of their homology to known allergens, or simple IgE cross-reactivity, which does not necessarily lead to clinical allergy. InformAll listed 71 sequences, of which 12 were removed because they were essentially allergen isoforms or not proven to be IgE-reactive. They included 4 of the 5 bovine IgG sequences and 6 of 8 frog (*Rana esculenta*) parvalbumin sequences with no demonstrable IgE binding. FARRP version 5.0 listed 62 animal food allergens. Sequences for 2 bovine dander allergens homologous to Der p 2 were excluded from our analysis because they have not been

From [a]the Institute of Food Research, Norwich, and [b]the Department of Pathophysiology, Medical University of Vienna.

Disclosure of potential conflict of interest: J. A. Jenkins and E. N. C. Mills have received grant support from the United Kingdom Biological and Biotechnological Sciences Research Council and the European Union. H. Breiteneder has declared that he has no conflict of interest.

Supported in part by the EU through the concerted action QLRT-2001-02284 (InformAll; QLRT-2001-02284) and the BBSRC through the competitive strategic grant to the Institute of Food Research. H.B. was supported by Austrian Science Fund grant SFB F01802.

Received for publication January 24, 2007; revised August 11, 2007; accepted for publication August 13, 2007.

Available online October 12, 2007.

Reprint requests: Heimo Breiteneder, PhD, Department of Pathophysiology, Medical University of Vienna, AKH-EBO-3Q, Waehringer Guertel 18-20, 1090 Vienna, Austria. E-mail: heimo.breiteneder@meduniwien.ac.at.

0091-6749/$32.00

doi:10.1016/j.jaci.2007.08.019

Abbreviation used
FARRP: Food Allergy Research and Resource Programme

characterized as food allergens. The InformAll set included 6 *Anisakis simplex* allergens, of which 2 were omitted because the sequences could not be assigned to a Pfam family, whereas FARRP included 5 *Anisakis simplex* allergens which were included despite being listed as worm rather than food allergens. Three tropomyosin entries without full deposited sequences from Indian prawn (*Fenneropenaeus indicus*), horned turban (*Turbo cornutus*), and squid (*Todarodes pacificus*), together with an arginine kinase from the Neptune rose shrimp (*Parapenaeus fissurus*), were included in the family count because the published peptide sequences allowed these allergens to be assigned to Pfam families. The final list of sequences included is given in this article's Table E1 in the Online Repository at www.jacionline.org.

Protein family assignment

The 2 allergen protein sets were assigned by using the Pfam version 21.0 database.[5] The Pfam database classifies protein sequences into families on the basis of sequence homology, which is related to conserved 3-dimensional structures and possible function. Pfam 21.0 contains 8957 families, with Pfam-A including more than 75% of all protein sequences reported in the SWISS-PROT and TReMBL databases. Where multiple domains were found within a single protein, they were counted only once, but nonidentical domains contributed to the scores of both families. The number of domains in genomes and the ranking of domains in genomes were taken from http://www.sanger.ac.uk/cgi-bin/Pfam/genome_dist.pl. Protein sequences were aligned by using T-Coffee,[6] and dendrograms were drawn with TREEVIEW (GubuSoft, Belmont, Mass).[7]

Structural bioinformatic analysis of allergen surfaces

Surface areas were calculated by MSMS[8] (http://molbio.info.nih.gov/structbio/basic.html). Conservation of surface atoms was displayed by using MOLMOL.[9] The surface areas calculated and displayed were solvent (ie, aqueous or nonaqueous) exposed areas, also termed the *contact surface*. The calculated surface area of carp β-parvalbumin was slightly smaller (5679 Å^2) than that of human α-parvalbumin (6005 Å^2). This difference arose because of the inclusion of an N-terminal methionine in the human structure (~100 Å^2) and because the surface of the nuclear magnetic resonance spectroscopy model of the human α-parvalbumin was slightly rougher than that of the carp β-parvalbumin model derived from an X-ray structure. This was supported by the fact that the X-ray structure of rat α-parvalbumin, which has 91% sequence identity with human α-parvalbumin, had a very similar surface area to carp β-parvalbumin (5757 Å^2).

RESULTS

Bioinformatic analysis of animal food allergens

The distribution of animal food allergens among Pfam-defined protein families is shown in Fig 1. The majority of animal food allergens could be classified as belonging to 1 of 3 families—tropomyosins, EF-hand proteins, and caseins—together with a tail of 14 families containing

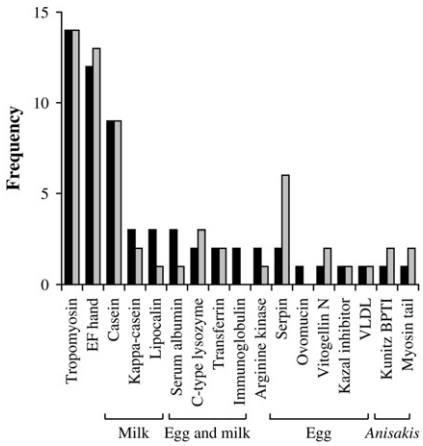

FIG 1. Distribution of animal food allergens into Pfam families. Allergens were from either version 2.1 of the InformAll (www.foodallergens.ifr.ac.uk; *dark bars*) or version 5.0 of FARRP (www.allergenonline.com; *gray bars*) databases. Pfam family names are given apart from very low density lipoprotein (VLDL), arginine kinase, and ovomucin, which represent proteins with multiple different domains. *BPTI*, Bovine pancreatic trypsin inhibitor.

only 1 to 3 reported allergens each. This situation is similar to that of plant food and pollen allergens.[2,3] Essentially the same distribution was found using the allergens from the InformAll or the Allergenonline (FARRP) search sets, which were composed of 57 and 60 sequences, respectively. The differences in composition of these search sets reflect differences in their inclusion criteria for each of these databases.[10] InformAll is an allergenic foods database that includes proteins as allergens on the basis of a clear demonstration of IgE binding (immunoblot or immunoassay) using sera from individuals with allergy to a particular food. In contrast, version 5 of the Allergenonline (FARRP) sequence database used in this study took a broader definition for inclusion of proteins without a strict requirement for demonstration of IgE binding. The smaller data set for animal food allergens reflects the facts that human beings use a less diverse range of animal-derived compared with plant-derived foods and that allergenic homologs found in milk or eggs from different species are not well characterized. Allergens from the herring worm, *Anisakis simplex*, were included in the search sets because they are consumed in raw fish and are recognized as triggering IgE-mediated allergies. However, it is a matter of debate whether this is a classic food allergic reaction or part of a normal immune reaction to a parasitic infection even though *Anisakis* is only partially infective in human beings.

As was observed for plant food allergens,[2] the ranking of allergenic protein families was not related to the frequency of a protein family in the human (*Homo sapiens*), zebra fish (*Danio rerio*), or *Drosophila* genomes (see this article's Table E1 in the Online Repository at www.jacionline.org). The highly restricted distribution of animal food, plant food, and pollen allergens is striking and emphasizes the fact that although in theory all proteins

have the potential to become allergens, in practice this is not the case.

Evolutionary relationships of major animal food allergen families

The ability of the immune system to discriminate self from nonself could mean animal food proteins with a high degree of similarity to human homologs would be poorly immunogenic in human beings and hence less likely to become allergens. We have therefore investigated the evolutionary relationships of the 3 main animal food allergen families, particularly in relation to human beings, and the allergenic potency of the different family members.

There are generally 4 types of muscle tropomyosins in mammals, of which tropomyosin-4 is the most divergent. This is probably because it is functionally different. It binds to actin filaments in nonmuscle cells such as fibroblasts as well as in muscle. In general, vertebrate tropomyosin sequences from other mammals, birds, and fish are at least 90% identical with at least 1 human tropomyosin. None of these have been reported to be allergenic[11] (Fig 2, *A*). In contrast, allergenic tropomyosins are confined to the invertebrate groups, arthropods (especially insects and crustaceans), and molluscs[12] and are at most only 55% identical to and equidistant from the closest human homologs (Table I). The arthropod tropomyosins from insects, crustacea, and nematodes (*Anisakis simplex*) are similar to each other and more closely related to mollusc tropomyosins than they are to vertebrate tropomyosins (Fig 2, *A*). Tropomyosins have also been identified as inhalant allergens, including those from 2 different species of cockroach (Per a 7 and Bla g 7) and house dust mite (Der p 10). These behave in a similar fashion to tropomyosins from crustacea with 77% to 78% sequence identity to shrimp (*Metapenaeus ensis*) tropomyosin but only 54% to 56% sequence identity to human tropomyosins. This explains the IgE and clinical cross-reactivity observed on occasion between these inhalant tropomyosins and those characterized as food allergens.[13]

The sharp division between allergenic and nonallergenic sequences probably reflects the way the human body discriminates self from nonself, because any antibody response to vertebrate tropomyosins might be expected to cause autoimmune disease. Tropomyosins with sequences greater than 90% identical to human tropomyosins are unlikely to be allergenic because they are unlikely to be immunogenic, as indicated by the reduction or loss of IgE binding of humanized shrimp tropomyosin IgE-epitopes modified to resemble vertebrate tropomyosins, on the basis of consensus sequences derived from a number of species but excluding human beings.[14]

The second largest animal food allergen family, the EF-hand, is composed almost entirely of parvalbumins, apart from a single troponin allergen from *Anisakis simplex*. Because troponins are functionally distinct from parvalbumins, subsequent analysis of EF-hand members was restricted to the latter. Parvalbumins can be subdivided into 2 distinct evolutionary lineages known as the

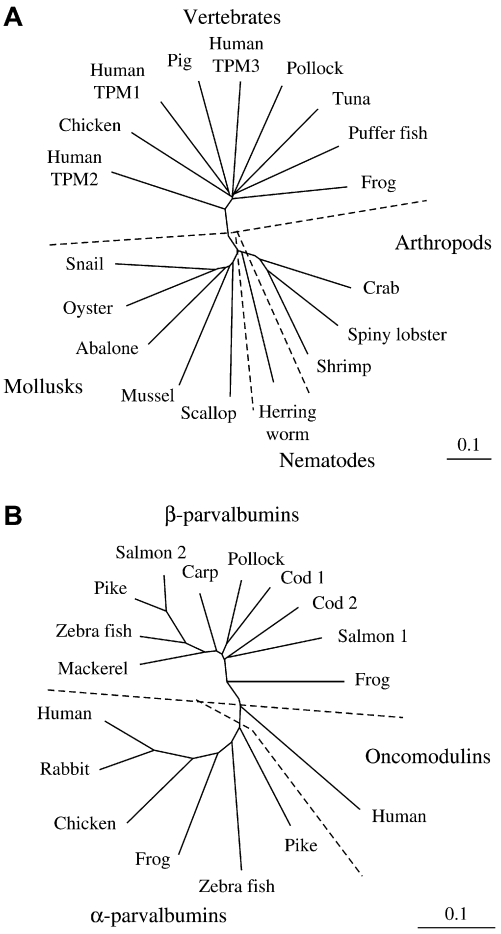

FIG 2. Unrooted dendrograms of tropomyosin **(A)** and parvalbumin **(B)** sequences. Representative allergens including the closest human homologs and homologs present in foods that have not been reported as allergens were selected for each family. The *scale bar* shows the branch length corresponding to 0.1 changes per residue, and identities were ordered by similarity to human α-parvalbumin.

α-parvalbumins and β-parvalbumins, although overall their folds are very similar (Table II; Fig 2, *B*). The α-parvalbumins are abundant in the muscle of fish and amphibians, rather less so in the fast twitch muscle of birds and mammals, and are not generally allergenic, with only 1 reported in frog.[15] In contrast, many β-parvalbumin allergens are found in a variety of fish species, and although they have a similar distribution to α-parvalbumins, they are absent from human muscle, showing only 56% identity to human α-parvalbumin (Fig 2, *B*).[12,15,16]

The differences in IgE cross-reactivity between β-parvalbumins from different fish species and the general lack of cross-reactivity with α-parvalbumins could result from conservation of surface structures, as has been suggested for certain plant food allergens,[2] especially because allergy to this protein involves the folded form.[17] The allergenic cod β-parvalbumin was compared with the only human parvalbumin, α-parvalbumin (Fig 3, *C and D*). When the sequences of different β-parvalbumins

TABLE I. Percentage identity of tropomyosins relative to those from human beings (*Homo sapiens*) and shrimp (*Metapenaeus ensis*), ordered by similarity to TPM1 from human being

Species	Allergenicity*	Sequence accession code	Percent identity	
			Human TPM1	Shrimp TPM Met e 1
Human TPM1 (*Homo sapiens*)	No	P09493	100	52
Chicken (*Gallus gallus*)	No	P04268	95	55
Frog (*Rana esculenta*)	No	P13105	94	53
Tuna (*Thunnus tonggol*)	No	BAD01050	93	53
Oyster (*Crassostrea gigas*)	Cra g 1	Q95WY0	44	54
Abalone (*Haliotis diversicolor*)	Hal d 1†	Q9GZ71	54	59
Snail (*Helix aspersa*)	Hel as 1	O97192	55	60
Spiny lobster (*Panulirus stimpsoni*)	Pan s 1	O61379	53	98
Crab (*Charybdis feriatus*)	Cha f 1	Q9N2R3	54	81
Greasybacked shrimp (*Metapenaeus ensis*)	Met e 1	Q25456	52	100
Herring worm (*Anisakis simplex*)	Ani s 3	Q9NAS5	58	69

*Official International Union of Immunological Sciences names (www.allergen.org) given unless marked†.

TABLE II. Percentage identity of α- and β-parvalbumins (PRVA, PRVB) ordered by similarity to human PRVA

Species	Allergenicity*	Accession code	Percent identity	
			Human (*Homo sapiens*) PARV	Cod (*Gadus morhua*) β-PARV
Human (*Homo sapiens*) PRVA	None	P20472	100	53
Chicken (*Gallus gallus*) PRVA	Not known	P80026	78	52
Frog (*Rana esculenta*) PRVA	No known	Q8JIU2	67	46
Pike (*Esox lucius*) PRVA	Not known	P02628	62	56
Frog (*Rana esculenta*) PRVB	Ran e 1	P02617	55	67
Carp (*Cyprinus carpio*) PRVB	Cyp c 1.02†	Q8UUS2	56	77
Salmon (*Salmo salar*) PRVB2	Sal s 1	Q91483	55	75
Pike (*Esox lucius*) PRVB	Not known	P02619	55	73
Cod (*Gadus morhua*) PRVB 1	Gad c 1	Q90YK9	53	100
Alaska pollack (*Theragra chalcogramma*) PRVB	The c 1†	Q90YK7	49	76

*Official International Union of Immunological Sciences names (www.allergen.org) given unless marked†.

were mapped onto the surface of either model, a nonrandom pattern of conservation was observed (Fig 3). One face was highly conserved (Fig 3, *F*), corresponding to the 2 calcium-binding segments, the other showing many substitutions corresponding to 1 of 3 repeating units (AB) and the loops linking the other repeated motifs (Fig 3, *C*). From this, IgE-binding sites located in the conserved calcium-binding face would be highly likely to show IgE cross-reactivity between α-parvalbumins and β-parvalbumins from most vertebrate muscles, including human beings. α-Parvalbumins have not been demonstrated as allergens in fish such as cod,[17] suggesting the potential for IgE to bind to this face is limited by clonal deletion as the human immune system matures. This would avoid setting up what could potentially be an autoimmune response. IgE epitopes have only been identified for cod parvalbumin,[18] 1 of which (residues 88-96) is highly conserved in both α-parvalbumins and β-parvalbumins and common to human beings, frog, and fish. Although representing a potential cross-reactive IgE binding site, it implies that the replacement of lysine 91 in human α-parvalbumin with a serine in fish β-parvalbumin is sufficient to stop IgE recognizing human parvalbumin, similar to the small differences between allergenic and human tropomyosins.[14] The variable AB surface is much more conserved between fish and frog (Fig 3, *B*) and might be anticipated to be the location of IgE epitopes responsible for IgE cross-reactivity observed across a wide range of fish species and frog.[16]

The third animal food allergen family, caseins, are exclusively mammalian proteins found in milk. β-Caseins from various mammals show 53% to 58% identity to human β-casein, although the identities between β-caseins, αS1-caseins, and αS2-caseins from goat (*Capra hircus*) and sheep (*Ovis aries*) relative to cow (*Bos taurus*) are more than 90% (Table III; Fig 4). Because there is no αS2-casein gene expressed in human beings, sequences from various other mammals were compared with the closer of the human αS1-casein or β-casein sequences. The known allergens are all less than 53% identical to corresponding human sequences, with the αS2-caseins least similar, with only ~16% identity (Table III; Fig 4). This

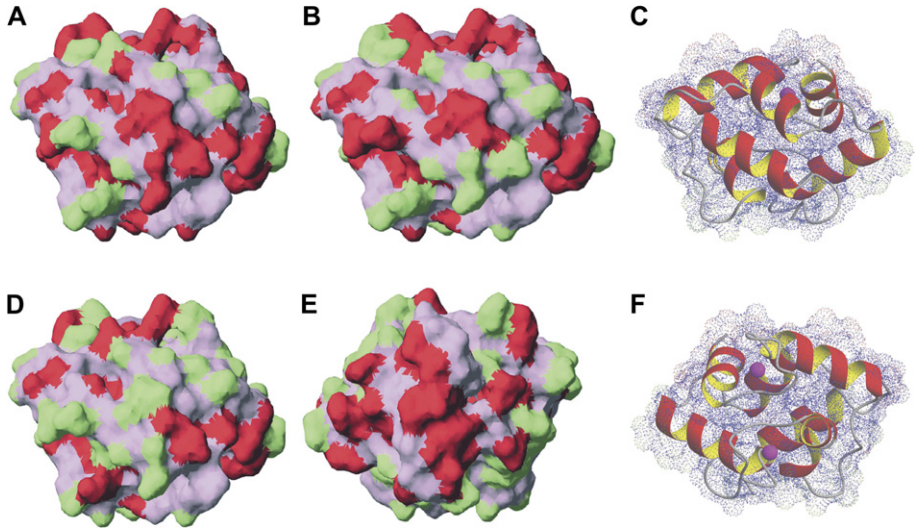

FIG 3. Conservation of residues on the contact surface of parvalbumins. **A, B, D, E,** The surface of carp β-parvalbumin colored to show conserved side *(red)* and main chain residues *(pink)*, substituted side chain residues *(green)*, and the C-terminal main chain *(green)*. The ribbon diagram of carp parvalbumin shows the less conserved **(C)** and more conserved faces **(F)** with the calcium atoms as *pink balls* and a dotted molecular contact surface. Models show the less conserved face comparing *(A)* cod (Q90YK9) to carp (P02618), *(B)* cod (Q90YK9) to frog (P02617), and *(D)* cod (Q90YK9) to human being (P20472). *E,* The rotation of 180° of *D* around a vertical axis shows the more conserved calcium binding face comparing cod (Q90YK9) to human being (P20472).

TABLE III. Percentage identity of caseins ordered by similarity to the closest bovine homolog

Casein	Accession code	Percent identity to closest: Bovine homolog	Percent identity to closest: Human homolog
αS1-Casein			
Cow	P02662	100	29
Goat	Q8MIH4	88	29
Sheep	P04653	88	28
Rat	P02661	22	27
Horse	Q8SPR1	39	44
Human being	P47710	29	100
Camel	O97943	41	36
Rabbit	P09115	37	37
αS2-Casein			
Cow	P02663	100	16
Goat 1	P33049	88	17
Goat 2	Q9TTQ7	87	17
Sheep	P04654	89	17
Camel	O97944	56	11
Rabbit	P50418	36	16
β-Casein			
Cow	P02666	100	53
Goat	Q712N8	91	54
Sheep	P11839	91	54
Horse	Q9GKK3	56	58
Human	P05814	53	100
Camel	Q9TVD0	66	58
Rabbit	P09116	52	55

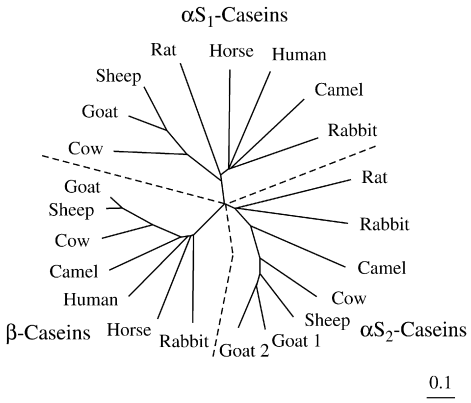

FIG 4. Unrooted dendrogram of α-caseins and β-caseins. The human αS2-casein sequence is not expressed and was not used in this analysis. The *scale bar* shows the branch length corresponding to 0.1 changes per residue.

appears to be linked to their IgE reactivity, with caseins least like human caseins being more reactive. Thus, 90% of a group of infants with cow's milk allergy had serum IgE against αS2-casein, and 55% against αS1-casein, with only 15% having IgE against β-casein, the bovine casein with highest identity to human casein.[19] This observation suggests that the 53% identity between bovine and human β-casein is already restricting IgE responses. These phylogenetic relationships are consistent with observations of IgE cross-reactivity between caseins and patterns of allergies to different species of milk, because anticasein IgE of patients with cow's milk allergy binds to homologous proteins from various animal species.[20,21]

This is consistent with *in vivo* observations that patients with cow's milk allergy generally react to goat's milk on oral challenge[22] but are able to tolerate mare's milk,[23] with horse caseins having sequence identities of 22% to 66% compared with cow. The ability of the immune system to discriminate between related proteins is also emphasized by the observation that individuals whose primary sensitization appears to be toward goat's milk caseins as a result of cheese consumption showed much greater IgE cross-reactivity with sheep's milk proteins than those from cow's milk to such an extent that they could generally tolerate cow's milk products.[24] Such observations reflect the fact that goat and sheep's milk proteins are more closely related to each other than either is to cow's milk proteins (Table III).

Minor animal food allergen families

The tail of minor allergens families shown in Fig 1 displays similar evolutionary relationships to those described in detail for the 3 main families. Several of these minor families have small numbers of members found exclusively in eggs[25,26] and milk[27] that generally show identities of no more than 76% with the closest human homolog and can include representatives of important inhalant allergen families such as the lipocalins.[28] The commonality of egg and milk allergens may originate from their role in providing nutrients to the young (Fig 1).

Three of the minor protein families contain allergens from species more distantly related to human beings, including the Kunitz inhibitors and paramyosin represented by *Anisakis* allergens,[29] and the arginine kinase family that contains important IgE cross-reactive inhalant allergens of a number of insects and mites.[30]

DISCUSSION

Previous analyses of plant food[2] and pollen allergens[3] have indicated that most belong to only a small number of protein superfamilies. Around 65% of plant food allergen sequences are members of just 4 protein families: the prolamin, cupin, Bet v 1–like, and profilin families.[2] Unlike plant allergens, almost all the animal food allergen families have homologs in the human proteome that may affect the way in which they are recognized by the human immune system. We have classified animal food allergens into 3 main families—tropomyosins, EF hand proteins, and caseins—together with a long tail of families containing only 1 to 3 reported allergens each. We have then considered the evolutionary relationships of the 3 main animal food allergen families, particularly in relation to human proteins and the allergenic potency of the different family members.

For all 3 main animal allergen families, their ability to act as allergens seems to be related to their closeness to human homologs. Thus, allergenic invertebrate tropomyosins are at most only 55% identical to the closest human homolog, whereas nonallergenic tropomyosins from mammals, birds, and fish are at least 90% identical.[12] This distinction possibly reflects the evolution of vertebrates from invertebrates that occurred with the appearance of the jawless fish around 500 million years ago.

The second largest animal food allergen family, the parvalbumins, has only been identified in vertebrates and shows slightly different evolutionary relationships. Of the 2 distinct evolutionary lineages of parvalbumins, only members of the β-parvalbumin family are allergenic, a form not found in human muscle.[12,15,16] The third largest animal food allergen family is even more confined, with caseins being exclusively mammalian proteins. Again, the extent of similarity between different caseins and their human counterparts appears to be linked with their allergenic reactivity, as is the case for the other allergen families.

The tropomyosin and EF-hand families appear to be intrinsically likely to be allergens. That being the case, it is legitimate to ask why vertebrate troponins and tropomyosins have not been reported as allergens at all and α-parvalbumins have been only very rarely reported as allergens. The most plausible answer is that the existence of closely related human homologs leads to these proteins being tolerated.

We have now demonstrated that the distribution into protein families of allergens of animal food, plant food, and pollen is highly restricted, emphasizing the fact that although, in theory, all proteins have the potential to become allergens, this is not usually the case. It may be that both the source of the allergens and routes of sensitization are important in determining which protein families predominate, with different families represented in pollen allergen families[3] and food allergen families. Further analysis is required to determine whether this pattern holds true for allergens involved in skin sensitization. It is intriguing that certain protein families appear to retain their importance irrespective of source or routes of sensitization, with for example the tropomyosins, EF-hands, and lipocalins appearing as both inhalant and food allergens.

In addition, we have demonstrated that the allergenicity of protein family members largely decreases as a function of their relatedness to human homologs, proteins closely related to human homologs not being allergenic. Thus, in the sequence set we analyzed, proteins with a sequence identity of $\leq$54% to human homologs were all allergenic, whereas those with a sequence identity greater than 63% to human homologs were rarely allergenic. This observation probably relates to the requirement for proteins to be recognized as nonself to mount an immune response, and it has been argued that a low degree of similarity to a host's proteome is required for immunogenicity.[31] The only exception is the cow's milk allergen bovine serum albumin that is 76% identical to the human homolog. It is generally thought this protein is a much less important cow's milk allergen than, for example, the caseins. It may be that as a consequence of the homology between human and bovine serum albumin, the latter stimulates a less effective immune response than other cow's milk proteins. The closeness between animal allergens and their human homologs has potential to be exploited in the development

of recombinant immunotherapeutics mutagenized to resemble human homologs more closely.[14] However, in certain individuals, immunologic responses mounted to certain animal food allergens seem close to becoming a form of autoimmune response, which could potentially be one side-effect of novel immunotherapy strategies. The *in silico* analysis we have undertaken provides a theoretical framework for future experimental studies to test the observation regarding the need for proteins to share sequence identity of less than around 62% with a host to be able to sensitize that host.

We thank Alan R. Mackie for carefully reading the manuscript.

REFERENCES

1. Sicherer SH, Sampson HA. 9. Food allergy. J Allergy Clin Immunol 2006;117:S470-5.
2. Jenkins JA, Griffiths-Jones S, Shewry PR, Breiteneder H, Mills EN. Structural relatedness of plant food allergens with specific reference to cross-reactive allergens: an in silico analysis. J Allergy Clin Immunol 2005;115:163-70.
3. Radauer C, Breiteneder H. Pollen allergens are restricted to few protein families and show distinct patterns of species distribution. J Allergy Clin Immunol 2006;117:141-7.
4. Spitzauer S. Allergy to mammalian proteins: at the borderline between foreign and self? Int Arch Allergy Immunol 1999;120:259-69.
5. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, et al. The Pfam protein families database. Nucleic Acids Res 2004;32:D138-41.
6. Notredame C, Higgins DG, Heringa J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 2000;302:205-17.
7. Page RD. TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 1996;12:357-8.
8. Sanner MF, Olson AJ, Spehner JC. Reduced surface: an efficient way to compute molecular surfaces. Biopolymers 1996;38:305-20.
9. Koradi R, Billeter M, Wuthrich K. MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 1996;14:51-5, 29-32.
10. Gendel SM, Jenkins JA. Allergen sequence databases. Mol Nutr Food Res 2006;50:633-7.
11. Ayuso R, Lehrer SB, Tanaka L, Ibanez MD, Pascual C, Burks AW, et al. IgE antibody response to vertebrate meat proteins including tropomyosin. Ann Allergy Asthma Immunol 1999;83:399-405.
12. Wild LG, Lehrer SB. Fish and shellfish allergy. Curr Allergy Asthma Rep 2005;5:74-9.
13. Fernandes J, Reshef A, Patton L, Ayuso R, Reese G, Lehrer SB. Immunoglobulin E antibody reactivity to the major shrimp allergen, tropomyosin, in unexposed Orthodox Jews. Clin Exp Allergy 2003;33:956-61.
14. Reese G, Viebranz J, Leong-Kee SM, Plante M, Lauer I, Randow S, et al. Reduced allergenic potency of VR9-1, a mutant of the major shrimp allergen Pen a 1 (tropomyosin). J Immunol 2005;175:8354-64.
15. Hamada Y, Nagashima Y, Shiomi K. Reactivity of serum immunoglobulin E to bullfrog Rana catesbeiana parvalbumins in fish-allergic patients. Fish Sci 2004;70:1137-43.
16. Hilger C, Thill L, Grigioni F, Lehners C, Falagiani P, Ferrara A, et al. IgE antibodies of fish allergic patients cross-react with frog parvalbumin. Allergy 2004;59:653-60.
17. Bugajska-Schretter A, Grote M, Vangelista L, Valent P, Sperr WR, Rumpold H, et al. Purification, biochemical. immunological characterisation of a major food allergen: different immunoglobulin E recognition of the apo- and calcium-bound forms of carp parvalbumin. Gut 2000;46:661-9.
18. Elsayed S, Apold J, Holen E, Vik H, Florvaag E, Dybendal T. The structural requirements of epitopes with IgE binding capacity demonstrated by three major allergens from fish, egg and tree pollen. Scand J Clin Lab Invest Suppl 1991;204:17-31.
19. Natale M, Bisson C, Monti G, Peltran A, Garoffo LP, Valentini S, et al. Cow's milk allergens identification by two-dimensional immunoblotting and mass spectrometry. Mol Nutr Food Res 2004;48:363-9.
20. Bernard H. IgE cross-reactivity with caseins from different species in humans allergic to cow's milk. Food Agricult Immunol 1999;11:101-11.
21. Restani P, Gaiaschi A, Plebani A, Beretta B, Cavagni G, Fiocchi A, et al. Cross-reactivity between milk proteins from different animal species. Clin Exp Allergy 1999;29:997-1004.
22. Bellioni-Businco B, Paganelli R, Lucenti P, Giampietro PG, Perborn H, Businco L. Allergenicity of goat's milk in children with cow's milk allergy. J Allergy Clin Immunol 1999;103:1191-4.
23. Businco L, Giampietro PG, Lucenti P, Lucaroni F, Pini C, Di Felice G, et al. Allergenicity of mare's milk in children with cow's milk allergy. J Allergy Clin Immunol 2000;105:1031-4.
24. Ah-Leung S, Bernard H, Bidat E, Paty E, Rance F, Scheinmann P, et al. Allergy to goat and sheep milk without allergy to cow's milk. Allergy 2006;61:1358-65.
25. Langeland T. A clinical and immunological study of allergy to hen's egg white, VI: occurrence of proteins cross-reacting with allergens in hen's egg white as studied in egg white from turkey, duck, goose, seagull, and in hen egg yolk, and hen and chicken sera and flesh. Allergy 1983;38:399-412.
26. Poulsen LK, Hansen TK, Norgaard A, Vestergaard H, Stahl Skov P, Bindslev-Jensen C. Allergens from fish and egg. Allergy 2001;56(suppl 67):39-42.
27. Bernard H, Creminon C, Yvon M, Wal JM. Specificity of the human IgE response to the different purified caseins in allergy to cow's milk proteins. Int Arch Allergy Immunol 1998;115:235-44.
28. Mantyjarvi R, Rautiainen J, Virtanen T. Lipocalins as allergens. Biochim Biophys Acta 2000;1482:308-17.
29. Perez-Perez J, Fernandez-Caldas E, Maranon F, Sastre J, Bernal ML, Rodriguez J, et al. Molecular cloning of paramyosin, a new allergen of Anisakis simplex. Int Arch Allergy Immunol 2000;123:120-9.
30. Binder M, Mahler V, Hayek B, Sperr WR, Scholler M, Prozell S, et al. Molecular and immunological characterization of arginine kinase from the Indianmeal moth, Plodia interpunctella, a novel cross-reactive invertebrate pan-allergen. J Immunol 2001;167:5470-7.
31. Kanduc D, Lucchese A, Mittelman A. Individuation of monoclonal anti-HPV16 E7 antibody linear peptide epitope by computational biology. Peptides 2001;22:1981-5.