7. Dust mite allergens

Biologists work very close to the frontier between bewilderment and understanding. Biology is complex, messy and richly various, like real life; it travels faster nowadays than physics or chemistry (which is just as well, because it has so much farther to go), and it travels nearer to the ground.

Peter Medawar,
Pluto's Republic (1982),
Oxford University Press, Oxford.

7.1 Why do dust mites produce allergens?

There can be few tasks in biological research messier than emptying a vacuum cleaner bag of dust, sieving it and preparing an aqueous suspension in order to make allergen extracts or remove mites. The fine grey powder gets everywhere and is not very wettable. It sticks to glassware, and the smell of old house dust and decomposing skin scales is like a mixture of the rancid fat of dirty kitchens and the mustiness of neglected libraries. In the 1920s and 1930s it was thought house dust contained a single allergen. Making aqueous extracts of house dust would have been a standard method. By the 1950s, it was becoming clear that house dust contained a complex mixture of allergens (see Voorhorst *et al.*, 1969). By the late 1960s dust mites were known to be a major component of this mix. By the early 1970s it had been determined that the faecal pellets of dust mites contained allergens. Halmai and Alexander (1971) isolated faecal pellets from the bodies of mites using a micromanipulator and used them in skin-prick tests to demonstrate their allergenicity. In 1981, Tovey *et al.* discovered that mite faeces were a major source of the dust mite allergen Der p 1. We now know of a large variety of allergens, most characterised to the last amino acid, many for which their original biological function is known, and several with their tertiary structure modelled.

Most allergens are biochemically active molecules and include enzymes, enzyme inhibitors, and proteins involved in molecular transport, regulation and cell and tissue structure. Reviews of mite allergens include those by Platts-Mills and Chapman (1987); International Workshop Report (1988); Platts-Mills *et al.* (1989b); Arlian (1991); Arruda and Chapman (1992); Stewart, Bird, Krska *et al.* (1992); Stewart (1994); Stewart and Thompson (1996); Colloff and Stewart (1997); Stewart and McWilliam (2001); Arlian (2002); Thomas *et al.* (2002); Kawamoto *et al.* (2002b); Stewart and Robinson (2003); and Nuttall *et al.* (2006). Most of this work deals with allergens produced by the pyroglyphid mites *Dermatophagoides pteronyssinus*, *D. farinae*, *D. microceras* and *Euroglyphus maynei*. More recently, data on storage mites such as *Lepidoglyphus destructor* and *Blomia tropicalis* have emerged, as well as the parasitic mites *Sarcoptes scabiei* and *Psoroptes ovis* which are related to the Pyroglyphidae. It is now clear that each of the major groups of dust

contains several faecally derived allergens, some of which have been shown to be hydrolases corresponding to those commonly associated with digestion in both vertebrates and non-arachnid invertebrates.

The enzymatic nature of many allergens begs the question, 'What are the most abundant sources of enzymes in dust mites?' The answer is the digestive system. So what facets of digestion are responsible for the production of large amounts of allergens? The digestive system of dust mites appears to be rather inefficient by mammalian standards, and is partially intracellular (see Chapter 2). Relatively large amounts of enzymes, which in an organism that digests extracellularly would be recycled and partly reclaimed, are voided in the faecal pellet, so they are the most abundant proteins derived from dust mites present in the domestic environment. We begin to get a clue as to why the digestive enzymes of dust mites are important allergens.

In general, the spectrum of digestive enzymes detected in an organism indicates the degree of trophic specialisation, and there are some data suggesting that this may be the case with mites (Bowman, 1984). The enzymes (and, therefore, allergens) detected in faecal extracts may reflect the diet of the mite (Colloff, 1994a). For pyroglyphid mites, diet includes protein, lipid and carbohydrate derived from sources such as human skin scales, bacteria and fungi, whereas in storage mites such nutrients may be derived from seed proteins and carbohydrates as well as microorganisms. There may be quantitative and qualitative differences in the spectrum of enzymes produced by mites occupying different ecological niches and, ultimately, recognised as allergenic by susceptible individuals. For example, skin contains a number of structural proteins usually resistant to conventional proteases found in the digestive tracts of both invertebrates and vertebrates. These include collagen and keratin, and it is possible that the mite Der p 1 cysteine protease evolved to act in concert with a reducing environment in the mite gut in order to digest such proteins. Collagen and keratin are not present in grain, suggesting that a similar enzyme will not be prominent in the faeces of storage mites.

Mites are assumed to digest bacteria and fungi, and the most likely taxa to be consumed will be the common commensals such as *Staphylococcus*, *Streptococcus*, Mycobacteria, Propionibacteria and *Aspergillus* species. Whether mites with different diets such as the stored products mites possess the same range of enzymes (and, therefore, allergens) as those produced by the pyroglyphid mites is not yet clear despite an overall qualitative similarity in the spectrum of enzymes found in whole mite extracts. A more definitive answer must await a detailed examination of mite-free faecal extracts. Despite this, interspecies similarities with regard to physicochemical and biochemical properties of allergens are present which account for the observed immunological cross-reactivity. However, the absence of cross-reactivity does not preclude the presence of a particular allergen in a mite species since functionally equivalent proteins may exist.

With new techniques in molecular biology and bioinformatics over the last decade, whole allergens have been sequenced (see Table 7.1) and tertiary structures modelled (e.g. Topham *et al.,* 1994; Mueller *et al.,* 1997; Derewenda *et al.,* 2002). Development and growth of gene and protein databases and bioinformatics tools greatly facilitated comparisons of the primary structure of allergens with other proteins of known function. Allergenic identity of different allergen groups in different mite species and complete or partial amino acid sequence data are available for over a hundred mite allergens (see Table 7.1).

As I show in this chapter, the focus on allergens as enzymes associated with mite faecal pellets (which when they become airborne are regarded as a primary vehicle for sensitisation) is central to an ageing paradigm that is in need of revision. Several major allergen groups are proteins involved in intracellular regulation and processing (e.g. groups 13, 14, 16 and 17) and are not associated with the gut at all. Others are components of muscle tissue (groups 10 and 11). Almost half of all known allergen groups are not associated with faecal pellets of dust mites (see Table 7.2). The fact that atopic patients become sensitised to these allergens means that they must be exposed not only to faecal pellets but also to particles of cell and muscle tissue. These particles, especially cellular debris, are likely to be considerably smaller than the 15–20 μm-diameter faecal pellets. Potentially, the nature of the exposure to those allergens not associated with the digestive system is quite different to that of the faecal hydrolases, which, in any event, are also carried on particles much smaller than faecal pellets (De Lucca *et al.,* 1999). The epidemiological implications of differential allergen exposure in relation to particle size are examined in more detail in Chapter 8 (see section 8.4.4).

In this chapter I have not gone into any detail on the mechanisms involved in the immune responses by

Table 7.1 Dust mite allergens for which full or partial amino acid sequences have been published and/or databased. Short N-terminal amino acid sequences (ca. ≤ 30 residues) are not included. ND = no data; *Databases: GenBank, EMBL, UniProt/SwissProt. GenBank numbers are entry/locus number (not accession numbers); (I) = incomplete sequence; N/A = not applicable (e.g. not databased). Der f = *Dermatophagoides farinae*; Der p = *D. pteronyssinus*; Der s = *D. siboney*; Eur m = *Euroglyphus maynei*; Pso o = *Psoroptes ovis*; Sar s = *Sarcoptes scabiei*; Aca s = *Acarus siro*; Ale o = *Aleuroglyphus ovatus*; Sui m = *Suidasia medanensis*; Tyr p = *Tyrophagus putrescentiae*; Gly d = *Glycyphagus domesticus*; Lep d = *Lepidoglyphus destructor*; Blo t = *Blomia tropicalis*.

Allergen name	No. amino acids (mature protein)	Calculated molecular weight (kDa)	Predicted isoelectric point	Predicted charge	*N*-glycosylation sites	NCBI *Entrez* accession no.*	References
Group 1							
Der f 1	223	25.191	6.14	–1.5	1	N/A	Dilworth *et al.*, 1991
	223	25.148	6.14	–1.5	1	ABO34946	Yasuhara *et al.*, 2001
	146 (I)	16.852	8.21	6.0	1	AF194431	Park *et al.*, unpublished
	107 (I)	12.277	8.21	3.5	0	AF194432	Park *et al.*, unpublished
	210 (I)	23.548	6.38	–0.5	1	AF285763	Hao *et al.*, unpublished
	196 (I)	22.037	6.77	1.0	1	DQ185509	Piboonpocanun *et al.*, unpublished
Der p 1	78 (I)	8.304	5.49	–1.5	1	DEPMHDA	Thomas *et al.*, 1988
	222	25.394	6.61	0.5	1	N/A	Chua *et al.*, 1988
	210 (I)	23.509	6.5	0.0	1	DPDERPIG	Kent *et al.*, 1992
	222	25.019	5.92	–2.0	1	DPU11695	Chua *et al.*, 1993
	117 (I)	13.421	5.5	–1.5	1	AY947536	Lucentini *et al.*, unpublished
	133 (I)	14.965	4.91	–4.5	1	AF145247	Park *et al.*, unpublished
	222	24.992	5.92	–2.0		DQ185508	Piboonpocanun *et al.*, unpublished
Eur m 1	217 (I)	24.375	6.61	0.5	1	EMB X60073	Kent *et al.*, 1992
	223	25.129	6.38	–0.5	1	AF047610	Smith *et al.*, 1999
Pso o 1	184	20.278	8.03	3.5	1	AF495854	Lee *et al.*, 2002
Sar s 1	226	25.326	9.1	12.5	0	AY525148	Holt *et al.*, 2004
Blo t 1	221	25.126	8.36	9.5	0	AF277840	Mora *et al.*, 2003
	239	27.232	6.0	–1.5	0	AY291322	Chew *et al.*, unpublished
Group 2							
Der f 2	129	14.021	6.9	1.0	0	N/A	Trudinger *et al.*, 1991
	129	14.044	6.9	1.0	0	DEPDER1	Yuuki *et al.*, 1991
	129	14.076	6.9	1.0	0	DEPDER2	Yuuki *et al.*, 1991
	129	14.034	6.9	1.0	0	DEPDER3	Yuuki *et al.*, 1991
	129	14.063	6.5	0.0	0	AB195580	Tsuuki *et al.*, unpublished
	129	14.081	7.0	1.5	0	AJ862836	Nandy *et al.*, 2003
	129	14.045	6.9	1.0	0	AJ862837	Nandy *et al.*, 2003
	158	17.628	7.8	5.0	0	AF346905	Hao *et al.*, unpublished
	129	14.035	6.1	–1.0	0	AY066008	Jin *et al.*, unpublished
	129	14.123	7.9	3.0	0	DQ185511	Piboonpocanun *et al.*, unpublished
	129	14.044	6.9	1.0	0	S70378	Okuhira, 1994
Der p 2	129	14.121	7.0	1.5	0	AF276239	Chua *et al.*, 1990a, b
Der p 2	129	14.008	7.0	1.5	0	N/A	Thomas & Chua, 1995
	129	14.088	7.4	2.5	0	DQ185510	Piboonpocanun *et al.*, unpublished
Der s 2	129	14.043	6.9	1.0	0	DQ367850	Jorge *et al.*, unpublished
Eur m 2	129	14.086	6.5	0.0	0	AF047613	Smith *et al.*, 1999
Pso o 2	126	13.468	6.5	0.0	0	AF187083	Temeyer *et al.*, 2002
Ale o 2	129	13.876	6.9	1.5	1	AY497898	Ramjan & Chew, unpublished
Sui m 2	126	13.493	7.1	2.0	0	AY497879	Reginald *et al.*, unpublished
Tyr p 2	128	13.54	8.1	3.5	0	EMB Y12690	Eriksson *et al.*, 1998
Gly d 2	128	13.79	4.7	–5.0	1	GDO249864	Gafvelin *et al.*, 2001
Gly d 2.02	125	13.366	8.4	4.0	0	GDO272216	Gafvelin *et al.*, 2001
Gly d 2.03	125	13.183	7.8	.02	0	AY288673	Chew *et al.*, unpublished
Lep d 2	125	13.108	6.5	0.0	0	N/A	Varela *et al.*, 1994
Lep d 2.0201b	125	13.108	6.5	0.0	0	LDRNAALL	Schmidt *et al.*, 1995
Lep d 2.0201a	125	13.147	6.9	1.0	0	LDLEPD1A	Schmidt *et al.*, 1995
Lep d 2.0202	125	13.175	6.9	1.0	0	AJ487973	Kaiser *et al.*, 2003
Lep d 2.0102	125	13.138	6.5	0.0	0	AJ487972	Kaiser *et al.*, 2003
Lep d 2.013	125	13.183	7.8	2.0	0	AY288142	Chew *et al.*, unpublished
Lep d 2.017	123	12.998	7.2	1.0	0	AY288143	Chew *et al.*, unpublished

continued

Table 7.1 Continued

Allergen name	No. amino acids (mature protein)	Calculated molecular weight (kDa)	Predicted isoelectric point	Predicted charge	*N*-glycosylation sites	NCBI *Entrez* accession no.*	References
Lep d 2.023	125	13.183	7.8	2.0	0	AY288144	Chew *et al.*, unpublished
Lep d 2.024	125	13.183	7.2	1.0	0	AY288145	Chew *et al.*, unpublished
Lep d 2.025	125	13.169	7.2	1.0	0	AY288146	Chew *et al.*, unpublished
Lep d 2.031	125	13.125	8.1	3.0	0	AY288147	Chew *et al.*, unpublished
Lep d 2.035	125	13.183	7.8	2.0	0	AY288148	Chew *et al.*, unpublished
Lep d 2.039	125	13.183	7.8	2.0	0	AY288149	Chew *et al.*, unpublished
Lep d 2.042	125	13.184	7.8	2.0	0	AY288150	Chew *et al.*, unpublished
Blo t 2	124	13.528	6.9	1.0	0	AY288139	Chew *et al.*, unpublished
	127	13.530	6.7	0.5	0	AY288140	Chew *et al.*, unpublished
	126	13.518	6.3	–0.5	0	AY288141	Chew *et al.*, unpublished
Group 3							
Der f 3	232	24.954	5.7	–2.0	1	DFU54781	Smith & Thomas, 1996
	232	24.914	5.3	–3.0	1	DEPPDF3	Nishiyama *et al.*, 1995
Der p 3	129 (I)	14.595	5.5	–2.5	1	AF145248	Park *et al.*, unpublished
	229	24.987	8.0	4.0	1	DPU11719	Smith *et al.*, 1994
Eur m 3	232	25.031	6.1	–1.0	1	AF047615	Smith *et al.*, unpublished
Sar s 3	231	25.77	8.1	5.0	2	AY333071	Holt *et al.*, 2003
Gly d 3	219	22.734	7.4	2.0	1	AY288674	Chew *et al.*, unpublished
Lep d 3	219	22.499	7.4	2.0	1	AY291571	Chew *et al.*, unpublished
Blo t 3	231	23.787	8.35	5.0	0	AY090091	Flores *et al.*, 2003
Blo t 3	231	23.824	8.8	7.0	0	N/A	Cheong *et al.*, 2003
	231	23.843	8.1	4.0	0	AY291323	Chew *et al.*, unpublished
Group 4							
Der p 4	496	57.15	7.2	8.5	1	AF144060	Mills *et al.*, 1999
Eur m 4	496	57.341	7.4	10.5	1	AF144061	Mills *et al.*, 1999
Blo t 4	484	55.055	7.02	7.5	0	AY291324	Chew *et al.*, unpublished
Group 5							
Der f 5	113	13.614	5.6	–2.5	1	BAE45865	Tsukui *et al.*, unpublished
Der p 5	113	13.587	4.9	–4.5	0	CAA35692	Tovey *et al.*, 1989
	113	13.587	4.9	–4.5	0	S76340	Lin *et al.*, 1994
	113	13.529	5.1	–3.5	0	P14004	Lin *et al.*, 1994
	117	14.178	5.0	–3.5	0	DQ354124	Weghofer *et al.*, unpublished
Gly d 5.01	239	27.517	5.0	–8	0	AY288675	Chew *et al.*, unpublished
Gly d 5.02	233	26.406	5.1	–6.5	1	AY288676	Chew *et al.*, unpublished
Lep d 5	110	12.550	6.4	–0.5	0	Q9U5P2	Eriksson *et al.*, 2001
5.02	115	13.057	6.4	–0.5	0	AY288151	Chew *et al.*, unpublished
5.04	115	13.057	6.4	–0.5	0	AY288152	Chew *et al.*, unpublished
Blo t 5	72 (I)	8.373 (I)	N/A	N/A	0	AAB49396	Caraballo *et al.*, 1996
	114	13.497	4.9	5.0	1	O96870	Arruda *et al.*, 1995, 1997a
Group 6							
Der f 6	230	25.034	6.1	–1.0	0	AF125187	Kawamoto *et al.*, 1999
Der p 6	231	24.885	5.3	–3.0	0	N/A	Bennett & Thomas 1996
Blo t 6	230	24.743	5.7	–2.0	0	AY291325	Chew *et al.*, unpublished
Group 7							
Der f 7	196	21.863	4.7	–10.0	1	S80655	Shen *et al.*, 1995b
Der p 7	198	22.179	4.6	–11.0	1	DPU37044	Shen *et al.*, 1993
Gly d 7	200	22.198	9.7	6.5	1	AY288677	Chew *et al.*, unpublished
Lep d 7	196	21.951	7.9	3.0	0	AJ271058	Eriksson *et al.*, 2001
Blo t 7	177 (I)	N/A	N/A	N/A	1	AY291326	Chew *et al.*, unpublished
Group 8							
Der p 8	219	25.589	6.8	0.5	2	S75286	O'Neill *et al.*, 1994a
	219	25.668	8.5	4.0	2	AY825938	Dougall *et al.*, 2005
	217	25.621	9.0	4.5	1	AY825939	Dougall *et al.*, 2005

continued

Table 7.1 Continued

Allergen name	No. amino acids (mature protein)	Calculated molecular weight (kDa)	Predicted isoelectric point	Predicted charge	*N*-glycosylation sites	NCBI *Entrez* accession no.*	References
Pso o 8	219	25.944	7.8	3.5	1	AF078684	Lee *et al.*, 1999
Sar s 8	219	25.841	8.3	2.5	1	AY825933	Dougall *et al.*, 2005
Gly d 8	214	24.763	9.0	5.5	3	AY288679	Chew *et al.*, unpublished
Lep d 8	214	24.793	9.0	5.5	3	AY291572	Chew *et al.*, unpublished
Blo t 8	236	27.630	4.8	–8.0	1	AY283284	Chew *et al.*, unpublished
Group 9							
Der f 9	132 (I)	14.374	9.5	5.0	0	AF194430	Park *et al.*, unpublished
Der p 9	220 224 228	23.756 23.766 24.485	6.8 9.0 8.5	0.5 8.0 5.0	2 1 1	AF409110 AY211952 AF409111	King *et al.*, unpublished Bi *et al.*, unpublished Mathaba *et al.*, unpublished
Blo t 9	224	23.598	8.6	6.5	0	AY291327	Chew *et al.*, unpublished
Group 10							
Der f 10	284	32.955	4.4	–23.0	0	DEPMAG44	Aki *et al.*, 1995
Der p 10	284 284 281	32.901 32.973 32.553	4.5 4.5 4.5	–22.5 –22.5 –22.5	0 0 0	DPY14906 AF016278 DQ247970	Asturias *et al.*, 1998 Smith *et al.*, unpublished Lucentini *et al.*, unpublished
Pso o 10	284	32.914	4.5	–20.5	0	AM114276	Nisbet *et al.*, 2006a
Tyr p 10	284	32.955	4.4	–27.5	1	AY623832	Lee *et al.*, unpublished
Gly d 10	284	32.674	4.3	–27.5	0	AY288684	Chew *et al.*, unpublished
Lep d 10	284	32.949	4.4	–23.0	0	AJ250096	Saarne *et al.*, 2003
Blo t 10	284	33.003	4.5	–22.5	0	N/A	Yi *et al.*, 2002
Der g 10	284	32.674	4.3	–27.5	2	AM167555	Nisbet *et al.*, 2006b
Group 11							
Der f 11	692 (I)	81.372	6.0	–8.0	2	AF352244	Tsai *et al.*, 1998
Der p 11	875	102.417	5.6	–16.5	4	AY189697	Lee *et al.*, 2004
Pso o 11	875	102.647	5.5	–16.5	3	AM114275	Nisbet *et al.*, 2006a
Sar s 11	679 (I) 876	79.784 102.454	5.8 5.7	–9.5 –14.0	2 4	DQ131648 AF317670	Zheng *et al.*, unpublished Mattsson *et al.*, 2001.
Blo t 11	875	102.028	5.7	–14.0	5	AF525465	Ramos *et al.*, 2001
Group 12							
Lep d 12	124	14.055	6.1	–3.5	0	AY293744	Chew *et al.*, unpublished
Blo t 12	125	14.337	5.6	–6.5	0	BTU27479	Puerta *et al.*, 1996a
Group 13							
Der f 13	132	14.980	6.8	0.0	0	2A0AA	Chan *et al.*, 2006
Aca s 13	130	14.297	5.3	–1.0	1	ASI6774	Eriksson *et al.*, 1999
Tyr p 13	132	14.573	6.9	0.0	1	AY710432	Jeong *et al.*, 2005a
Gly d 13	132	14.82	7.0	0.0	0	AY288679	Chew *et al.*, unpublished 2003
Lep d 13	132	14.723	7.0	0.0	0	AJ250279	Eriksson *et al.*, 2001
Blo t 13	131 130 130 131	14.8 14.814 14.814 14.862	5.4 5.4 5.4 5.3	–1.0 –1.0 –1.0 –1.5	0 0 0 0	BTU58106 AY283286 AY283287 AY283294	Caraballo *et al.*, 1997 Chew *et al.*, unpublished Chew *et al.*, unpublished Chew *et al.*, unpublished
Group 14							
Der f 14	341 (I)	39.668	7.5	4.5	0	DEPMAG	Aki *et al.*, 1994a
Der f 14	349 (I)	40.545	7.6	6.0	1	D17686	Fujikawa *et al.*, 1996
Der p 14	1788	205.737	7.9	28.5	1	AF373221	Epton *et al.*,1999, 2001a
Eur m 14	1650	189.541	8.2	31.5	1	AF149827	Epton *et al.*, 1999
Sar s 14	115 (I) 330 (I) 719 (I)	13.174	9.2	4.5	2	BM176880 AF462196 DQ109676	Harumal *et al.*, 2003 Fischer *et al.*, 2003 Ljunggren *et al.*, 2006
Blo t 14	338 (I)	39.353	9.6	9.0	1	AY090092	Flores *et al.*, unpublished
Group 15							
Der f 15	535	61.111	5.3	–10.5	1	AF178772	McCall *et al.*, 2001
Der p 15	538	61.413	5.1	–13.0	1	DQ078740	O'Neill *et al.*, 2006

continued

Table 7.1 Continued

Allergen name	No. amino acids (mature protein)	Calculated molecular weight (kDa)	Predicted isoelectric point	Predicted charge	*N*-glycosylation sites	NCBI *Entrez* accession no.*	References
Der p 15s	512	58.868	5.1	–12.0	1	DQ078741	O'Neill *et al.*, 2006
Group 16							
Der f 16	480	55.131	6.2	–2.5	3	AF465625	Kawamoto *et al.*, 2002a
Group 18							
Der f 18	437	49.468	6.0	–7.5	2	AYO93656	Weber *et al.*, 2003
Der p 18	437	49.227	5.8	–8.5	2	DQ078739	O'Neill *et al.*, 2006
Blo t 18	437	49.231	5.8	–8.0	2	AY291330	Chew *et al.*, unpublished
Group 19							
Blo t 19	70	7.226	8.4	5.0	0	N/A	Nge & Chua, unpublished
Group 20							
Der p 20	356	40.477	8.5	7.5	0	N/A	Thomas, unpublished
Group 21							
Der p 21	125	14.909	5.0	–3.5	0	ABC73706	Weghofer *et al.*, unpublished
Blo t 21	129	14.944	5.6	–2.5	0	AY800348	Gao *et al.*, 2007
Other allergens							
D. farinae							
Alt a homologue	490	54.197	5.4	–5.5	2	AY283296	Chew *et al.*, unpublished
Mag 29 (HSP 70)	145 (I)	15.594	4.4	–10.5	0	DEPMAG29	Aki *et al.*, 1994b
CPW1 (cathepsin F)	127	14.057	4.5	–5.5	2	AF145249	Park *et al.*, unpublished
D. pteronyssinus							
13.8 kDa enzyme	134	14.064	8.6	8.0	0	AF409109	Mathaba *et al.*, 2002
T. putrescentiae							
α-tubulin	450	50.037	4.8	–15.5	1	AY986760	Jeong *et al.*, 2005b
S. medanensis							
Profilin	130	14.127	6.5	0.0	1	AY803196	Reginald *et al.*, unpublished
G. domesticus							
Jun a 3 (thaumatin)	213	23.096	6.6	1.0	0	AY288680	Chew *et al.*, unpublished
Mal d 3 (lipid transfer protein)	90	9.476	7.3	2.0	0	AY288681	Chew *et al.*, unpublished
Mala s 6 (cyclophilin)	197	21.609	8.7	4.0	1	AY288682	Chew *et al.*, unpublished
Enolase homologue	362	39.639	5.5	–4.5	4	AY288683	Chew *et al.*, unpublished
L. destructor							
α-tubulin	450	50.038	4.8	–15.5	1	AJ428050	Saarne *et al.*, 2003
Cystatin	98	11.293	5.5	–2.5	0	AJ428051	Kaiser, unpublished
B. tropicalis							
Mag 29 (HSP 70)	650	71.495	5.2	–9.0	5	AY291333	Chew *et al.*, unpublished
Aldehyde dehydrogenase	416	45.241	5.6	–3.0	2	AY291328	Chew *et al.*, unpublished
Gal d 1 (IGF-binding protein)	248	27.828	4.4	–12.5	1	AY291331	Chew *et al.*, unpublished
Profilin	130	14.127	6.5	0.0	0	AY291334	Chew *et al.*, unpublished
Alt a 6 (ribosomal protein P2)	114	11.674	4.3	–8.0		AY291329	Chew *et al.*, unpublished
Hexosaminidase	341	39.432	5.3	–6.0	3	AY291332	Chew *et al.*, unpublished
Ves m 1 (lipase)	260	28.941	7.2	5.0	1	AY291335	Chew *et al.*, unpublished
Cathepsin D homologue	384 384	42.516 42.574	7.6 7.6	5.5 6.5	4 4	AY792951 AY792952	Chew *et al.*, unpublished Chew *et al.*, unpublished

*some are nucleotide and some are protein database numbers

atopic patients to mite allergens. The literature on this topic is vast and complex and its inclusion would greatly increase the length of this chapter, probably to no great end. This book is mainly about the biology of dust mites. I have tried therefore to remain focused on the biological role of allergens within the mites, with some consideration of those properties that render them allergenic.

Table 7.2 Functions of mite allergen groups. Loc. = location within mites; G = gut digestive cells/faecal pellets; M = muscle cells; I = intracellular (other than midgut cells). Cysteine residues: refers to number of conserved residues in aligned sequences (see Figures 7.9–7.28). Unpublished references refer to entries of sequence data on databases, see Table 7.1.

Group	Mite spp.	Identity/function	Loc.	References
1	Df, Dm, Dp, Ds, Em, Po, Ss, Bt	Cysteine peptidase	G	Chapman & Platts-Mills, 1980; Stewart & Turner, 1980; Lind, 1986b; Stewart & Fisher, 1986; Chua *et al.,* 1988; Kent *et al.,* 1992; Ferrándiz *et al.,*1995b; Smith *et al.,* 1999; Mora *et al.,* 2003
2	Df, Dp, Ds, Em, Po, Ao, Tp, Sm, Gd, Ld, Bt	Lipid binding protein	I	Heymann *et al.,* 1986, 1989; Yuuki *et al.,* 1991; Trudinger *et al.,* 1991; Valera *et al.,* 1994; Schmidt *et al.,* 1995; Eriksson *et al.,* 1998; Gafvelin *et al.,* 2001; Temeyer *et al.,* 2002; Kaiser *et al.,* 2003; Ichikawa *et al.,* 2005; Keber *et al.,* 2005
3	Df, Dp, Ds, Em, Ss, Gd, Ld, Bt	Trypsin	G	Heymann *et al.,* 1989; Stewart *et al.,* 1989, 1992, 1994; Ando *et al.,*1993; Smith *et al.,* 1994; Nishiyama *et al.,* 1995; Smith & Thomas, 1996; Ferrándiz *et al.,* 1997; Flores *et al.,* 2003; Holt *et al.,* 2003; Cheong *et al.,* 2003
4	Dp, Em, Bt	Alpha-amylase	G	Lake *et al.,* 1991; Mills *et al.,* 1999
5	Df, Dp, Gd, Ld, Bt	Structural protein		Tovey *et al.,* 1989; Lin *et al.,* 1994; O'Neill *et al.,* 1994b; Caraballo *et al.,* 1996; Arruda *et al.,* 1997a; Eriksson *et al.,* 2001; Liaw *et al.,* 2001; Kuo *et al.,* 2003; Yi *et al.,* 2004
6	Df, Dp, Bt	Chymotrypsin	G	Yasueda *et al.,* 1993; King *et al.,* 1996; Bennett & Thomas 1996; Kawamoto *et al.,* 1999
7	Df, Dp, Gd, Ld, Bt	Function unknown		Shen *et al.,* 1993, 1995a, 1995b, 1996; Eriksson *et al.,* 2001
8	Dp, Po, Ss, Gd, Ld, Bt	Glutathione-*s*-transferase	?I	O'Neill *et al.,* 1994; Dougall *et al.,* 2005; Lee *et al.,* 1999; C. Huang *et al.,* 2006
9	Df, Dp, Bt	Collagenase	G	King *et al.,*1996
10	Df, Dp, Po, Tp, Gd, Ld, Dg	Tropomyosin	M	Aki *et al.,* 1995; Asturias *et al.,* 1998; Yi *et al.,* 2002; Saare *et al.,* 2003; Huntley *et al.,* 2004; Nisbet *et al.,* 2006a, b
11	Df, Dp, Po, Ss, Bt	Paramyosin	M	Tsai *et al.,* 1998; Ramos *et al.,* 2001; Mattson *et al.,* 2001; Huntley *et al.,* 2004; Lee *et al.,* 2004; Teo *et al.,* 2006
12	Ld, Bt	Contains peritrophin A domain: chitin binding	I	Puerta *et al.,* 1996a
13	Df, As, Tp, Gd, Ld, Bt	Fatty acid-binding protein	I	Caraballo *et al.,* 1997; Eriksson *et al.,* 2001; Jeong *et al.,* 2005a; Chan *et al.,* 2006
14	Df, Dp, Em, Po, Ss, Bt	Vitellogenin: egg yolk storage protein	I	Aki *et al.,* 1994a; Fujikawa *et al.,* 1996; Mattsson *et al.,* 1999; Epton *et al.,* 1999, 2001a; Harumal *et al.,* 2003; Huntley *et al.,* 2004; Ljundgren *et al.,* 2005
15	Df, Dp	Family 18 chitinase	I	McCall *et al.,* 2001; O'Neill *et al.,* 2006
16	Df	Gelsolin: actin-binding	I	Kawamoto *et al.,* 2002a
17	Df	Calcium-binding protein	I	Tategaki *et al.,* 2000
18	Df, Dp, Bt	Chitinase	I	Weber *et al.,* 2003; O'Neill *et al.,* 2006
19	Bt	Anti-microbial peptide	G	Thomas *et al.,* 2002
20	Dp	Arginine kinase	?I	Thomas, unpublished
21	Bt, Dp	Structural protein	G	Gao *et al.,* 2007; Weghofer, unpublished
Dpt4	Dp	Lipoprotein		Stewart *et al.,* 1983
13.8 kDa bacteriolytic enzyme	Dpt, Df	P60 hydrolase		Mathaba *et al.,* 2002

continued

Table 7.2 Continued

Group	Mite spp.	Identity/function	Loc.	References
Major allergen 1	Ss			Fischer *et al.,* 2003
Alpha-tubulin	Tp, Ld	Alpha-tubulin	I	Jeong, 2005b; Saarne *et al.,* 2003
Mag 29	Df, Bt	Heat shock protein 70	I	Aki *et al.,* 1994b
CPW1	Df	Cathepsin F homologue		Park *et al.,* unpublished
39 kDa allergen	Ld	Function unknown		Ansotegui *et al.,* 1991
79/93 kDa allergen	Ld	Function unknown		
Cystatin	Ld	Cystatin A		Kaiser *et al.,* unpublished
Aldehyde dehydrogenase	Df, Bt	Aldehyde dehydrogenase		Chew *et al.,* unpublished
Gal d 1 homologue	Bt	Ovomucoid allergen; insulin growth factor-binding protein		Chew *et al.,* unpublished
Enolase	Gd	Phosphopyruvate hydratase		Chew *et al.,* unpublished
Mal d 3 homologue	Gd	Type 1 non-specific lipid transfer protein		Chew *et al.,* unpublished
Mala s 6 homologue	Gd	Peptidylprolyl isomerase B		Chew *et al.,* unpublished
Jun a 3 homologue	Gd	Thaumatin homologue		Chew *et al.,* unpublished
BTP1	Bt			Chew *et al.,* unpublished
Profilin	Sm, Bt	Plant pan-allergen		Chew *et al.,* unpublished
Hexosaminidase	Bt	Beta-*n*-acetyl hexosaminidase		Chew *et al.,* unpublished
Alt a 6 homologue	Bt	Ribosomal protein P2		Chew *et al.,* unpublished
Ves m 1 homologue	Bt	Lipase		Chew *et al.,* unpublished
Cathepsin D homologue	Bt	Cathepsin D homologue		Chew *et al.,* unpublished

Ao = *Aleuroglyphus ovatus*; As = *Acarus siro*; Bt = *Blomia tropicalis*; Df = *Dermatophagoides farinae*; Dm = *D. microceras*; Dp = *D. pteronyssinus*; Ds = *D. siboney*; Em = *Euroglyphus maynei*; Gd = *Glycyphagus domesticus*; Ld = *Lepidoglyphus destructor*; Po = *Psoroptes ovatus*; Sm = *Suidasia medanensis*; Ss = *Sarcoptes scabiei*; Tp = *Tyrophagus putrescentiae*; Dg = *Dermanyssus gallinae*.

7.2 Which mites produce clinically important allergens?

Probably any mite, or indeed virtually any organism, is capable of producing proteins and peptides to which humans could become allergic, given the right circumstances. In fact, only a few hundred organisms, out of millions, are known to have allergenic properties. The ones that tend to cause problems are ones that humans are likely to come into contact with on a regular basis. There has always been an intimate association between surveys of mites in homes and assessment of mite allergenicity. The mites that were investigated for their allergenicity were the ones that were recorded repeatedly in houses. The establishment of the distribution and abundance of species provided the basis for investigations of their

allergenicity. For example, some 35 species of mites had been recorded from house dust in Japan, with *D. farinae* known to be a source of allergens similar to those isolated from house dust, having a mean relative abundance of only 4% (Oshima, 1970). This led Miyamoto *et al.* (1969) to conclude that in order to account for the clinical significance of allergens of *D. farinae*, they must cross-react with those of more abundant mites. They did, with those of *D. pteronyssinus* which had a relative mean abundance of about 30%. Later, when the high prevalence and abundance of *Blomia tropicalis* was recognised in the Caribbean and Latin America, patterns of sensitisation and cross-reactivity of this and other dust mites, such as *D. siboney*, began to be investigated (Puerta *et al.,* 1991; Ferrándiz *et al.,* 1995a, b).

What follows is a summary of research on those superfamilies, families and genera of mites known to cause allergies (see Table 7.3).

7.2.1 Pyroglyphidae

a Dermatophagoides

The allergens of members of this genus are by far the most thoroughly investigated. Sensitisation to *D. farinae* (then called *D. culinae*, a junior synonym of *D. farinae*, see Appendix 1) using skin-prick testing with aqueous extracts of the mites was reported by Pepys *et al.* (1968), Spieksma and Voorhorst (1969) and to *D. pteronyssinus* by Spieksma and Voorhorst (1969) and Miyamoto *et al.* (1969). In the 1970s and 1980s, scores of publications detailing allergic reactions to *D. pteronyssinus* and *D. farinae* were published. There are too many to mention here, but the relevant abstracts can be found in the *Review of Applied Entomology* (*Series B, Medical and Veterinary*), volumes 58–78 (1970–1990).

During the 1970s there were several reports of trials of hyposensitisation of allergic patients with crude extracts of *Dermatophagoides* spp. (Maunsell *et al.,* 1971; Gaddie *et al.,* 1976; Gabriel *et al.,* 1977) as well as measurements of anti-*Dermatophagoides* IgE antibodies for diagnosis of mite allergy, for epidemiological studies (Turner *et al.,* 1975) and for monitoring the course of immunotherapy (Nakamura and Yoshida, 1977). It became apparent that the crude extracts of dust mites contained a mixture of allergens, and research turned to attempts at isolation and purification (Ishii *et al.,* 1973; Biliotti *et al.,* 1975; Baldo *et al.,* 1977; Nakagawa *et al.,* 1977; Chapman and Platts-Mills, 1978; Kabasawa and Ishii, 1979).

The first major allergen, now known as Der p 1, was purified and characterised independently by Stewart and Turner (1980) and Chapman and Platts-Mills (1980), shortly followed by Der f 1 by Dandeu *et al.* (1982), Der p 2 (Lind, 1985) and Der f 2 (Yasueda *et al.,* 1986). These allergens were given informal names before current nomenclature was introduced (see Table 7.4). There followed a period of intense research activity involving further physicochemical characterisation of the allergens, production of monoclonal antibodies and mapping of epitopes, characterisation of further allergens and defining their effects on the immune system. Most of this research has been summarised in the proceedings of the first two international workshops on dust mite allergens and asthma (International Workshop Report, 1988; Platts-Mills *et al.,* 1992).

Extensive cross-reactivity between *D. pteronyssinus* and *D. farinae* has long been known (Miyamoto *et al.,* 1969; Dasgupta and Cunliffe, 1970). Platts-Mills *et al.* (1986a) found that a major allergen of *D. pteronyssinus* (Der p 1) possessed both species-specific epitopes and ones that cross-reacted with those of Der f 1 of *D. farinae*. Further, Der m 1 from *D. microceras* and Der s 1 from *D. siboney* show close physicochemical and immunological similarities with Der p 1 and Der f 1 (Lind, 1986a, b; Ferrándiz, 1997; Ferrándiz *et al.,* 1998).

b Euroglyphus

There is increasing recognition that *Euroglyphus maynei* is common in house dust throughout the world and is a major source of mite-derived allergens in house dust (reviewed by Colloff, 1991c). Voorhorst *et al.* (1969) first reported that extracts *of E. maynei* produced positive skin-prick test reactions and that *E. maynei* and *D. pteronyssinus* contained the same or a similar allergen. Platts-Mills *et al.* (1986a) found that extracts of *E. maynei* incubated with anti-Der p 1 sera inhibited more than 90% of Der p 1 binding activity, indicating that *E. maynei* produces an allergen which cross-reacts with Der p 1. Mumcuoglu (1977b, c) found *E. maynei* extracts elicited the most frequent and severe positive scratch test reactions of nine mite species and noted some cross-reactivity between *D. pteronyssinus, D. farinae* and *E. maynei* as well as species-specific allergens. Using the Schultz-Dale test of smooth-muscle reaction of guinea pigs, Mumcuoglu (1977c) found that *E. maynei* extracts evoked positive responses only in those animals that had been

Table 7.3 Genera of mites known or suspected of producing allergens. Abbreviations for evidence of the presence of allergens in mites: S = skin test or RAST with mite extract; C = clinical evidence; P = purification and/or amino acid sequencing of allergens. See text for references.

Order, family and genus	Major habitat	Evidence of allergens	Most important setting for exposure
ASTIGMATA			
Pyroglyphidae			
Dermatophagoides	House dust	S,C,P	Domestic
Euroglyphus	House dust	S,C,P	Domestic
Malayoglyphus	House dust	S	Domestic
Sturnophagoides	House dust/nests of birds	S	Domestic
Acaridae			
Acarus	Stored products	S,C,P	Occupational, esp. bakers and farmers
Tyrophagus	Stored products	S,C,P	Occupational
Aleuroglyphus	Stored products	S,P	Occupational
Suidasia	House dust/stored products	S,C,P	Domestic/occupational
Thyreophagus	Stored products	S,C	Occupational
Glycyphagidae			
Glycyphagus	House dust/stored products	S,P	Domestic/occupational
Lepidoglyphus	House dust/stored products	S,C,P	Domestic/occupational
Gohieria	House dust/stored products	S	Domestic/occupational
Austroglycyphagus	House dust/nests of birds and mammals	P	Domestic
Hemisarcoptidae			
Hemisarcoptes	Predator of agricultural pests	S,C	Occupational
Echymyopodidae			
Blomia	House dust/stored products	S,C,P	Domestic/occupational
Chortoglyphidae			
Chortoglyphus	House dust/stored products	S	Domestic/occupational
Sarcoptidae			
Sarcoptes	Ectoparasite of mammals	S,C,P	Domestic
Psoroptidae			
Psoroptes	Ectoparasite of mammals	S,C,P	Occupational
PROSTIGMATA			
Cheyletidae			
Cheyletus	House dust/stored products	S,C	Domestic/occupational
Pyemotidae			
Pyemotes	Stored products	S,C	Occupational
Tetranychidae			
Panonychus	Pest of horticultural crops	S,C	Occupational
Tetranychus	Pest of horticultural crops	S,C	Occupational
MESOSTIGMATA			
Dermanyssus gallinae	Nests of birds in houses	C,P	Domestic/occupational
Phytoseiulus persimilis	Predator of phytophagous mites	S,C	Occupational: greenhouse workers
Hypoaspis miles	Predator of phytophagous mites	S,C	Occupational: greenhouse workers
Amblyseius cucumeris	Predator of phytophagous mites	S,C	Occupational: greenhouse workers

Table 7.4 Synonyms of informal names of mite allergens.

Informal or original name	Formal name/new designation	Reference
Ag 11	Der f 1	Le Mao *et al.,* 1981; Dandeu *et al.,* 1982
Bt-M	Blo t 5	Caraballo *et al.,* 1996
DF1	Der f 1	Yasueda *et al.,* 1986
DF2	Der f 2	Yasueda *et al.,* 1986
DF5	Der f 6	Yasueda *et al.,* 1993
Df6	Der f 1	Lind, 1986b
Df 11	Der f 1	Le Mao *et al.,* 1985
Df642	Der f 11	Tsai *et al.,* 1998
Dm6	Der m 1	Lind, 1986b
Dp1	Der p 1	Yasueda *et al.,* 1989b
DP2	Der p 2	Yasueda *et al.,* 1989b
DP5	Der p 6	Yasueda *et al.,* 1993
λ Dp 15	Der p 8	O'Neill *et al.,* 1994a
Dp42	Der p 1	Lind, 1985, 1986b
Dpt 12	Der p 1	Stewart, 1982; Stewart & Turner, 1980
Dpt 22	Der p 2	Chua *et al.,* 1988
DpX	Der p 2	Lind, 1985
Lep d 1	Lep d 2	Varela *et al.,* 1994; Schmidt *et al.,* 1995
M-177	Der f 14	Fujikawa *et al.,* 1998
Mag 1	Fragment of Der f 14	Aki *et al.,* 1994a
Mag 3	Fragment of Der f 14	Fujikawa *et al.,* 1996
Mag 15	Der f 16	Tategaki *et al.,* 2000; Kawamoto *et al.,* 2002a
Mag 44	Der f 10	Aki *et al.,* 1995
Mag 50	Der f 17	Tategaki *et al.,* 2000
MSA1 (not ASA1)	Sar s 14	Mattsson *et al.,* 1999
Me1	Der f 2	Haida *et al.,* 1985
Me2	Der f 1	Yamashita *et al.,* 1989
P_1	Der p 1	Chapman & Platts-Mills, 1980
Po16k	Pso o 2	Temeyer *et al.,* 2002
Ssag1	Sar s 14	Harumal *et al.,* 2003

sensitised with *D. pteronyssinus* extract, and not those sensitised with extracts of the stored products mites *Tyrophagus putrescentiae, Chortoglyphus arcuatus* and *Lepidoglyphus destructor.* The likelihood of common cross-reacting allergens between *E. maynei* and *D. pteronyssinus* was reinforced by the finding that *D. pteronyssinus,* which was the numerically dominant species in 71% of homes and was present in 89% of homes, gave positive skin tests in 75% of the patients, while *E. maynei,* which caused 85% of positive skin tests, was numerically dominant in only 17% of homes and was present in 37% (Mumcuoglu, 1977b). A similar lack of correlation between skin-prick test positivity and frequency of occurrence of *E. maynei* was noted by Charpin *et al.* (1986). In marked contrast to Mumcuoglu's findings, Nannelli *et al.* (1983) found only 0.6% positive skin-prick tests to *E. maynei* in atopic individuals.

Van Hage-Hamsten and Johansson (1989) investigated the cross-reactivity of extracts *of E. maynei* with those of *Acarus siro, Lepidoglyphus destructor* and *D. pteronyssinus* using sera from Swedish farmers. They found positive reactions to *E. maynei* in 4.5% of sera that were mite-positive by RAST (radioallergosorbent test), compared with 5.2% for *D. pteronyssinus.* Some 8% of farmers had IgE against *E. maynei*

only, indicating *E. maynei* has species-specific allergens. This was confirmed by RAST inhibition studies whereby *D. pteronyssinus* extract failed to inhibit *E. maynei*-positive sera from binding to paper discs coated with *E. maynei* extract. An allergen seemingly homologous with Der p 1 was found to be present in *E. maynei*-rich dust. Van Hage-Hamsten and Johansson (1989) found moderate cross-reactivity with *D. pteronyssinus* but none with *Lepidglyphus destructor* or *Acarus siro*. Although *E. maynei* possesses a group 1 allergen homologue, as evidenced by Kent *et al.* (1992), who reported 85% amino acid sequence homology with Der p1 and Der f 1, Arruda and Chapman (1992) found no evidence for *E. maynei* group 2 allergens. However, a 14 kDa allergen was identified by immunoblotting by Colloff *et al.* (1992a) and group 2 allergens have been purified (Smith *et al.,* 1999). Major *E. maynei*-specific IgE-binding components of 16, 38 and 62 kDa have been identified by immunoblotting, as well as components common to *D. pteronyssinus* of 14, 24, 97 and 175 kDa using sera from atopic people who had domestic exposure to both species.

c *Malayoglyphus and Sturnophagoides*

In a survey of dust mites in Singapore, Chew *et al.* (1999a) found *Malayoglyphus intermedius* and *Sturnophagoides brasiliensis* in 10–31% and 42–84% of dust samples. They found 50–70% of atopics were sensitised to these species. A more detailed study on frequency of sensitisation found 72% of patients with allergic asthma or rhinitis were skin-prick test positive to *S. brasiliensis* (Chew *et al.,* 1999b), fifth in rank frequency after *Blomia tropicalis*, *Dermatophaogides farinae*, *D. pteronyssinus* and *Austroglycyphagus malaysiensis*.

7.2.2 Acaroidea

a *Acarus*

The cosmopolitan genus *Acarus* contains some of the most important pests of stored products (Hughes, 1976), although its members occur in a wide range of habitats including house dust and agricultural soils, grasslands, poultry houses and birds' nests. Exposure to these mites can therefore occur in domestic, agricultural or occupational settings. Sensitisation to *A. farris* and *A. siro* was reported among farmers by Ingram *et al.* (1979) and grain workers by Blainey *et al.* (1989). Allergy to *Acarus siro* was detailed among asthmatics by Maunsell *et al.* (1968) and by Pepys *et al.* (1968); among asthmatic children versus non asthmatic controls by Morrison Smith *et al.* (1969); atopic outpatients attending an allergy clinic by Spieksma and Voorhorst (1969); cheese-makers by Molina *et al.* (1975, 1977); patients with asthma or rhinitis by Wraith *et al.* (1979); farmers by van Hage-Hamsten *et al.* (1987); asthmatic children by Boner *et al.* (1989); and bakers by Revsbech and Dueholm (1990). *Acarus siro* has been implicated in cases of anaphylaxis following ingestion of contaminated food in France (Dutau, 2002). Cross-reactivity of *A. siro* with dust mites and other storage mites was investigated by Mumcuoglu (1977b), Griffin *et al.* (1989) and Johansson *et al.* (1994). Aca s 13, a fatty-acid binding protein, shows significant sequence homology with other group 13 allergens and the recombinant protein was bound by 23% of patients who were RAST-positive to *A. siro* extract (Eriksson *et al.,* 1999).

b *Tyrophagus*

Exposure to species of *Tyrophagus* can, like *Acarus*, occur in domestic and occupational circumstances. Farmers, grain handlers, bakers, cheese-makers and small-goods processors have become sensitised (summarised by Eriksson *et al.,* 1998), mostly to what has been reported as *Tyrophagus putrescentiae*, although occupational exposure is also likely to *T. longior*, *T. nieswanderi*, *T. palmarum* and *T. perniciosus*, species that are significant pests of stored products (Hughes, 1976). Czernecki and Kraus (1978) reported an outbreak of dermatitis in a butcher's family, exposed to *T. dimidiatus* growing on mouldy bacon.

Sensitisation to *Tyrophagus putrescentiae* was reported by Spieksma and Voorhorst (1969); Miyamoto *et al.* (1969); Araujo-Fontaine *et al.* (1974); Green and Woolcock (1978); Ingram *et al.* (1979); Blainey *et al.* (1989); and Heyraud *et al.* (1989). Allergenic properties were investigated by Arlian *et al.* (1984a). Cross-reactivity of *Tyrophagus* spp. with dust mites and other storage mites was investigated by Arlian *et al.* (1984b); Griffin *et al.* (1989); and Johansson *et al.* (1994). Rufli (1970) reported skin-prick test cross-reactivity between *D. pteronyssinus* and *T. putrescentiae*, as did Mumcuoglu (1977b). The group 2 allergen was characterised from *T. putrescentiae* by Eriksson *et al.* (1998).

c *Aleuroglyphus*

The only allergenically important species in this genus, *Aleuroglyphus ovatus*, is found in stored products, mammal nests, farms and human dwellings. Sensitisation to *A. ovatus* was reported by Miyamoto

et al. (1969) and Puerta *et al.* (1993). Prevalence of specific IgE against this species was reported by Silton *et al.* (1991). *Aleuroglyphus* sp. has been implicated in cases of anaphylaxis following ingestion of contaminated food in France (Dutau, 2002).

d Suidasia

The only known species of this genus of allergenic significance is *Suidasia medanensis*. It is probably a commensal species associated with bees (Hughes, 1976), but also crops up in stored food and houses. Sensitisation to *S. medanensis* was reported by Miyamoto *et al.* (1969). Some 73% of asthmatic patients in Cartagena, Colombia had positive IgE reactivity to *S. mendanensis*, and cross-reactivity was demonstrated between *S. medanensis*, *Blomia tropicalis* and *Dermatophagoides farinae* (Puerta *et al.,* 2005). An un-named *Suidasia* sp. was found in large numbers in flour eaten by Venezuelan patients who had suffered anaphylaxis (Sánchez-Borges *et al.,* 1997, 2001) and *Suidasia nesbitti* has been implicated in cases of anaphylaxis following ingestion of contaminated food in France (Dutau, 2002).

e Rhizoglyphus

Rhizoglyphus contains several species that are pests of tubers and corms such as potatoes, onions and tulip bulbs. Occupational atopic dermatitis caused by *Rhizoglyphus* sp. was reported by Pigatto *et al.* (1991).

f Thyreophagus

Exposure of German farmers to *Thyreophagus* sp. was documented by Musken *et al.* (2003 and earlier papers cited therein). *Thyreophagus entomophagus* has been implicated in cases of anaphylaxis following ingestion of contaminated food in France (Dutau, 2002).

g Hemisarcoptes

This genus contains several egg predators of scale insects. Arlian *et al.* (1999a) found *H. cooremani* to be allergenic and, using immunoblotting, identified two putative allergens with molecular weights of 16 and 19 kDa.

7.2.3 Glycyphagoidea

a Glycyphagus

Sensitisation to *Glycyphagus domesticus* has been reported by Maunsell *et al.* (1968); Pepys *et al.* (1968); Voorhorst *et al.* (1969); Ingram *et al.* (1979); van Hage-Hamsten *et al.* (1987); Blainey *et al.* (1989); and to *Glycyphagus privatus* by Boner *et al.* (1989). Cross-reactivity of *G. domesticus* and *G. privatus* with dust mites and other storage mites was investigated by Mumcuoglu (1977b). Some 11 allergens from *G. domesticus* have been characterised and sequenced (see Tables 7.1, 7.2).

b Lepidoglyphus

Lepidoglyphus destructor is one of the major species responsible for both occupational allergies due to exposure to contaminated stored products mite and exposure within the home. Sensitisation to this species has been reported by Spieksma and Voorhorst (1969); Araujo-Fontaine *et al.* (1974); Ingram *et al.* (1979); van Hage-Hamsten *et al.* (1987); Boner *et al.* (1989); Blainey *et al.* (1989); and Heyraud *et al.* (1989). Allergen components of an extract of *L. destructor* were investigated by immunoblotting by Johansson *et al.* (1988). Cross-reactivity with dust mites and other storage mites was investigated by Mumcuoglu (1977b); Griffin *et al.* (1989); and Johansson *et al.* (1991). Monoclonal antibodies *to L. destructor* allergens were produced by Ventas *et al.* (1991) and Härfast *et al.* (1992). A 39 kD allergen was identified by Ansotegui *et al.* (1991), and cDNA analysis of Lep d 1 was reported by Schmidt *et al.* (1995). To date, about 12 allergens from *L. destructor* have been characterised and sequenced (see Tables 7.1, 7.2).

c Gohieria

Gohieria fusca is a common stored-products species found in house dust worldwide. Cross-reactivity of *G. fusca* with dust mites and other storage mites was investigated by Mumcuoglu (1977b). Sensitisation to this species has been reported by Boner *et al.* (1989).

d Blomia

Blomia kulagini is probably the same species as *Blomia tropicalis* (see Chapter 1). However, they have been treated as separate species in studies of allergens. *B. tropicalis* is widely distributed in the tropics and subtropics (see Chapter 4). Sensitisation to *Blomia* was reported by Miyamoto *et al.* (1969) using skin-prick tests. A high prevalence of sensitisation to this mite occurs in tropical countries (Chew *et al.,* 1999b; Kuo *et al.,* 1999, 2003). Multiple, species-specific allergens were detected by Arlian *et al.* (1993). Immune responses to *B. kulagini* were investigated by van Hage-Hamsten *et al.* (1990a, b). Positive nasal challenge responses to an extract of *B. tropicalis* were recorded by Stanaland *et al.* (1996). Allergic cross-reactivity with *Lepidoglyphus*

destructor and clinical significance were investigated by van Hage-Hamsten *et al.* (1990a) and Puerta *et al.* (1991), and cross-reactivity of Blo t 5 with Der p 5 by Kuo *et al.* (2003). Arruda *et al.* (1995) purified and identified the allergen Blo t 5; Arruda *et al.* (1997a) identified Der p 5 as a major sensitising allergen. Y. Gao *et al.* (2007) purified Blo t 21. Some 24 examples of allergens from *B. tropicalis* have been characterised and sequenced (Tables 7.1, 7.2), and the solution structure of Blo t 5 has been detailed by Naik *et al.* (2008; see Figure 7.6b).

e Chortoglyphus

Sensitisation to *Chortoglyphus arcuatus* was reported by Miyamoto *et al.* (1969) and Puerta *et al.* (1993). Cross-reactivity of *C. arcuatus* with dust mites and other storage mites was investigated by Mumcuoglu (1977b) and Garcia-Robaina *et al.* (1998).

f Austroglycyphagus malaysiensis

In a survey of dust mites in Singapore, Chew *et al.* (1999a) found *Austroglycyphagus malaysiensis* in 20–54% of dust samples and 50–70% of atopics were sensitised to this species. A more detailed study on frequency of sensitisation found 72% of patients with allergic asthma or rhinitis were skin-prick test positive to *S. brasiliensis* (Chew *et al.,* 1999b), the fourth most important mite in frequency of allergenicity for this patient group.

7.2.4 Psoroptoidea

a Sarcoptes

The genus *Sarcoptes* contains *S. scabiei*, the causative agent of scabies in humans and other mammals. Infestation with *S. scabiei* causes an allergic skin reaction, lesions and itching. Falk and Bolle (1980a, b) found scabies infection stimulated production of IgE antibodies that cross-reacted with those of *Dermatophagoides* and that *S. scabiei* also produced specific allergens, as did Arlian *et al.* (1988). *S. scabiei* homologues of *Dermatophagoides* allergens have been identified (groups 1, 3, 8 and 11; Table 7.2).

b Psoroptes

Psoroptes is a genus that includes several species that cause mange in mammals, including *P. ovis*, the agent of sheep scab. Psoroptic mange is a form of allergic dermatitis affecting livestock. It is a fatal disease of considerable economic significance that requires stock to be quarantined if there is any likelihood they are infected. Cross-reactivity between *Psoroptes* spp. and *D. pteronyssinus* allergens was detailed by Stewart and Fisher (1986). A search for potential vaccines led to the identification, cloning and sequencing of group 2 allergens (Pruett, 1999; Temeyer *et al.,* 2002). Homologues of group 10, 11 and 14 allergens have been identified by Huntly *et al.* (2004) and Nisbet *et al.* (2006a).

7.2.5 Cheyletidae

a Cheyletus

This genus is a member of the Prostigmata. *Cheyletus* spp. have piercing-sucking mouthparts (see Chapter 2, Figure 2.1) and are usually predators of other mites (Wharton and Arlian, 1972b). Sensitisation to *C. malaccensis* was documented by Morita *et al.* (1975). *C. eruditus* and *C. malaccensis* and a member of a related genus, *Chelacaropsis*, have been recorded as biting humans, feeding on tissue fluid (Yoshikawa, 1980, 1987). These mites cause itchy papular skin lesions (Yoshikawa, 1980, 1985; Yoshikawa *et al.,* 1983; Yamada *et al.,* 1988). Liquid-feeding mites do not produce faecal pellets, so the source of their allergens is likely to be the salivary glands. Over half of a group of asthmatics skin-prick tested with an extract of *Cheyletus* were found to be positive to it, and extracts of house dust and a *Dermatophagoides* sp. (Pérez Lozano, 1979). Some preliminary isolation and characterisation of the allergens of *Cheyletus eruditus* using immunoblotting was reported by Musken *et al.* (1996), including a possible dust mite group 4 allergen.

7.2.6 Pyemotidae

a Pyemotes

This genus, also a member of the Prostigmata and having piercing-sucking mouthparts, was identified as the cause of asthmatic and allergic symptoms in grain workers by numerous authors from the mid-19th century onwards (reviewed by Alexander, 1984). *Pyemotes ventricosus* causes allergic symptoms of what is known as 'Grain itch' and was identified as the agent responsible for epidemic allergic asthma among grain workers by Ancona (1923; Figure 4.5d).

7.2.7 Tetranychoidea

Tetranychoid mites include the plant-feeding spider mites and false spider mites. Occupational allergy to these mites is usually associated with fruit growing and harvesting. Bernecker (1970) found that half the

asthmatics who were skin-prick test positive to *Dermatophagoides pteronyssinus* were also positive to *Panonychus ulmi*, and two-thirds were also positive to *Tetranychus urticae*. Rufli (1970) also reported on skin-prick test cross-reactivity between *D. pteronyssinus* and *T. urticae*. Michel *et al.* (1977) reported asthma, conjunctivitis and dermatitis apparently due to *Panonychus ulmi* among apple growers; Reunala *et al.* (1983) reported similar symptoms, plus skin-prick test and RAST results, in two greenhouse workers exposed to the two-spotted spider mite, *Tetranychus urticae*. Other reports of allergy to spider mites include those of Kroidl *et al.* (1992), Kim *et al.* (1999), Lee *et al.* (2000) and Kronqvist *et al.* (2005).

7.2.8 Trombiculidae

Chiggers, also known as itch mites, are the parasitic larvae of trombiculid mites, and include species that transmit rickettsial diseases of mammals. Over 20 species have been reported as responsible for itchy papular eruptions in humans (Wharton and Fuller, 1952).

7.2.9 Mesostigmatid mites

Most mesostigmatid mites are predators or parasites, with piercing chelicerate mouthparts. Alexander (1984) reviewed skin eruptions caused by mesostigmatid mites. Bernecker (1970) found that a fifth of the asthmatics who were skin-prick test positive to *Dermatophagoides pteronyssinus* were also positive to *Dermanyssus gallinae*. Rufli (1970) also reported on skin-prick test cross-reactivity between these two species. Frenken (1962) and Sexton and Haynes (1975) reported allergic symptoms due to bites of *Dermanyssus gallinae*, and Kronqvist *et al.* (2005) documented IgE sensitisation and related asthma and rhinoconjunctivitis by greenhouse workers to *Phytoseiulus persimilis* and *Hypoaspis miles*, two biocontrol agents for the red spider mite, *Tetranychus urticae*. De Jong *et al.* (2004) reported on a group of bell pepper growers of whom 23% had occupational allergy to *Amblyseius cucumeris*, a predatory mite and biocontrol agent. A group 10 tropomyosin allergen from the poultry mite *Dermanyssus gallinae* (Der g 10) has been cloned, sequenced and databased (Nisbet *et al.,* 2006b).

7.3 Localisation of allergens within the mites

Long before it was known that mite allergens were mostly digestive enzymes, various investigations were made to localise within the mites the sites of synthesis and storage of the allergens. The first attempt was that of Halmai and Alexander (1971) that indicated that the isolated faeces were allergenic. Mumcuoglu and Rufli (1979) used indirect immunofluorescence microscopy of sections of dust mites to detect sites of allergen localisation. They used polyclonal rabbit anti-sera raised against a partially purified extract of *Dermatophagoides pteronyssinus* to screen some 1500 sections and found that fluorescence was most intense in the posterior midgut, anterior hindgut and cuticular regions. During preparation the mite extract was heated to 80°C for 4 hours, followed by reduction by boiling for an unspecified period. This procedure would have virtually removed all of the binding activity of the heat-labile Der p 1 (Lombardero *et al.,* 1990) and the resulting antiserum that was used for immunolabelling would have been directed against a high proportion of heat-stable allergen epitopes such as those of Der p 2. Additionally, the mites were fixed in 25% glutaraldehyde, which may denature the allergen and thus interfere with allergen-antibody binding. Nevertheless, the results greatly strengthened previous indications that the allergens of dust mites were associated with the digestive system. Group 2 allergens were subsequently localised by Jeong *et al.* (2002).

Stewart and Turner (1980) demonstrated that fluorescent-labelled polyclonal anti-Der p 1 antibody stained the gut wall of *Dermatophagoides pteronyssinus,* and Thompson and Carswell (1988) estimated rates of incorporation of radiolabelled culture medium into mite tissues. They suggested the long time course (>14 days) for the process was indirect evidence that Der p 1 is a protein synthesised by the mite and secreted into the digestive system rather than a by-product of digestion, in which case a far shorter time course would have been anticipated. Subsequently, Thomas *et al.* (1991) localised Der p 1 in the gut of *D. pteronyssinus* using immunogold labelling. Tovey and Baldo (1990) used fluorescent-labelled rabbit polyclonal antibodies raised against whole mite extract and against Der p 1, as well as a murine monoclonal anti-Der p 1 antibody to isolate the sites of deposition of the allergen in fresh frozen sections of *D. pteronyssinus*. With the polyclonal anti-whole mite extract, fluorescent antibody binding was most intense in the gut wall and gut contents as well as in the gnathosoma and cuticular-subcuticular regions – a similar pattern to that demonstrated by Mumcuoglu and Rufli (1979). With the polyclonal anti-Der p 1 antibody,

fluorescence was localised to the anterior midgut and its contents and the posterior midgut. The same pattern was observed with the monoclonal anti-Der p 1 with additional labelling in the oesophageal region. Tovey and Baldo (1990) suggested that Der p 1 is secreted from the wall of the anterior midgut and becomes incorporated into the faecal pellet. Fluorescent-labelled soybean lectin (specific for D-galactose) bound to the gut wall, contents and the peritrophic membrane, whereas wheat germ agglutinin (specific for *N*-acetyl-*D*-glucosamine) was more generally distributed, with strong fluorescence from the walls of the anterior midgut. Rees *et al.* (1992) used polyclonal monospecific sheep anti-Der p 1 antibody to localise Der p 1 to the gnathosoma, faecal pellets and midgut, visualised by peroxidase staining, in paraffin sections of mites.

Van Hage-Hamsten *et al.* (1992a) used murine monoclonal antibodies raised against an extract of *Lepidoglyphus destructor* to localise the 39 kDa allergen in frozen sections of the mite using immunoperoxidase staining. They found labelling of the gut, and faecal pellets as well as cuticular and subcuticular tissues, though sera from *L. destructor*-allergic patients only bound to the gut and faecal pellets. Subsequently van Hage-Hamsten *et al.* (1995) localised Lep d 2 and a 79 and 93 kDa complex in *Lepidoglyphus destructor* and visualised them using confocal microscopy. Anti-Lep d 2 monoclonals bound most intensely to the gnathosomal region, with fainter staining in the gut and faecal pellets, with no staining of the cuticle, whereas anti-79/93 kDa antibodies bound to the gnathosomal region and the cuticle but not the faecal pellets. Despite all of the above work, we are not much further forward in understanding precisely where the sites of synthesis of mite allergens are located. However, this can be postulated by a closer examination of ultrastructural data together with a consideration of the enzymic properties of the allergens (see section 2.2.3a).

Der f 15, a major allergen for atopic dogs, has been identified as a chitinase. It has been localised mainly in the chitin-lined parts of the gut of *D. farinae* – the oesophagus and hindgut – but not within faecal pellets (McCall *et al.,* 2001). The same pattern of localisation was found for Der f 18, which also shows homology with chitinases (Weber *et al.,* 2003).

Blo t 21 is not an enzyme but a small alpha-helical protein of unknown function related to groups 5 and 7 allergens. Gao *et al.* (2007) found it was present in the gut and faecal pellets of *Blomia tropicalis*. Most intense staining could be seen along the epithelium of the midgut caecae as well as in the semi-digested food balls.

The regions of the body that have stained positive for allergen in these studies, namely the buccal-pharyngeal-oesophageal region, the anterior and posterior midgut and the cuticular and subcuticular tissues, indicate the presence of four groups of allergen-enzymes. First, a group of salivary-gland associated carbohydrases which may well be group 4 allergens; second, the cysteine proteinase group 1 allergens associated with the midgut and faecal pellets and probably localised within midgut digestive cells; third, an enzyme associated with moulting, most probably localised within epidermally derived haemocytes associated with ecdysis (see section 2.5.1); and fourth, some chitinases associated with the anterior midgut.

7.4 Groups of mite allergens and their classification

7.4.1 Allergen nomenclature

The International Union of Immunological Societies (IUIS) Subcommittee on Allergen Nomenclature and Standardisation decided that allergens should be abbreviated using the first three letters of the generic name and the first letter of the specific name of the organism from which they are derived (Marsh *et al.,* 1987; King *et al.,* 1994). These are written in Roman characters with a capital for the first letter of the generic abbreviation, hence Der p. There follows a number, indicating the group to which the allergen belongs (Der p 1, Der p 2), which historically has reflected the chronology of their purification and characterisation, but now is intended to reflect functional biochemical and amino acid sequence similarities of allergens within a major taxon (i.e. all mite group 3 allergens are trypsins, but group 3 allergens of cats are cystatins). Before this the letters were written in italics and the numbers in Roman numerals, hence *Der p* I, *Der p* II. Prior to the IUIS standardisation, allergens were given informal names, which can be confusing for the reader when dealing with older literature. A list of commonly used informal names and their current synonyms is given in Table 7.4.

7.4.2 Classification of mite allergens

Now that the amino acid sequences of over 20 groups of mite allergens have been databased, it has become

possible to explore similarities in primary and secondary protein structure using molecular bioinformatics tools such as Basic Local Alignment Search Tool (BLAST) searches, sequence alignment and peptide statistics programs. These tools have been particularly important in providing comparative data on proteins of known function with similar amino acid sequences to mite allergens.

Mite allergens have been divided into groups on the basis of physicochemical properties and functions where these are known or can be deduced. Thus, group 3 allergens are all trypsins; group 4, amylase. This section deals with the higher order classification of allergen groups, to attempt to determine what other groups of allergens the group 3 trypsins and group 4 amylases are most similar to and place them in 'families'. The purpose of such classifications is that they might be of predictive value in determining, for example, the common molecular and structural features of the groups that make them allergenic.

The allergen classification detailed by Aalberse (2000) covered allergens from all sources, not just mites, and was based on secondary protein structure. It consisted of four families:

- those with anti-parallel β-strands including the immunoglobulin fold family (mite group 2) and serine proteases (groups 3, 6 and 9);
- anti-parallel β-sheets closely associated with α-helices (lipocalin – mite group 13);
- α and β structures that are not closely associated one with the other, including mite group 1 allergens; and
- α-helical lipid transfer proteins, e.g. Fel d 1 chain 1.

A new classification of mite allergens described in this chapter in Table 7.5 is based on:

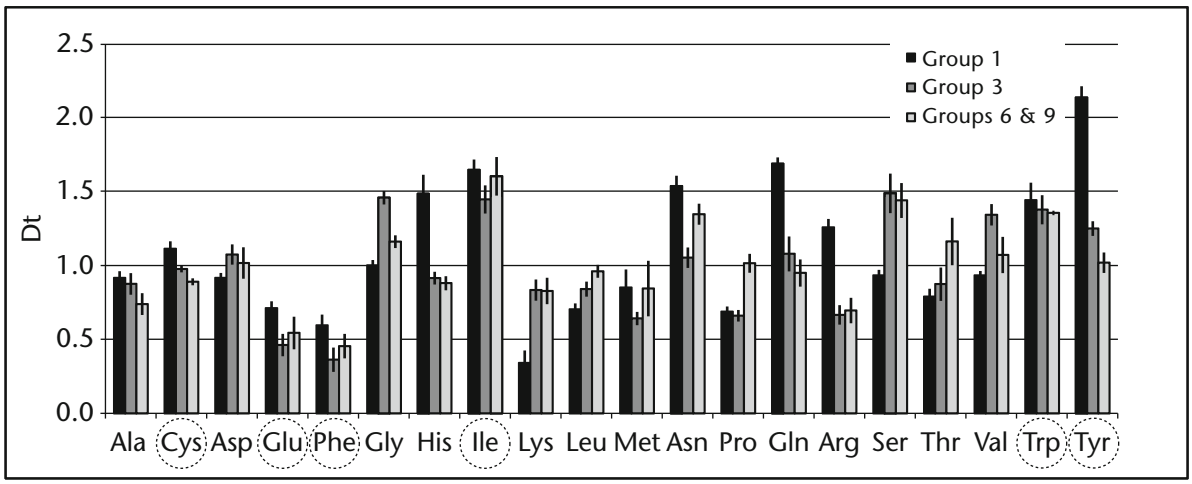
- comparisons of amino acid compositions;
- secondary structure as represented by topological maps;
- tertiary structure of allergens or other homologous proteins;
- known functional similarities;
- amino acid sequence homology; and
- shared conserved structural domains (Table 7.7).

The outline of the classification is given in Table 7.5. It includes six 'families': the peptidases (mite allergen groups 1, 3, 6 and 9), glycosidases (groups 4, 12, 15 and 18), transferases (groups 8 and 20), small alpha-helical proteins (groups 5, 7 and 21), muscle proteins (10 and 11) and the lipid-binding proteins (2, 13 and 14). This leaves groups 16, 17 and 19 unclassified. Group 16 are gelsolins, involved in actin binding and capping. Amino acid sequences of group 17 allergen have not been published or databased as far as I can discover. Group 19 is represented only by Blo t 19 which is very small (7 kDa) as mite allergens go, and about which not much is known, other than its amino

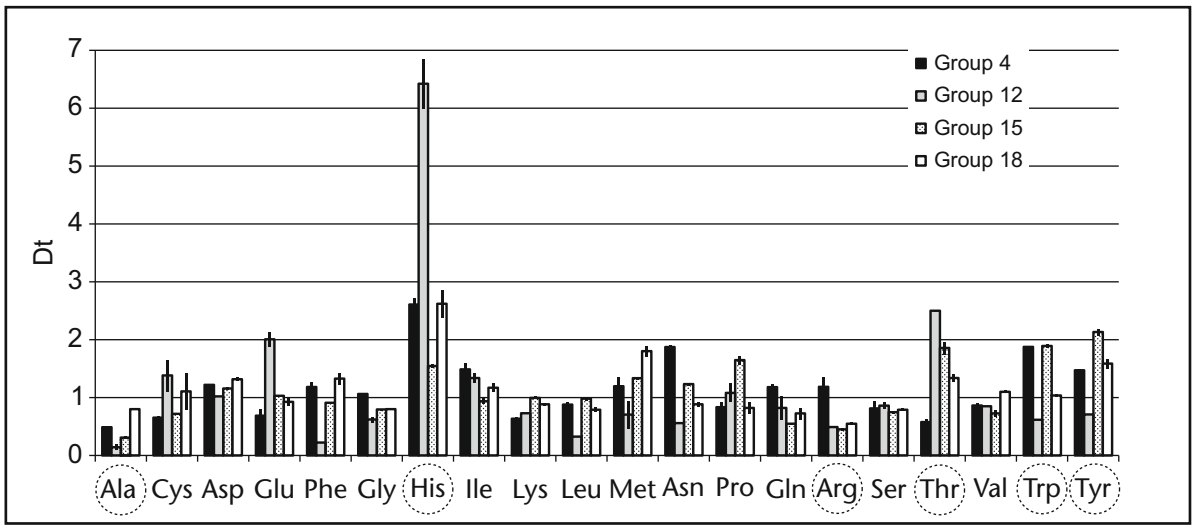
Table 7.5 Classification of dust mite allergen groups into families with shared characteristics.

Family and groups	Most frequent amino acids	Least frequent amino acids	Mean % aromatic amino acids	Mean % polar amino acids	Mean % charged amino acids	Tertiary structure
Peptidases 1, 3, 6, 9	Ile, Tyr, Trp	Phe, Glu	10.6	45.1	21.6	Globular, in two halves; central cleft containing active site
Glycosidases 4, 12, 15, 18	His, Thr, Trp, Tyr	Ala, Arg	16.7	47.9	28.7	4, 15 and 18: globular $(\alpha/\beta)_8$ barrel
Transferases 8, 20	Met, Tyr, Phe, Leu	Cys, Ser	13.7	46.8	28.8	Globular, two domains separated by a surface cleft
Small alpha-helical proteins 5, 7, 21	Met, Glu, Lys, Ile	Cys, Trp	10	47.3	36.6	5 and 21: coiled-coil alpha-helical bundle
Muscle proteins 10, 11	Glu, Gln, Arg	Cys, Gly Phe, Pro, Trp	4.2	62.3	43.4	Elongated alpha-helix
Lipid-binding proteins 2, 13, 14	Lys, Val, Asp, Ile	Trp, Pro	8.0	33.9	30.3	2 and 13: globular beta-barrels

(a) The peptidases



(b) The glycosidases



(c) The transferases

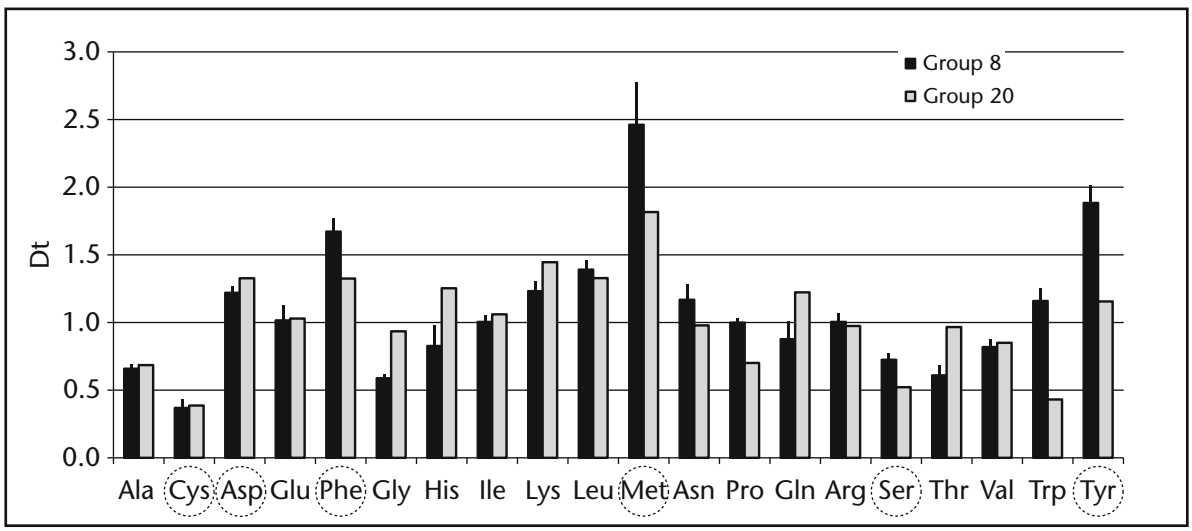
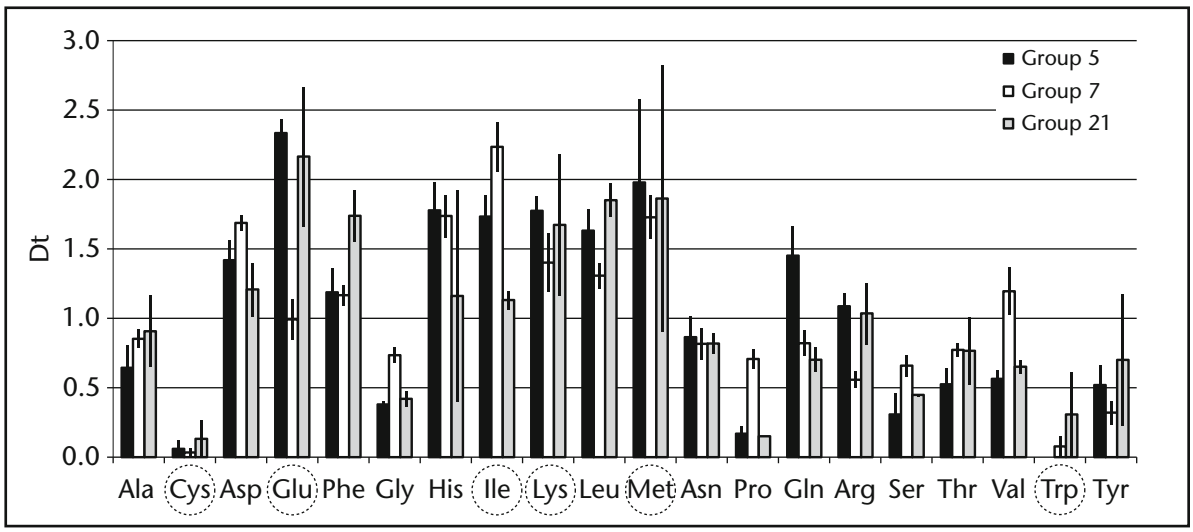
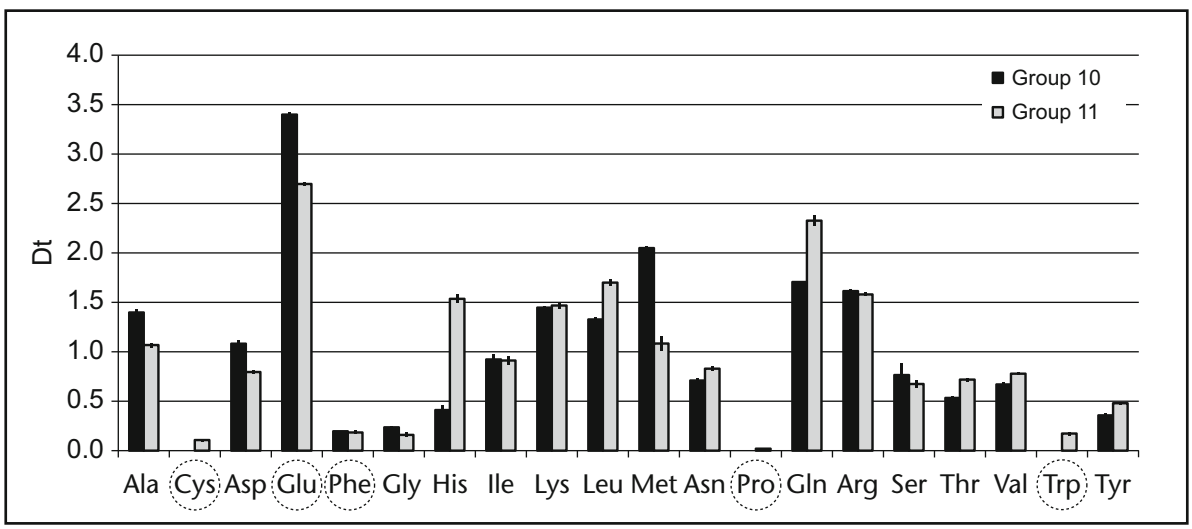


Figure 7.1 Amino acid composition of the various groups of dust mite allergen sequences in Table 7.1, represented as the transformed Dayhoff statistic (D_t), where D_t = molar percentage of each amino acid divided by the Dayhoff statistic (D), which is the relative occurrence of an amino acid per 1000 residues, normalised to occurrence per 100 residues. Figures represent means ± standard error. All calculations were done using the mature protein sequence only. Dotted circles highlight the signature patterns.

(d) The small alpha-helical proteins



(e) The muscle proteins



(f) The lipid-binding proteins

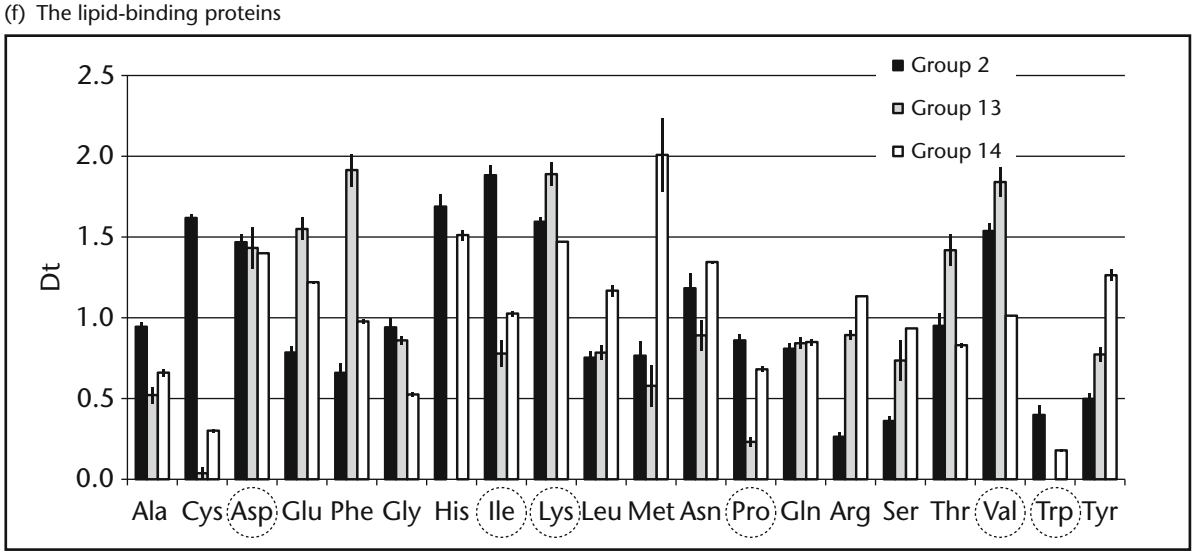


Figure 7.1 Continued

acid sequence, and its possible anti-microbial function (deduced from amino acid sequence comparisons).

Amino acid compositions represent the relative proportion of the different amino acids that make up the protein. Direct comparisons of proportions tend to have relatively high levels of variation, so to eliminate 'noise' I used the transformed Dayhoff value, D_t. This is the molar percentage of each amino acid divided by the Dayhoff statistic (D), which is the relative occurrence of an amino acid per 1000 residues, normalised to occurrence per 100 residues. Figure 7.1 shows D_t values for each of the six allergen groups. I used the program PEPSTATS (Australian National Genomic Information Service, http://www.angis.org.au/).

Secondary structure, as represented by topological maps of the proteins, is simply a method of predicting the arrangement of the structural elements in a linear map (Figure 7.5), detailing how much and which parts of the molecule are made up of alpha-helices, beta-strands and random coils. I used the Network Protein Sequence Analysis (Combet *et al.,* 2000), http://npsa-pbil.ibcp.fr; 3-D jigsaw (http://www.bmm.icnet.uk); and PeptideStructure (GCG). Some secondary structure predictions were found to be of limited use because they gave predictions contrary to x-ray crystallography-based tertiary structure. I have used them sparingly. For details of tertiary structure of allergens and related proteins I used the RCSB Protein Data Bank (www.pdb.org).

Known functional similarities are based either on direct experimental evidence of functional attributes of the allergen molecule (e.g. group 3 allergens shown to have trypsin activity by artificial substrate specificity) or are inferred from sequence comparisons with proteins of known function. Much of the experimental evidence for digestive enzyme-allergens has been covered in the section on digestion in Chapter 2 (see section 2.2.3).

Amino acid sequence homology (see Figures 7.10–7.27; Table 7.3) was determined using BLAST protein-protein searches (Altschul *et al.,* 1997). Amino acid sequences were aligned using sequence alignment software such as MUSCLE (Edgar, 2004; http://phylogenomics.berkeley.edu/) and T-COFFEE (Notredame *et al.,* 2000; http://tcoffee.vital-it.ch/cgi-bin/Tcoffee/tcoffee_cgi/index.cgi). Shared structural domains were determined using conserved domain searches in the NCBI database (Marchler-Bauer and Bryant, 2004) and Pfam, the Protein Families Database (http://www.sanger.ac.uk/Software/Pfam/).

7.4.3 The peptidases

This includes all allergens that are peptidase enzymes (EC 3.4), including groups 1 (cysteine proteinases), 3 (trypsins), 6 (chymotrypsins) and 9 (collagenases) (see Table 7.5). Characteristic of the amino acid composition of the peptidases (shown in Figure 7.1a) are high frequencies of isoleucine, tyrosine and tryptophan (mean transformed Dayhoff value, D_t of 1.6, 1.5 and 1.4 respectively), and the lowest are glutamate and phenylalanine (D_t 0.6 and 0.5). The peptidases also have a relatively high cysteine content (D_t 1.0) with six conserved residues. Predicted secondary structures indicate proteins with a series of 3–6 short alpha-helices (18–24% of the molecule), and 6–12 beta-strands (20–25%) with no clear patterns of association. About 53–60% of the molecule is made up of random coil.

Allergen groups 3, 6 and 9 share trypsin-like serine protease domains and are members of the chymotrypsin family S1. These are secretory proteins, synthesised with a signal peptide which targets it to the secretory pathway, and a pro-enzyme which prevents activation until the enzyme is required. Proteolytic activation occurs extracellularly, or within cellular vesicles or other storage organelles (Rawlings and Barrett, 1994). Group 1 allergens have peptidase C1 domains (papain family cysteine protease). Cysteine protease activity of group 1 allergens was confirmed experimentally by Stewart *et al.* (1991), as was serine protease activity for group 3 (Stewart, Ward *et al.,* 1992), group 6 (Yasueda *et al.,* 1993) and group 9 allergens (King *et al.,* 1996).

There is mounting evidence that the enzymatic activity of some mite allergens plays a significant role in enhancing their allergenicity (e.g. Hewitt *et al.,* 1995). The effects of dust mite peptidases on the initiation and enhancement of allergic responses have been reviewed by Robinson *et al.* (1997), Caughey and Nadel (1997), Shakib *et al.* (1998), Sharma *et al.* (2003) and Reed and Kito (2004), and it appears mite peptidases may augment or facilitate the action of endogenous peptidase (such as mast cell tryptase and elastase) in allergic inflammatory reactions, or act as proteolytic triggers of inflammation in their own right. The important point here is that not all mite allergens need to be proteases for such mechanisms to be important, because protease allergens can facilitate the allergenicity of presentation of non-enzymic allergens. Some of the effects of mite proteases are as follows:

- Proteoloytic activity of Der p 1, 3, 6 and 9 causes disruption of the lung epithelium, allowing the

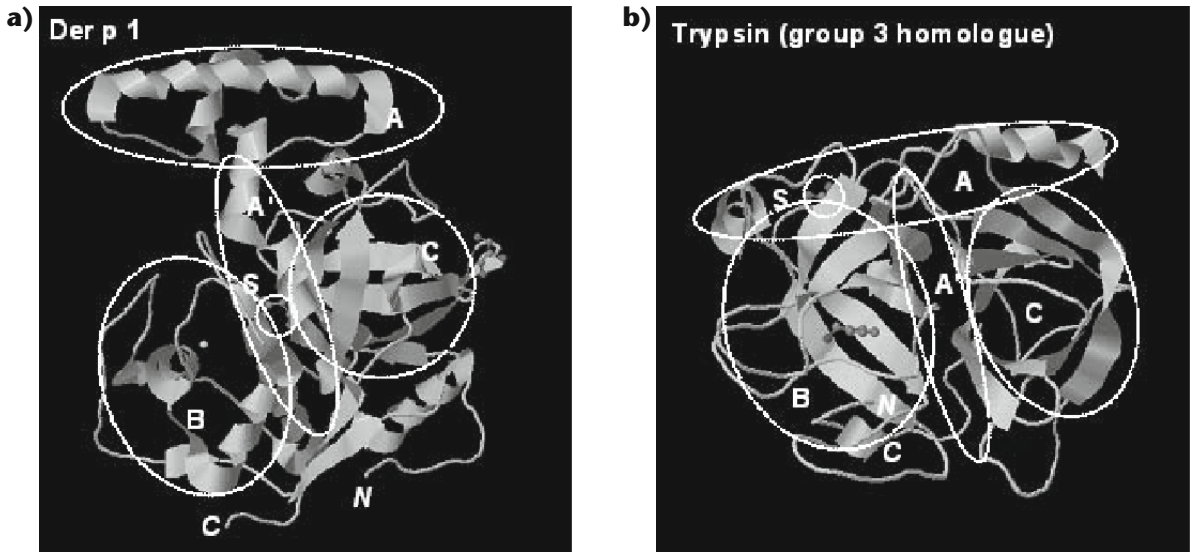


Figure 7.2 Tertiary structures of members of the peptidase family of mite allergens (groups 1, 3, 6 and 9) and related proteins. **a)** Der p 1 in monomeric form with pro-region attached (circled area, A and A') (Meno *et al.*, 2005; RCSB Protein Data Bank entry (PDB) 1XKG). The left-hand region is predominantly alpha-helix, the right-hand region beta-strands. The cleft-like region between them contains the catalytic site (S) and in the inactivated form of the molecule is covered by the A′ region of the pro-enzyme. **b)** Trypsin (group 3 allergens) (Leiros *et al.*, 2004; PDB 1UTN). A is the alpha-helical region, B and C are the left-hand and right-hand beta-strand regions, A' is the central cleft between the two and S is the active site.

allergens access to underlying antigen-presenting cells (Herbert *et al.*, 1995; Wan *et al.*, 1999, 2001).

- Der p 1 can cleave immunomodulators such as CD23, the human low affinity IgE receptor, and CD25, the alpha-subunit of the human interleukin 2 receptor (reviewed by Shakib *et al.*, 1998).
- Gough *et al.* (1999) found enzymically active Der p 1 tripled the amount of Der p 1-specific IgE produced by mice compared with inactivated control Der p 1. Active Der p 1 also enhanced the IgE response to ovalbumin (Gough *et al.*, 2001). This adjuvant effect was IgE-specific. There was no similar effect on IgG antibody production.
- The proteoloytic activity of Der p 1, 3 and 9 is responsible for triggering the release of pro-inflammatory mediators such as interleukins. Der p 3 can switch on protease-activated receptors (PARs) on respiratory epithelial cells resulting in the release of interleukin-8 (IL-8). Der p 1 can also induce IL-8 but via a different mechanism (Adam *et al.*, 2006).
- *D. farinae* serine protease (probably Der f 3) has kallikrein activity, causing liberation of kinins from kininogens (Takahashi *et al.*, 1990). Kinins are important inflammatory mediators.
- Der p 3 and Der f 3 were found to activate the complement system (Mauro *et al.*, 1997).

a Group 1 allergens

The group 1 allergens are polymorphic, 25 kDa, acidic/neutral proteins. They belong to the cysteine group of proteolytic enzymes which include the mammalian enzymes cathepsin B and H, and the plant enzymes actinidin and papain. Group 1 allergens are recognised by at least 70% of mite allergic individuals. They have been identified in *Dermatophagoides pteronyssinus*, *D. farinae*, *D. microceras*, *D. siboney*, *Euroglyphus maynei* and *Blomia tropicalis* (Chapman and Platts-Mills, 1980; Stewart and Turner, 1980; Lind, 1986b; Chua *et al.*, 1988; Kent *et al.*, 1992; Ferrándiz *et al.*, 1995b; Mora *et al.*, 2003), as well as in the mange mites *Psoroptes cuniculi* and *P. ovis* (Stewart and Fisher, 1986; Lee *et al.*, 2002) and the scabies mite *Sarcoptes scabiei* (Holt *et al.*, 2004). The allergens are found in both whole body and faecal extracts, and are synthesised by cells lining the gastrointestinal tract (Tovey and Baldo, 1990; see section 7.3).

Group 1 allergens have not been isolated from *Acarus*, *Tyrophagus*, *Glycyphagus* and *Lepidoglyphus* spp. and it seems unlikely that these mites possess them or they would have been detected by now. One reason might be the diet of storage mites (seed storage proteins and the moulds that live on them) is a rich source of protease inhibitors. Barber *et al.* (1996) found cereal prolamins inactivated Der p 1. It may be that

stored products mites use alternative proteinases that are not inhibited by compounds in grain or moulds.

The amino acid sequences of Group 1 allergens from the pyroglyphids (*Dermatophagoides* spp. and *Euroglyphus maynei*) comprise a signal peptide of 18/19 residues in *Dermatophagoides* spp., followed by a pro-peptide of 79/80 residues and the mature protein of 222–223 residues (Chua *et al.,* 1988; Dilworth *et al.,* 1991; Kent *et al.,* 1992; Table 7.1; Figure 7.9). Other cysteine proteases also have transient pre- and pro-form intermediates (Chua *et al.,* 1988; Dilworth *et al.,* 1991). The enzyme is produced in an inactive form and becomes active following cleavage of the pro-peptide by auto-catalysis. A similar situation probably applies to Pso o 1 (Lee *et al.,* 2002), as well as for other peptidase allergens.

The crystal structure of the pro-enzymic form of Der p 1 (Meno *et al.,* 2005; RCSB Protein Data Bank entry: 1XKG) shows the active site is located in a large surface cleft. The pro-peptide, which consists of four alpha-helices, is folded in such a way that it covers the cleft and active site, rendering the enzyme inactive. The mature Der p 1 is a globular molecule folded into two domains separated by the surface cleft. The left-hand domain (consisting of the N-terminal half of the mature protein: residues 21–116 in Figure 7.9) is predominantly alpha-helix and the right-hand domain (the C-terminal part: residues 117–223, Figure 7.9) is mostly beta-sheet. This overall structure is concordant with the model of the 3-D structure of Der p 1 by Topham *et al.* (1994) and is illustrated in Figure 7.2a. The active site consists of Q29 and C35 on the left and H171 and N190 on the right.

de Halleux *et al.* (2006) showed the mature, active form of Der p 1 undergoes dimerisation at pH 8 but is monomeric at pH 7.5. X-ray crystallography showed the monomers bound via H73 on the first momomer to D147, D149 and Y166 on the second. N169 binds to its counterpart on the other monomer. This region would be blocked by the pro-enzyme, so only the mature protein can dimerise. The authors suggest that the concentration of Der p 1 in faecal pellets is sufficiently high to maintain dimerism under natural conditions and infer that the dimer is likely to be more allergenic than the monomer. Interestingly the midgut of dust mites, where digestive hydrolysis takes place, is slightly acid and Der p 1 would likely exist there as a monomer. But the hindgut, where faecal pellets accumulate, is neutral or slightly alkaline, allowing for dimers to form.

Meno *et al.* (2005) and de Halleux *et al.* (2006) found that Der p 1 contains a metal binding site near the catalytic site, though its biological function is not known.

The pyroglyphid Group 1 allergens have the highest sequence identity (ca. 78% between the three allergens). Pso o 1 has ca. 58% identity with the pyroglyphid allergens, with which it shares a potential glycosylation site at N53, and only 42% with Sar s 1. Blo t 1 has ca. 31% identity with the pyroglyphids and 29% each with Sar s 1 and Pso 1. Sar s 1 has 34% identity with the pyroglyphids. Highest homology is in the region of the active site and the cysteine residues (Figure 7.9). The pairing pattern of cysteines in disulphide bridge formation was identified by Memo *et al.* (2005) as C4–C118, C32–C72 and C66–C104. C118 is substituted by asparagine in Blo t 1. Also H73, the crucial residue for dimerisation in Der p 1, is substituted for aspartate in Sar s 1, glycine in Pso o 1 and glutamate in Blo t 1.

Der p 1 has only one gene (W.A. Smith *et al.,* 2001), but since isoallergens exist, they must be alleles. By contrast, the scabies mite, *Sarcoptes scabiei*, has at least 10 group 1 genes, with signal peptides of 19–26 amino acids and mature proteins of 225–251 amino acids (Holt *et al.,* 2004). Some of the genes have mutations that would render the protein inactive. A similar multi-gene family exists for Sar s 3 (see below). This phenomenon is thought to be an adaptation to parasitism involving evasion of the host immune response (Holt *et al.,* 2003).

b Group 3 allergens

The group 3 allergens are trypsin-like serine protease enzymes with calculated molecular weights of 23–25 kDa, although on SDS-PAGE gels they migrate as a 30 kDa band. They have been isolated from bodies and faecal pellets of *Dermatophagoides farinae*, *D. pteronyssinus*, *D. siboney*, *Euroglyphus maynei*, *Sarcoptes scabiei*, *Glycyphagus domesticus*, *Lepidoglyphus destructor* and *Blomia tropicalis* (see Tables 7.1 and 7.3).

The group 3 allergens have a 15–20 residue signal peptide followed by a pro-enzyme sequence of 9–26 residues and a mature protein of 219–232 residues (Figure 7.11). Recombinant Der f 3 expressed with the intact pro-sequence was enzymatically inactive, but became active on removal of the pro-sequence (Nishiyama *et al.,* 1995). This is similar to the situation with group 1 allergens where the pro-enzyme prevents enzymatic activity by occluding the active

site. The group 3 allergens belong to the S1 family of serine proteases, as do group 6 and 9 allergens. All of them possess the catalytic residues H41, D85 and S185 characteristic of the S1 chymotrypsin family of serine proteases, as well as a series of conserved sequences associated with the active sites (Figure 7.2b; Rawlings and Barrett, 1994; their Figure 2).

The group 3 allergens are polymorphic and in Der p 3 are allelic, since the allergen is encoded by a single gene (Smith and Thomas, 1996). *Sarcoptes scabiei* has a multi-gene family for group 3 allergens (Holt *et al.,* 2003).

Pyroglyphid and *Blomia* group 3 allergens are not glycosylated. Gly d 2 and Lep d 2 have a single putative glycosylation site and Sar s 3 has two sites. Conserved cysteine residues are paired to form disulphide bonds as in trypsin (numbering for Der f 3; Figure 7.11): C26–C42, C152–C169 and C181–C207. In the pyroglyphid allergens there is also an unpaired conserved cysteine (C12). Amino acid sequence homology with Der f 3 shows 82% identity with Eur m 3, 80% with Der p 3, 49% with Blo t 3, and 45% with trypsin from the mosquito *Anopheles gambiae*. Sar s 3 shows ca. 43–44% identity with the pyroglyphid allergens.

The IgE binding frequency of group 3 allergens is relatively high (mean 67%, range 40–100%; Table 7.6). Members of group 3 are considered to be major allergens. The enzymatic activity of serine protease allergens may enhance their allergenicity though initiating non-allergic inflammatory responses, as with group 1 enzyme-allergens. Der p 3 and Der p 9 activated the receptor PAF_2 on lung epithelia which triggers the release of the pro-inflammatory cytokines GM-CSF and eotaxin (Sun *et al.,* 2001). Der p 3 and Der f 3 were found to activate the complement system, cleaving the complement components C3 and C5 to produce the anaphylatoxins C3a and C5a (Mauro *et al.,* 1997).

c Group 6 allergens

The group 6 allergens are 25 kDa serine proteases that have chymotrypsin activity based on their substrate affinity (Yasueda *et al.,* 1993). They are known from *Dermatophagoides farinae*, *D. pteronyssinus* and *Blomia tropicalis*. Group 6 allergens appear to be recognised by a markedly lower proportion of mite allergic individuals than the other peptidases (31–65%; Table 7.6; Yasueda *et al.,* 1993; King *et al.,* 1996; Bennett and Thomas, 1996; Kawamoto *et al.,* 1999).

Der 6 allergens have no N-glycosylation sites on the mature protein and consist of 16–17 hydrophobic residues, indicating a signal peptide (Kawamoto *et al.,* 1999), followed by a 32–34 pro-enzyme and a 230–231 mature protein. The active sites and conserved residues around the active sites are the same as in groups 3 and 9: typical of S1 peptidases. The six conserved cysteines are paired similarly. Der p 6 and Der f 6 have 75% identity. They share about 35–37% sequence identity with the group 3 allergens and about 27–30% with group 9 (see Figures 7.2, 7.14, 7.17). There is little cross-reactivity between groups 3 and 6 (Yasueda *et al.,* 1993). Bennett and Thomas (1996) identified serine 178 as conserved in all chymotrypsins and essential for substrate specificity. In Blo t 6 (GenBank accession no. AY291325, Chew *et al.,* unpublished) this serine is substituted by a glutamine.

d Group 9 allergens

The group 9 allergens migrate to positions around 27–30 kDa on SDS-PAGE and have a calculated molecular weight of 23.8–24.5 kDa (Table 7.1). They are serine proteases with collagenase activity and have been isolated and characterised from *Dermatophagoides farinae*, *D. pteronyssinus* and *Blomia tropicalis*. Collagenase activity was also detected in *Euroglyphus maynei* (Stewart *et al.,* 1994). Der p 9 was isolated from faeces-enriched extracts and was found to cleave collagen (King *et al.,* 1996). It showed IgE-binding to 92% of sera from mite-allergic patients, and had limited cross-reactivity with Der p 3 but not Der p 6. The relatively high concentration of Der p 9 isolated from the faeces (2.8% w/w) suggests group 9 allergens are not only important digestive enzymes but are likely to be present in substantial concentrations in homes. This may account for the high IgE-binding frequency (Table 7.6). The proteolytic activity of Der p 9 (along with Der p 3) activated PAF_2 to induce release of the cytokines GM-CSF and eotaxin (Sun *et al.,* 2001).

Der p 9 has a signal peptide of 16 residues (based on Der p 3), a pro-enzyme region of nine residues and a mature protein of 219 residues. A BLAST search showed the Der p 9 sequence in Figure 7.17 had 70% sequence identity with Der p 3 and 61% with Der f 3. The active sites and conserved residues around the active sites are the same as in groups 3 and 6. There are six conserved cysteines and no N-glycosylation sites. In Der p 9, the equivalent of serine 178, conserved in all chymotrypsins and essential for substrate specificity, is aspartate 166, although it is still serine in the Der f 9 and Blo t 9 sequences.

Table 7.6 IgE binding frequency to enzyme allergens and non-enzyme allergens. * = sera from children.

Enzyme allergens. Native [n] or recombinant [r]	% frequency of IgE response	Raw frequency data	Reference	Non-enzyme allergens. Native [n] or recombinant [r]	% frequency of IgE response	Raw frequency data	Reference
nDer f 1	79	18/23	Ando *et al.*, 1993	nDer f 2	83	19/23	Ando *et al.*, 1993
nDer f 1	89	78/88	Yasueda *et al.*, 1993	nDer p 2	93	82/88	Yasueda *et al.*, 1993
nDer f 1	90	28/31	Aki *et al.*, 1995	nDer f 2	93	82/88	Yasueda *et al.*, 1993
nDer p 1	100	51/51	Stewart *et al.*, 1992	nDer p 2	100	35/35	King *et al.*, 1996
nDer p 1	90	79/88	Yasueda *et al.*, 1993	rDer p 2	100	10/10	Weghofer *et al.*, 2005
nDer p 1	96	34/35	King *et al.*, 1996	rDer p 2	88	36/41	Shen *et al.*, 1996
nDer p 1	100	10/10	Weghofer *et al.*, 2005	rDer p 2	90	27/30	Shen *et al.*, 1993
nDer p 1	93	28/30	Shen *et al.*, 1993	rDer p 2	82	14/17	Chua *et al.*, 1990b
rBlo t 1	62	17/27	Mora *et al.*, 2003	nGly d 2	94	16/17	Gafvelin *et al.*, 2001
nDer f 3	70	16/23	Ando *et al.*, 1993	rGly d 2	94	16/17	Gafvelin *et al.*, 2001
nDer f 3	42	37/88	Yasueda *et al.*, 1993	nLep d 2	95	21/22	Ventas *et al.*, 1991
nDer p 3	100	55/55	Stewart *et al.*, 1992	rLep d 2	59		Kronqvist *et al.*, 2000 (cf. Eriksson *et al.*, 2001)
nDer p 3	51	45/88	Yasueda *et al.*, 1993	nDer f 2	74	23/31	Aki *et al.*, 1995
nDer p 3	97	34/35	King *et al.*, 1996	rDer p 2	92	22/24	Tsai *et al.*, 1998
rBlo t 3	51	23/45	Cheong *et al.*, 2003	Der p 5	52	13/25	Lin *et al.*, 1994
nDer p 4	46	12/27	Lake *et al.*, 1991	Lep d 5	9	4/45	Eriksson *et al.*, 2001
nDer p 4	25	5/20*	Lake *et al.*, 1991	Der p 5	50	5/10	Weghofer *et al.*, 2005
rDer p 4	30	3/10	Mills *et al.*, 1999	Der p 5	55	21/38	Tovey *et al.*, 1989
nDer p 4	50	5/10	Weghofer *et al.*, 2005	Blo t 5	47	47/100	Caraballo *et al.*, 1996
nDer f 6	31	27/88	Yasueda *et al.*, 1993	Lep d 7	62	28/45	Eriksson *et al.*, 2001
rDer f 6	40	15/38	Kawamoto *et al.*, 1999	Der p 7	30	3/10	Weghofer *et al.*, 2005
nDer p 6	39	34/88	Yasueda *et al.*, 1993	Der p 7	46	19/41	Shen *et al.*, 1996
nDer p 6	65	24/35	King *et al.*,. 1996	rDer p 7	53	16/30	Shen *et al.*, 1993
nDer p 8	96	53/55	C. Huang *et al.*, 2006	Der p 10	0	0/10	Weghofer *et al.*, 2005
nDer p 8	75	15/20	C. Huang *et al.*, 2006	rDer p 10	6	4/71	Asturias *et al.*, 1998
nDer p 8	65	11/17	C. Huang *et al.*, 2006	nDer f 10	81	25/31	Aki *et al.*, 1995
rDer p 8	84	46/55	C. Huang *et al.*, 2006	rBlo t 10	29	27/93	Yi *et al.*, 2002
rDer p 8	20	2/10	Weghofer *et al.*, 2005	rBlo t 10	20	7/35	Yi *et al.*, 2002
rDer p 8	40	77/193	O'Neill *et al.*, 1994	rBlo t 11	52	33/63	Ramos *et al.*, 2001
nDer p 9	92	32/35	King *et al.*, 1996	nDer f 11	88	21/24	Tsai *et al.*, 1998
rDer p 15	70	19/27	O'Neill *et al.*, 2006	rBlo t 11	5	1/22	Teo *et al.*, 2006
nDer f 18	54	13/24	Weber *et al.*, 2003	nBlo t 11	77	17/22	Teo *et al.*, 2006
rDer p 18	63	17/27	O'Neill *et al.*, 2006	rBlo t 11	50	11/22	Teo *et al.*, 2006
				rDer p 11	50	50/100	Lee *et al.*, 2004
				rBlo t 12	50	16/32	Puerta *et al.*, 1996a
				Lep d 13	13	6/45	Eriksson *et al.*, 2001
				rAca s 13	23	3/13	Eriksson *et al.*, 1999
				rBlo t 13	11		Caraballo *et al.*, 1997
				Der p 14	0	0/10	Weghofer *et al.*, 2005
				rDer f 14	30	13/43	Aki *et al.*, 1994a
				nDer f 14	70	16/23	Fujikawa *et al.*, 1996
				rDer f 14	39	9/23	Fujikawa *et al.*, 1996
				rDer f 16	35	ND	Tategaki *et al.*, 2000
				rDer f 16	47	8/17	Kawamoto *et al.*, 2002a
				rDer f 17	35	ND	Tategaki *et al.*, 2000
				rBlo t 19	10	ND	Thomas *et al.*, 2002
				rHSP70	10	4/41	Aki *et al.*, 1994a

7.4.4 The glycosidases

This group of allergens belongs to the enzyme class EC 3.2, the glycosidases (Table 7.5; cf also Chapter 2, 2.2.3b). It includes allergen group 4 (alpha-amylase) and groups 15 and 18 (chitinases). Blo t 12 is a chitin-binding protein possessing a peritrophin A domain. It is included here because it shows some sequence homology with the chitinases though it is probably not an enzyme.

With the exception of group 12 (146 residues) this family consists of comparatively large molecules as mite allergens go (437–555 residues). With regard to their amino acid composition, the highest frequency is histidine (mean D_t 3.3), then threonine (1.6), tyrosine (1.5) and tryptophan (1.4). The lowest is alanine (0.4) then arginine (0.7). The glycosidases have a relatively high cysteine content (mean D_t 1.0), a feature they share with the peptidases (Figure 7.1b).

Alpha-amylases and chitinases share basically the same tertiary structure, an $(\alpha/\beta)_8$ barrel which is detailed below. The tertiary structure of the group 12 allergens are not known, but they are too small to be an $(\alpha/\beta)_8$ barrel. The secondary structure prediction indicates two major N-terminal alpha-helices and six major beta-strands.

a Group 4 allergens

The group 4 allergens are alpha-amylases (Lake *et al.*, 1991), migrating on SDS-PAGE at 60 kDa, and with a calculated molecular weight of 55–57 kDa. They have

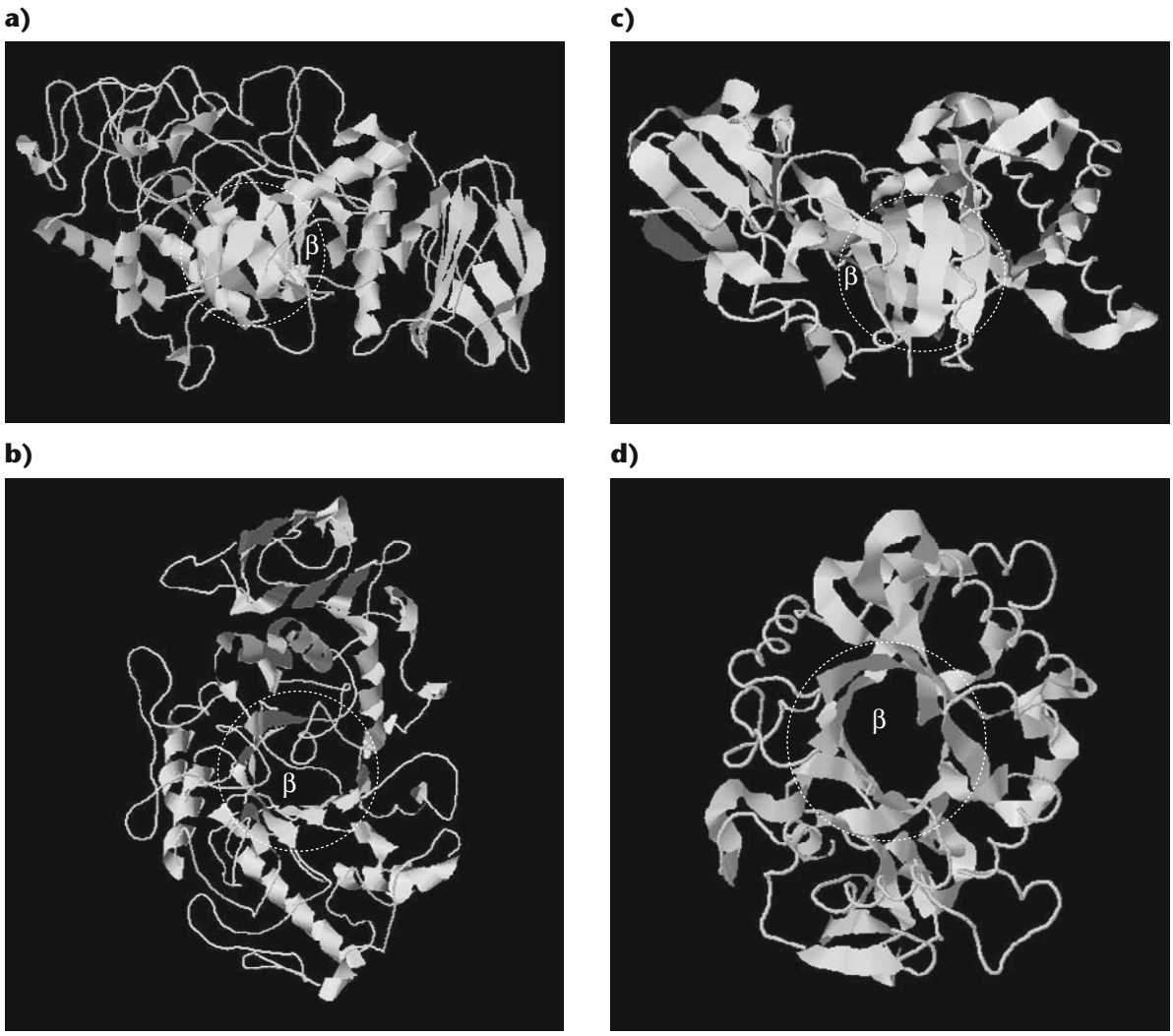


Figure 7.3 Tertiary structures of proteins related to members of the glycosidase family of mite allergens (groups 4, 12, 15 and 18). **a)** Lateral view; and **b)** end-on view of alpha-amylase homologue (mealworm, PDB 1CLV) of group 4 allergens showing an $(\alpha/\beta)_8$ barrel structure (β), which Janeček (1997) referred to as 'this charming fold'. **c)** Lateral view; and **d)** end-on view of chitinase homologue (human, PDB 1HKK) of groups 15 and 18.

Table 7.7 Structural domains identified within dust mite allergen molecules, cf. SDAP – Structural Database of Allergenic Proteins (http://fermi.utmb.edu/SDAP/index.html) and Pfam, the Protein Families Database (http://www.sanger.ac.uk/Software/Pfam/).

Allergen group	Domain/family	Examples of non-mite allergens with same domain
1	Peptidase C1 (Papain family cysteine protease)	Kiwi fruit: Act c 1; pineapple: Ana c 2; papaya: Car p 1
2	1) ML: MD-2-related lipid recognition domain 2) Immunoglobulin E-set domain	None
3	Tryp-SPc: trypsin-like serine protease	Bee: Api m 7
4	1) Alpha-amylase, catalytic domain 2) Alpha-amylase, C-terminal all-beta domain	Fungi: Asp o 21
5	No conserved domains detected	None
6	Tryp-SPc: trypsin-like serine protease	None
7	No conserved domains detected	None
8	1) Glutathione S-transferase, N-terminal domain 2) Glutathione S-transferase, C-terminal domain	Cockroach: Bla g 5 Fungi: Pen c 24
9	Tryp-SPc: trypsin-like serine protease	None
10	Tropomyosin	Nematode: Ani s 3; insects: Chi k 10, Lep s 1, Per a 7; Crustacea: Pen a 1, Cha f 1; molluscs: Cra g 1, Per v 1, Hal d 1, Mim n 1
11	1) Myosin tail 2) Smc, chromosome segregation ATPases	Nematode: Ani s 2
12	Chitin binding Peritrophin-A domain	None
13	Lipocalin/cytosolic fatty-acid binding protein family	Cow: Bos d 2; horse: Equ c 1; dog: Can f 1, Can f 2; cat: Fel d 4; mouse: Mus m 1; rat: Rat n 1
14	LPD-N Lipoprotein N-terminal domain	Chicken: Gal d vitellogenin
15	1) Glycosyl hydrolases family 18 2) Chitin binding peritrophin-A domain	None None
16	GEL: 4 gelsolin homology domains	None
17	Calcium-binding EF hand	None
18	Glyco 18 domain	None
19	No conserved domains detected	None
20	1) ATP guanido phospotransferase, N-terminal domain 2) ATP guanido phospotransferase, C-terminal catalytic domain 3) Actin-binding domain	None
21	No conserved domains detected	None
Aldehyde dehydrogenase	Aldedh, Aldehyde dehydrogenase	Fungi: Alt a 10
13.8 kDa allergen	NLPC_P60	
Lipase homologue	Pancreatic lipase-like	Wasp: Ves m 1
Hexosaminidase	Glycosyl hydrolase family 20, catalytic domain	Plants: Ole e 10
Cathepsin F	Peptidase_C1A	None
Cathepsin D	Eukaryotic aspartyl protease	Insects: Bla g 2; fungi: Asp f 10
Enolase	Enolase	Fungi: Alt a 6, Alt a 11, Asp f22, Cla h 6, Pen c 22; plants: Hev b 9
Cyclophilin	cyclophilin_ABH_like	Fungi: Asp f 11, Mala s 6; plants: Bet v 7
Thaumatin	THN, Thaumatin family	Plant: Act c 2, Jun a 3, Mal d 2
Cystatin A	Cystatin domain	Cat: Fel d 3
Heat shock protein 70	HSP70	Fungi: Alt a 3
IGFP	1) IB, insulin growth factor-binding protein homologues 2) KAZAL, Kazal type serine protease inhibitors 3) IGcam, immunoglobulin cell adhesion molecule	Chicken: Gal d 1
Alpha-tubulin	alpha_tubulin	None
Profilin	PROF, Profilin	Plants: Amb a 8, Ara h 5, Cyn d 12, Hev b 8, Mal d 4,Ole e 2, Phl p 12
Lipid transfer protein	AAI_LTSS; Alpha Amylase Inhibitor	Plants: Amb a 6, Gly m 1, Mal d 3
Ribosomal protein P2	RPP1A, Ribosomal protein	None

been isolated and characterised from *Dermatophagoides farinae*, *D. pteronyssinus*, *Euroglyphus maynei* and *Blomia tropicalis* (Tables 7.2, 7.3; Lake *et al.,* 1991; Mills *et al.,* 1999; Chew *et al.,* unpublished), and alpha-amylases have been detected in several stored products species (Bowman, 1984; Stewart *et al.,* 1991; cf. 2.2.3b).

The group 4 allergens have a 22–25 residue signal sequence followed by a 496 residue mature protein (484 for Blo t 4; Figure 7.12). Blo t 4 has no N-glycosylation sites. Der p 4 has one at N10, Eur m 4 has one at N147. There are eight conserved cysteines that form disulphide bonds.

Alpha-amylases from animals are highly conserved along the entire sequence (Janeček, 1997). They have been used as 'molecular clocks' to estimate the timing of evolutionary divergence of animal groups (Hickey *et al.,* 1987). Amino acid sequences of Der p 4 and Eur m 4 were 94% identical, and Der p 4 shows 55–60% identity with insect alpha-amylases such as those from the flies *Drosophila* spp. and *Bibio marci* and the beetle *Tenebrio molitor*. Several arthropods have multiple genes for amylases, and isoforms of mite amylases have been noted (Bowman and Lessiter, 1985; Lake *et al.,* 1991). Cereals contain amylase inhibitors, which are important occupational allergens (Sánchez-Monge *et al.,* 1996). In stored products mites, multi-genes for amylase isozymes may be an evolutionary response to dietary inhibitors.

The three-dimensional structure of amylase consists of an $(\alpha/\beta)_8$ barrel – similar but not identical to a TIM barrel – consisting of eight parallel beta-strands forming an inner cylindrical beta-sheet, surrounded by eight alpha-helices, thus the barrel consists of eight repeated alpha/beta units (Janeček, 1997; Figure 7.3).

Group 4 allergens have intermediate to low frequency of IgE binding (mean 38%; Table 7.6). Frequency of IgE binding to a panel of sera from children was only 25%, compared with 46% with sera from adults (Lake *et al.,* 1991). This may reflect the relative abundance of this allergen: it was isolated from spent mite growth medium in far lower concentrations than were serine proteases (Lake *et al.,* 1991; King *et al.,* 1996), though Mills *et al.,* (1999) pointed out that levels could potentially vary with differential gene expression.

b Group 12 allergens

This group consists of the 14 kDa Blo t 12 (Puerta *et al.,* 1996a) and Lep d 12 (AY293744; Chew *et al.,* unpublished; Table 7.1). The composition of the 123–125 amino acid residues of the mature proteins fits best with the chitinase groups 15 and 18 (Figure 7.1b), although the mean transformed Dayhoff value for histidine in Blo t 12 is markedly higher than for the others.

A BLAST search revealed some limited homology with a hypothetical protein from the fruit fly *Drosophila melanogaster*, a peritrophic membrane binding protein from the cabbage looper moth *Trichoplusia ni* and Der f 15 (Figure 7.20). The peritrophic membrane (PM) chitin binding protein, CBP1 from *Trichoplusia ni*, isolated by Wang *et al.* (2004), is much larger than Der p 12, containing 1171 amino acid residues (124 kDa). The sequence similarity is most apparent in the C-terminal end of the CBP1 protein, predominantly in the cysteine-rich, chitin-binding, peritrophin-A domain which has a general consensus signal $CX_{13–20}CX_{5–6}CX_{9–19}CX_{10–14}CX_{4–14}C$ (Tellam *et al.,* 1999). The terminal cysteine appears to be absent from the Blo t 12 amino acid sequence entered in Genbank (BTU27479). However, if the nucleotide sequence for Blo t 12 is run through a translation program (e.g. *Translate* [GCG]), there is an extra Cys residue at position 151 ($N_{146}XLINC_{151}$). The peritrophin-A domain, underlined for CBP1 in Figure 7.20, corresponds to domain no. 11 of Wang *et al.* (2004; their Figure 1).

The high cysteine content indicates that a key structural property of the peritrophin-A domain is resistance to proteases in the arthropod midgut. It is involved in the binding of PM proteins to chitin fibrils and forming the structural basis of the peritrophic envelope (see Chapter 2). It appears probable that Blo t 12 is involved in this process.

Puerta *et al.* (1996a) found recombinant Blo t 12 bound IgE in sera of patients allergic to *Blomia tropicalis* with a frequency of 50% (Table 7.6).

c Group 15 allergens

Group 15 allergens have been characterised from *Dermatophagoides farinae* (McCall *et al.,* 2001) and *D. pteronyssinus* (O' Neill *et al.,* 2006). Der f 15 is a protein of 63.2 kDa and 555 amino acids which on SDS-PAGE has an apparent molecular weight of 98/109 kDa due to extensive *O*-glycosylation (almost 50% is carbohydrate) along a proline/serine/threonine-rich *O*-glycosylation domain of 84 residues (McCall *et al.,* 2001). It shows homology with insect chitinases which belong to family 18 of the glucohydrolase superfamily (Figure 7.23). The highly conserved sequence YXFDG-LDLDWEYP is typical of insect chitinases (Kramer

and Muthukrishnan, 1997; Weber *et al.*, 2003; Merzendorfer and Zimoch, 2003) and is consistent with the active site signature of the family 18 chitinases, including the glutamate residue (E) that is essential for catalysis (Figure 7.23). As well as this catalytic domain, family 18 chitinases possess a PEST-region (proline [P]/glutamate [E]/serine [S]/threonine [T]-rich) associated with *O*-glycosylation and a cysteine-rich region (Kramer and Muthukrishnan, 1997).

Manduca sexta chitinase was the first insect chitinase to be cloned and sequenced. The enzyme is expressed during moulting and is up-regulated by ecdysteroid and down-regulated by juvenile hormone (Kramer *et al.*, 1993). The sequence is aligned with Der p 15 and Der f 15 in Figure 7.23 and shows 35% identity. Der p 15 and Der f 15 have ca. 90% identity. Der f 15 also shows some homology with Blo t 12 (Puerta *et al.*, 1996a) and Der f 18. Like Der f 18, Der f 15 was found to be a major allergen for atopic dogs sensitised to dust mites, with 80% frequency of IgE-binding of sera. O' Neill *et al.* (2006) found 70% IgE-binding frequency with mite-sensitive human sera. Prior to this, the only published data on IgE-binding frequency of group 15 allergens was a 2-D immunoblot study by Le Mao *et al.* (1998). The amino acid sequence of part of one of the proteins (AFEPHGYLLTAAV) shows 100% identity with Der f 15 (residues 166–178). The amino acid sequences for two isoforms of Der p 15 have been published (O' Neill *et al.*, 2006; Table 7.1).

d Group 18 allergens

Group 18 allergens are known from *Dermatophagoides farinae*, *D. pteronyssinus* and *Blomia tropicalis* (Table 7.2). Der f 18 is a 60 kDa (50 kDa for the mature protein) homologue of chitinase (Weber *et al.*, 2003). Some 54% of human sera that were positive to extracts of *D. farinae* were bound by Der f 18, and some 57–77% of canine sera. The amino acid sequence for Der p 18 has been published (O'Neill *et al.*, 2006; Table 7.1) and shows a high degree of homology with Der f 18 (Figure 7.25). Both allergens consist of a 25 amino acid signal sequence and a 437 amino acid mature protein. A BLAST search of the Der p 18 sequence showed highest homology to other invertebrate chitinases. Homology with the group 15 allergens (cf. above) is only about 30%; highest in the regions of the conserved cysteine residues and the two chitinase catalytic domains (Figures 7.23, 7.25).

7.4.5 The transferases

The transferases belong to EC group 2 and include allergen groups 8 (glutathione transferase) and 20 (arginine kinase). The amino acid composition of the transferase mite allergen family is fairly uniform between the two groups and quite different from the other families (Figure 7.1c; Table 7.5). It has relatively high frequencies of methionine, tyrosine, phenylalanine and leucine (mean Dt of 2.1, 1.5, 1.5 and 1.4 respectively). Lowest mean frequencies are for cysteine (0.4) and serine (0.6).

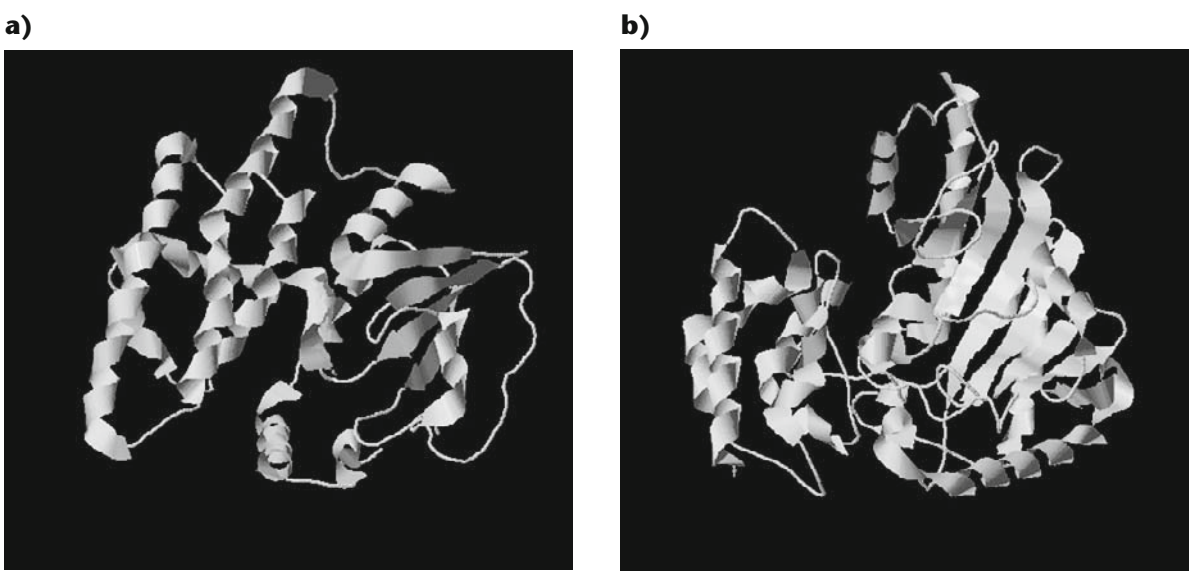


Figure 7.4 Tertiary structures of proteins related to members of the transferase family of mite allergens (groups 8 and 20). **a)** Glutathione s-transferase (human, PDB 1EEM); **b)** arginine kinase (PDB 1P52).

Tertiary structure based on X-ray crystallography of glutathione transferase and arginine kinase (Figure 7.4) indicates they are somewhat similar to the proteases: globular molecules folded into two domains separated by a surface cleft.

a Group 8 allergens

The group 8 allergens have been isolated from *Dermatophagoides pteronyssinus*, *Psoroptes ovis*, *Sarcoptes scabiei*, *Glycyphagus domesticus*, *Lepidoglyphus destructor* and *Blomia tropicalis* (Tables 7.1–7.3). Sequence identity is moderate (Figure 7.16).

The group 8 allergens are 26 kDa glutathione transferases (EC 2.5.1.18; Figure 7.4a). EC 2.5.1 enzymes are involved in transferring alkyl or aryl groups other than methyl groups. (RX + glutathione = HX + R-S-glutathione, where R can be an aromatic, aliphatic or heterocyclic group and X can be a sulphate, nitrile or halide.) These broad-spectrum enzymes are thought to play a major role in detoxification (Armstrong, 1997). They can also act as peroxidases, isomerases, or in the sequestration of hydrophobic molecules. Substrates can range from toxic exogenous molecules to by-products of cellular metabolism.

Der p 8 has at least eight isoforms and may be encoded by a multi-gene family (Thomas *et al.,* 2002; C. Huang *et al.,* 2006). O' Neill *et al.* (1994a) found recombinant Der p 8 bound to 40% of sera. C. Huang *et al.* (2006) showed 65–96% IgE reactivity to nDer p 8 and 84% to rDer p 8 (Table 7.6).

Allergen Bla g 5 of the cockroach *Blattella germanica* is a glutathione S-transferase (Arruda *et al.,* 1997b). C. Huang *et al.* (2006) reported high cross-reactivity between Bla g 5 and Der p 8.

b Group 20 allergens

Der p 20 is an arginine kinase (Thomas, unpublished; EC 2.7.3.3; Figure 7.4b). These enzymes are involved in transferring phosphorous-containing groups. Arginine kinase catalyses the reaction involved in arginine metabolism in the urea cycle: ATP + *l*-arginine = ADP + *n*-phospho-*l*-arginine. Der p 20 shares 78% amino acid sequence homology with the allergen Pen m 2 from the prawn, *Penaeus monodon* (Figure 7.27) as well as several other arginine kinases from decapod crustaceans (Yu *et al.,* 2003).

7.4.6 The small alpha-helical proteins

This group includes the allergen groups 5, 7 and 21. Although their function is unknown, they appear not to be enzymes. Characteristic of their amino acid composition are relatively high frequencies of methionine, glutamate, lysine and isoleucine (Figure 7.1d). Lowest mean frequencies are for tryptophan, which is absent from group 5, and cysteine, which is absent from group 21 and almost absent from groups 5 and 7. Predicted secondary and tertiary structure (Figure 7.6b) shows proteins consisting predominantly of bundles of alpha-helices.

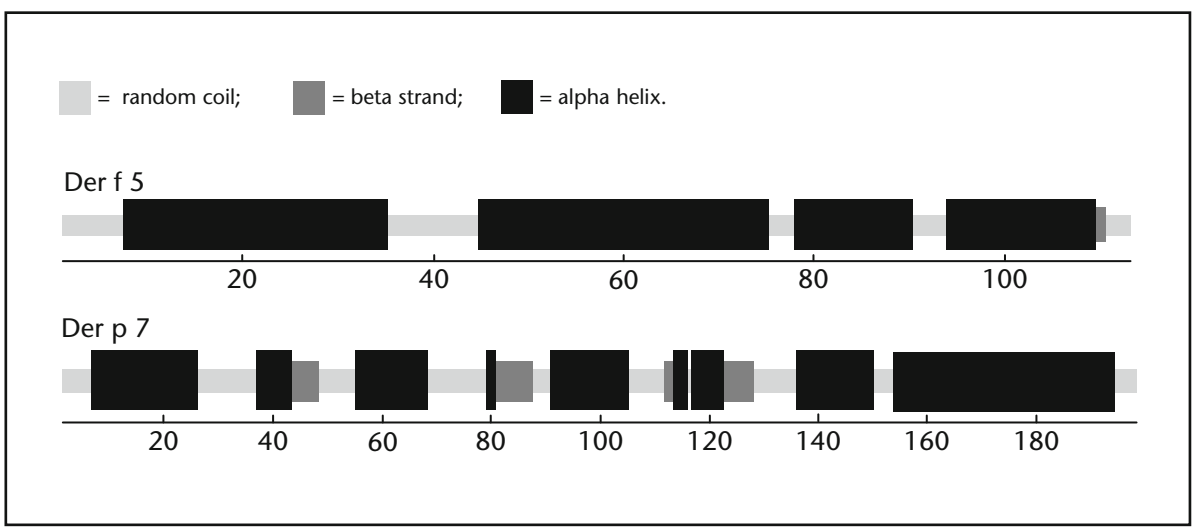


Figure 7.5 Topological map showing predicted secondary structure of two mite allergens of the small alpha-helical protein group. Scale bar = amino acid sequence length.

a Group 5 allergens

Group 5 allergens are known from *Dermatophagoides farinae*, *D. pteronyssinus*, *Glycyphagus domesticus*, *Lepidoglyphus destructor* and *Blomia tropicalis* (Table 7.2). The function of these allergens remains elusive, even though Der p 5 was one of the first allergens to be fully sequenced (Tovey *et al.,* 1989). On a BLAST search the highest matches of Der p 5 and Der f 5 are with other group 5 allergens, then relatively low (22–25%) identity (but ca. 50% similarity) with phospholipase A1 from *Shewanella* sp. (Bacteria), anthranilate isomerise from *Helicobacter pylori* (Bacteria) and glutathione peroxidase from *Tetrahymenia* (Protista). None of these matches is particularly convincing and no residue matches are longer than pairs.

The group 5 allergens have derived molecular weights of 12.5–14.2 kDa. They consist of 20–36 residue leader sequences and mature proteins of 112–113 amino acids (Figure 7.13). Two Gly d 5 isotypes, Gly d 5.01 and 5.02 (Table 7.1), are 26–27 kDa molecules but they show limited sequence identity with other group 5 allergens (data not shown). Isoforms are known for Der p 5 (Lin *et al.,* 1994), Blo t 5 (Yi *et al.,* 2004) and Lep d 5 (Eriksson *et al.,* 2001). The mature protein of Der p 5 in the native form starts at the aspartate residue (O'Neill *et al.,* 1994b) rather than the preceding glutamate, identified as residue 1 by Lin *et al.* (1994).

There is 76% sequence identity between Der p 1 and Der f 1, 44% between Blo t 5 and Lep d 5 and 45% between Blo t 5 and Der p 5 (Figure 7.13). There is only low cross-reactivity between Der p 5 and Blo t 5 (Kuo *et al.,* 2003). Der p 5 is recognised by serum IgE from about 50% of mite allergic individuals (Tovey *et al.,* 1989; Lin *et al.,* 1994). Blo t 5 is a major allergen, predominantly in tropical regions, with 70–92% IgE-binding frequency (Arruda *et al.,* 1995, 1997; Kuo *et al.,* 1999, 2003).

Group 5 allergens have no conserved cysteines and only Blo t 5 has a potential N-glycosylation site. Secondary structure prediction indicates alpha-helices separated by short coils (Figure 7.6b), confirmed by NMR (Naik *et al.,* 2008) for Blo t 5, which showed the tertiary structure consisted of a bundle of three coiled-coil alpha-helices each at least 28 residues long, showing a pattern of four heptad repeats. The heptad repeat pattern, denoted $(abcdefg)_n$), relates to the arrangement of hydrophobic residues in the core of the helix (Liaw *et al.,* 2001) at positions *a* and *d*. Blo t 5 polymerises irreversibly under acid conditions (pH 4–5) to form long filaments. The leader peptide sequence of Blo t 5 is predictive of a protein that is secreted extra-cellularly through acidic vesicles, indicating that polymerisation occurs during secretion and is part of protein function (Liaw *et al.,* 2001). But what function? Group 5 and group 21 allergens are closely related and Blo t 21 has been localised in the midgut epithelium hindgut and faecal pellets of *Blomia tropicalis* (Gao *et al.,* 2007). The evidence points to a highly charged structural protein, secreted from the midgut epithelium as a polymer. These characteristics suggest that group 5 and group 21 allergens may be involved in the formation of the peritrophic envelope, possibly comprising the protein matrix in which chitin fibrils are embedded (Chapter 2). This possible function would be consistent with the presence of Blo t 21 in the hindgut and faecal pellets.

b Group 7 allergens

Group 7 allergens are known from the same mite species as for group 5 (Table 7.2). Der p 7 is a 22 kDa protein recognised by about 50% of mite allergic individuals, consisting of a 17 residue signal peptide and a mature protein of 196–200 residues (Figure 7.15; Shen *et al.,* 1993; Shen *et al.,* 1995a). The biochemical identity of this allergen is unknown.

All the group 7 allergens have an N-glycosylation site, though in Blo t 7 and Gly d 7 it is in different positions from the *Dermatophagoides* allergens. The amino acid sequence of Der f 7, Der p 7 and Gly d 7 indicates a potential glycosylation site, and indeed Shen *et al.* (1995) found Der p 7 to be 16% carbohydrate and that it had several isoforms. Secondary structure prediction indicates seven extended alpha-helices separated by short coils and three short beta-strands (Figure 7.5).

IgE-binding frequency to Der p 7 was 37–50% (Shen *et al.,* 1993, 1995a, 1996), to Der f 7 46% (Shen *et al.,* 1995b) and to rLep d 7 it was 67% (Eriksson *et al.,* 2001) (Table 7.6).

c Group 21 allergens

Der p 21 is known only from a database entry (DQ354124; Weghofer *et al.,* unpublished). Interestingly, the signal peptide is quite similar to those of the group 2 allergens, particularly Ale o 2. Der p 21 has relatively high sequence identity to some of the group 5 allergens (Figure 7.13). This relationship was recently confirmed by Gao *et al.* (2007), who isolated and characterised Blo t 21. This protein consists of 129 amino acid residues and has 39% sequence homology with Blo t 5. It is the product of a single-

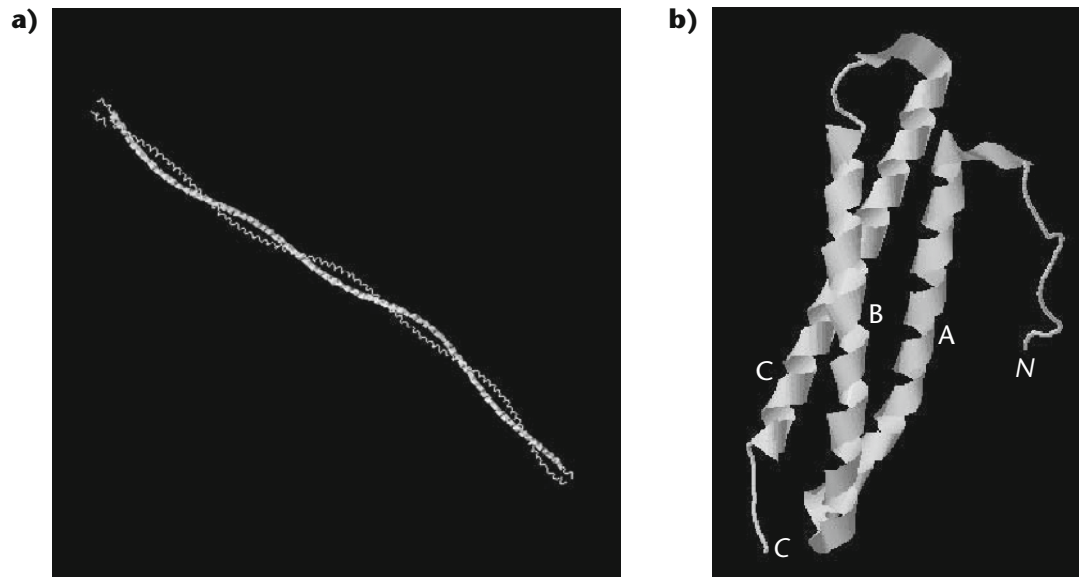


Figure 7.6 Tertiary structures of helical allergens. **a)** Tropomyosin – a homologue, the muscle protein family of mite allergens (groups 10 and 11); **b)** a member of the small alpha-helix protein family (groups 5,7 and 21), Blo t 5 (PDB 2JMH), showing three coiled-coil triple helices (A–C).

copy gene, and has an alpha-helical structure like the other group 5 allergens. Positive skin-prick tests to Blo t 21 were found in 60–95% of patients with asthma and/or rhinitis, and there was little cross-reactivity between Blo t 21 and Blo t 5. The localisation of Der p 21 and its possible function in the gut and faecal pellets has been mentioned above (see Group 5 allergens).

7.4.7 The muscle proteins

Included here are the allergen groups 10 (tropomyosin) and 11 (paramyosin). Characteristic of their amino acid composition are high frequencies of glutamate, glutamine, arginine and methionine (mean Dt of 3.1, 2.0, 1.6 and 1.6 respectively) (Figure 7.1e). Lowest mean frequencies are for cysteine (0.1) and proline (0.01). The tertiary structure of these molecules is a departure from the typical small, globular structures encountered so far. These are long, strand-like alpha-helices.

a Group 10 allergens

Group 10 allergens are known from *Dermatophagoides pteronyssinus*, *D. farinae*, *Psoroptes ovis*, *Tyrophagus putrescentiae*, *Glycyphagus domesticus*, *Lepidoglyphus destructor* and *Dermanyssus gallinae* (Table 7.2). This is another allergen group (along with group 20) for which there is homology with crustacean allergens. Tropomyosins are muscle-binding proteins involved in the regulation of the contraction of actin. They are highly conserved and are found in all animals. Tropomyosin allergens from crustaceans, molluscs and insects all cross-react with those of dust mites (Reese *et al.*, 1999; Jeong *et al.*, 2006), and they are regarded as so-called pan-allergens (Reese *et al.*, 1999).

Group 10 allergens, e.g. Blo t 10 (Yi *et al.*, 2002) and Der p 10 (Asturias *et al.*, 1998), have a derived molecular weight of about 33 kDa (Table 7.1) consisting of a 15 residue signal peptide and a mature protein of 285 residues. On SDS-PAGE, tropomyosins appear as a band at about 37 kDa. There are no glycosylation sites and no conserved cysteines, consistent with its tertiary structure: a dimeric coiled-coil, consisting entirely of alpha-helix (Figure 7.6a).

Sequence identity within the seven mite tropomyosins in Figure 7.18 is 77%, higher than for any other allergen except paramyosin (cf. below). IgE-binding frequency was reported as 81% for nDer f 10 (Aki *et al.*, 1995). Group 10 allergens can be abundant in nature and they probably can be regarded as major allergens. Recombinant allergens show much lower binding, e.g. 13% for Lep d 10 (Saarne *et al.*, 2003) (Table 7.6).

b Group 11 allergens

Group 11 allergens are paramyosins, another ubiquitous group of muscle proteins, identified from *Dermatophagoides pteronyssinus*, *D. farinae*, *Psoroptes ovis*, *Sarcoptes scabiei* and *Blomia tropicalis*

(Table 7.2). Like tropomyosin they are highly conserved and cross-reactive (Figure 7.19).

They are large molecules, with a derived molecular weight of 102 kDa (ca. 98 kDa on SDS-PAGE gels) and show 42–80% IgE binding frequency (Tsai *et al.,* 1998, 2005; Ramos *et al.,* 2001; Lee *et al.,* 2004). Like group 10 allergens, IgE reactivity to the recombinant form of Blo t 11 was significantly lower than to the natural form (Teo *et al.,* 2006).

7.4.8 The lipid-binding proteins

Included here are the allergen groups 2, 13 (fatty acid-binding proteins) and 14 (vitellogenin). The amino acid composition of the lipid-binding allergens is somewhat less concordant than it is for other mite allergen families. They have lysine and asparagine at consistently high frequencies (mean D_t of 1.7 and 1.5 respectively), followed by aspartate (1.4) and isoleucine (1.2). Lowest frequencies are for tryptophan (0.2) and proline (0.6). Tryptophan is absent from group 13 allergens. Similarly, cysteine occurs at low frequency in groups 13 and 14 but is common in group 2. Histidine is common in groups 2 and 14 but absent in group 13 (Table 7.5; Figure 7.1f).

Tertiary structure based on X-ray crystallography shows groups 2 and 13 allergens consist of beta-barrels (Figure 7.7a, b). Group 14 is much larger and more complex, and is likely to present allergenically as a series of subunits (cf. below).

a Group 2 allergens

The group 2 allergens are polymorphic, neutral to basic 14 kDa non-glycosylated allergens recognised by the majority of mite allergic individuals (Tables 7.1 and 7.6). They are probably involved in lipid binding and transport, although their precise function remains a mystery. They appear not to be enzymes. Group 2 allergens have been identified in a higher diversity of mites than any other allergen, from *Dermatophagoides pteronyssinus*, *D. farinae*, *D. siboney*, *Euroglyphus maynei*, *Psoroptes ovis*, *Aleuroglyphus ovatus*, *Tyrophagus putrescentiae*, *Suidasia medanensis*, *Glycyphagus domesticus*, *Lepidoglyphus destructor* (previously designated Lep d 1) and *Blomia tropicalis*.

The complete amino acid sequences of group 2 allergens (the sequence of Der s 2 is only partial) indicate the allergens are synthesised as preproteins with signal peptides of 16–17 residues and mature proteins of 125–129 residues (Yuuki *et al.,* 1991; Chua *et al.,* 1990a; Trudinger *et al.,* 1991; Varela *et al.,* 1994; Schmidt *et al.,* 1995). Group 2 allergens are quite highly conserved (Figure 7.10). Homologous sequences are not concentrated around active sites and cysteine residues as they are in group 1 allergens but are evenly distributed along the molecule. Differences that reflect the phylogenetic relationships of the mite species are evident, and the sequences in Figure 7.10 have been arranged according to families or superfamilies to emphasise this characteristic.

The tertiary structure of Der p 2 and Der f 2 have been modelled using nuclear magnetic resonance spectroscopy (NMR; Mueller *et al.,* 1998; Ichikawa *et al.,* 2005). The crystal structure of Der p 2 and Der f 2 was derived empirically by X-ray diffraction (Derewenda *et al.,* 2002; Johannessen *et al.,* 2005) and found to be very similar. The NMR models differ from the crystal structures in not showing a central cavity. The crystal structure shows a cylindrical molecule containing 10 beta-strands and a short apha-helix. It shows a characteristic immunoglobulin-like fold comprising two anti-parallel beta-sheets, one consisting of three beta-strands, the other of five, overlying each other at an angle of about 30° to form a beta-sandwich or beta-barrel (Figure 7.7a, b). Between the sheets is a large central cavity lined with hydrophobic and aromatic residues, suggesting a lipid-binding function.

The cysteine-cysteine disulphide bond pairings of Der f 2 are C8–C119, C21–C27 and C73–C78 (Nishiyama *et al.,* 1993). This arrangement is likely to be conserved in all group 2 allergens as well as in the 16 kDa mammalian epididymal secretory proteins, human HE1 (Kirchoff *et al.,* 1996) and its porcine homologue (Okamura *et al.,* 1999), which have ca. 25–35% sequence identity with Der p 2. Thomas and Chua (1995) suggested that the group 2 allergens are secretions of the male mite reproductive system. But it is unlikely they are confined to males. Localisation of group 2 allergens in thin sections of mites shows they were mainly in the gnathosomal region, and no concordance with the male reproductive system (section 7.5 below). Males constitute only about 10–15% of natural dust mite populations (Colloff, 1992a; Chapter 5) and would have to be working overtime to produce the amounts of group 2 allergens present in house dust (which are comparable to the levels of group 1 allergens; Chapter 8).

HE1 is the same as Niemann-Pick C2 protein (NPC2; Friedland *et al.,* 2003), involved in cholesterol binding and transport via the low-density lipoprotein

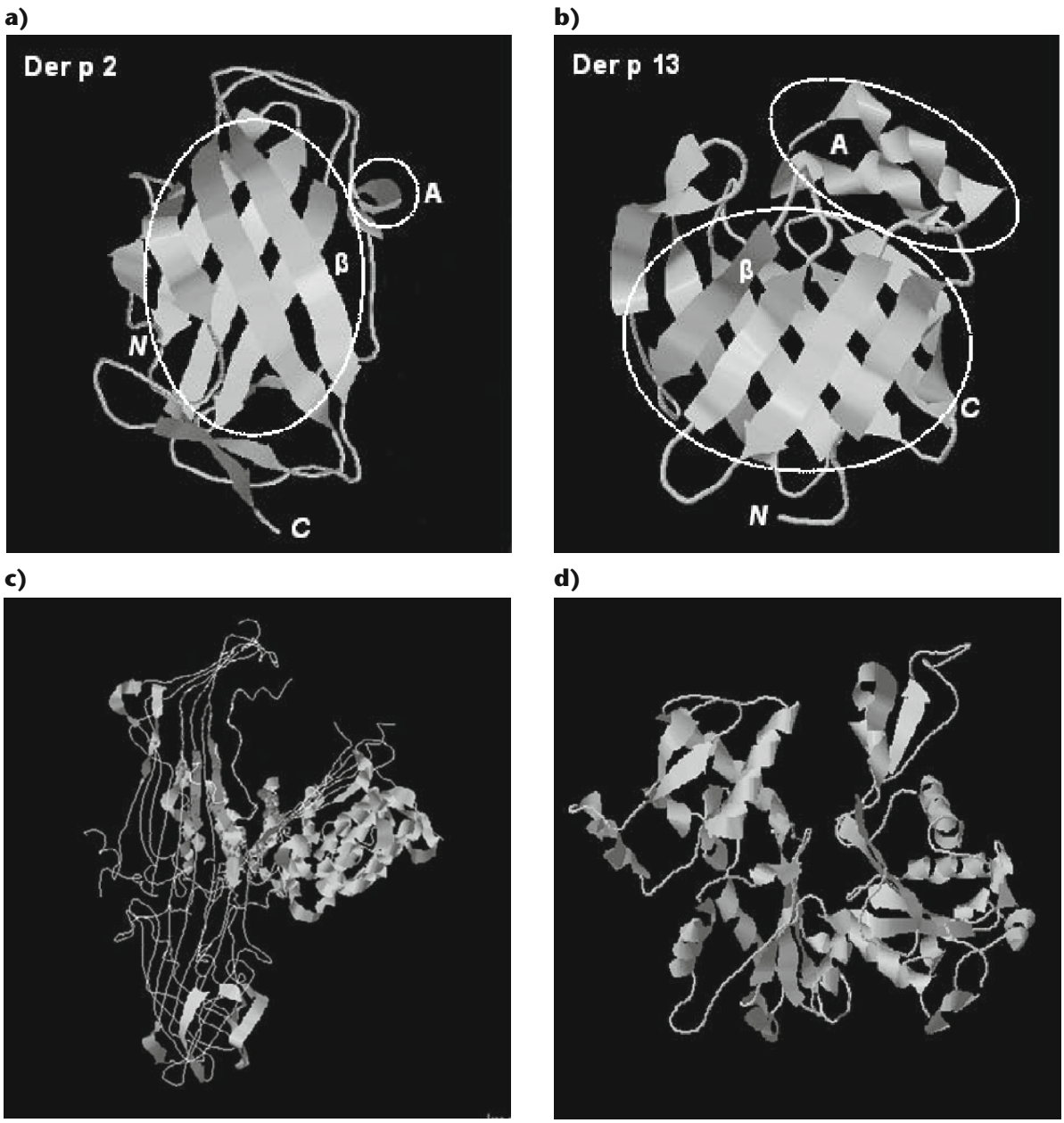


Figure 7.7 Tertiary structures of members of the lipid-binding family of mite allergens (groups 2, 13, 14 and 16) and related proteins. **a)** Der p 2 (PDB 1KTJ) showing a beta-barrel of two anti-parallel beta-sheets, one consisting of three beta-strands, the other of five. A = singly-turn alpha-helix; **b)** Der p 13 (PDB 2A0A) showing a beta-barrel (β) consisting of two sets of five beta-strands. A = helix-turn-helix motif; **c)** vitellogenin homologue of group 14 allergens; **d)** gelsolin homologue of group 16 allergens.

pathway. Absence of NPC2 in humans causes Niemann-Pick disease, a hereditary, fatal neurodegenerative disease which involves the accumulation in mutant cells of cholesterol-filled lysosomes and subsequent cell death. When superimposed on each other, the crystal structures of NPC2 and Der p 2 are quite similar (Friedland *et al.*, 2003, their Figure 5b) although the hydrophobic core differs in some important respects.

There is some sequence and structural homology between group 2 allergens and the human Rho-specific guanine dissociation inhibitor, RhoGDI (Derewenda *et al.*, 2002) and MD-2 that binds to lipopolysaccharides and the lipophilic GM2-activator protein (Keber *et al.*, 2005; Johannessen *et al.*, 2005). These are lipid-binding molecules belonging to the ML family of proteins involved in lipid recognition, metabolism and

transport. A sequence homology search by Mueller *et al.* (1998) revealed 26% identity with the protein esr16 (which had 36% sequence identity with HE1) from tracheal epithelial cells of a moth, the tobacco hornworm, *Manduca sexta* (Mészáros and Morton, 1996a). Levels of mRNA for the esr16 gene increase prior to moulting and the expression of the esr16 gene can be regulated with ecdysteroids, the moulting hormones (Mészáros and Morton, 1996a, b). The residues that form the lipid-binding cavity in Der p 2 are conserved in esr16 (Mueller *et al.,* 1998). Ecdysteroids are lipid hormones found in arthropods with very similar structure to cholesterol (which does not function as a hormone in arthropods). It may be that ecdysteroids as well as dietary cholesterol are the natural lipid ligands of group 2 allergens. The binding of hormones such as ecdysone is associated with regulation of moulting (Chapter 2). Although the site of ecdysteroid production in dust mites is not known, it is interesting that Lep d 2 was localised in the gnathosomal region (van Hage-Hamsten *et al.,* 1995), corresponding with the location of the tracheal epithelial cells in *Manduca sexta*.

b Group 13 allergens

Group 13 allergens have not been isolated from one pyroglyphid species, *Dermatophagoides farinae*, as well as from *Acarus siro*, *Tyrophagus putrescentiae*, *Glycyphagus domesticus*, *Lepidoglyphus destructor* and *Blomia tropicalis*. All these allergens have been fully sequenced (Table 7.1). These allergens are 15–17 kDa proteins showing high homology with fatty acid-binding proteins (FABP) belonging to the lipocalin family. Lipocalins are involved in the binding of small hydrophobic molecules such as fatty acids, glycerol, steroids, arachidonic acid, histamine and many other molecules with critical cellular physiological function (Flower, 1996). Group 13 allergens show relatively high sequence homology with retinoic acid binding proteins of decapod Crustacea (Gu *et al.,* 2002) and insects (Mansfield *et al.,* 1998).

Amino acid sequences of group 13 indicate an absence of conserved cysteine residues, signal or propeptide sequences, and only Aca s 13 and Tyr p 13 have potential glycosylation sites. The degree of sequence homology is extremely high, with residue identities of between 58 and 82% (Figure 7.21), but the degree of allergenicity is very low: only 10–23% of sera from patients allergic to the relevant mite showed IgE binding activity (Table 7.6).

The tertiary structure of Der f 13, modelled using NMR spectroscopy (Chan *et al.,* 2006), shows an N-terminal beta-strand followed by a helix-turn-helix motif and nine anti-parallel beta-strands forming a beta-barrel, with the two alpha-helices at one end of the barrel (Figure 7.7b). This structure is very similar to vertebrate FABPs such as human brain FABP (Balendiran *et al.,* 2000). It is broadly similar to group 2 allergens which also appear to be involved in lipid binding.

c Group 14 allergens

Group 14 allergens have been found in *Dermatophagoides farinae*, *D. pteronyssinus*, *Euroglyphus maynei*, *Psoroptes ovis*, *Sarcotes scabiei* and *Blomia tropicalis* (Tables 7.1, 7.2). Complete amino acid sequences are available for Eur m 14 (M-177; Epton *et al.,* 1999) and Der p 14 (Epton *et al.,* 1999, 2001a, b). They are 190–206 kDa homologues of vitellogenin and/or the related protein apolipophorinin. Mag 1 (Aki *et al.,* 1994a) and Mag 3 (Fujikawa *et al.,* 1996) are fragments of Der f 14 (Epton *et al.,* 1999), as is Ssag1 (Harumal *et al.,* 2003) and MSA1 (Mattsson *et al.,* 1999). MSA1 (Mattsson *et al.,* 1999) is different from the antigen ASA1 (Ljunggren *et al.,* 2006) even though it is this latter paper that is cited against the GenBank entry for MSA1 (DQ109676). Fujikawa *et al.* (1996) reported that Mag 3 is not the same as the large lipoprotein allergen Dpt4 (Stewart *et al.,* 1983). IgE-binding frequency of Group 14 allergens was 30–39% for rDer f 14 (Aki *et al.,* 1994; Fujikawa *et al.,* 1996) and 70% for nDer f 14 (Fujikawa *et al.,* 1996).

Vitellogenins are large (200–700 kDa), complex carotenoid lipoproteins that constitute the egg yolk of invertebrates and vertebrates. They can be heavily glycosylated and phosphorylated, and may bind with lipids, metals and vitamins. In athropods they are expressed in the invertebrate equivalent of the liver – the fat body in insects, or the hepatopancreas in crustaceans (as well as the ovaries) – secreted into the haemolymph, taken in to the ovary and processed further (Chen *et al.,* 1997; Tsang *et al.,* 2003). Consistent with its role as an egg storage protein secreted into the haemolymph, Harumal *et al.* (2003) showed immunohistochemical localisation of Sar s 14 in internal organs, cuticle and in eggs. Fujikawa *et al.* (1996) found Mag 3 localised in the oesophagus, gut and other internal organs.

Group 14 allergens show relatively high homology (Figure 7.22). Der p 14 shows next

highest identity (20–22%) with the N-terminal regions of several vitellogenins from prawns and shrimps (Chen *et al.,* 1997; Tsang *et al.,* 2003; Avarre *et al.,* 2003; Phiriyangkul and Utarabhand, 2006). Although sequence identity is relatively low, it is consistent and evenly distributed along the Der p 14 molecule. Crustacean vitellogenin is encoded by multiple genes (Tsang *et al.,* 2003), present in females and males but only expressed by females. Gravid females constitute only about 10% of individuals in dust mite populations, and eggs about 30% (Colloff, 1992a; Chapter 5). Even accounting for eggs containing fully developed larvae, somewhere between a third and a half of the individuals in a mite population are likely to contain some group 14 allergens. In eggs at an early stage of development, the bulk of the protein present is likely to be vitellogenin, so this allergen is likely to be present in relatively large amounts in natural populations. However, as Epton *et al.* (1999) point out, it is likely to be underestimated in aqueous allergen extracts because it is not water soluble.

Cleavage sites are present in vitellogenins corresponding to the consensus motif RXRR or RXKR which are recognised by furin convertases. These subtilisin-like enzymes cleave the protein into subunits (vitellins) prior to secretion (Chen and Raikhel, 1996). Two groups of cleavage site motifs are highlighted on the sequences in Figure 7.22 at positions 835 and 1137 (Der p 14 numbering), which do not correspond to those of crustaceans (Avarre *et al.,* 2003) or insects (Chen *et al.,* 1997). Cleavage of Der p 14 at these points would result in three subunits of 95, 36 and 76 kDa (fragments I, II and III). Group 14 allergens are large lipoproteins and seem at first sight to be rather unlikely candidates for allergens. But it appears they could be exposed to the human immune system in smaller, subunit forms. This fits with the observation by Fujikawa *et al.* (1998) that Der f 14 broke down on storage and that the fragments were more allergenic than the whole protein. The fragments they obtained on SDS-PAGE were 144, 125 and 52 kDa, which corresponds with M177 minus fragment II, M177 minus fragment III, and M 177 minus fragments I and II, adjusting for differences between derived molecular weight of Der p 14 (206 kDa) and SDS-PAGE molecular weight of Der f 14 (177 kDa). Epton *et al.* (2001a, b) also found western blotting of mite eggs with antisera to a rDer p 14 fragment showed three major bands, consistent with subunit cleavage.

7.4.9 Unclassified allergens

a Group 16 allergens

Der f 16 (Mag 15; Tategaki *et al.,* 2000) is a 53 kDa gelsolin-like protein that bound IgE from allergic patients with a frequency of 47% (Kawamoto *et al.*, 2002b). Gelsolins (Figure 7.7d) are actin-binding proteins that sever and cap actin filaments and are involved in gel-to-sol (hence the name) transformations within the cell, allowing for cell movement (Silacci *et al.,* 2004). The process is regulated by Ca^{++} and the gelsolin molecule has several calcium-binding domains. Der f 16 showed ca. 34–38% amino acid sequence homology with crustacean, insect and human gelsolins (Kawamoto *et al.,* 2002a).

The amino acid composition of Der f 16 (Figure 7.24) shows some similarity with the methionine-rich alpha-helical proteins (groups 5, 7 and 21), but topology mapping indicates only 31% of the molecule is alpha-helix, compared with >50% for the other allergens in this group (Figure 7.5).

b Group 17 allergens

Der f 17 (a.k.a. Mag50) is a 30 kDa calcium-binding protein homologue, deduced on the basis of the allergen possessing Ca^{++}-binding capacity and a helix-loop-helix (EF-hand) motif (Tategaki *et al.,* 2000; Kawamoto *et al.,* 2002b). The full amino acid sequence has not been published and I have been unable to find a database record. Tategaki *et al.* (2000) stress the importance of the Ca^{++}-binding component of the EF-hand motif in IgE-binding activity, but the partial amino acid sequence shows no homology with other EF-hand Ca^{++}-binding allergens such as fish parvalbumins or the pollen allergens Bet v 4 and Ole e 3. Der f 17 bound IgE from mite-allergic patients with a frequency of 35%.

c Group 19 allergens

Blo t 19 is a 7 kDa peptide (Nge and Chua, IUIS database). Few details about this allergen have been published. The amino acid sequence is given in Figure 7.26 courtesy of Cheong Nge (pers. comm., 2006). The recombinant allergen was cited by Thomas *et al.* (2002) as having only 10% frequency of IgE-binding activity.

A BLAST search revealed 76% homology between Blo t 19 and the antibacterial factor, ASABF-a from the nematode, *Ascaris suum*. ASABF (*Ascaris suum* antibacterial factor) peptides are induced as part of the immune response to the presence of bacteria in the pseudocoelom of the nematode (Pillai *et al.,* 2003). Blot 19 has

limited hololology with CSaβ-type antimicrobial peptides, including insect defensins (Zhang and Kato, 2003). Blo t 19 has eight conserved cysteine residues, most of which are part of an invertebrate defensin consensus (Figure 7.26; Dimarcq *et al.,* 1998). Like the group 12 allergens, the high cysteine content suggests that a key structural property of Blo t 19 is resistance to proteases in the midgut, which is one of the most obvious sites for induction of an immune response against bacteria. Zhang and Kato (2003) detail the cysteine-pairing for ASABF (C1–C5, C2–C6, C3–C7, C4–C8), whereby disulphide bridges are formed, indicating a tightly folded, highly protease-resistant molecule.

7.4.10 Other mite allergens not included in the IUIS classification

There are many other allergens that have been isolated but not assigned to groups 1–21. For some, especially those that are known only from database entries, the evidence of their allergenicity may be based only on amino acid sequence homology with known allergens. For others, data on frequency of IgE-binding has been published.

a Oxidoreductases

Oxyreductases (EC 1.2) act on the aldehyde or oxo-group of donors and include aldehyde dehydrogenase (EC 1.2.1.3) which catalyses the oxidation of aldehydes to carboxylic acids in the presence of NAD or NADP as a co-factor. Two aldehyde dehydrogenase (ALDH) homologue allergens have been databased, from *Blomia tropicalis* (AY291328; Chew *et al.,* unpublished) and *Dermatophagoides farinae* (AY283296, Chew *et al.,* unpublished; Figure 7.28). Other aldehyde dehydrogenases include the mould allergens Alt a 10 from *Alternaria alternata* and Cla h 3 from *Cladosporium herbarium*. The mite allergens have a conserved cysteine (Cys230), found in all aldehyde dehydrogenases with catalytic activity (Perozich *et al.,* 1999).

b Hydrolases

A 13.8 kDa bacteriolytic enzyme was isolated and characterised from *Dermatophagoides pteronyssinus* (Mathaba *et al.,* 2002). The protein shows homology with the P60 family of bacterial cell wall-associated hydrolases. The signal peptide is eukaryotic rather than bacterial, suggesting the protein may be derived from bacteria within the mites and incorporated into the mite genome via horizontal gene transfer. The allergenicity of this molecule has not been determined.

A *Blomia tropicalis* protein (AY291335; Chew *et al.,* unpublished), showing homology with the allergen Ves m 1 (Hoffman, 1994) from the wasp *Vespula maculifrons*, is a lipase-like enzyme homologue similar to pancreatic triglyceride lipase and phospholipase A1 (EC 3.1.1.4). Lipases are esterases that hydrolyze long-chain acyl-triglycerides into di- and monoglycerides, glycerol, and free fatty acids. The active site of a lipase contains a catalytic triad consisting of Ser-His-Asp/Glu.

A putative allergen from *Blomia tropicalis* (AY291332; Chew *et al.,* unpublished) shows homology with beta-*N*-hexosaminidase (EC 3.2.1.52). This enzyme catalyses the hydrolysis of terminal non-reducing acetylhexosamine residues in *N*-acetyl-beta-D-hexosaminides. The enzyme has the same active site as for *N*-glucosaminidase and *N*-galactosaminidase activities and therefore could be involved in moulting via the hydrolysis of *N*-acetylglucosamine residues of chitin.

A cathepsin F (EC 3.4.22.41) homologue has been identified from *Dermatophagoides farinae* (CPW1 cysteine proteinase AF145249; Park *et al.,* unpublished; Table 7.1).

The cathepsin D homologue allergen and its isoform from *Blomia tropicalis* (EC 3.4.23.5; isoform AY792952; Chew *et al.,* unpublished) show high homology with other arthropod aspartic proteases, including that from the tick *Haemaphysalis longicornis* (Boldbaatar *et al.,* 2006; Figure 7.29).

Carbon-oxygen lyases (EC 4.2): Enolase, phosphopyruvate hydratase (aka 2-phospho-D-glycerate hydrolase) EC 4.2.1.11. Enolases are homodimeric enzymes that catalyse the reversible dehydration of 2-phospho-D-glycerate to phosphoenolpyruvate as part of the glycolytic and gluconeogenesis pathways. The reaction is facilitated by the presence of metal ions. Thus these enzymes are involved in CHO transport and metabolism. They include *Glycyphagus domesticus* enolase homologue AY288683 (Chew *et al.,* unpublished).

cis-trans-*Isomerases* (EC 5.2): Peptidylprolyl isomerase B (Cyclophilin; EC 5.2.1.8). These enzymes are involved in protein folding and assembly and include *Glycyphagus domesticus* Mala s 6 homologue (AY288682; Chew *et al.,* unpublished).

c Enzyme inhibitors

Thaumatins are protease inhibitors, functioning as anti-fungal agents. *Glycyphagus domesticus* Jun a

3 homologue (AY288680; Chew *et al.,* unpublished) appears to be a thaumatin.

Cystatin A (Stefin A; Cystatin AS) is a cysteine protease inhibitor. These proteins regulate endogenous protease activity and protect against exogenous protease activity. *Lepidoglyphus destructor* cystatin is claimed to be IgE-binding (AJ428051; Kaiser *et al.,* unpublished).

Mite heat shock protein 70 homologues (Figure 7.30) include *Dermatophagoides farinae* Mag29 (DEPMAG29; Aki *et al.,* 1994b) and *Blomia tropicalis* Mag29 (AY291333; Chew *et al.,* unpublished). Heat shock proteins inhibit the release of lysosomal proteases that trigger cell death, and are expressed in cells stressed by heat, radiation or toxins (Nylandsted *et al.,* 2004). They are highly conserved and found in bacteria, plants and animals. The *B. tropicalis* protein showed over 90% homology with a 70 kDa heat shock cognate protein (HSC70) from the midge *Chironomus tentans* (Figure 7.30; Karouna-Renier *et al.,* 2003). A recombinant Mag29 fragment from *D. farinae* showed a positive IgE-binding frequency (by immunoblotting) of 4 out of 41 sera (9.8%) from mite-allergic patients (Aki *et al.,* 1994b). The putative allergenicity of Mag29 from *Blomia tropicalis* is presumably based on its amino acid sequence homology with Mag29 from *D. farinae*.

d Insulin growth factor–binding protein homologue

These proteins show homology with:

- high affinity binding partners of insulin-like growth factors;
- Kazal-type serine protease inhibitors and follistatin-like domains;
- the immunoglobulin domain cell adhesion molecule (cam) subfamily – members are components of neural cell adhesion molecules (N-CAM L1), Fasciclin II and the insect immune protein hemolin (Sun *et al.,* 1990).

A putative IGF-BP allergen is the Blo t homologue of allergen Gal d 1 (AY291331; Chew *et al.,* unpublished).

e Cytoskeletal proteins

Alpha-tubulins are involved in cytoskeletal microtubule formation. There is published experimental evidence of IgE binding activity to alpha-tubulins of *Lepidoglyphus destructor* (Saarne *et al.,* 2003) and *Tyrophagus putrescentiae* (Jeong *et al.,* 2005b). These are highly conserved molecules (Figure 7.31). The amino acid sequence of the *L. destructor* alpha-tubulin shows >90% identity with those from mouse, chicken and *Xenopus laevis*.

Profilin binds actin monomers, membrane polyphosphoinositides and poly-L-proline. Profilin can inhibit actin polymerisation into F-actin by binding to monomeric actin (G-actin) and terminal F-actin subunits, but – as a regulator of the cytoskeleton – it may also promote actin polymerisation. It plays a role in the assembly of branched actin filament networks. Profilins are considered as pan-allergens in plants, and profilins have been identified from *Blomia tropicalis* (AY291334; Chew *et al.,* unpublished) and *Suidasia medanensis* (AY803196; Reginald *et al.,* unpublished). The mite profilins (Figure 7.32) show ca. 38–40% sequence homology with profilin allergens from plants such as those from brazil nut, banana, mango and pineapple.

f Proteins involved in cell maintenance

Type 1 non-specific lipid transfer proteins (LTPs) can enhance *in vitro* transfer of phospholipids between membranes and can bind acyl chains. *Glycyphagus domesticus* Mal d 3 homologue appears to be an LTP (AY288681; Chew *et al.,* unpublished).

Ribosomal protein P2 plays an important role in the elongation step of protein synthesis. P1 and P2 exist as dimers at the large ribosomal subunit. P2 belongs to the ribosomal protein L12P family and is involved in translation, ribosomal biogenesis and structure. *Blomia tropicalis* Blo t homologue of allergen Alt a 6 appears to be a P2 protein (AY291329; Chew *et al.,* unpublished).

g Others

- Allergens of 39 kDa from *D. farinae* (Aki *et al.,* 1994b) and *L. destructor* (Ansotegui *et al.,* 1991), both of unknown function; the latter with IgE-binding activity in eight out of eight sera positive to *L. destructor*.
- A glucoamylase-like allergen (Muzinich *et al.,* 1997) from *D. pteronyssinus*.
- Dpt4, a high molecular weight phospholipid-containing lipoprotein allergen, with an apparent molecular weight of 244 kDa, from *D. pteronyssinus* (Stewart and Turner, 1980b; Stewart *et al.,* 1983).
- An allergen designated Dpt 36 from *D. pteronyssinus* (Stewart *et al.,* 1988).
- Two allergenic components of 79 and 93 kDa of unknown function from *L. destructor*, which

bound IgE from *L. destructor*-positive sera with a frequency of 60% (Olsson *et al.,* 1994).

7.5 Cross-reactivity and sequence polymorphisms of mite allergens

Allergens are said to be cross-reactive when IgE antibodies that are specific to a particular allergen will also bind to another. This implies there is a similarity in the structure and conformation of the part of the allergen molecule to which the antibody molecule binds. Cross-reactivity has some significant diagnostic implications. People who are sensitised to a particular allergen may give a positive skin-prick test reaction to a different allergen (or allergen source) to which they are not exposed, as was found for an urban population who were skin-prick test positive to *D. pteronyssinus*, but also to several storage mites to which they had little or no exposure (Luczynska *et al.,* 1990b).

The cross-reactivity between Der p 2, Der f 2 and Eur m 2 was found to be due to a relatively high degree of amino acid sequence homology, and the surface of the molecule where antibody binding occurs was relatively conserved between the allergens. Lep d 2 and Tyr p 2 have multiple amino acid substitutions across the antigenic surface and do not cross react with the others (A.M. Smith *et al.,* 2001).

This finding is consistent with the greater phylogenetic distance between the pyroglyphid mites than between the glycyphagoid and acaroids mites (Chapter 1, Figure 1.2).

The source of all the allergens that have been purified and characterised has been mites grown in culture. Sometimes these mites have been grown in the laboratory for years, and it is not surprising that their allergens might differ from mites from the wild to which people are naturally exposed. Polymorphisms have been investigated in wild and cultured *Dermatophagoides farinae* by Thomas *et al.* (1992) and Yuuki *et al.* (1997), in *D. pteronyssinus* by Chua *et al.* (1996), W.A. Smith *et al.* (2001), Hales *et al.* (2002) and Nandy *et al.* (2003) as well as in *Lepidoglyphus destructor* by Schmidt *et al.* (1995) and Kaiser *et al.* (2003). The paper by W.A. Smith *et al.* (2001) deals with group 1 allergens, the others with group 2 allergens, and Smith and Thomas (1996) examine polymorphisms of Der p 3.

Group 1 and 2 allergens are highly polymorphic or genetically variable. Other mite allergens have relatively low levels of polymorphism, such as groups 4 and 5, 10 and 11, and 15 and 18. The implications of high levels of polymorphism (allelic variation) are:

- high frequency of IgE binding due to increased opportunity for T-cell receptor binding and increased overall responsiveness of the allergic population;
- IgE binding to local or regional variants (alleles), with implications for accuracy of diagnostics and epidemiological investigations;
- evolution of distinct polymorphisms that vary in their allergenicity – different polymorphisms may activate low-affinity T-cell interactions that drive Th2-type responses.

The nature of the polymorphism differs between group 1 and group 2 allergens. For Der p 1, polymorphism is represented primarily by seemingly random substitutions of amino acids within the allergen molecule (W.A. Smith *et al.,* 2001). There is little evidence that such polymorphism is cumulative and has become fixed within mite populations. For Der p 2, there is a distinct evolutionary pattern, with substitutions focussed on amino acids 40, 47, 111 and 114 (W.A. Smith *et al.,* 2001; Hales *et al.,* 2002). Further, substitutions 47 and 114 usually co-occur. For Lep d 2, substitutions were found at positions 55 and 102. The difference between Der p 2 and Lep d 2 is that in the latter, there is more than one copy of the gene, whereas in the former, there appears to be only a single gene. This is evidenced by the presence of an intron in one Lep d 2 amplicon but not in another (Kaiser *et al.,* 2003). However, unlike Der p 2, no major difference in IgE binding or T-cell responses could be attributed to Lep d 2 polymorphism.

7.6 Conclusions

Why are some proteins allergens and others not? The proteolytically active allergens appear to be able to do a lot of enzymic damage to the human immune system – cleaving IgE receptors, making the lung epithelium leaky, enhancing specific IgE production, releasing inflammatory mediators and generally bringing about the up-regulation of the allergic response. So why can non-enzymes (like group 2) also be major allergens? The answer may lie in the timing of allergen presentation and exposure. If enzymes and non-enzymes are presented to the human immune system at about the same time, then the enzymic activity facilitates the allergic response to the non-enzymic allergens, as was seen in the example of the adjuvant effect of Der p 1 on production of specific IgE to the bystander allergen, ovalbumin (Gough *et al.,* 2001). So the presence in homes of high levels of a diverse cocktail of airborne

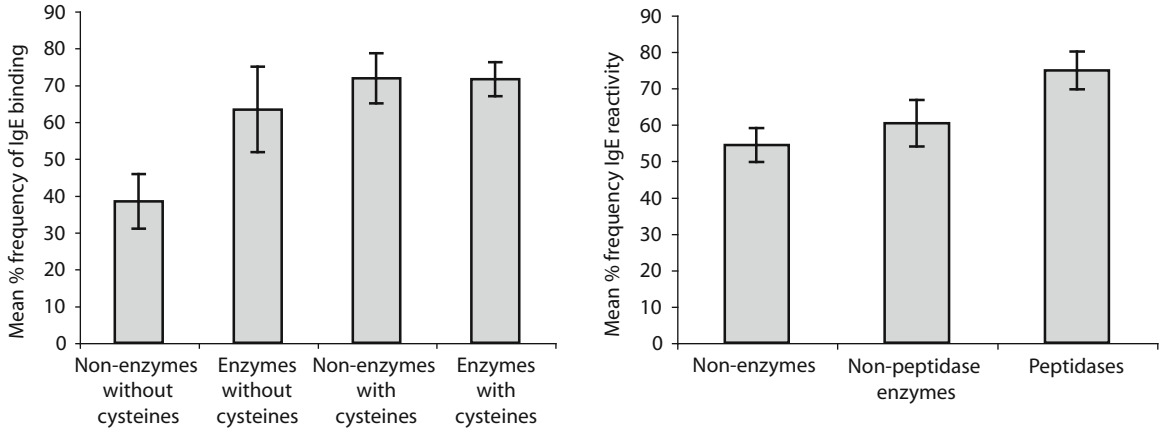


Figure 7.8 IgE binding frequency (± standard error) of major dust mite allergens in relation to their enzymatic function and the presence of conserved cysteine residues. Based on data in Table 7.6.

dust mite allergens, whether borne in faecal pellets, fragments thereof or on other types of particle, is likely to result in co-exposure of the airways to enzymes that pre-condition the immune environment for the bystander non-enzymes.

Whether allergens are major or minor in terms of frequency of IgE responses in a clinical population, may depend not just on the absolute amounts to which people are exposed but whether they are co-presented with enzyme-allergens. To examine this further I plotted frequency of IgE responses to particular allergens against whether the allergens were enzymes or non-enzymes. The enzymes had slightly higher IgE binding frequencies, but the difference was not statistically significant. If the enzyme groups were divided into peptidases and others and the analysis re-run, mean IgE binding to proteases was 75%, significantly higher than the non-peptidase enzymes at 60% and non-enzymes at 54% (Table 7.6; Figure 7.8). There is no IgE-binding data available for allergens 20 and 21. That the peptidase allergens (groups 1, 3, 6 and 9) are on average more likely to elicit IgE responses than non-peptidase enzyme allergens (groups 4, 8, 12, 15, 18) or non-enzyme allergens (groups 2, 5, 7, 10, 11, 13, 14, 16) needs to be qualified. First, the degree of variation between allergen groups within these categories is quite high (e.g. group 6 allergens show much lower frequency of IgE reactivity than other peptidases, group 2 allergens show higher frequency than all other non-enzyme allergens), as it is between mite species (Der p 1: 93–100%, Blo t 1: 62%). Second, the only IgE binding data available for some groups (12, 13, 15, 16 and 19) is from recombinant allergens, which may have lower IgE-binding activity than their native counterparts. Third, the data on IgE binding frequency may not be representative of what happens in the population generally. In cases where the IgE-binding data comes from studies where it has been used to demonstrate the allergenicity of a newly isolated allergen, it may be more representative of what sera happen to be available in the laboratory freezer at the time. Nevertheless, the relationship is intriguing, but it does need to be tested more rigorously.

In a search for common properties of allergens, Furmonaviciene and Shakib (2001) analysed 3-D structures of those allergens showing a common motif, but they really only dealt with group 1 allergens. Chan *et al.* (2006) reckoned that the frequency of charged and polar residues on the allergens is an important factor in determining the allergenicity of the molecule. They found that the non-allergic human homologues of Der f 13 (human FABPs 1–8), Der p 2 (NPC2) and the horse allergen Equ c 1 (lipocalin 9) had significantly fewer charged residues than the allergens, indicating that charged residues play an important role in IgE binding and allergenicity, either because the molecules are more soluble or because charged residues may form more suitable epitopes. Liaw *et al.* (2001) observed that Der p 5 had a high content of charged residues (40%), that non-enzyme allergens such as groups 5, 7, 10, 11 and Mag 29 contain coiled-coil helices consisting of multiple heptad repeats, and that many coiled-coil proteins are known to be strongly antigenic.

Appendix – amino acid sequences

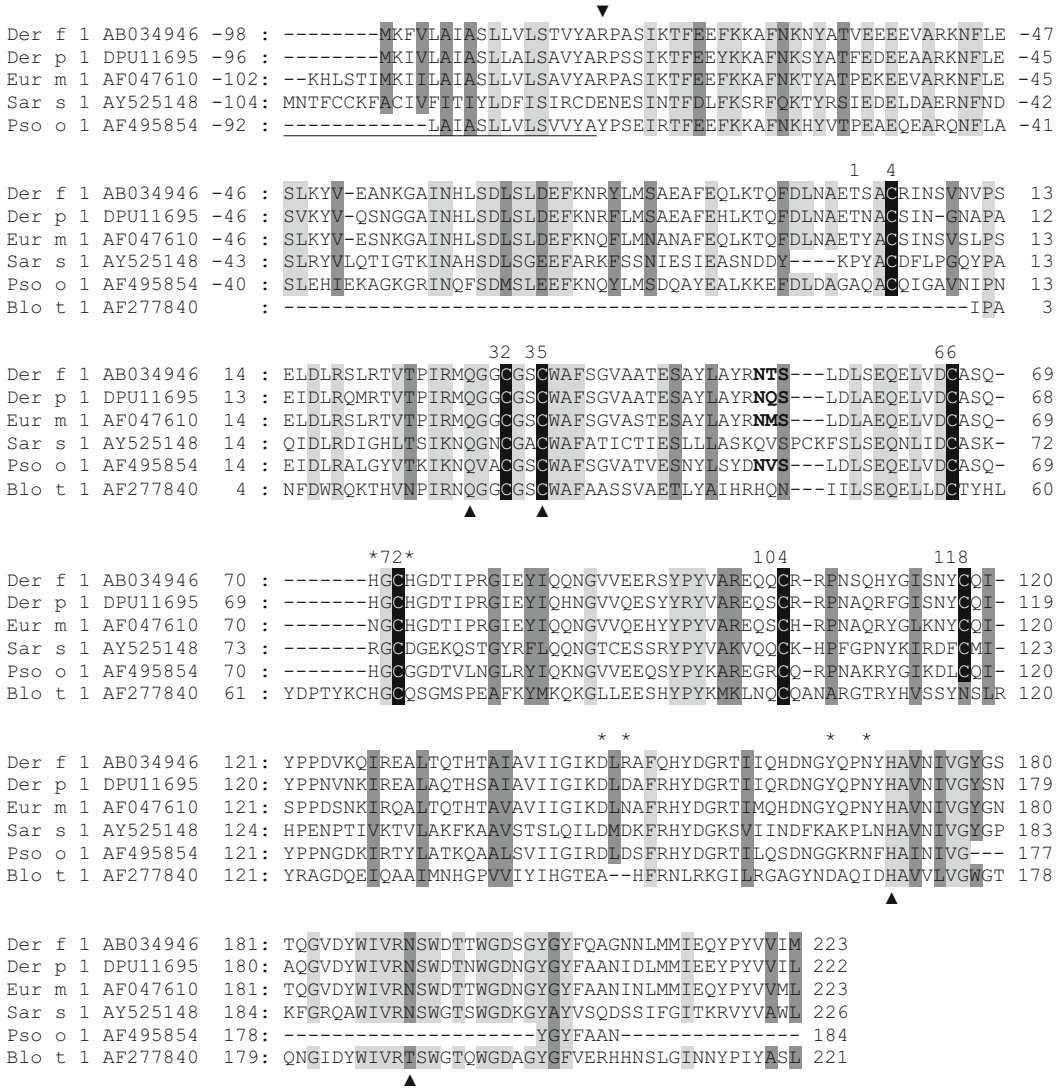


Figure 7.9 Multiple amino acid sequence alignment of group 1 allergens (cysteine proteinase). **C** = conserved cysteine residues (numbered); **L** = invariant sequences; **L** conservative substitutions; __ = signal (pre-) peptide; ▼ = starting position of pro-peptide (Der 1); 1 = starting position of mature protein; **NXS NXT** = *N*-glycosylation sites (mature protein); ▲ = residues comprising the active site (from Meno *et al.*, 2005); residues involved in dimerisation of Der p 1 (from de Halleux *et al.*, 2006).

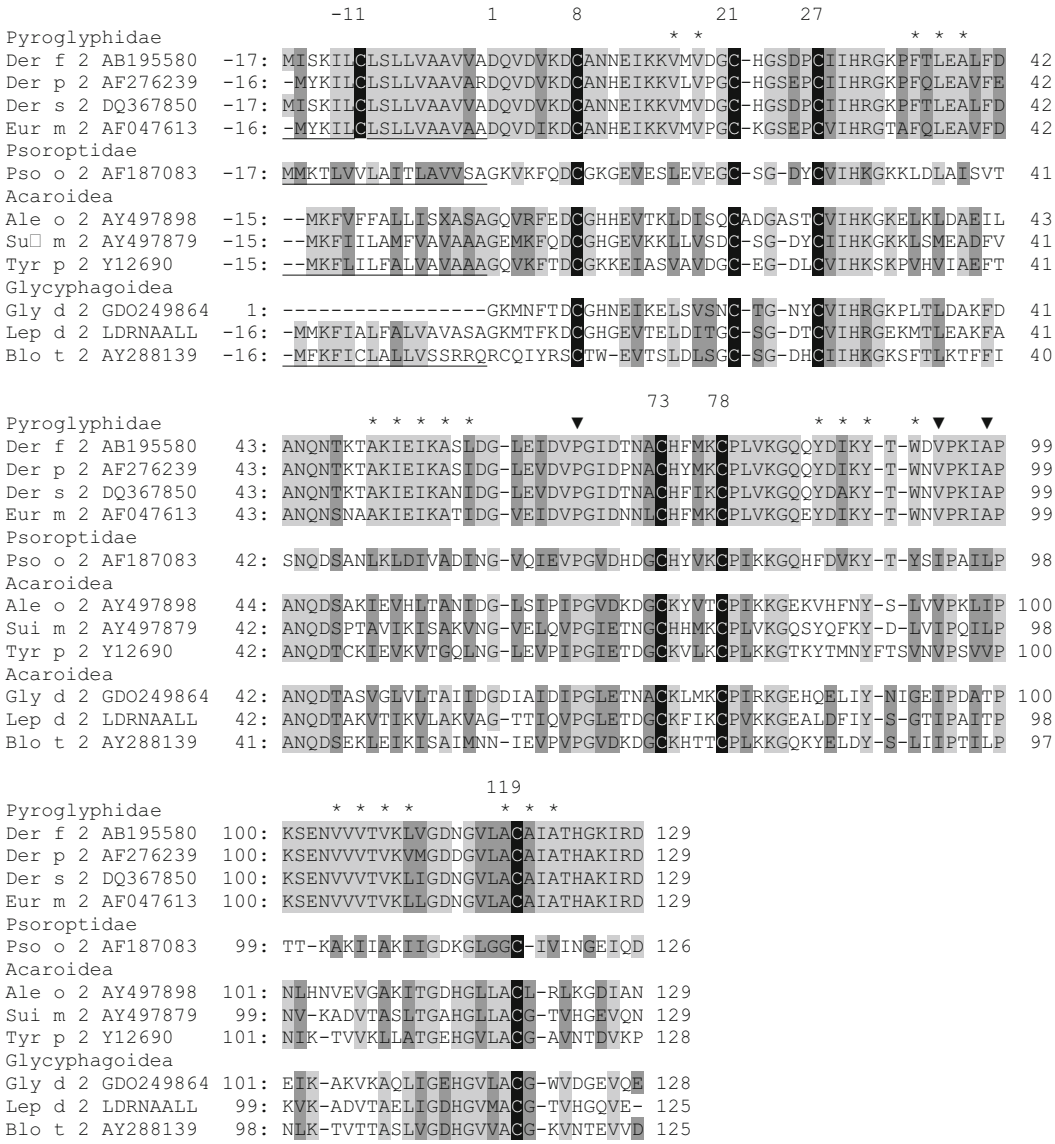


Figure 7.10 Multiple amino acid sequence alignment of group 2 allergens (function unknown). **C** = conserved cysteine residues (numbered); **L** = invariant sequences; **L** conservative substitutions; __ = signal peptide; 1 = starting position of mature protein. ▼ = residues essential to cholesterol binding in NPC2 protein (Friedland *et al.*, 2003); * = hydrophobic and aromatic residues lining central cavity identified by Derewenda *et al.* (2002).

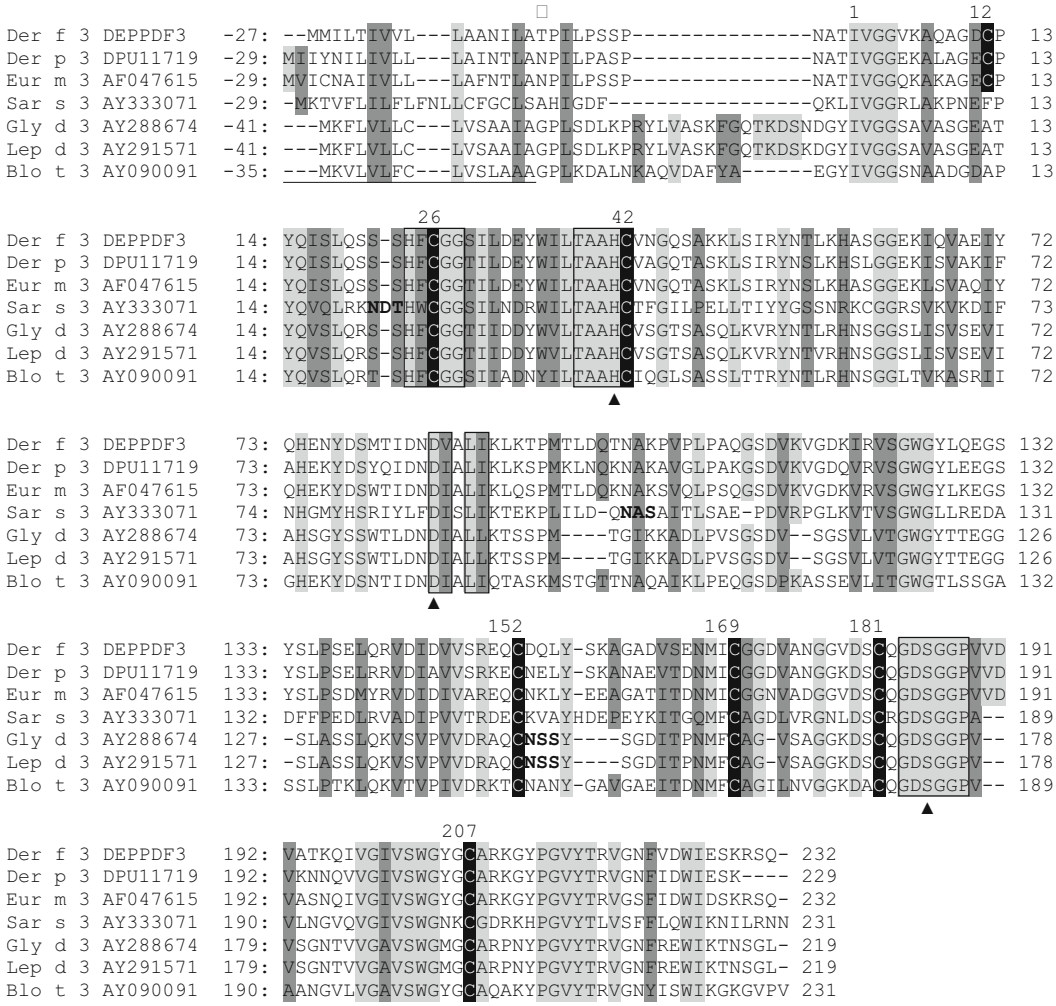


Figure 7.11 Multiple amino acid sequence alignment of group 3 allergens (trypsin). **C** = conserved cysteine residues (numbered against Der f 3); **L** = invariant sequences; **L** conservative substitutions; __ = signal peptide; ▼ = starting position of pro-peptide; **NXS NXT** = *N*-glycosylation sites (mature protein, where X = any amino acid except proline); 1 = starting position of mature protein; ▲ = catalytic residues typical of the serine peptidase family S1; boxed sequences = conserved domains of S1 serine peptidases.

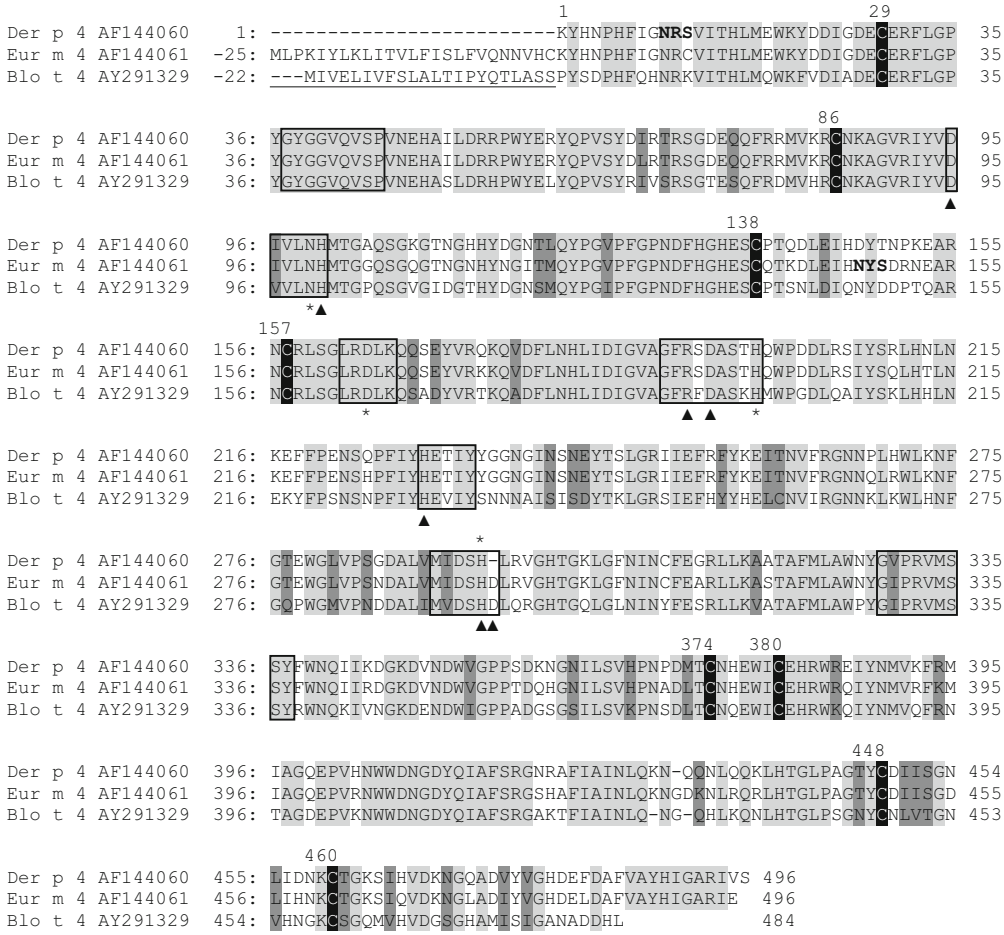


Figure 7.12 Multiple amino acid sequence alignment of group 4 allergens (amylase). **C** = conserved cysteine residues (numbered); **L** = invariant sequences; **L** conservative substitutions; 1 = starting position of mature protein; * = the first position is not the N-terminus of the complete molecule; __ = signal peptide; **NXS NXT** = *N*-glycosylation sites (mature protein, where X = any amino acid except proline); boxed regions = conserved regions typical of alpha-amylases (see Janeček, 1997); ▲ = catalytic residues; * = calcium-binding residues.

Figure 7.13 Multiple amino acid sequence alignment of allergens of groups 5 and 21 (alpha-helical proteins). **L** = invariant sequences; **L** conservative substitutions; 1 = starting position of mature protein; = signal peptide; **NXS NXT** = *N*-glycosylation sites (mature protein, where X = any amino acid except proline).

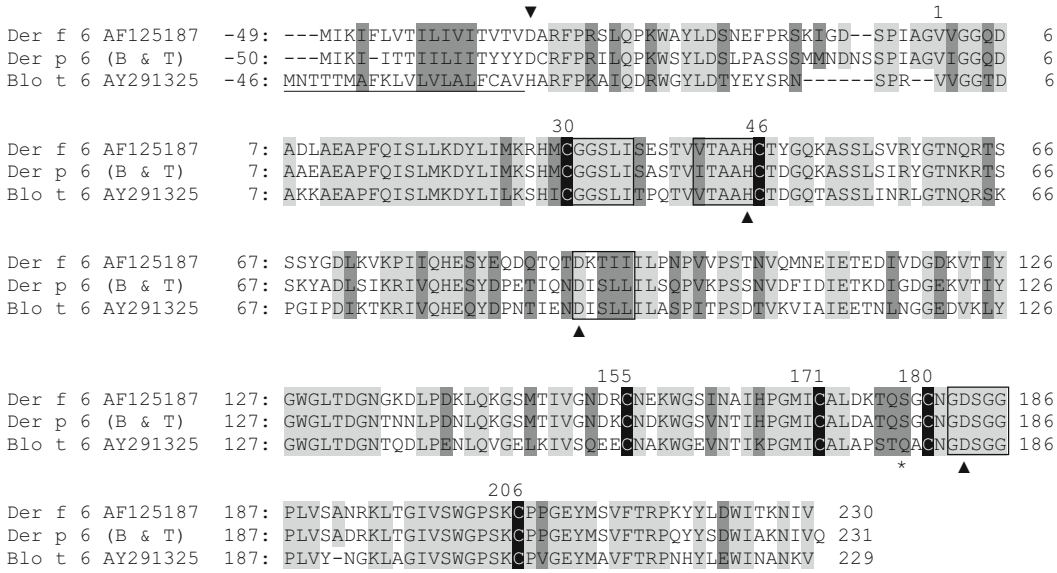


Figure 7.14 Multiple amino acid sequence alignment of group 6 allergens (chymotrypsin). Der p 6 (B & T) = Bennett & Thomas (1996). **C** = conserved cysteine residues (numbered); **L** = invariant sequences; **L** conservative substitutions; __ = signal peptide; ▼ = putative starting position of pro-peptide; 1 = starting position of mature protein; ▲ = catalytic residues typical of the serine peptidase family S1; boxed sequences = conserved domains of S1 serine peptidases; * = serine residue (Der 6) conferring substrate specificity.

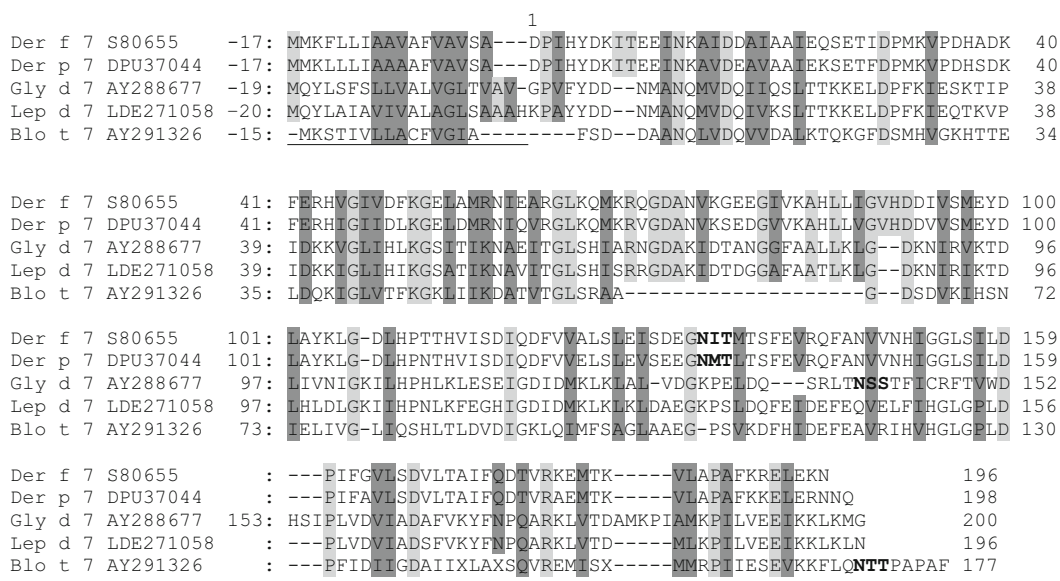


Figure 7.15 Multiple amino acid sequence alignment of group 7 allergens (function unknown). **L** = invariant sequences; **L** conservative substitutions; __ = signal peptide; **NXS NXT** = *N*-glycosylation sites (mature protein, where X = any amino acid except proline); 1 = starting position of mature protein.

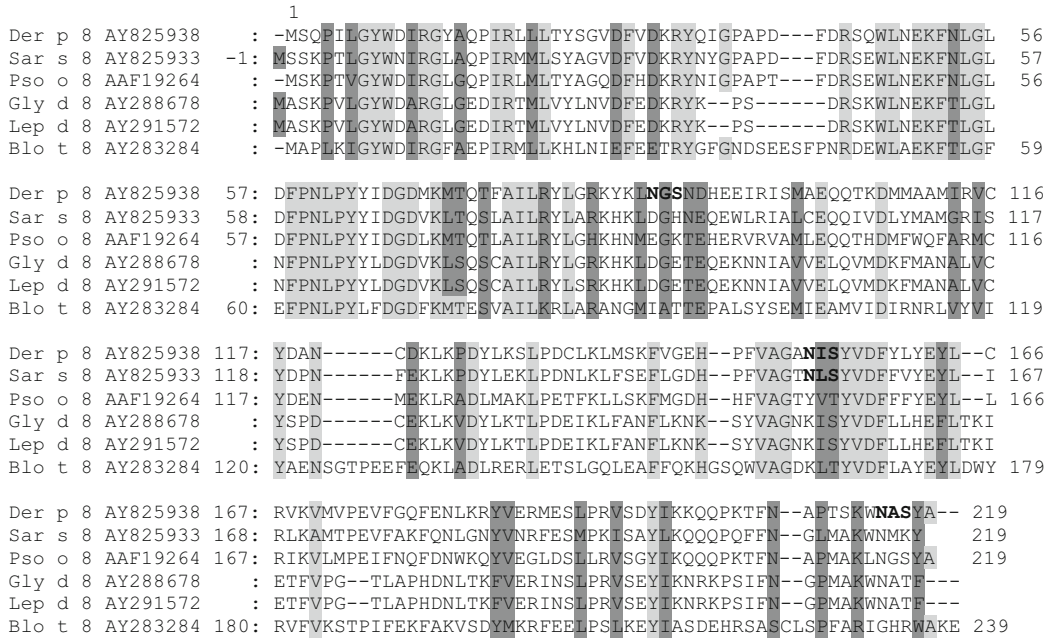


Figure 7.16 Multiple amino acid sequence alignment of group 8 allergens (glutathione-s-transferase). **L** = invariant sequences; **L** conservative substitutions; **NXS NXT** = *N*-glycosylation sites (mature protein, where X = any amino acid except proline); 1 = starting position of mature protein.

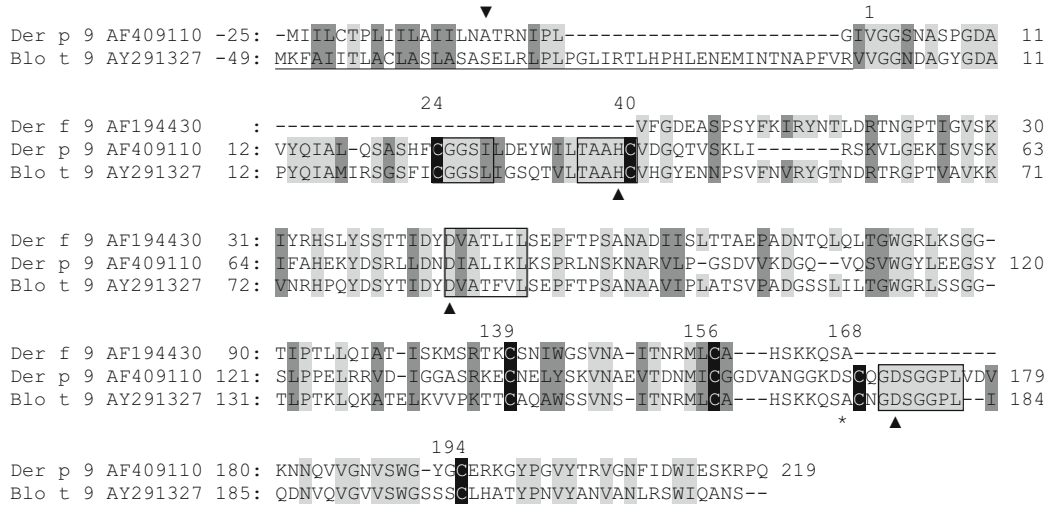


Figure 7.17 Multiple amino acid sequence alignment of group 9 allergens (collagenase). **L** = invariant sequences; **L** conservative substitutions; __ = signal peptide; ▼ = putative starting position of pro-peptide (based on Der p 3); 1 = starting position of mature protein; **C** = conserved cysteine residues (numbered from Der p 9); ▲ = catalytic residues typical of the serine peptidase family S1; boxed sequences = conserved domains of S1 serine peptidases. * = equivalent residue to serine 178, the residue in Der 6 allergens conferring substrate specificity.

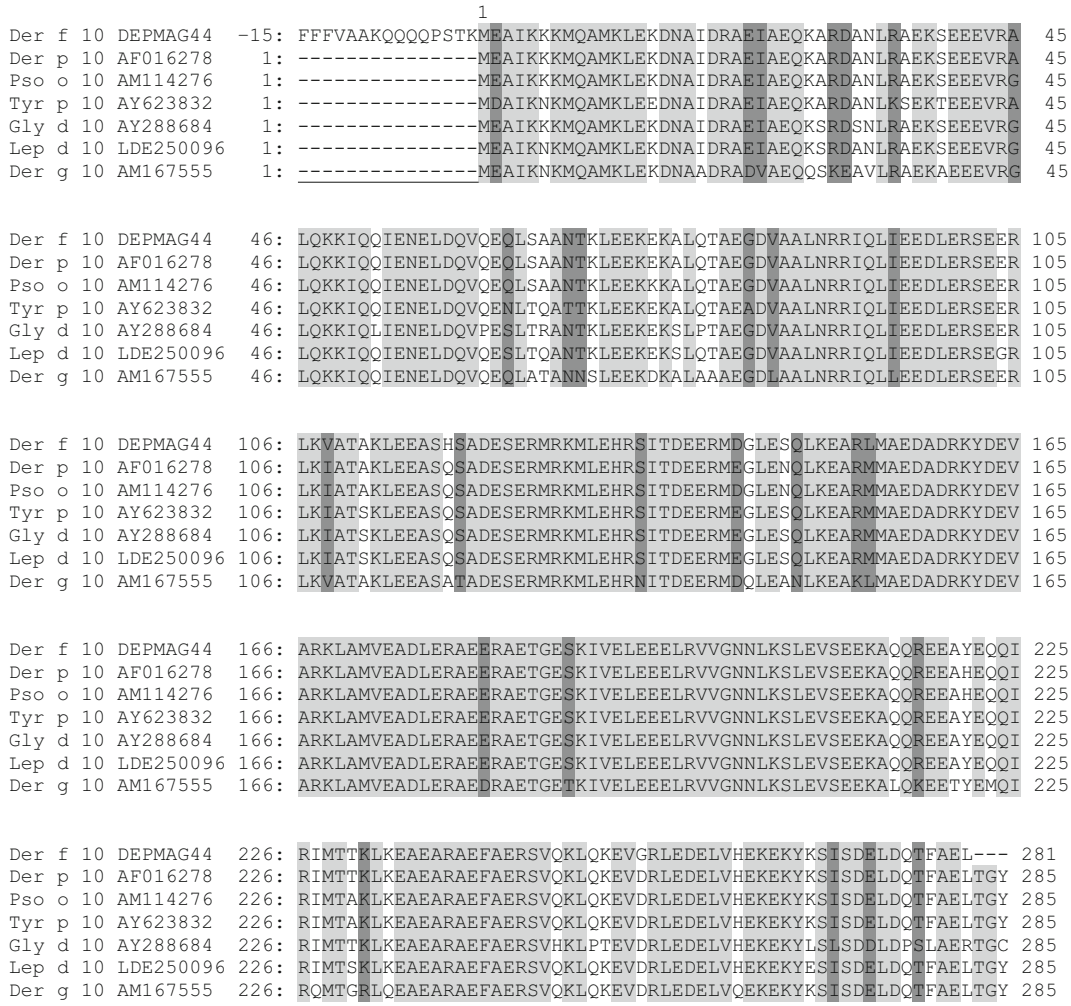


Figure 7.18 Multiple amino acid sequence alignment of group 10 allergens (tropomyosin). **L** = invariant sequences; **L** conservative substitutions; __ = signal peptide; 1 = starting position of mature protein.

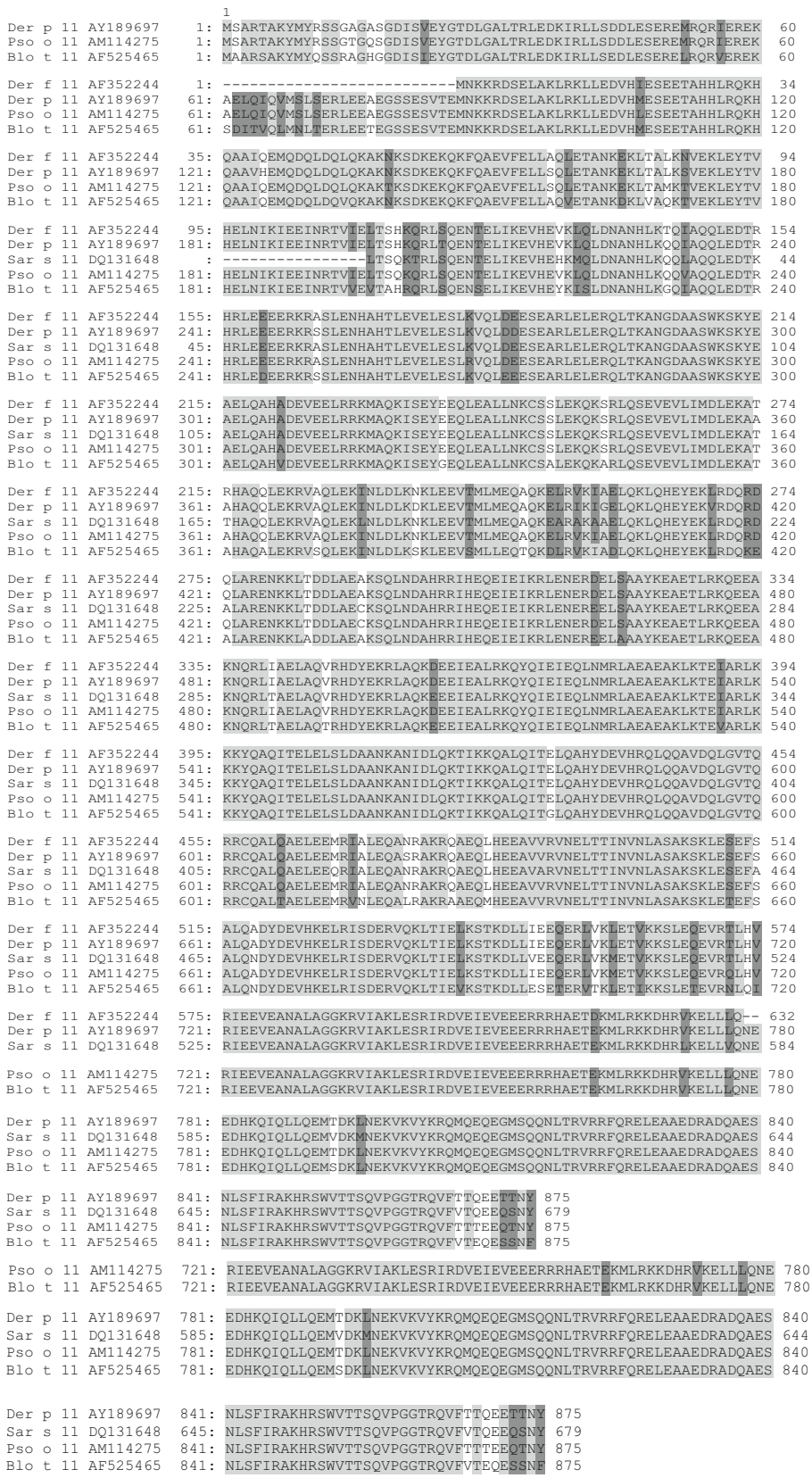


Figure 7.19 Multiple amino acid sequence alignment of group 11 allergens (paramyosin). L = invariant sequences; L conservative substitutions; 1 = starting position of mature protein.

```
                                                    1
Blo t 12 BTU27479  -21: MKSVLIFLVAIALFSSANIVSADEEQTTRGRHTEPDDHHEKPTTQCTHEETTSTQHHHEE   39
Lep d 12 AAQ55550  -20: MKSVLIFLVAIALFS-ANIVSAD-EQTTRGRHTEPDDHHEKPTTHATHEETTSTQHHHEE   38
CBP1     AY345124  929: EEAPAICAAEGSSGVLVAHENCNQFYKCANGVPVAFTCSASLLYNPYRGDCDWPSNVECG  988

Blo t 12 BTU27479   40: VVTTQTPHHEEKTTTEETHHSDDLIVHEGGKIYHVVCHEEGPIHIQ---EMCNKYIICSK   96
Lep d 12 AAQ55550   39: -VTTQTPHHEEKTTTEETHHSDDLIVHEGGKTYHVVCHEEGPIPHP---GNVHKYIICSK   94
CBP1     AY345124  989: NRPISVPDDNNVGTSTTTMPDDNQVINDDPSQAPSICAENGSSGVLVAHENCNQYYICSA 1048

Blo t 12 BTU27479   97: SGSLWYITVMPCSIGTKFDPISRNCVLDN      125
Lep d 12 AAQ55550   95: SGSLWYITVMPCSIGTKFDPISRNCVLDN---   123
CBP1     AY345124 1049: GRPV----PMPCSSGLLFNPVNRACDWPQNVVC 1077
```

Figure 7.20 Amino acid sequence alignment of group 12 allergens, with the peritrophic membrane protein CBP1 (chitin binding protein 1) from cabbage looper caterpillar (*Trichplusia ni*). L = invariant sequences; L conservative substitutions; C = conserved cysteine residues; 1 = starting position of mature protein; __ = signal peptide. **CAEN…CSA** = the peritrophin-A domain of CBP1.

```
                          1
Der f 13 2A0AA       1: MASIEGKYKLEKSEKFDEFLDKLGVGFMVKTAAKTLKPTFEVAIENDQYIFRSLSTFKNT  60
Tyr p 13 AY710432    1: MVQLNGSYKLEKSDNFDAFLKELGVNFVTRNLAKSASPTVEVIVDGDSYTIKTSSTLKNS  60
Aca s 13 (E, 1999)   1: --QINGSYKLEKSDNFDAFLKELGLNFVTRNLAKSATPTVEVSVNGDSYTIKTASTLKNT  58
Gly d 13 AAQ54609    1: MANIVGQYKLEKSENFDQFLDKLGVGFLVKTAAKTVKPTLEVAVDGDTYIFRSLSTFKNT  60
Lep d 13 CAB62213    1: MANIAGQYKLDKSENFDQFLDKLGVGFLVKTAAKTVKPTLEVAVDGDTYIFRSLSTFKNT  60
Blo t 13 BTU58106    1: -MPIEGKYKLEKSDNFDKFLDELGVGFMVKTAAKTLKPTLEVDVQGDTYVFRSLSTFKNT  59

Der f 13 2A0AA      61: TEAKFKLGEEFEEDRADGKRVKTVIQKEGDNKFVQTQFGDKEVKIIREFNGDEVVVTASC 120
Tyr p 13 AY710432   61: SEIKFKLGEEFEEDRADGKKVQTSVTKEGDNKLVQVQKGDKPVTIVREFSEEGLTVTATV 120
Aca s 13 (E, 1999)  59: TEISKFLGEEFEEARADGKTVKTVVNKESDTKFVQVQQGDKEVTIVREFSDEGLTVTATV 118
Gly d 13 AAQ54609   61: TEIKFKLGEEFEEDRADGKRVKTVVVKEGDNKFVQTQFGDKEVKVVREFKGDEVEVTASV 120
Lep d 13 CAB62213   61: TEIKFKLGEEFEEDRADGKRVKTVIVKDGDNKFVQTQYGDKEVKVVREFKGDEVEVTASV 120
Blo t 13 BTU58106   60: TEIKFKLGEEFEEDRADGKRVKTVVNKEGDNKFIQTQYGDKEVKIVRDFQGDDVVVTASV 119

Der f 13 2A0AA     121: DGVTSVRTYKRI 132
Tyr p 13 AY710432  121: NGVTSVRFYKRQ 132
Aca s 13 (E, 1999) 119: SGVTSVRFYKRQ 130
Gly d 13 AAQ54609  121: DGVNSVRLYKRL 132
Lep d 13 CAB62213  121: DGVTSVRPYKRA 132
Blo t 13 BTU58106  120: GDVTSVRTYKRI 131
```

Figure 7.21 Multiple amino acid sequence alignment of group 13 allergens (fatty acid binding protein). L = invariant sequences; L conservative substitutions; 1 = starting position of mature protein; **NXS NXT** = *N*-glycosylation sites (where X = any amino acid except proline); __ = cytosolic fatty acid-binding protein signature. E, 1999 = Eriksson *et al.*, 1999.

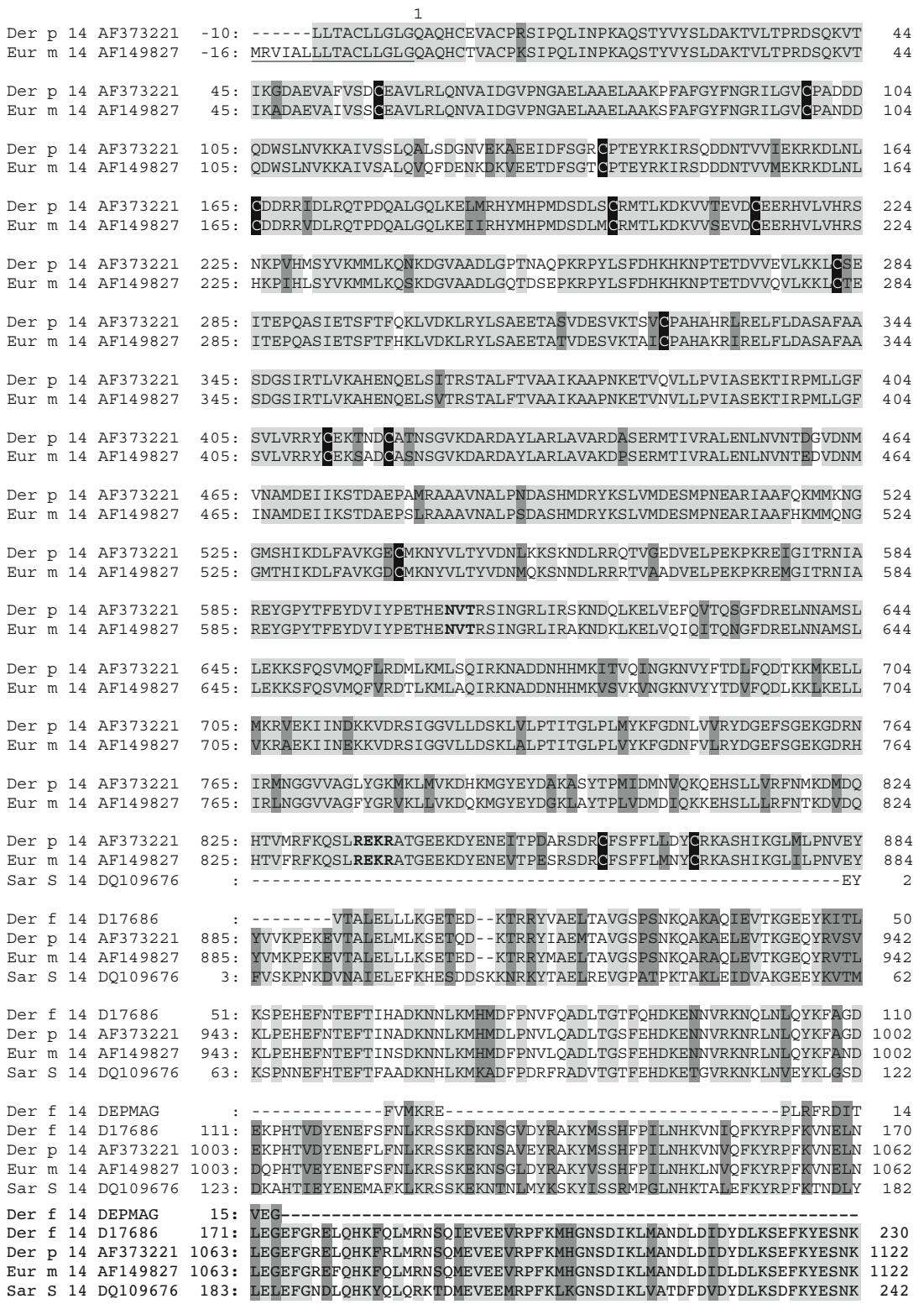


Figure 7.22 Multiple amino acid sequence alignment of group 14 allergens (apolipophorin or vitellogenin). DEPMAG = Mag 1; D17686 = Mag 3. **L** = invariant sequences; **L** conservative substitutions; **C** = conserved cysteine residues; __ = signal peptide; 1 = starting position of mature protein; **NXS NXT** = *N*-glycosylation sites (mature protein, where X = any amino acid except proline); **RX(K/R)R** = subtilisin cleavage sites.

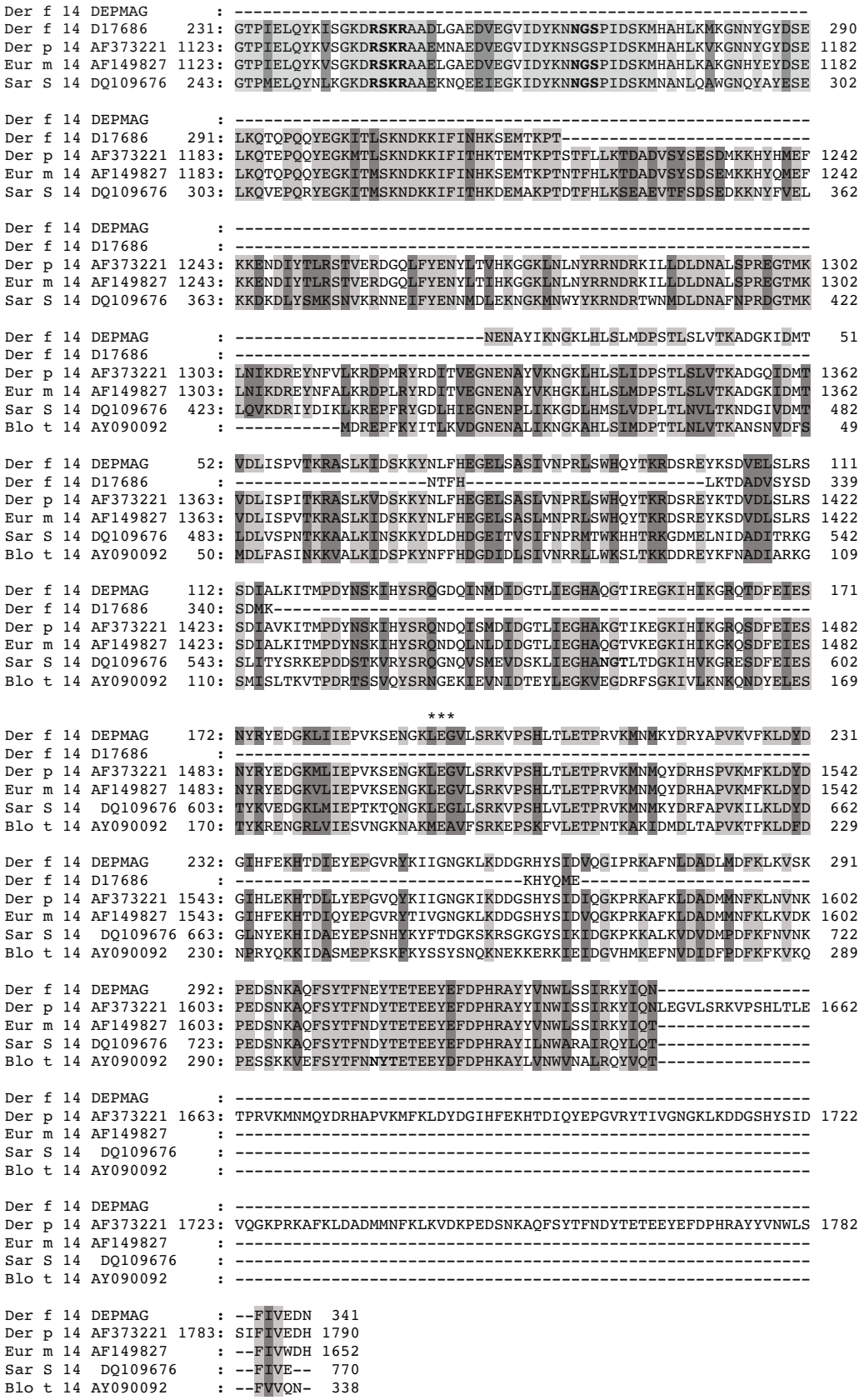


Figure 7.22 Continued

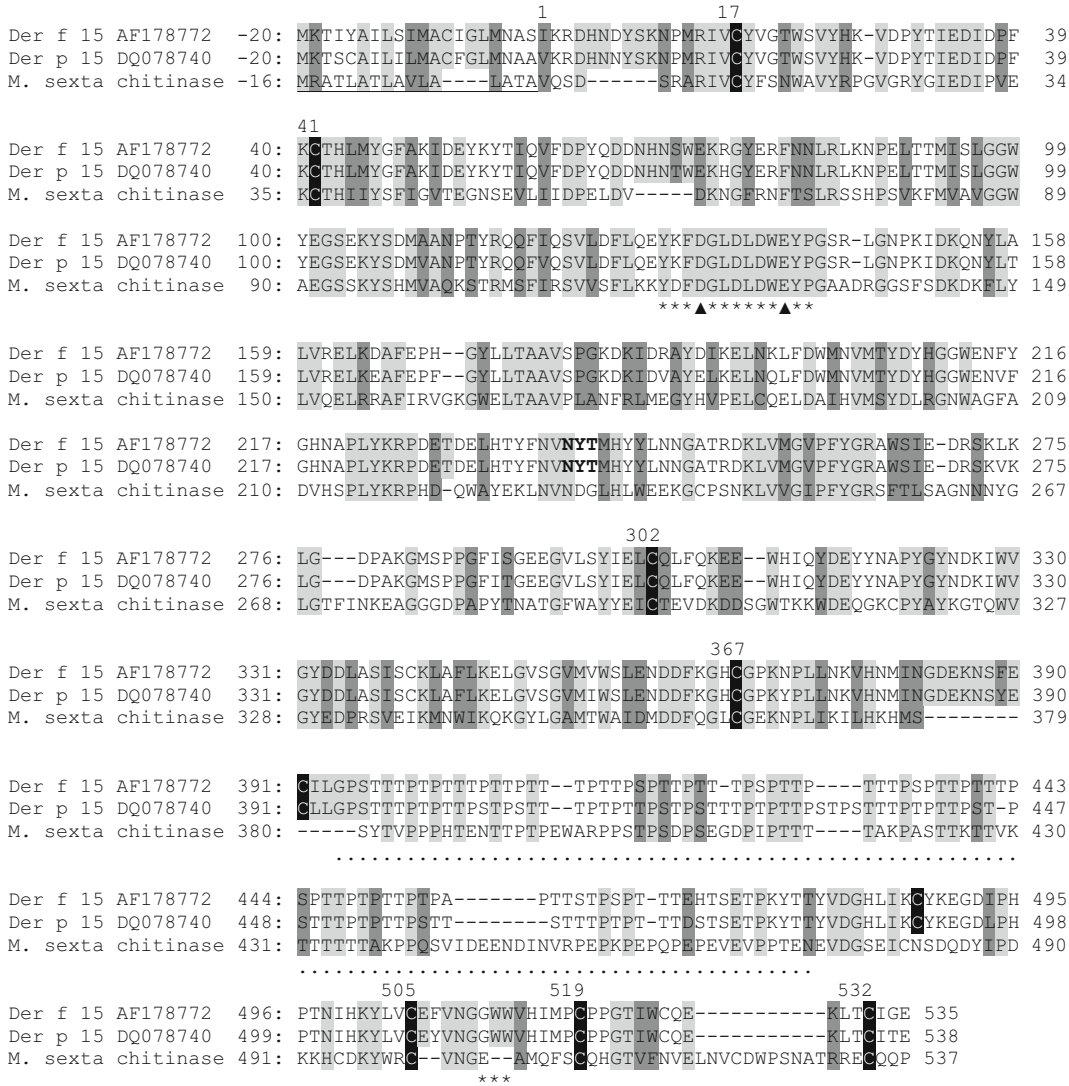


Figure 7.23 Multiple amino acid sequence alignment of group 15 allergens with chitinase from the moth *Manduca sexta* (sequence from Kramer and Muthukrishnan, 1997). L = invariant sequences; L conservative substitutions; **NXS NXT** = *N*-glycosylation sites (mature protein); __ = signal peptide; 1 = starting position of mature protein; C = conserved cysteine residues (numbered); *** = conserved sequences typical of arthropod chitinase domains; ▲ = catalytic residues; ….. = proline/serine/threonine-rich *O*-glycosylation domain.

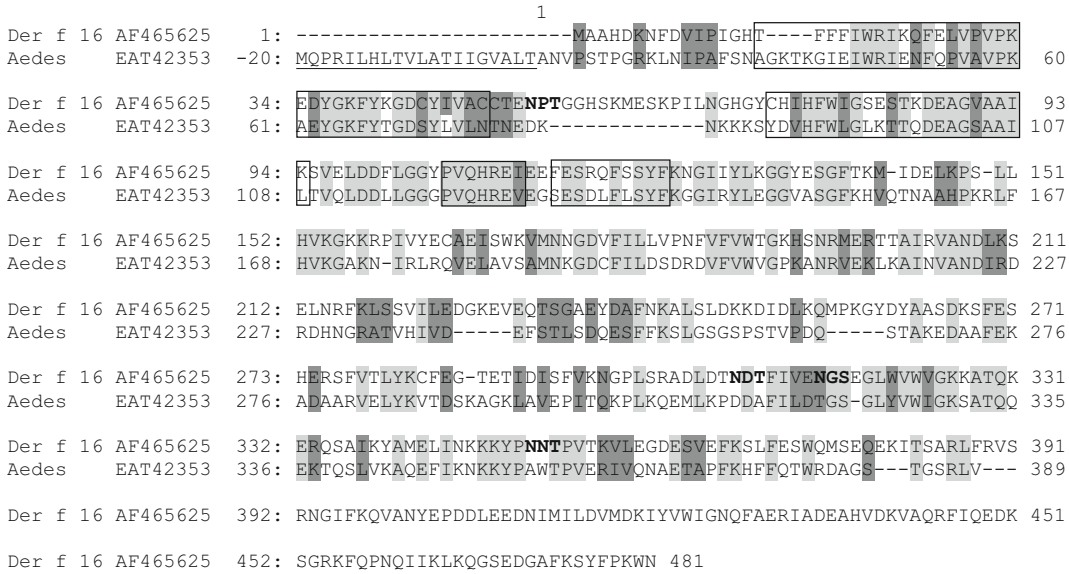


Figure 7.24 Amino acid sequence alignment of Der f 16, the only group 16 allergen (gelsolin) with *Aedes aegypti* gelsolin precursor (EAT42353). L = invariant sequences; L conservative substitutions; __ = signal peptide; 1 = starting position of mature protein; **NXS NXT** = *N*-glycosylation sites (mature protein); boxed = gelsolin homology domain.

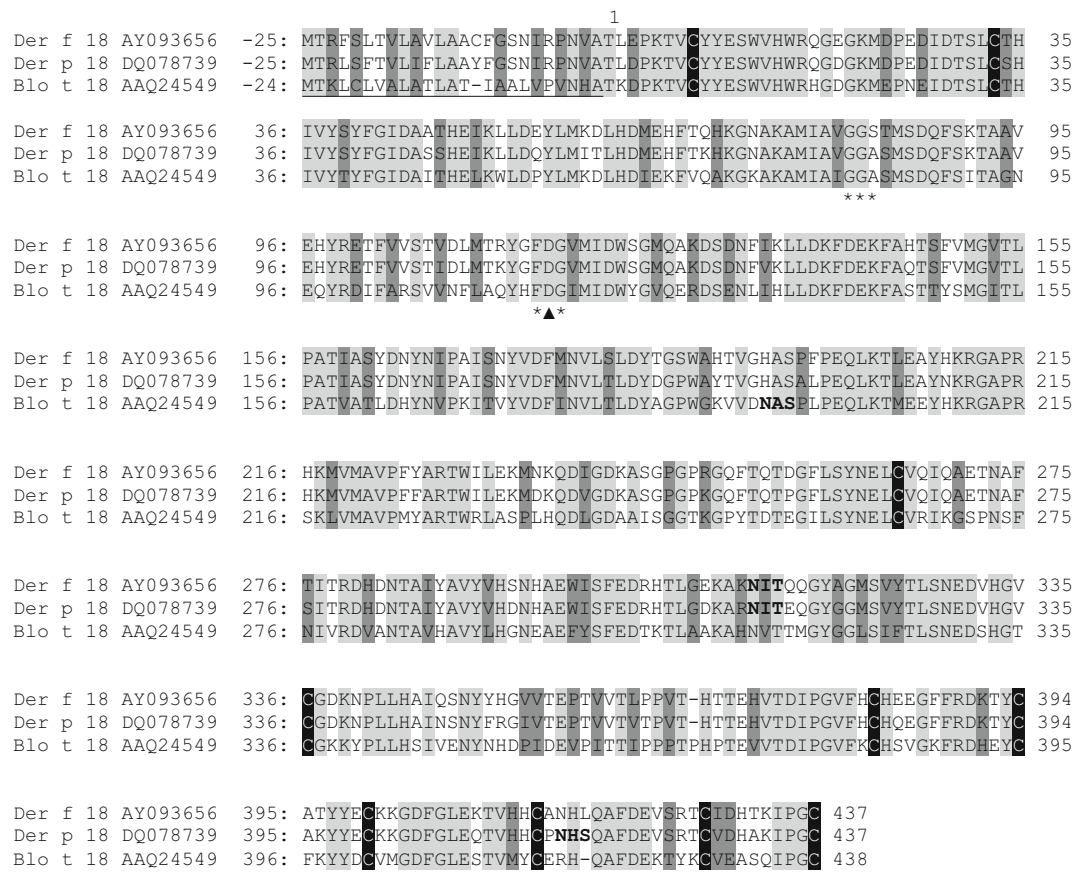


Figure 7.25 Multiple amino acid sequence alignment of group 18 allergens (chitinase). L = invariant sequences; L conservative substitutions; **NXS NXT** = *N*-glycosylation sites (mature protein); __ = signal peptide; 1 = starting position of mature protein; C = conserved cysteine residues; *** = conserved sequences typical of arthropod chitinases; ▲ = catalytic residues.

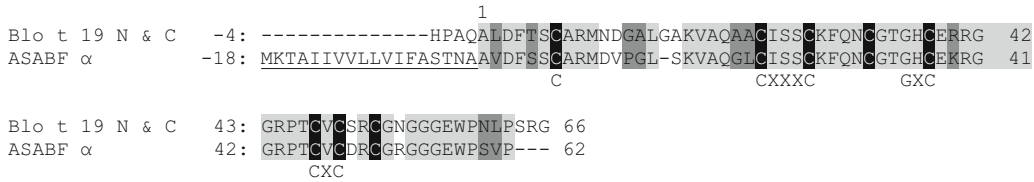


Figure 7.26 Amino acid sequence alignment of Blo t 19, the only known group 19 allergen, with the antibacterial factor ASABF- α from the roundworm, *Ascaris suum*. **C** = conserved cysteine residues; **L** = invariant sequences; **L** conservative substitutions; __ = signal peptide; 1 = starting position of mature protein; CXXC = invertebrate defensin consensus; N & C = Nge and Chua, unpublished data; Nge, personal communication.

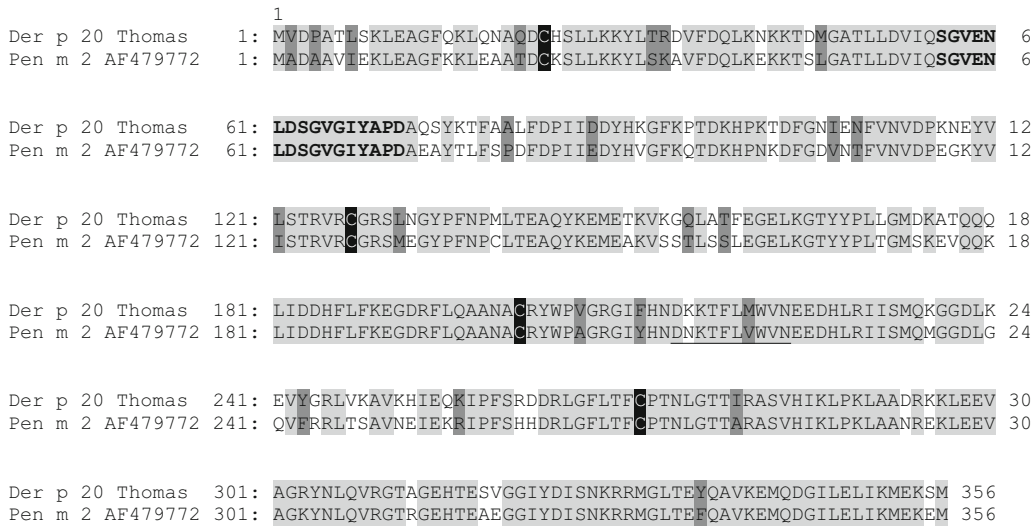


Figure 7.27 Amino acid sequence alignment of Der p 20 (Thomas, unpublished), the only known group 20 allergen, with the arginine kinase allergen, Pen m 2, from the prawn, *Penaeus monodon* (Yu *et al.*, 2003). **C** = conserved cysteine residues; **L** = invariant sequences; **L** conservative substitutions; 1 = starting position of mature protein; **SGV...APD** = active site; __ = putative actin-binding domain.

```
Der f DF0793       1: -------------MAQVEVKYTQIFINNEWHDSISGKTFETINPFTEEKLANVQEGDKAD  47
Blo t AY291328     1: ---------------MVQVKYTKIFINNEWHDSLSGKTFPTINPSTEEKLADIQAGEKAD  45
X.l.  AB016718     1: MSEKTILHGLPPPLSNLEIKYTKIFINNEWHNSISGKKFPVYNPATGEKVCEVEEGDKED  60

Der f DF0793      48: IDRAVVAAVDAFRFDSPWRQMDASQRGHLLYRLADLIERDQDYIASLESMDNGKPKTMAL 107
Blo t AY291328    46: VDKAVEAAVKAFKIGSEWRKMDASARGRLLNKLAELVERDQKYLAELETLDNGKPLPMAM 105
X.l.  AB016718    61: VNKAVKAAREAFQIGSPWRRLDASERGRMLNKLADLVERDRLILSTMESIDGGKPYTASY 120

Der f DF0793     108: F-DVDLAIKVFRYYAGYADKIHGKTIPADGKVFAFTRIEPVGICGQIVPWNFPFLMASWK 166
Blo t AY291328   106: N-DIHAAIRGFEYYAGWADKIHGKTIPADGNVCSFTKIEPVGVCGQIIPWNFPLVMLAWK 164
X.l.  AB016718   121: FGDIPGAIKSLRYCAGWADKVQGRTIPMDGDYFTFTRHEPVGVCGQIIPWNFPLVMFAWK 180

Der f DF0793     167: FGPALCAGNTVVLKPAEQTPLSALYLASLTKEGGFPPGVVNVVPGFGETAGAALVDNPKV 226
Blo t AY291328   165: WAPALATGCTVVLKPAEQTPLTALYMGSLIIEAGFPPGVVNIVPGYGPTAGAALVDNPLV 224
X.l.  AB016718   181: IAPALCCGNTVVIKPAEQTPLTALYMGSLIKEAGIPPGVVNIVPGYGPTAGAAISYHMDI 240

Der f DF0793     227: DKIAFTGSTEIGKLIMR-NGSHSMKRITLELGGKSPLVVTENVEDIAQAARTAQDSCFLN 285
Blo t AY291328   225: DKIAFTGSTEVGKLISK-NASATLKRVSLELGGKSPLVVTENF-DVTMAAEIASAGCMVN 282
X.l.  AB016718   241: DKVAFTGSTEVGKLIKEAAGKSNLKRVTLELGGKSPNIIFADA-DLDLAVEHAHNGLFFH 299

Der f DF0793     286: MGQCCCAGTRTFVHESIYDEFVKHSVEYCQSHVFGNPFDSKTAFGPQVDKIQMNRILEMI 345
Blo t AY291328   283: QGQACCAATRTFVHESIYDEFVKQSAIIAQRRVVGCPFEAATVQGPQIDATQTEKILELI 342
X.l.  AB016718   300: QGQCCIAGSRIFVEEPIYDEFVRKSVERAKKRVLGDPFAPCVNQGPQIDKEQYDKILELI 359
                           ▲

Der f DF0793     346: ESGKQEGARCVAGGNRMDKRGYFVEPTVFADVTDGMRIAREEIFGPVQQILKYKTLDEVI 405
Blo t AY291328   343: ESGKKQGARLVTGGSRMKQKGYFIEPTVFADVDDSMRIAKEEIFGPVQQLFKYKTLDEAI 402
X.l.  AB016718   360: ESGKKEGAKLQCGGSAWGEKGFYISPTVFSDVKDDMRIAKEEIFGPVQQILKFKTIDEVI 419

Der f DF0793     406: ERCNDTNYGLGSAILTNDINEAMKFSRSIRAGSVWINIPYMIPVSVQTPFGGFKESGVGR 465
Blo t AY291328   403: ERSNATEY-------------------------VW--LPAF                   416
X.l.  AB016718   420: KRANNTKYGLAAGVFTKDMDKAILMSTALQAGTVW--INCYSAMSPQSPFGGFKMSGNGR 477

Der f DF0793     466: ELGEDGLRGYGEIKTVVIMDREKKM 490
X.l.  AB016718   478: EMGEYGLHEYTEVKTVIMKISQKNS 502
```

Figure 7.28 Amino acid sequence alignment of aldehyde dehydrogenase homologue allergens with *Xenopus laevis* aldehyde dehydrogenase. **C** = conserved cysteine residues; **L** = invariant sequences; ▲ = catalytic cysteine; **L** conservative substitutions; **NXS NXT** = *N*-glycosylation sites (mature protein).

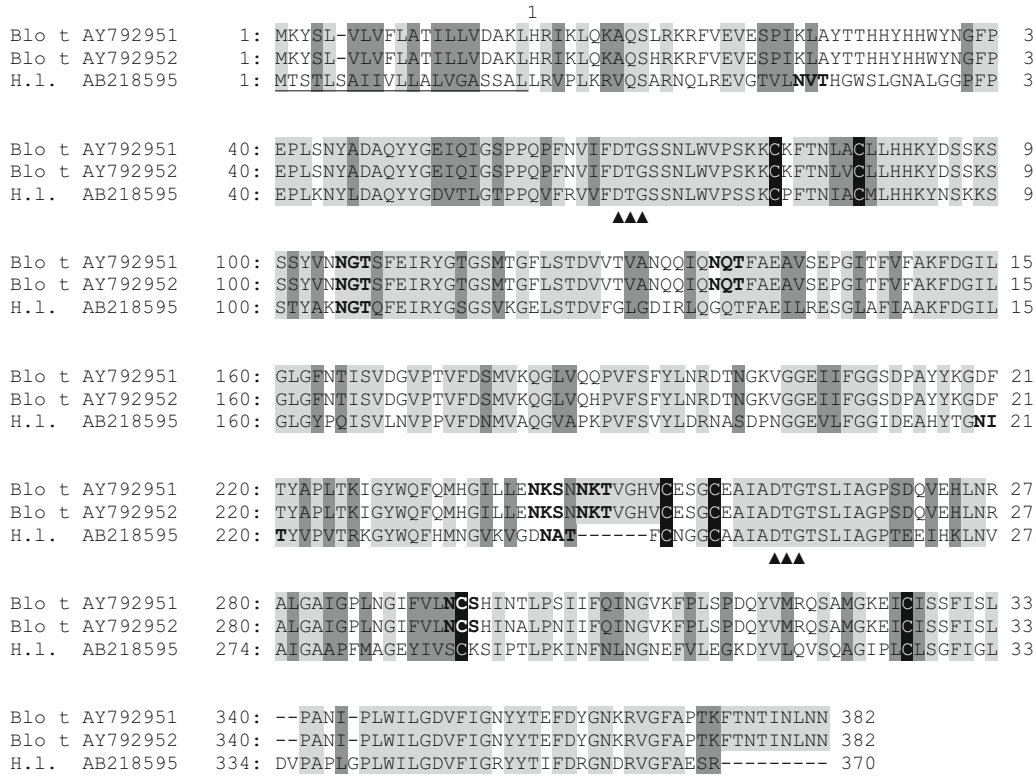


Figure 7.29 Amino acid sequence alignment of cathepsin D homologue allergens with *Haemaphysalis longicornis* aspartic protease (Boldbaatar *et al.*, 2006). **C** = conserved cysteine residues; **L** = invariant sequences; **L** conservative substitutions; 1 = starting position of mature protein; __ = signal peptide; ▲ = catalytic residues of aspartic proteases.

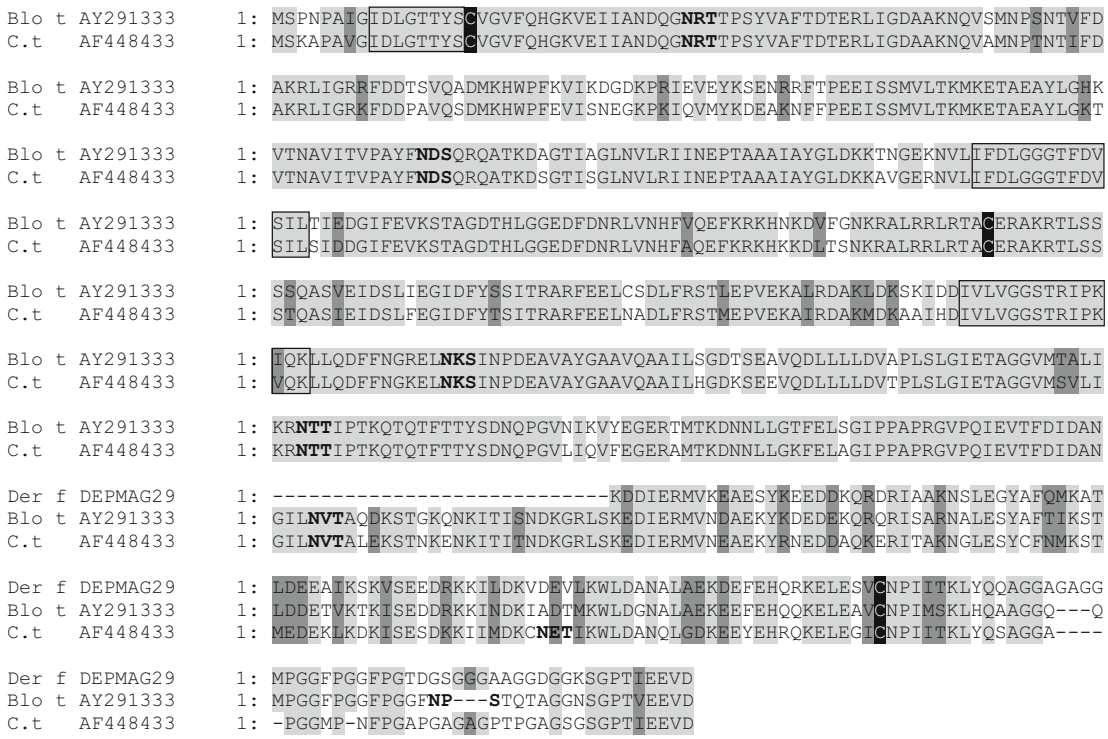


Figure 7.30 Amino acid sequence alignment of Mag29 heat shock protein 70 allergens from *Dermatophagoides farinae* and *Blomia tropicalis* with heat shock cognate protein 70 from the midge, *Chironomus tentans* (Karouna-Renier *et al.*, 2003). **C** = conserved cysteine residues; **L** = invariant sequences; **L** conservative substitutions; **NXS NXT** = *N*-glycosylation sites (mature protein); boxed residues = heat shock protein 70 family signatures.

```
Lep d  AJ428050   1: MRECISVHVGQAGVQIGNACWELYCLEHGIQPDGQMPSDKTIGTGDDSFNTFFSETGSGK  60
Tyr p  AY986760   1: MRECISVHVGQAGVQIGNACWELYCLEHGIQPDGQMPSDKTIGTGDDSFNTFFSETGSGK  60

Lep d  AJ428050  61: HVPRAVYVDLEPTVVDEVRTGTYRQLFHPEQLITGKEDAANNYARGHYTIGKEIVDLVLD 120
Tyr p  AY986760  61: HVPRAVYVDLEPTVVDEVRTGTYRQLFHPEQLITGKEDAANNYARGHYTIGKEIVDVVLD 120

Lep d  AJ428050 121: RIRKLADQCTGLQGFLIFHSFGGGTGSGFTSLLMERLSVDYGKKSKLEFAVYPAPQVSTA 180
Tyr p  AY986760 121: RIRKLSDQCTGLQGFLIFHSFGGGTGSGFTSLLMERLSVDYGKKSKLEFAVYPAPQVSTA 180

Lep d  AJ428050 181: VVEPYNSILTTHTTLEHSDCAFMVDNEAIYDICRRNLDIERPTYTNLNRLIGQIVSSITA 240
Tyr p  AY986760 181: VVEPYNSILTTHTTLEHSDCAFMVDNEAIYDICRRNLDIERPTYTNLNRLIGQIVSSITA 240

Lep d  AJ428050 241: SLRFDGALNVDLTEFQTNLVPYPRIHFPLVTYAPVISSEKAYHEQLTVSEITNTCFEPAN 300
Tyr p  AY986760 241: SLRFDGALNVELTEFQTNLVPYPRIHFPLVTYSPVISAEKAYHEQLTVAEITNTCFEPQN 300

Lep d  AJ428050 301: QMVKCDPRHGKYMACCLLYRGDVVPKDVNAAIAAIKTKRSIQFVDWCPTGFKVGINYQPP 360
Tyr p  AY986760 301: QMVKCDPRHGKYMACCLLYRGDVVPKDVNAAIAGIKTKRSIQFVDWCPTGFKVGINYQPP 360

Lep d  AJ428050 361: TVVPGGDLAKVQRAVCMLSNTTAIAEAWARLDHKFDLMYAKRAFVHWYVGEGMEEGEFSE 400
Tyr p  AY986760 361: TVVPGGDLAKVQRAVCMLSNTTAIAEAWARLDHKFDLMYAKRAFVHWYVGEGMEEGEFSE 400

Lep d  AJ428050 401: AREDLAALEKDYEEVGLDSTEADDTAGEEF 430
Tyr p  AY986760 401: AREDLAALEKDYEEVGLDSTEAEGGDGEEF 430
```

Figure 7.31 Amino acid sequence alignment of alpha-tubulin allergens from *Lepidoglyphus destructor* and *Tyrophagus putrescentiae*. **L** = invariant sequences; **L** conservative substitutions.

```
Blo t  AY291334   1: MSWQSYVDNQICQHVECRLAVIAG-LDGSVWAKFEKDIPKQVSQQELKTIADAIRTNPNS 59
Sui m  AY803196   1: MSWQSYVDNQICQHVDCSLAVIASNQDGAIWAQFERE-NQSISPNELKTIAETIRQNPAG 59

Blo t  AY291334  60: FLEGGIHLGGEKYICIQADNSLVRGRKGSSALCIVATNTCLLAAATVDGFPPGQLNNVVE 109
Sui m  AY803196  60: FLDNGIHIGGSKYICIQADNTLVRGRKGSSALCIVATNTCLLIAATVDGFPPGQLNNVIE 109

Blo t  AY291334 110: KLGDYLKANNY 120
Sui m  AY803196 110: KLGDYLRSNNY 120
```

Figure 7.32 Amino acid sequence alignment of profilin allergens from *Blomia tropicalis* and *Suidasia medanensis*. **C** = conserved cysteine residues; **L** = invariant sequences; **L** conservative substitutions; **NXS NXT** = *N*-glycosylation sites (mature protein).