

Comparative molecular modelling identifies a common putative IgE epitope on cysteine protease allergens of diverse sources

R. FURMONAVICIENE, H. F. SEWELL and F. SHAKIB

Division of Molecular and Clinical Immunology, Faculty of Medicine and Health Sciences, University of Nottingham, Nottingham, UK

Summary

Background Previous approaches for studying common allergenic epitopes have mainly focused on sequence comparisons, which unfortunately yield little or no information on the shape of the epitope which is the most important determinant of cross-reactivity.

Objectives The aim of this study was to investigate the structural basis for cross-reactivity between a previously identified immunodominant epitope of the house dust mite allergen Der p 1 (Leu147-Gln160) and the corresponding epitopes on other allergens that are either taxonomically closely related (i.e. cysteine proteases of other mite species) or representing evolutionary conserved structures (i.e. plant, human and parasite cysteine proteases).

Methods We carried out comparative molecular modelling on a range of cysteine proteases, including those of other mite species (Der f 1 and Eur m 1), human (cathepsins B, K, L, S and O), plants (papain, chymopapain and actinidin) and parasites (cruzain, cathepsin L-like Leishmania protease, Entamoeba ACP1 protease and Schistosoma Q26534, Q11003 and cathepsin L proteases).

Results Our study shows that all the cysteine proteases investigated here display an epitope corresponding to that previously identified on Der p 1, but with varying shapes and degree of accessibility. It appears that the core of the epitope on these homologous cysteine proteases consists of a centrally located conserved Tyr residue flanked on either sides by accessible amino acids.

Conclusion Therefore, these cysteine proteases seem to use similar accessible structures, which may form the basis for the rational design of generic epitope-directed treatment strategies for controlling allergic diseases.

Keywords: Allergen, cysteine proteases, Der p 1, IgE

Clinical and Experimental Allergy, Vol. 30, pp. 1307–1313. Submitted 8 September 1999; revised 27 January 2000; accepted 16 February 2000.

Introduction

House dust mite allergens are a major cause of bronchial asthma, allergic rhinitis and atopic dermatitis, and as such have been strongly implicated in the increasing prevalence of these diseases. Der p 1, a major allergen of the house dust mite *Dermatophagoides pteronyssinus*, accounts for over

50% of the total human IgE antimite allergen response [1,2]. More than 80% of individuals who are sensitive to this mite species produce IgE antibodies to Der p 1 [3,4].

These observations have therefore recently prompted us to determine the epitope recognized by a mouse monoclonal anti-Der p 1 antibody (mAb 2C7) representative of a major component of the human IgE anti-Der p 1 response [5]. The immunodominant epitope identified by mAb 2C7 is Leu147-Gln160, which according to the 3-dimensional model of Der p 1 is contained in a loop (Val140-Tyr169) located on the surface of the domain comprising mainly

Correspondence: Dr F. Shakib, Division of Molecular and Clinical Immunology, Faculty of Medicine and Health Sciences, Queen's Medical Centre, Nottingham, NG7 2UH, UK.

β -sheets [6]. We have also shown that the crystallographically determined 3-dimensional structures of other homologous cysteine proteases, namely chymopapain, papain and actinidin have accessible loops corresponding to that containing the Der p 1 epitope recognized by mAb 2C7, thereby suggesting that this antibody may cross-react with these molecules. This was indeed confirmed by our ELISA inhibition tests, which also showed that the degree of cross-reactivity was directly related to the shape of the epitope (i.e. the superimposition of the main chain and the position and length of accessible side chains) [6].

Our recent comparative structural studies [6] and those conducted earlier [7,8] suggest that allergens use a restricted set of structural features to elicit IgE responses. Clearly, such observations would bring the prospect of developing generic allergen-derived peptides for controlling allergic diseases much closer. As part of these ongoing efforts, we sought to extend our comparative molecular modelling studies to other cysteine proteases, including those of other mite species (Der f 1 and Eur m 1), human (cathepsins B, K, L, S and O), plants (papain, chymopapain and actinidin) and parasites (cruzain, cathepsin L-like Leishmania protease, Entamoeba ACP1 protease and Schistosoma Q26534, Q11003 and cathepsin L proteases). Here we demonstrate that homologous cysteine proteases of diverse sources share a common putative IgE epitope.

Materials and methods

Comparative molecular modelling

The 3-dimensional model of Der p 1 was reproduced from the previously published co-ordinates [9].

The 3-dimensional structures of papain [10], chymopapain [11], actinidin [12], cruzain [13] and human cathepsin B [14] were reproduced according to the co-ordinates of the corresponding crystal structures available from the Protein Data Bank (PDB). Theoretical co-ordinates of human cathepsins K [15] and S [16], as well as the cathepsin L-like Leishmania protease (code 1BMJ) [17], were obtained from the PDB database.

Theoretical co-ordinates of human cathepsin L [18], Schistosoma Q26534, Q11003 [19] and cathepsin L proteases [20], as well as the cathepsin-L like protease from Entamoeba [21], were obtained from the SWISS-MODEL Repository (<http://www.expasy.ch/swissmod/SWISS-MODEL.html>).

Models of Der f 1 and Eur m 1, as well as that of human cathepsin O, were generated using the Internet server SWISS-MODEL [22,23]. The Der f 1 model was built using the following templates: 11YAL.pdb, 11MEG.pdb, 11PPO.pdb and 11AEC.pdb, as well as the co-ordinates of Der p 1. The Eur m 1 model was generated using the most

similar templates from the ExNRL-3D database (11AEC.pdb, 12ACT.pdb, 11MEG.pdb and 11PPO.pdb), as well as the co-ordinates of Der p 1. For the model of human cathepsin O the following templates were used: 11BXF.pdb, 18PCH.pdb, 11BMJ.pdb, 11AEC.pdb and 12ACT.pdb.

All molecular models were viewed and examined for the epitopes corresponding to that of Der p 1 (Leu147-Gln160) using the program Swiss Pdb Viewer [22,23]. Corresponding epitopes were examined for accessibility, mobility, charge, conserved amino acids and epitope conformation by superimposing the 3-dimensional models of the homologous cysteine proteases on that of Der p 1.

Reactivity of mAb 2C7 with Der p 1 and homologous cysteine proteases

Der p 1 (1 μ g/mL; prepared as described in our recent paper [24]) was bound directly onto microtitre plates and the ability of various concentrations of soluble Der p 1, chymopapain (Sigma, Poole, UK), papain (BDH Darmstadt, Germany), actinidin and human cathepsins B (Sigma) and L (Calbiochem-Novabiochem, Nottingham, UK) to inhibit the binding of 1:1000 dilution of anti-Der p 1 mAb 2C7 (prepared as described in our recent paper [5]) was assessed using 1:1000 dilution of alkaline phosphatase conjugated monoclonal antimouse IgG (Sigma).

Results

Cross-reactivity between Der p 1 and plant cysteine proteases

Comparison of the 3-dimensional model of Der p 1 (Fig. 1a) with the crystallographically determined structures of chymopapain (Fig. 1b), papain (Fig. 1c) and actinidin (Fig. 1d) revealed a similarity in the shape of the cross-reactive epitope, which occupy a protruding loop on the surface of the domains mainly comprised of β -sheets. The main chains of these homologous cysteine proteases could be superimposed (Fig. 1e), with the exception of the C-terminal region (Asp154-Thr157) of the Der p 1 epitope. These observations are very much in line with our ELISA inhibition data showing that anti-Der p 1 mAb 2C7 also binds to these homologous cysteine proteases in the order chymopapain > papain > actinidin (Fig. 2). The corresponding sequences of Der p 1, chymopapain, papain and actinidin are shown in Table 1.

These data clearly show that immunological reactivity with anti-Der p 1 mAb 2C7 is mostly dependent on shape complementarity, rather than the charge distribution, of the residues comprising the epitope (Table 1). Based on this observation, we extended our comparative molecular

modelling efforts to other cysteine proteases, namely those of other mite species, human and parasites.

Cross-reactivity between Der p 1 and cysteine proteases of other mite species

The shape of the corresponding epitopes in the homologous mite cysteine proteases Der f 1 (Fig. 1f) and Eur m 1 (Fig. 1g) was very similar to that of Der p 1, which is to be expected given their close taxonomic relationship. The slight differences between the α -helical part of the main chain and those at the bent part of the loop are compensated for by the protruding amino acid side chains (Fig. 1h). All these cysteine proteases appear to have a characteristic conserved motif, namely His152-Thr157, towards the C-terminus of the cross-reactive epitope (Table 1).

Cross-reactivity between Der p 1 and human cathepsins

The shape of the corresponding epitopes in the homologous cathepsins L (Fig. 1i), K (Fig. 1j) and S (Fig. 1k) was very similar to that of Der p 1. As in plant cysteine proteases (Fig. 1e), the main chains of these homologous cysteine proteases could be superimposed, with the exception of the C-terminal region (Asp154-Thr157) of the Der p 1 epitope. In the case of cathepsin O, the main chain differed slightly as the α -helical part did not form a complete turn at the N-terminus of the epitope (Fig. 1l). The shape of the epitope in cathepsin O suggests a lower degree of cross-reactivity than with cathepsins L, K and S (Fig. 1m). The basic antigenic motif in these cysteine proteases is apparently formed by the most protruding residues flanking the centrally located conserved Tyr residue (Table 1).

Interestingly, our ELISA inhibition data showed that cathepsin L, but not cathepsin B, is detectable with anti-Der p 1 mAb 2C7 and that its order of reactivity is as follows: chymopapain > cathepsin L > papain > actinidin (Fig. 2). This is not surprising since we found that the main chain of the Der p 1 epitope could be superimposed on that of cathepsin L, but not on that of cathepsin B (Fig. 1n). Despite sequence similarity (Table 1), the epitope in cathepsin B assumes a different disposition in the 3-dimensional structure and is shielded by neighbouring loops. Thus, the reactivity of the cysteine proteases in our ELISA inhibition (Fig. 2) directly reflects the shape of their epitopes compared to that of Der p 1 (chymopapain > cathepsin L > papain > actinidin), which can best be viewed sideways (Fig. 1z').

Cross-reactivity between Der p 1 and parasite cysteine proteases

Nematode cysteine proteases are represented in this study by the two cathepsin L-like Schistosomal cysteine proteases

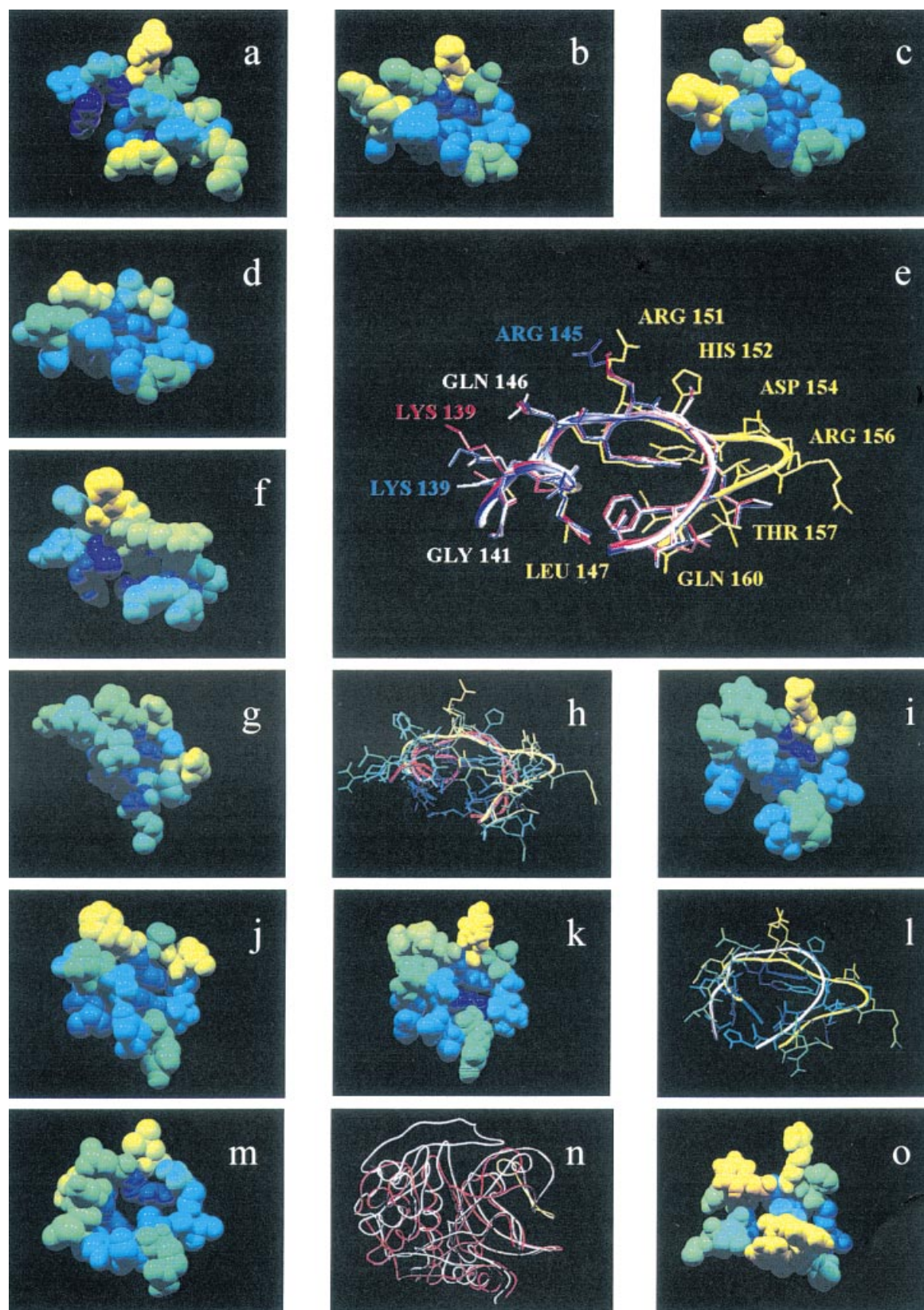
Q26534 (Fig. 1o) and Q11003 (Fig. 1p) and Schistosomal cathepsin L (Fig. 1q). The central part of the main chain of these cysteine proteases could be superimposed on the corresponding part of the main chain of Der p 1, with the shape of the corresponding epitope in cathepsin L-like cysteine protease Q26534 being the most similar to that of Der p 1 (Fig. 1r,s,t). The shape of the epitope on this protease is similar to the corresponding epitope on chymopapain, with LXSXXXD/E being a common motif (Table 1).

The backbone of the *Entamoeba* cysteine protease ACP1 (Fig. 1u) is the one that could best be superimposed on that of Der p 1, with the other two intracellular parasite cysteine proteases, namely cruzain (Fig. 1v) and *Leishmania* cathepsin L-like protease (Fig. 1w), showing a lower degree of overlap at the α -helical and at the C-terminal parts of the epitope. The shape of the epitopes suggests the following order of cross-reactivity: ACP1 protease (Fig. 1x) > cruzain (Fig. 1y) > *Leishmania* cathepsin L-like protease (Fig. 1z). The ACP1 protease has a structural motif common to papain (QLXR) and chymopapain (QLXK) (Table 1).

Discussion

Cross-reactivity amongst allergens is of considerable scientific and clinical interests. Previous attempts to study common allergenic epitopes have mainly focused on sequence comparisons [7], which identified conserved residues within regions important for maintaining the secondary and tertiary structures, thereby forming the hydrophobic core of the molecule. Unfortunately, such approaches yield little or no information on the shape of the epitope which is the most important determinant of cross-reactivity. For instance, our recent studies have demonstrated that the binding of the mouse monoclonal anti-Der p 1 antibody mAb 2C7 to chymopapain, papain and actinidin was directly related to the shape of the cross-reactive epitope (i.e. the superimposition of the main chain and the position and length of accessible side chains) [6].

In this study, we used comparative molecular modelling to investigate the structural basis for cross-reactivity between a previously identified Der p 1 epitope (Leu147-Gln160) [6] and the corresponding epitopes on other allergens that are either taxonomically closely related (i.e. cysteine proteases of other mite species) or representing evolutionary conserved structures (i.e. plant, human and parasite cysteine proteases). The study revealed that all the cysteine proteases investigated here (two from other mite species, three from plants, five from human and six from parasites) displayed the corresponding epitope, but with varying shapes and degree of accessibility. This has indeed been directly confirmed by our ELISA inhibition tests, using chymopapain, papain, actinidin, and human



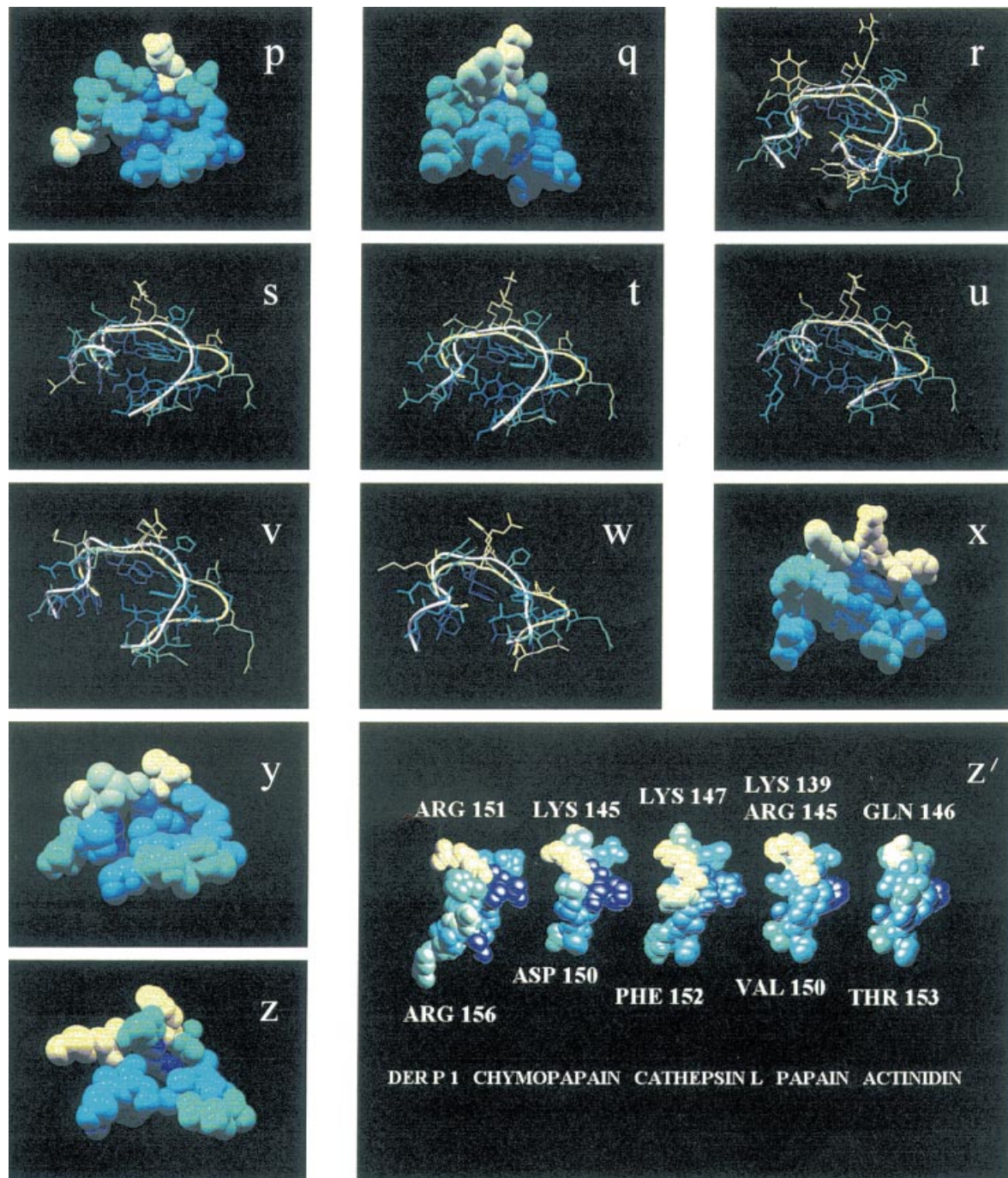


Fig. 1. The shape of the Der p 1 epitope (Leu147-Gln160) recognized by mAb 2C7 [6] (a) and those of the corresponding part in chymopapain (b), papain (c), actinidin (d), Der f 1 (f), Eur m 1 (g), human cathepsins L (i), K (j), S (k) and O (m), Schistosoma cysteine proteases Q11003 (o), Q26534 (p) and cathepsin L (q), Entamoeba cysteine protease ACP1 (x), cruzain (y) and cathepsin L-like Leishmania cysteine protease (z); amino acids constituting the epitope are coloured according to their solvent accessibility (decreasing in the order yellow, green, light blue and dark blue). Superimposition of the main chain of the Der p 1 epitope (yellow ribbon) on those (white ribbons) of human cathepsin O (l), Schistosoma cysteine proteases Q11003 (r), Q26534 (s) and cathepsin L (t), Entamoeba cysteine protease ACP1 (u), cruzain (v) and cathepsin L-like Leishmania cysteine protease (w); those of Der f 1 and Eur m 1 (h) are shown in different colours (magenta and red, respectively). The main chain of the Der p 1 epitope (the yellow part of the red ribbon) cannot be superimposed on that of human cathepsin B (white ribbon) (n). Superimposition of the Der p 1 epitope (yellow) on those of the plant cysteine proteases chymopapain (red), papain (blue) and actinidin (white) is shown (e). The reactivity of the cysteine proteases in ELISA inhibition (see Fig. 2) directly reflects the shape of their epitopes compared to that of Der p 1 (chymopapain > cathepsin L > papain > actinidin), which can best be viewed sideways (z').

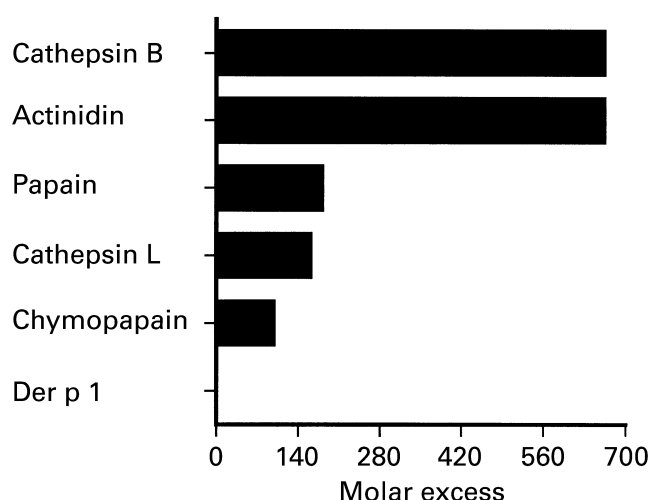


Fig. 2. The ability of anti-Der p 1 mAb 2C7 to recognize chymopapain, papain, actinidin and human cathepsins B and L was assessed by ELISA inhibition using Der p 1 coated microtitre plates, and results were compared with that obtained using a concentration of Der p 1 giving 50% inhibition ($0.6 \mu\text{g/mL}$). To achieve this degree of inhibition, 100-fold molar excess of chymopapain, 163-fold molar excess of human cathepsin L and 183-fold molar excess of papain were needed. Actinidin and human cathepsin B showed no inhibition even when tested at about 700-fold molar excess.

cathepsins B and L, which showed that the degree of cross-reactivity was directly related to the shape of the accessible side chains, rather than the charge distribution of the residues, comprising the epitope.

The identification of similarity in shape between such common putative IgE epitopes should help our understanding of the molecular basis for cross-reactivity amongst cysteine proteases of the papain family. For example, it appears that the core of the epitope on these cysteine proteases consists of a centrally located conserved Tyr residue flanked on either sides by accessible amino acids forming the motif R/QHYDGR in mite cysteine proteases and Q/F/ILYK/R in chymopapain, papain, human cathepsin S, *Schistosoma* cysteine protease Q26534, *Schistosoma* cathepsin L and *Entamoeba* cysteine protease ACP1. Indeed, looking at the backbone of the corresponding epitopes, this particular region represents the part that can most easily be superimposed (with the exception of human cathepsin B).

The novelty of the present study is the use of comparative molecular modelling to determine the molecular basis of immunological cross-reactivity between cysteine proteases of the papain family. Given the demonstration that these cysteine proteases use similar accessible structures, our approach may form the basis for the rational design of generic epitope-directed treatment strategies for controlling allergic diseases. Clearly, the future availability of crystal structures of other allergens will contribute to the further refinement of this approach.

Table 1. Cysteine protease sequences corresponding to the Der p 1 sequence recognized by anti-Der p 1 mAb 2C7 (Leu147-Gln160) [6]; solvent accessible residues are shown in bold.

Protease	Source	Acc/PDB code	aa	Homologous sequence
Der p 1	mite	P08176	147–160	L D A F R H Y D G R T I I Q
Chymopapain	plant	PDB:1YAL	138–151	G K P F Q L Y K S G V F D G
Papain	plant	PDB:1PPN	138–151	G K D F Q L Y R G G I F V G
Actinidin	plant	PDB:2ACT	141–154	G D A F K H Y S S G I F T G
Der f 1	mite	P16311	138–151	L R A F Q H Y D G R T I I Q
Eur m 1	mite	P25780	138–151	L N A F R H Y D G R T I M Q
Cathepsin L	human	PDB:1CJL	140–153	H E S F L F Y K E G I Y F E
Cathepsin K	human	PDB:1ATK	139–152	L T S F Q F Y S K G V Y Y D
Cathepsin S	human	PDB:1BXF	142–155	H P S F F L Y R S G V Y Y E
Cathepsin O	human	P43234	144–157	A V S W Q D Y L G G I I Q H
Cathepsin B	human	PDB:1CSB	177–190	Y S D F L L Y K S G V Y Q H
<i>Schistosoma</i> protease	parasite	Q26534	139–152	L D S L I L Y K S G I Y E S
<i>Schistosoma</i> protease	parasite	Q11003	136–152	A L L L Q F Y R H G I S H P
<i>Schistosoma</i> cathepsin L	parasite	Q26564	144–157	L D D L I L Y K S G I Y E S
<i>Entamoeba</i> protease	parasite	P36184	135–149	R P S F Q L Y K K G W I Y S
Cruzain	parasite	PDB:1AIM	138–151	A S S W M T Y T G G V M T S
<i>Leishmania</i> protease	parasite	PDB:1BMJ	142–155	A S S F M S Y H S G V L W S

Acknowledgements

RF is supported by a University of Nottingham Postgraduate Scholarship and a Jack and Pat Mallabar Charitable Foundation grant. For enquiries concerning the co-ordinates of the molecular models please contact Dr Ruta Furmonaviciene (E-mail: nixrf@nottingham.ac.uk).

References

- 1 Chapman MD, Platts-Mills TAE. Purification and characterisation of the major allergen from *Dermatophagoides pteronyssinus*-antigen P1. *J Immunol* 1980; 125:587–96.
- 2 van der Zee JS, van Swieten P, Jansen HM, Aalberse RC. Skin tests and histamine release with P1-depleted *Dermatophagoides pteronyssinus* body extracts and purified P1. *J Allergy Clin Immunol* 1988; 8:884–96.
- 3 Chapman MD, Rowntree SR, Di Prisco De Fuenmajor MS *et al.* Quantitative assessments of IgG and IgE antibodies to inhalant allergens in patients with atopic dermatitis. *J Allergy Clin Immunol* 1983; 72:27–33.
- 4 Krillis S, Baldo BA, Sutton R, Basten A. Antigens and allergens from the common house dust mite *Dermatophagoides pteronyssinus*. I. Demonstration of multiple allergens by immunochemical and biological analysis. *J Allergy Clin Immunol* 1984; 74:132–41.
- 5 McElveen JE, Clark MR, Smith SJ *et al.* Primary sequence and molecular model of the variable region of a mouse monoclonal anti-Der p 1 antibody showing a similar epitope specificity as human IgE. *Clin Exp Allergy* 1998; 18:371–9.
- 6 Furmonaviciene R, Tighe PJ, Clark MR *et al.* The use of phage-peptide libraries to define the epitope specificity of a mouse monoclonal anti-Der p 1 antibody representative of a major component of the human immunoglobulin E anti-Der p 1 Response. *Clin Exp Allergy* 1999; 29:1563–71.
- 7 Mohapatra SS. Structural motifs as a basis for cross-reactivity among pollen allergens. In: *Molecular Biology and Immunology of Allergens* (eds Kraft D, Sehon A). London: CRC Press, 1993: 69–81.
- 8 Metzler WJ, Mueller L. Structural considerations of antigen recognition: three-dimensional structure determination of the ragweed allergens Amb t 5 and Amb a 5 by nucleic magnetic resonance. In: *Molecular Biology and Immunology of Allergens* (eds Kraft D, Sehon A). London: CRC Press, 1993: 53–62.
- 9 Topham CM, Srinivasan N, Thorpe CJ *et al.* Comparative modelling of major house dust mite allergen Der p 1: Structure validation using an extended environmental amino acid propensity table. *Protein Eng* 1994; 7:869–94.
- 10 Harris GW, Pickersgill RW, Howlin B, Moss DS. The segmented anisotropic refinement of monoclinic papain by the application of the rigid-body TLS model and comparison to bovine ribonuclease A. *Acta Crystallogr* 1992; 48:67–75.
- 11 Maes D, Bouckaert J, Poortmans F *et al.* Structure of chymopapain at 1.7 angstroms resolution. *Biochemistry* 1996; 35:16292–8.
- 12 Baker EN. Structure of actinidin, after refinement at 1.7 Å resolution. *J Mol Biol* 1980; 141:441–84.
- 13 Gillmor SA, Craik CS, Fletterick RJ. Structural determinants of specificity in the cysteine protease cruzain. *Protein Sci* 1997; 6:1603–11.
- 14 Turk D, Podobnik M, Popovic T *et al.* Crystal structure of cathepsin B inhibited with Ca 030 at 2.1 angstroms resolution: a basis for the design of specific epoxysuccinyl inhibitors. *Biochemistry* 1995; 34:4791–7.
- 15 Thao B, Janson CA, Amegadzie BY *et al.* Crystal structure of human osteoclast cathepsin K complex with E-64. *Nat Struct Biol* 1997; 4:109–11.
- 16 Fengler A, Brandt W. Three-dimensional structures of the cysteine proteases cathepsins K and S deduced by knowledge based modelling and active-site characteristics. *Prot Eng* 1998; 11:1007–13.
- 17 Selzer PM, Chen X, Chan VJ *et al.* Leishmania major: molecular modelling of cysteine proteases and prediction of new nonpeptide inhibitors. *Exp Parasitol* 1997; 87:212–21.
- 18 Gal S, Gottesman MM. Isolation and sequence of a cDNA for human pro-cahepsin L. *Biochem J* 1988; 253:303–6.
- 19 Day SR, Dalton JP, Clough KA *et al.* Characterisation and cloning of the cathepsin L proteinases of *Schistosoma japonicum*. *Biochem Biophys Res Commun* 1995; 217:1–9.
- 20 Michael A, Ghoneim H, Resto M *et al.* Sequence, characterisation and isolation of a cysteine proteinase cathepsin L in *Schistosoma mansoni*. *Mol Biochem Parasitol* 1995; 73:7–18.
- 21 Tannich E, Scholze H, Nickel R, Horstmann RD. Homologous cysteine proteinases of pathogenic and nonpathogenic *Entamoeba histolytica*. *J Biol Chem* 1991; 266:4798–803.
- 22 Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb viewer: an environment for comparative protein modelling. *Electrophoresis* 1997; 18:2714–23.
- 23 Peitsch MC. ProMod and Swiss-Model: internet-based tools for automated comparative protein modelling. *Biochem Soc Trans* 1996; 24:274–9.
- 24 Schulz O, Sewell HF, Shakib F. A sensitive fluorescent assay for measuring the cysteine protease activity of Der p 1, a major allergen from the dust mite *Dermatophagoides pteronyssinus*. *J Clin Pathol: Mol Pathol* 1998; 51:222–4.