Available online at www.sciencedirect.com

ScienceDirect

Mendeleev Commun., 2023, **33**, 796–798

Mendeleev Communications

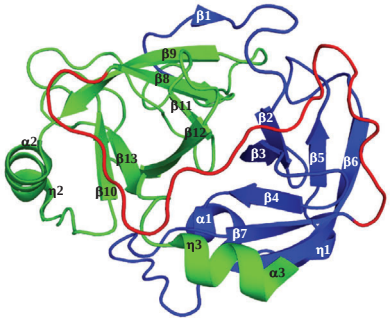
X-ray structure of recombinant house dust mite allergen Der p 3

Vladimir I. Timofeev,*[a] Yulia A. Abramchik,[b] Mikhail B. Shevtsov,[c] Maria A. Kostromina,[b] Sergey K. Zavriev,[b] Evgeny A. Zayats,[b] Roman S. Esipov[b] and Inna P. Kuranova[a]

[a] *A. V. Shubnikov Institute of Crystallography, FSRC 'Crystallography and Photonics', Russian Academy of Sciences, 119333 Moscow, Russian Federation. E-mail: tostars@mail.ru*
[b] *M. M. Shemyakin–Yu. A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 117997 Moscow, Russian Federation*
[c] *Institute of Physics and Technology, 141700 Dolgoprudny, Moscow Region, Russian Federation*

DOI: 10.1016/j.mencom.2023.10.019

The X-ray structure of the Der p 3 allergen was solved at 2.25 Å resolution using the molecular replacement method. The most remarkable difference in conformation of the polypeptide chains between the Der p 3 allergen and bovine trypsin is observed in the disordered parts of the polypeptide chain where the degree of homology is lower. The active site of the Der p 3 allergen which involves the residues of catalytic triad Ser197, His 52, and Asp97 (Ser195, His57, and Asp102 in bovine trypsin) responsible for the specific binding of the positively charged substrates Asp191 (Asp189 in bovine trypsin) is in the cleft between the domains.



Keywords: allergen, X-ray structure, mite, *Dermatophagoides pteronyssinus*, crystal structure, serine protease.

House dust mite (HDM) allergens are one of the major sensitization factors in the development of allergy and asthma worldwide.[1,2] Such allergens are divided into distinct groups based on their biochemical composition, molecular weight and homologous sequences.[3] An important role in the allergy play Der p serine proteases, which belong to the main allergens of house dust mite *Dermatophagoides pteronyssinus*.[4,5] The key allergen of these groups, Der p 3 protease, according to amino acid sequence homology, belongs to the trypsin family of serine proteases.[5] It is known that Der p 3 proteases interact with the respiratory epithelium.[6,7] The Der p 3 serine protease takes part in the stimulation of Orai1 channels. It has been shown[8] that therapy involving simultaneous inhibition of the Der p 3 protein and Orai1 channels suppressed the activation of mast cells by house dust mites. The main problem in diagnosing of allergy is the search for epitopes to predict allergenic activity.[9,10] The arrangement of the antigenic regions of proteins is necessary for the creation of synthetic vaccines, immunodiagnostic tests and antibody production. Knowledge of the spatial structure of target proteins contributes to the development of more effective anti-allergic drugs.

Here, a three-dimensional structure of the recombinant Der p 3 allergen was determined at 2.25 Å resolution by molecular replacement. The cloning, expression and purification of the recombinant Der p 3 allergen is described elsewhere.[11] Suitable for X-ray study[†] crystals of the recombinant protein have been grown by hanging-drop vapor-diffusion technique from a protein solution in sodium citrate buffer using ammonium sulfate as a precipitant.[11] Amino acid sequence alignment of the allergen molecule with two other proteins of the trypsin family, bovine trypsin (PDB ID: 4I8H) and human mesotrypsin (PDB ID: 3P95), as well as the elements of a secondary structure of the molecules are presented in Figure 1.

It is known that that Der p 3 protease of house dust mite *Dermatophagoides pteronyssinus* is synthesized as a zymogen and contains a propeptide sequence at its N-terminal.[15] The sequence homology between the Der p 3 allergen and bovine trypsin and between Der p 3 and human mesotrypsin is equal to

Table 1 The refinement statistics of the structure of the dust mite allergen.

PDB ID	7PZO
Resolution range/Å	30.0–2.25
R(cryst)	0.203
R(free)	0.279
RMS	
Bonds/Å	0.008
Angles/deg	1.752
Ramachandran plot	
Favoured (%)	94
Allowed (%)	6
Outliers (%)	0

[†] The set of X-ray diffraction data to 2.25 Å resolution was collected at 100 K at synchrotron ESRF (France), station ID23-1. The crystals belonging to the space group *C*121 had unit cell parameters: $a = 134.66$, $b = 44.65$ and $c = 72.50$ Å; $\alpha = \gamma = 90°$, $\beta = 102.82°$ and contained two enzyme monomers per asymmetric unit. The structure was solved by molecular replacement using BALBES Program.[12] Coordinates of catalytically inactive S195A mutant human mesotrypsin complexed with bovine pancreatic trypsin inhibitor mutant BPTI-K15R/R17D (PDB_ID: 3P95)[13] were used as a starting model. The structure was refined to 2.25 Å using REFMAC.[14] The refinement statistics is presented in Table 1.

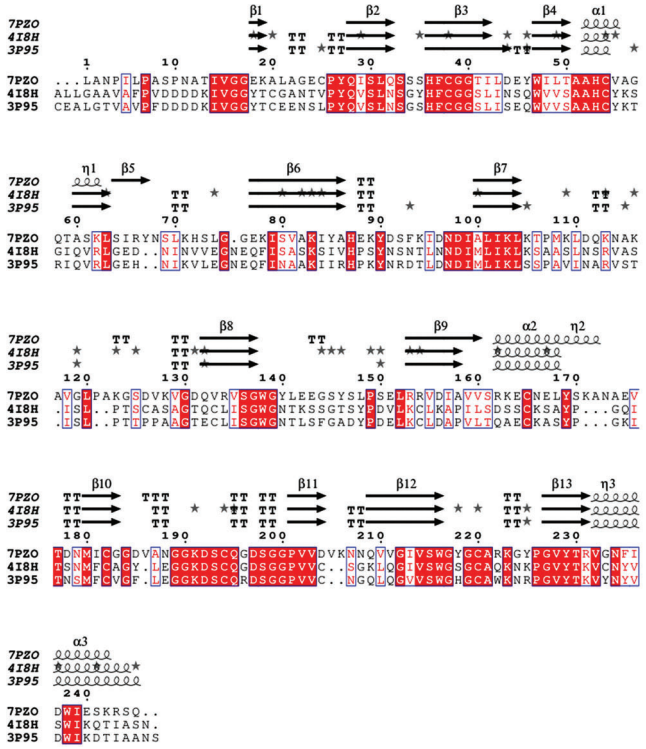


Figure 1 Amino acid sequence alignment of the Der p 3 allergen (PDB ID: 7PZO), bovine trypsin (PDB ID: 4I8B), and human mesotrypsin (PDB ID: 3P95). The elements of the secondary structure are shown as helices and arrows above the alignment. The identical amino acids are colored red.

36.8 and 31.1%, respectively, whereas the degree of the sequence homology between bovine trypsin and human mesotrypsin is higher and reaches 56.2%.[16,17] The superposition on Cα-atoms of Der p 3 allergen and bovine trypsin (PDB ID: 4I8H) molecules gives an RMSD equal to 0.578 Å. This confirms the similarity of the polypeptide chain folding in the Der p 3 allergen molecule and other typical representative of the trypsin family. The superposition on Cα atoms of two enzyme molecules situated in the asymmetric unit shows an RMSD of 0.189 Å. The fold of the polypeptide chain in the allergen molecule and elements of the secondary structure of the molecule are shown in Figure 2.

As in other proteins of the trypsin family, the polypeptide chain in the allergen molecule forms two domains: the N-terminal domain includes the amino acid residues Ile**13**–Leu**103**, and the C-terminal domain includes residues Gln131–Thr242. Long disordered loop Leu104–Asp130 connects the strands β7 and β8 of both domains. Each domain contains a β-barrel of seven (N-terminal) and six (C-terminal) β-strands, respectively. Residues **60**–**64** in the N-terminal domain form η1 helices

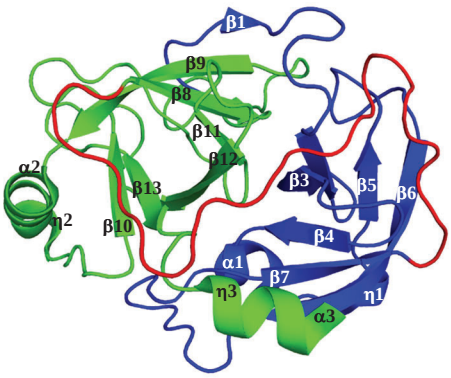


Figure 2 Polypeptide fold in the molecule of the Der p 3 allergen. The N-terminal domain is colored blue, the C-terminal domain is colored green, and the intermediate loop is colored red. The elements of the secondary structure are denoted by Greek letters.

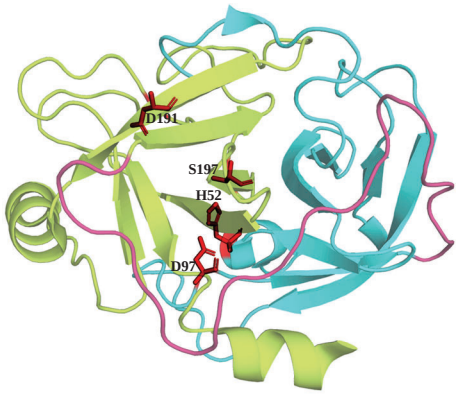


Figure 3 Residues of the catalytic triad Ser 197, His 52, Asp 97, and Asp191, responsible for the binding of a positively charged substrate, in the active site of the Der p 3 molecule.

instead of strands in other trypsins. Situated in the C-terminal domain, α2η2 and η3α3 helices are shorter in an allergen molecule compared to other trypsins. There are three disulfide bridges in the allergen molecule: one, Cys**37**–Cys**53**, in the N-terminal and two, Cys164–Cys181 and Cys193–Cys219, in the C-terminal domain. They occupy the same space positions as in trypsins from other sources. However, most proteins of the trypsin family contain six disulfide bridges.

The most remarkable difference in conformation of the polypeptide chains between the Der p 3 allergen and bovine trypsin is observed in the disordered parts of the polypeptide chain where the degree of homology is lower. Some homologous loops have a different length in both molecules. Compared with bovine trypsin, the Der p 3 allergen molecule contains an additional η helix (Ala59–Leu62), and the loop between β5 and β6 strands is shorter than in bovine trypsin. The long loop connecting two domains has different conformations in trypsins from different sources. The difference in the conformation of some flexible loops is found even between two allegen molecules situated in the asymmetric unit (residues 141–146, 184–187, and 220–226).

The active site of the Der p 3 allergen which involves the residues of the catalytic triad Ser197, His 52, and Asp97 (Ser195, His57, and Asp102 in bovine trypsin) responsible for the specific binding of the positively charged substrates Asp191 (Asp189 in bovine trypsin) is in the cleft between the domains (Figure 3).

The segment 193–199 of long curved loop 182–200 between strands β10 and β11, comprising catalytic Ser197 and Asp 191, restricts the active site area from the side of the C-domain. Two other residues of the catalytic triad, His52 and Asp97, are in the N-terminal domain. The helix α1, containing His52 along with the loop β6β7 (residues 85–99) containing Asp97, limits the active site area from the side of the N-terminal domain. It should be noted that not only position of the residues of the catalytic triad, but also conformation of their side chains is similar in both proteins.

The superposition on Cα atoms of the Der p 3 molecule and the bovine trypsin molecule complexed with inhibitor benzamidine (PDB ID 4I8H)[16] allowed us to compare the ligand binding areas in both enzymes. It was found that the nearest surrounding of the bound ligand is similar in both enzymes.

In conclusion, the obtained data will help to clarify the mechanism of allergy occurrence and the possibility of allergy therapy.

This work was performed within the State assignment of Federal Scientific Research Center ‘Crystallography and Photonics’ of the Russian Academy of Sciences.

References

1 P. Demoly and P. J. Bousquet, *Allergy*, 2008, **63**, 251.
2 T. A. Platts-Mills, D. Vervloet, W. R. Thomas, R. C. Aalberse and M. D. Chapman, *J. Allergy Clin. Immunol.*, 1997, **100**, S2.
3 W. R. Thomas, W.-A. Smith and B. J. Hales, *Chang Gung Med. J.*, 2004, **27**, 563.
4 L. G. Gregory and C. M. Lloyd, *Trends Immunol*., 2011, **32**, 402.
5 G. A. Stuart, L. D. Ward, R. J. Simpson and P. J. Thompson, *Immunology*, 1992, **75**, 29.
6 C. King, S. Brennan, P. J. Thompson and G. A. Stewart, *J. Immunol.*, 1998, **161**, 3645.
7 G. Sun, M. A. Stacey, M. Schmidt, L. Mori and S. Mattoli, *J. Immunol.*, 2001, **167**, 1014.
8 Y.-P. Lin, C. Nelson, H. Kramer and A. B. Parekh, *Mol. Cell.*, 2018, **70**, 228.
9 B. J. Hales, A. C. Martin, L. J. Pearce, I. A. Laing, C. M. Hayden, J. Goldblatt, P. N. Le Souëf and W. R. Thomas, *J. Allergy Clin. Immunol.*, 2006, **118**, 361.
10 O. Ivanciuc, C. H. Schein and W. Braun, *Bioinformatics*, 2002, **18**, 1358.
11 V. I. Timofeev, Yu. A. Abramchik, N. E. Zhukhlistova, O. O. Mikheeva, M. B. Shevtsov, E. A. Zayats, D. D. Lykoshin, M. A. Kostromina, R. S. Esipov and I. P. Kuranova, *Crystallogr. Rep.*, 2023, **68**, 52 (*Kristallografiya*, 2023, **68**, 46).
12 F. Long, A. Vagin, P. Young and G. N. Murshudov, *Acta Crystallogr.*, 2008, **D64**, 125.
13 M. A. Salameh, A. S. Soares, A. Hockla, D. C. Radisky and E. S. Radisky, *Biochem J.*, 2011, **440**, 95.
14 R. A. Nicholls, F. Long and G. N. Murshudov, *Acta Crystallogr.*, 2012, **D68**, 404.
15 M.-E. Dumez, N. Teller, F. Mercier, T. Tanaka, I. Vandenberghe, M. Vandenbranden, B. Devreese, A. Luxen, J.-M. Frère, A. Matagne, A. Jacquet, M. Galleni and A. Chevigné, *J. Biol. Chem.*, 2008, **283**, 30606.
16 *The PyMOL Molecular Graphics System, Version 1.8,* Schrödinger LLC, New York, 2015.
17 D. Liebschner, M. Dauter, A. Brzuszkiewicz and Z. Dauter, *Acta Crystallogr.*, 2013, **D69**, 1447.

Received: 23rd May 2023; Com. 23/7177