Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic analysis of the major horse allergen Equ c 1

Christophe Grégoire,[a] Gisele A. Tavares,[b] Hans K. Lorenzo,[b] Jean-Pierre Dandeu,[a] Bernard David[a] and Pedro M. Alzari[b]*

[a]Unité d'Immuno-Allergie, Institut Pasteur, 25 Rue du Dr Roux, 75724 Paris CEDEX 15, France, and [b]Unité de Biochimie Structurale, Institut Pasteur, 25 Rue du Dr Roux, 75724 Paris CEDEX 15, France

Correspondence e-mail: alzari@pasteur.fr

Received 22 May 1998
Accepted 23 November 1998

The secreted protein Equ c 1 is the major component responsible for the induction of specific IgE antibodies in patients sensitized to horse allergens. Equ c 1 belongs to the lipocalin superfamily of hydrophobic ligand-binding proteins, which also includes other known allergens. Equilibrium sedimentation and gel-filtration studies demonstrate that both the glycosylated form of Equ c 1 purified from horse salivary glands and the non-glycosylated recombinant form expressed in bacteria exist predominantly as dimers in solution. As observed for other dimeric lipocalins, acidic pH and low protein concentration favour dimer dissociation. The recombinant form of Equ c 1 has been crystallized using ammonium sulfate as a precipitant. The crystals belong to the tetragonal space group $P4_12_12$ with cell parameters $a = b = 84.0$, $c = 56.1$ Å, and contain a single molecule in the asymmetric unit. A complete data set from native crystals was collected at the synchrotron source in Hamburg to 2.9 Å resolution using a frozen crystal, and structure determination is in progress.

1. Introduction

Humans can develop IgE antibody responses owing to environmental allergens derived from different sources such as pollens, dust mites, animal danders, insects and fungi. The extracellular protein Equ c 1, one of the three important horse allergens, is present in the extract of horse hair and dander and is responsible for the production of specific anti-Equ c 1 IgE antibodies and subsequent development of asthma in atopic patients who are sensitized to horse allergens (Dandeu *et al.*, 1993).

Equ c 1 is a glycoprotein member of the lipocalin superfamily (Grégoire *et al.*, 1996). Lipocalins form a large group of small transport proteins involved in a variety of functions such as retinol, odoron and pheromone transport, invertebrate cryptic colouration and prostaglandin synthesis (for a review, see Flower, 1997). The three-dimensional structures of several lipocalins [retinol-binding protein (RBP), Newcomer *et al.*, 1984; β-lactoglobulin, Papiz *et al.*, 1986; Brownlow *et al.*, 1997; MUP, Bocskei *et al.*, 1992; and odorant-binding protein (OBP), Tegoni *et al.*, 1996; Bianchet *et al.*, 1996] revealed that these proteins share a common three-dimensional structure. The lipocalin fold essentially consists of an eight-stranded β-barrel formed by two perpendicular antiparallel sheets and followed by a C-terminal α-helix. A central hydrophobic cavity within the barrel encloses the ligand-binding site in RBP and presumably in other members of the family.

Equ c 1 is a single polypeptide with a relative molecular mass of 21.5 kDa and a p*I* of 3.9, as shown by SDS-polyacrylamide gel electrophoresis (SDS–PAGE) and isoeletric focusing PAGE (Dandeu *et al.*, 1995). The corresponding cDNA was cloned from the sublingual salivary gland and sequenced, and the non-glycosylated recombinant allergen was expressed in a bacterial system. A search of sequence databases revealed a homology with members of the lipocalin superfamily (Grégoire *et al.*, 1996). The highest similarity was found with the major urinary proteins cLac1 (MUP4), cSmx1 (MUP5), cLiv1 (MUP1) and cLiv6 (MUP2), all of which share 48–51% of identical amino acids with Equ c 1. The actual biological role of this allergen is unknown, although it might be related to the binding and transport of small hydrophobic molecules.

Several lipocalins, generally expressed in the liver and in secretory glands, may have allergenic properties (Arruda *et al.*, 1995; Flower, 1997). These include the mouse major urinary protein mMUP (Lorusso *et al.*, 1986), rat α-2-microglobulin (rA2U) (Walls & Longbottom, 1985), bovine β-lactoglobulin (Ball *et al.*, 1994), cockroach allergen Bla g4 (Arruda *et al.*, 1995) and the bovine dander allergen Bos d2 (Mantyjarvi *et al.*, 1996). At present it is unclear whether the allergenic properties of lipocalins arise from common shared IgE

© 1999 International Union of Crystallography
Printed in Denmark – all rights reserved

epitopes or whether these properties are related in some way with their primary functional role as transport proteins of small hydrophobic ligands. As a first step towards the understanding of the physiological role and antigenic (or allergenic) properties of Equ c 1, we report here the crystallization and preliminary crystallographic analysis of this major horse allergen.

2. Methods and results

2.1. Protein expression and purification of recombinant Equ c 1

Equ c 1 cloned from the horse cDNA sublingual salivary glands (SLG) was inserted in pET28a (Novagen) for expression in *Escherichia coli* (Grégoire *et al.*, 1996). This expression vector is under control of the T7 *lac* promoter and includes the kanamycin-resistance gene and a His6 tag at the N-terminus of the recombinant protein. Competent *E. coli* strain BL21 DE3 cells harbouring the plasmid pET28a with the SLG Equ c 1 cDNA insert were grown at 310 K in LB medium supplemented with 30 µg ml^{-1} kanamycin. The recombinant protein was expressed after induction by 1 m*M* IPTG at A_{600} = 3 and grown for 20 h at 310 K. Cells were disrupted by osmotic shock and the inclusion bodies pelleted by centrifugation for 1 h at 40000*g*. The pellet was resuspended and solubilized in 6 *M* guanidine–HCl solution and refolded by dialysis against the equilibrating buffer for immobilized metal-ion affinity chromatography (PBS, 0.5 *M* NaCl, 1 m*M* imidazole; Porath *et al.*, 1975). Specific elution of rSLG Equ c 1 was performed as described (Grégoire *et al.*, 1996). The fractions were concentrated to 10 mg ml^{-1} and equilibrated against the Factor Xa proteolysis buffer (50 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl and 1 mM $CaCl_2$) using stirred-cell ultrafiltration (Millipore) with a cut-off of 5000 Da. The His-tag proteolysis was performed overnight at room temperature with an enzyme:substrate ratio of 1:100(*w*/*w*). The N-terminal His-tag of rSLG Equ c 1 and Factor Xa enzyme were removed from the digest using a thyophilic chromatography purification (Goubran Boutros *et al.*, 1998). The digest product equilibrated with 20 m*M* phosphate buffer pH 7.0, 1 *M* $(NH_4)_2SO_4$ was loaded into the column and the absorbed proteins were eluted using a linear gradient of 1–0 *M* $(NH_4)_2SO_4$. The specific elution of rSLG Equ c 1 was obtained at 0.4 *M* $(NH_4)_2SO_4$. The fractions containing the rSLG Equ c 1 without His-tag were pooled, dialyzed against 50 m*M* Tris–HCl buffer pH 8.0, and concentrated at 6.0 mg ml^{-1} (assuming a specific extinction coefficient ε = 0.754 ml cm^{-1} mg^{-1}).

2.2. Evidence for dimer formation

The recombinant rSLG Equ c 1 protein, as well as the natural protein purified from hair and dander as described (Dandeu *et al.*, 1993), were subjected to equilibrium ultracentrifugation studies. The experiments were carried out at 293 K with a Beckman XL-A analytical ultracentrifuge using the protocol described by Kataeva *et al.* (1997), and working with protein concentrations of ~0.5 mg ml^{-1} in each case. A partial specific protein volume of 0.73 ml g^{-1} was assumed in all calculations. Fitting to the high concentration data with a single-component model indicated that both the natural and recombinant allergens are primarily found as a dimer under the tested conditions. The inclusion of a monomer component resulted in a slight improvement of the fit, suggesting that a dimer–monomer equilibrium could occur in solution. According to these experiments, the molecular weights of natural (glycosylated) and recombinant Equ c 1 monomers were found to be 23.5 and 20.3 kDa, respectively.

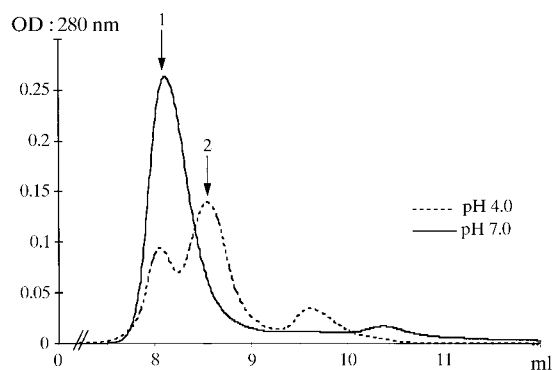

Figure 1
Elution profile of rSLG Equ c 1. The protein (100 µl at 1 mg ml^{-1}) equilibrated with 0.15 *M* NaCl, citrate/phosphate buffer at pH 7 or pH 4 was loaded into a TSK G3000SW (Ultropac) column pre-equilibrated with the same buffer. The elution peaks corresponding to the dimeric (1) and monomeric (2) species are indicated.

Gel-filtration chromatography experiments confirmed the above results. Both natural and recombinant Equ c 1 elute from the column (TSK G3000SW, Ultropac, equilibrated with 0.15 *M* NaCl, citrate/phosphate buffer) at the same time as size standards of comparable molecular weights (~45 kDa). Furthermore, these experiments showed that rSLG Equ c 1 dimers can be reversibly dissociated by lowering the protein concentration or by changing the buffer to acidic pH. As shown in Fig. 1, at neutral pH the protein elutes as a single peak at 8.1 ml, which corresponds to the dimeric species. However, at pH 4 a mixture of dimer (~8 ml) and monomer (8.6 ml) is observed, indicating that the protein dissociates at acidic pH. Independent gel-filtration experiments carried out using a Superdex 75 column on a SMART System (Pharmacia) confirmed that dilution of the protein also induces dimer dissociation (data not shown).

2.3. Crystallization and data collection

Crystals of Factor Xa-treated rSLG Equ c 1 were grown at 291 K using the hanging-drop vapour-diffusion technique. The best crystals were obtained when the protein solution at a concentration of 6 mg ml^{-1} was mixed with an equal volume of the reservoir solution containing 2 *M* $(NH_4)_2SO_4$ and 0.1 *M* Tris–HCl buffer pH 8.5. Tetragonal bipyramidal crystals appeared overnight and grew to full size (~0.3 × 0.3 × 0.5 mm) in a few days. Similar crystallization trials were also carried out for the natural Equ c 1 protein. However, these attempts have been unsuccessful, even after seeding the drops with crystals of rSLG Equ c 1, thus suggesting that protein micro-heterogeneity and/or protein glycosylation interfere with the intermolecular interactions necessary to stabilize the crystal.

Preliminary diffraction analysis was carried out using a MAR Research image-plate scanner mounted on a Rigaku RU-200 rotating-anode generator. The crystals diffract X-rays to 3 Å resolution and belong to the tetragonal space group $P4_12_12$ (or $P4_32_12$) with cell dimensions *a* = *b* = 84.0, *c* = 56.1 Å. Assuming one molecule in the asymmetric unit, the specific crystal volume is V_m = 2.47 Å^3 Da^{-1}, corresponding to a solvent content of ~50%, which falls within the normal range observed in most protein crystals (Matthews, 1968). A complete diffraction data set to 2.9 Å resolution has been collected from a flash-frozen crystal at 110 K using synchrotron radiation at EMBL/DESY in Hamburg (beamline BW7A) and processed with programs *DENZO* and *SCALEPACK* (Otwinowski, 1993). A total of 76026 measured intensities was reduced to 4647 unique reflections (99.9% complete to 2.9 Å resolution), with an overall R_{merge} value of 8.9% (30.1% in the resolution range 2.97–2.9 Å). Preliminary molecular-replacement calculations were carried out with the

program *AMoRe* (Navaza, 1994) using diffraction data between 10 and 4 Å resolution and the atomic coordinates of the mouse MUP1 protein (Bocskei *et al.*, 1992; PDB code 1MUP) as a search model. Assuming space group $P4_12_12$, one clear solution was obtained with a correlation coefficient of 38.9% (next highest value 28.4%) and a crystallographic R factor of 48.7%. Optimization of crystal-growth conditions to achieve higher resolution, co-crystallization trials with potential specific ligands and structure determination are currently in progress.

This work was supported by grants from Institut Pasteur, the Association pour la Recherche contre le Cancer (ARC), France and the European Union for the work at EMBL Hamburg through the HCMP Access to Large Installations Project (Contract Number CHGE-CT93-0040). The authors would like to thank M. E. Goldberg for performing the equilibrium sedimentation analysis.

References

Arruda, L. K., Vailes, L. D., Hayden, M. L., Benjamin, D. C. & Chapman, M. D. (1995). *J. Biol. Chem.* **270**, 31196–31201.

Ball, G., Shelton, M. J., Walsh, B. J., Hill, D. J, Hosking, C. S. & Howden, M. E. (1994). *Clin. Exp. Allergy*, **24**, 758–764.

Bianchet, M., Bains, G., Pelosi, P., Pevsner, J., Snyder, S. H., Monaco, H. L. & Amzel, L. M. (1996). *Nature Struct. Biol.* **3**, 934–939.

Bocskei, Z., Groom, C. R., Flower, D. R., Wright, C. E., Phillips, S. E., Cavaggioni, A., Findlay, J. B. & North, A. C. (1992). *Nature (London)*, **360**, 186–188.

Brownlow, S., Morais Cabral, J. H., Cooper, R., Flower, D. R., Yewdal, S. J., Polikarpov, I., North, A. C. T. & Sawyer, L. (1997). *Structure*, **5**, 481–495.

Dandeu, J. P., Rabillon, J., Carmi-Leroy, A., Divanovic, A., Carmoin, L. & David, B. (1995). *J. Allergy Clin. Immunol.* **95**, 348.

Dandeu, J. P., Rabillon, J., Divanovic, A., Carmi-Leroy, A. & David, B. (1993). *J. Chromatogr.* **621**, 23–31.

Flower, D. R. (1997). *Biochem. J.* **318**, 1–14.

Goubran Boutros, H., Rabillon, J., Grégoire, C., David, B. & Dandeu, J.-P. (1998). *J. Chromatogr. B*, **710**, 57–65.

Grégoire, C., Rosinski-Chupin, I., Rabillon, J., Alzari, P. M., David, B. & Dandeu, J. P. (1996). *J. Biol. Chem.* **271**, 32951–32959.

Kataeva, I., Guglielmi, G. & Béguin, P. (1997). *Biochem. J.* **326**, 617–624.

Lorusso, J. R., Moffat, S. & Ohman, J. L. (1986). *J. Allergy Clin. Immunol.* **78**, 928–937.

Mantyjarvi, R. A., Rytkonen, M., Pentikainen, J., Rautiainen, J., Virtanen, T., Santa, H. & Laatikainen, R. (1996). *J. Allergy Clin. Immunol.* **97**, 212.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Newcomer, M. E., Jones, T. A., Aqvist, J., Sundelin, J., Eriksson, U., Rask, L. & Peterson, P. A. (1984). *EMBO J.* **3**, 1451–1454.

Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Papiz, M. Z., Sawyer, L., Eliopoulos, E. E., North, A. C., Findlay, J. B., Sivaprasadarao, R., Jones, T. A., Newcomer, M. E. & Kraulis, P. J. (1986). *Nature (London)*, **324**, 383–385.

Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. (1975). *Nature (London)*, **258**, 598–599.

Tegoni, M., Ramoni, R., Bignetti, E., Spinelli, S. & Cambillau, C. (1996). *Nature Struct. Biol.* **3**, 863–867.

Walls, A. F. & Longbottom, J. L. (1985). *J. Allergy Clin. Immunol.* **75**, 242–251.