# crystallization communications

Complex assembly, crystallization and preliminary X-ray crystallographic studies of the swine major histocompatibility complex molecule SLA-1*1502

Xiaocheng Pan,[a] Jianxun Qi,[b,c] Nianzhi Zhang,[a] Qirun Li,[a] Chunsheng Yin,[a] Rong Chen,[a] Feng Gao[d] and Chun Xia[a,b]*

[a]Department of Microbiology and Immunology, College of Veterinary Medicine, China Agricultural University, Bejing 100094, People's Republic of China, [b]CAS Key Laboratory of Pathogenic Microbiology and Immunology (CASPMI), Institute of Microbiology, Chinese Academy of Sciences (CAS), Beijing 100101, People's Republic of China, [c]Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, People's Republic of China, and [d]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

Correspondence e-mail: xiachun@cau.edu.cn

In order to illustrate the structure of the swine MHC class I (SLA-I) molecule and to evaluate the cytotoxic T lymphocyte (CTL) response against porcine reproductive and respiratory syndrome virus (PRRSV), the ternary complex of the SLA-I molecule termed SLA-1*1502 with β_2-microglobulin and the CTL epitope TMPPGFELY (PRRSV-NSP9$_{TY9}$) derived from PRRSV nonstructural protein 9 (residues 198–206) was assembled and crystallized. The crystal diffracted X-rays to 2.2 Å resolution and belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 66.1$, $b = 74.1$, $c = 98.6$ Å; it contained one molecule in the asymmetric unit. The Matthews coefficient and the solvent content were calculated to be 2.74 $\mathrm{Å^3\,Da^{-1}}$ and 55.17%, respectively. The results will be helpful in obtaining insight into the structural basis of the presentation of viral epitopes by SLA-I.

1. Introduction

Major histocompatibility complex (MHC) class I molecules play an important role in the elimination of intracellular pathogens (Klein *et al.*, 1993). MHC class I molecules are membrane-surface proteins and their main function is defined as the binding and presentation of viral epitopes on the surface of virally infected cells. Subsequently, cytotoxic T lymphocytes (CTLs) first recognize the viral epitopes *via* specific T-cell receptors (TCRs) and then lyse the infected cells (Garboczi *et al.*, 1996). The MHC class I molecule consists of three components: a polymorphic heavy chain (α chain), a monomorphic light chain (β_2-microglobulin; β_2m) and an epitope of 8–11 amino-acid residues positioned in a cleft formed by the α1 and α2 domains of the heavy chain (Madden, 1995). Analysis of MHC class I structures is essential in order to understand the details of antigen presentation in CTL immunity. Swine MHC class I genes termed swine leukocyte antigen class I (SLA-I) are located in the 7p1.1 band of the short arm of chromosome 7 (Smith *et al.*, 1995). There are three constitutively expressed classical swine MHC class I genes in the genome, namely SLA-1, SLA-2 and SLA-3. As in humans and mice, MHC class I molecules and SLA-I molecules are critical to antiviral responses in swine (Ho *et al.*, 2009).

PRRSV is a single-stranded positive RNA virus of the Arteriviridae family which can result in reproductive failure in pregnant sows or respiratory complications. Porcine reproductive and respiratory syndrome (PRRS) emerged in North America in 1987 (Elazhary *et al.*, 1991) and the disease was subsequently found in Europe and Asia. Currently, PRRS is spreading worldwide in swine-cultivating countries and has resulted in devastating economic losses (Tian *et al.*, 2007). Previous research on PRRSV suggested that CTL immunity may play a central role in viral clearance (Lamontagne *et al.*, 2003). However, little is known about the three-dimensional structure of SLA-I and antiviral studies for viruses such as PRRSV in swine are limited. In this article, we report the refolding, purification and crystallization of SLA-I (SLA-1*1502) with a CTL epitope.
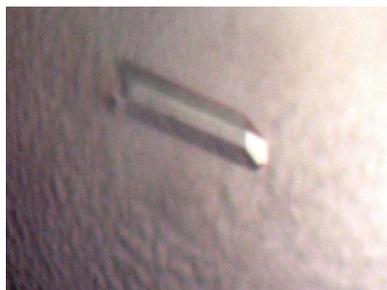
2. Materials and methods

2.1. Preparation of SLA-1*1502 and β_2m proteins

SLA-1*1502 heavy-chain cDNA (GenBank accession No. HQ909439, residues 1–275) was amplified by RT-PCR from a specific pathogen-free (SPF) swine. The SLA-1*1502 gene was ligated into the pET21a expression vector and transformed into *Escherichia coli* strain BL21 (DE3). The expression vector pET21a/swine β_2m (sβ_2m; residues 1–98; GenBank accession No. BAG32341) was constructed previously in our laboratory and was also transformed into *E. coli* strain BL21 (DE3). The recombinant proteins were both expressed as inclusion bodies (Zhang *et al.*, 2010). The bacteria were harvested and suspended in cold phosphate-buffered saline (PBS). After sonication, the samples were centrifuged at 16 000g and the pellets were washed three times with a solution consisting of 20 mM Tris–HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 1 mM DTT and 0.5% Triton X-100. SLA-1*1502 heavy chain and sβ_2m inclusion bodies were dissolved in guanidinium chloride (Gua–HCl) buffer [6 M Gua–HCl, 50 mM Tris–HCl pH 8.0, 10 mM EDTA, 100 mM NaCl, 10%(v/v) glycerine, 10 mM DTT].

2.2. Preparation of the SLA-1*1502 complex

The SLA-1*1502–PRRSV-NSP9$_{TY9}$ complex was prepared essentially as described previously by Garboczi *et al.* (1992) with modifications introduced in our laboratory (Zhang *et al.*, 2010). Firstly, the epitope PRRSV-NSP9$_{TY9}$ (TMPPGFELY) derived from PRRSV nonstructural protein 9 (PRRSV-NSP9) residues 198–206 was dissolved in dimethyl sulfoxide (DMSO). The PRRSV-NSP9$_{TY9}$ eptitope and the SLA-1*1502 and sβ_2m inclusion bodies were then refolded in a 3:1:1 molar ratio using the gradual solution method. After incubation for 24 h at 277 K, the soluble portion was concentrated and purified by chromatography on a Superdex 200 16/60 HiLoad (GE Healthcare) size-exclusion column followed by Resource Q (GE Healthcare) anion-exchange chromatography.

2.3. Crystallization of the SLA-1*1502 complex

The purified complex (~44 kDa) was dialyzed against crystallization buffer (20 mM Tris–HCl pH 8.0, 50 mM NaCl) and concentrated to 12 mg ml^{-1}. Crystallization trials were set up with Index and Crystal Screen (Hampton Research) at 277 and 291 K using the hanging-drop method (Zhang *et al.*, 2010). Two drops containing equal volumes of protein solution (at 6 and 12 mg ml^{-1}) and reservoir crystallization buffer (1 µl each) were placed over a well containing 200 ml reservoir solution using a VDX plate (HR3-142, Hampton Research). Using a protein concentration of 12 mg ml^{-1}, crystals were obtained in 10–14 d using Index solution No. 65 (0.1 M ammonium acetate, 0.1 M Bis-Tris pH 5.5, 17% PEG 10 000) at 277 K.

2.4. Data collection and processing

Diffraction data were collected to 2.2 Å resolution using an in-house X-ray source (Rigaku MicroMax-007 desktop rotating-anode X-ray generator with a Cu target operated at 40 kV and 30 mA) and an R-AXIS IV^{++} imaging-plate detector at a wavelength of 1.5418 Å. The crystal was first soaked in reservoir solution containing 25% glycerol as a cryoprotectant for several seconds and then flash-cooled in a stream of gaseous nitrogen at 100 K (Parkin & Hope, 1998). The collected intensities were indexed, integrated, corrected for absorption, scaled and merged using *HKL*-2000 (Otwinowski & Minor, 1997).

2.5. Alignment of SLA-1*1502 with human and mouse MHC class I molecules

The amino-acid sequence of SLA-1*1502 was aligned with those of typical SLA-1, SLA-2, SLA-3, human and mouse MHC class I molecules. The GenBank accession Nos. are as follows: SLA-1*1502, HQ909439; SLA-1*0101, FJ952938; SLA-2*0201, EU432086; SLA-3*0302, EU432097; HLA-A2, U56825; H-2Kb, V00746. Alignment was performed using the *DNAMAN* v.5.2.2 program (Lynnon Biosoft).

3. Results and discussion

Based on the *NetMHCpan* 2.0 server (http://www.cbs.dtu.dk/services/NetMHCpan; Hoof *et al.*, 2009), we synthesized a series of potential epitopes for the SLA-1*1502 allele and validated the CTL function of PRRSV-NSP9$_{TY9}$. SLA-1*1502 could be refolded in the presence of sβ_2m and PRRSV-NSP9$_{TY9}$. Refolding resulted in an approximately 12% yield of the correctly folded complex (~44 kDa), which could be purified to homogeneity by Superdex 200 16/60 HiLoad size-exclusion chromatography and Resource Q anion-exchange chromatography (Fig. 1). The chromatographic elution profile
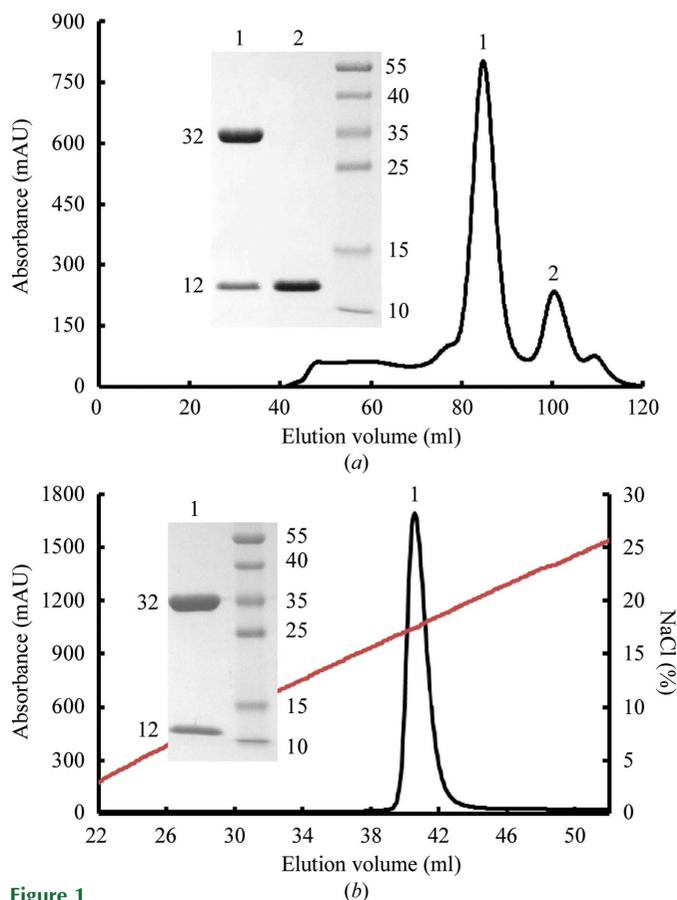
Figure 1
Purification of the refolded SLA-I complex (SLA-1*1502, sβ_2m and the epitope PRRSV-NSP9$_{TY9}$) by FPLC Superdex 200 16/60 Hi-Load gel-filtration and Resource Q anion-exchange chromatography (GE Healthcare). (*a*) Gel-filtration profile of the refolded products. Peak 1 represents the correctly refolded complex (~44 kDa) and peak 2 represents sβ_2m. Inset: reduced SDS–PAGE gel (15%) for peak 1 and peak 2. The right column contains molecular-weight markers (kDa). (*b*) Results of further purification of the refolded products by anion-exchange chromatography. Peak 1 represents the SLA-I complex, which was eluted at an NaCl concentration of 16.6–20.2%. Inset: reduced SDS–PAGE gel (15%) for peak 1.

Figure 2
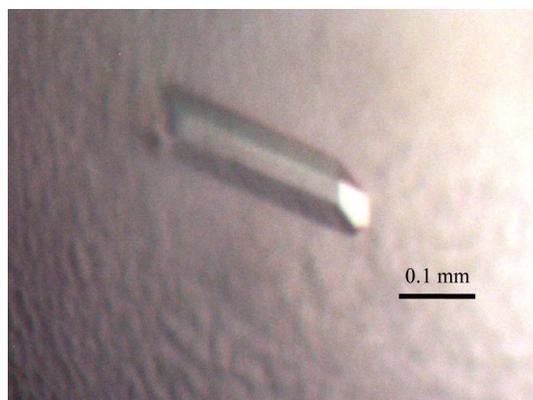Photograph of the crystal used in the diffraction analysis.

Table 1
X-ray diffraction data and processing statistics.

Values in parentheses are for the highest resolution shell.

Space group	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 66.1$, $b = 74.1$, $c = 98.6$
Resolution range (Å)	50.00–2.20 (2.28–2.20)
Total No. of reflections	197524
No. of unique reflections	24678
Completeness (%)	99.5 (98.9)
Average $I/\sigma(I)$	27.4 (7.5)
R_{merge}† (%)	8.3 (28.6)
Average multiplicity	7.9 (7.9)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity from multiple measurements.

showed two peaks corresponding to the refolded complex (peak 1, ∼44 kDa) and sβ_2m (peak 2) (Fig. 1a). The refolded complex was further purified by Resource Q anion-exchange chromatography and eluted at an NaCl concentration of 16.6–20.2% (Fig. 1b). The refolded complex was confirmed by SDS–PAGE (see inset in Fig. 1). SDS–PAGE analysis showed two bands corresponding to the expected molecular weights of SLA-1*1502 (∼32 kDa) and sβ_2m (∼12 kDa).

After purification and concentration, the SLA-1*1502–PRRSV-NSP9$_{TY9}$ complex was set up for crystal screening. A large single crystal (Fig. 2) appeared in 10–14 d under the initial conditions. The crystal belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 66.1$, $b = 74.1$, $c = 98.6$ Å, and diffracted to 2.2 Å resolution. The crystal was estimated to contain one molecule in the asymmetric unit and the Matthews coefficient value V_M was 2.74 Å^3 Da^{-1}, with a calculated solvent content of 55.17%. Selected data statistics are shown in Table 1.

A multiple amino-acid sequence alignment of SLA-1*1502 with other known swine, human and mouse class I MHCs is shown in Fig. 3. SLA-1*1502 has 92.36, 85.82 and 87.27% sequence identity to the other SLA-Is (SLA-1, SLA-2 and SLA-3, respectively) and has 72.73 and 66.18% sequence identity to human and mouse class I MHCs,
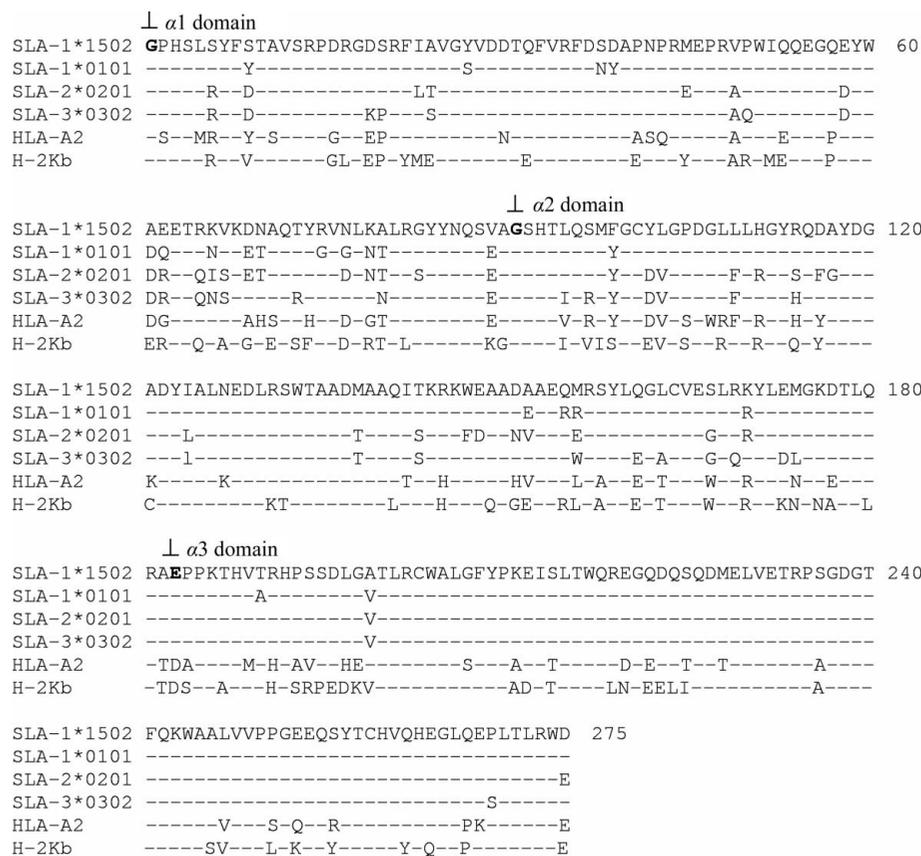
```
              ⊥ α1 domain
SLA-1*1502 GPHSLSYFSTAVSRPDRGDSRFIAVGYVDDTQFVRFDSDAPNPRMEPRVPWIQQEGQEYW  60
SLA-1*0101 --------Y-----------------S----------NY--------------------
SLA-2*0201 -----R--D-------------LT--------------------E---A--------D--
SLA-3*0302 -----R--D---------KP---S--------------------AQ------D--
HLA-A2     -S--MR--Y-S----G---EP---------N----------ASQ-----A---E---P---
H-2Kb      -----R--V------GL-EP-YME-------E--------E---Y---AR-ME---P---

                                        ⊥ α2 domain
SLA-1*1502 AEETRKVKDNAQTYRVNLKALRGYYNQSVAGSHTLQSMFGCYLGPDGLLLHGYRQDAYDG 120
SLA-1*0101 DQ---N--ET----G-G-NT-------E----------Y------------
SLA-2*0201 DR--QIS-ET------D-NT--S-----E---------Y--DV-----F-R--S-FG---
SLA-3*0302 DR--QNS-----R------N--------E-----I-R-Y--DV-----F----H------
HLA-A2     DG------AHS--H--D-GT--------E-----V-R-Y--DV-S-WRF-R--H-Y----
H-2Kb      ER--Q-A-G-E-SF--D-RT-L------KG----I-VIS--EV-S--R--R--Q-Y----

SLA-1*1502 ADYIALNEDLRSWTAADMAAQITKRKWEAADAAEQMRSYLQGLCVESLRKYLEMGKDTLQ 180
SLA-1*0101 ----------------------------E--RR------------R----------
SLA-2*0201 ---L-------------T----S---FD--NV---E----------G--R-------
SLA-3*0302 ---l-------------T----S---------W----E-A---G-Q---DL------
HLA-A2     K-----K-------------T--H-----HV---L-A--E-T---W--R---N--E---
H-2Kb      C---------KT-------L---H---Q-GE--RL-A--E-T---W--R--KN-NA--L

              ⊥ α3 domain
SLA-1*1502 RAEPPKTHVTRHPSSDLGATLRCWALGFYPKEISLTWQREGQDQSQDMELVETRPSGDGT 240
SLA-1*0101 ---------A--------V------------------------------
SLA-2*0201 ------------------V------------------------------
SLA-3*0302 ------------------V------------------------------
HLA-A2     -TDA----M-H-AV--HE--------S---A--T-----D-E--T--T-------A----
H-2Kb      -TDS--A---H-SRPEDKV----------AD-T---LN-EELI----------A----

SLA-1*1502 FQKWAALVVPPGEEQSYTCHVQHEGLQEPLTLRWD  275
SLA-1*0101 -----------------------------------
SLA-2*0201 ----------------------------------E
SLA-3*0302 ---------------------------S------
HLA-A2     -----V---S-Q--R---------PK------E
H-2Kb      -----SV---L-K--Y-----Y-Q--P------E
```

Figure 3
Amino-acid alignment of SLA-1*1502 and other MHC class I molecules. GenBank accession Nos. are as follows: SLA-1*1502, HQ909439; SLA-1*0101, FJ952938; SLA-2*0201, EU432086; SLA-3*0302, EU432097; HLA-A2, U56825; H-2Kb, V00746. '–' indicates identity with SLA-1*1502 and '⊥' indicates the starting locations of the α1, α2 and α3 domains.

respectively. The sequence identities show that there is an obvious species difference between the swine, human and mouse molecules.

References

Elazhary, Y., Weber, J., Bikour, H., Morin, M. & Girard, C. (1991). *Vet. Rec.* **129**, 495–496.
Garboczi, D. N., Ghosh, P., Utz, U., Fan, Q. R., Biddison, W. E. & Wiley, D. C. (1996). *Nature (London)*, **384**, 134–141.
Garboczi, D. N., Hung, D. T. & Wiley, D. C. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 3429–3433.
Ho, C.-S., Lunney, J. K., Ando, A., Rogel-Gaillard, C., Lee, J.-H., Schook, L. B. & Smith, D. M. (2009). *Tissue Antigens*, **73**, 307–315.
Hoof, I., Peters, B., Sidney, J., Pedersen, L. E., Sette, A., Lund, O., Buus, S. & Nielsen, M. (2009). *Immunogenetics*, **61**, 1–13.
Klein, J., Satta, Y., O'hUigin, C. & Takahata, N. (1993). *Annu. Rev. Immunol.* **11**, 269–295.
Lamontagne, L., Pagé, C., Larochelle, R. & Magar, R. (2003). *Viral Immunol.* **16**, 395–406.
Madden, D. R. (1995). *Annu. Rev. Immunol.* **13**, 587–622.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Parkin, S. & Hope, H. (1998). *J. Appl. Cryst.* **31**, 945–953.
Smith, T. P., Rohrer, G. A., Alexander, L. J., Troyer, D. L., Kirby-Dobbels, K. R., Janzen, M. A., Cornwell, D. L., Louis, C. F., Schook, L. B. & Beattie, C. W. (1995). *Genome Res.* **5**, 259–271.
Tian, K. *et al.* (2007). *PLoS One*, **2**, e526.
Zhang, J., Chen, Y., Gao, F., Chen, W., Qi, J. & Xia, C. (2010). *Acta Cryst.* F**66**, 99–101.