Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic analysis of *Mycoplasma arthritidis*-derived mitogen complexed with peptide/MHC class II antigen

Yiwei Zhao,[a] Zhong Li,[a] Sandra Drozd,[a] Yi Guo,[a] Robert Stack,[a] Charles Hauer[a,b] and Hongmin Li[a,b]*

[a]Wadsworth Center, New York State Department of Health, Empire State Plaza, PO Box 509, Albany, New York 12201-0509, USA, and [b]Department of Biomedical Sciences, School of Public Health, University at Albany, State University of New York, Empire State Plaza, PO Box 509, Albany, New York 12201-0509, USA

Correspondence e-mail: lih@wadsworth.org

Received 9 October 2003
Accepted 2 December 2003

Mycoplasma arthritidis-derived mitogen (MAM), a bacterial superantigen, has been crystallized in complex with its human receptor, major histocompatibility complex (MHC) class II antigen, by the hanging-drop vapor-diffusion method. Crystals were obtained under three conditions, with ammonium sulfate, phosphate salt and PEG 8000 as the precipitant. The crystals grown under these conditions all belong to space group $I222$, with the same unit-cell parameters: $a = 137.4$, $b = 178.2$, $c = 179.6$ Å. Diffraction data were collected to 3.3 and 3.4 Å resolution from crystals of native and selenomethionylated MAM–MHC complexes, respectively. Self- and cross-rotation function calculations suggest the presence of two complex molecules in the asymmetric unit, resulting in a V_M of 4.0 and a solvent content of 69%. An interpretable electron-density map was produced using a combination of molecular replacement and SAD phasing.

1. Introduction

Superantigens (SAgs) are functionally related immunoregulatory proteins generally produced by bacteria and viruses (Kappler *et al.*, 1988; Kotzin *et al.*, 1993; Li *et al.*, 1999; White *et al.*, 1989). By simultaneously binding to the T-cell antigen receptor (TCR) on T cells and to the major histocompatibility complex (MHC) antigen on antigen-presenting cells, SAgs can stimulate the activation of a large fraction (up to 20%) of T cells bearing particular TCR Vβ elements. In contrast, conventional peptide antigens can trigger T-cell activation only at the level of one per 10^5 cells. The recognition of SAgs by T cells may have several consequences for the responding T cells, including proliferation and expansion, the induction of nonresponsiveness (anergy) or even cell death (deletion) (MacDonald *et al.*, 1993). Owing to their unique features and biological effects, SAgs have been hypothesized to play important roles in a number of human diseases, including food poisoning, toxic shock syndrome and autoimmune diseases such as multiple sclerosis and rheumatoid arthritis (RA) (Abe *et al.*, 1992; Conrad *et al.*, 1994; Kotzin *et al.*, 1993; McCormick *et al.*, 2001; Renno & Acha-Orbea, 1996).

The best-characterized group of SAgs belongs to the pyrogenic toxin SAg family from *Staphylococcus aureus* and *Streptococcus pyogenes* (Choi *et al.*, 1990; Dellabona *et al.*, 1990; Marrack & Kappler, 1990). Some of these SAgs have been structurally characterized both in unbound forms and in complexes with the TCR and MHC molecules; the structures show apparent diversity in the interactions between SAgs and their immunoreceptors (Fraser *et al.*, 2000; Li *et al.*, 1999; Mitchell *et al.*, 2000; Papageorgiou & Acharya, 2000; Sundberg *et al.*, 2002; Swaminathan *et al.*, 1992). Among the superantigenic proteins of viral origin, only mouse mammary tumor virus-encoded SAgs have been elucidated in detail (Acha-Orbea & MacDonald, 1995). Other viral proteins have been reported to exhibit superantigenic activity (Li *et al.*, 1999), including proteins encoded by an endogenous human retrovirus reported recently (Stauffer *et al.*, 2001; Sutkowski *et al.*, 2001).

Mycoplasma arthritidis-derived mitogen (MAM) is produced by *M. arthritidis*, a member of the eubacterial class Mollicutes (Cole, 1991). MAM functions like a conventional SAg and can induce spontaneous chronic arthritis, which resembles human RA, in genetically susceptible strains of rodents (Cole & Atkin, 1991; Cole & Griffiths, 1993). Mature MAM is a basic thermostable and acid-labile protein of 213 amino acids (Cole *et al.*, 1996). Although the molecular weight of MAM is within the range of values seen for staphylococcal and streptococcal SAgs, the primary sequence of MAM does not exhibit significant global homology to any other SAg. Although MAM, like other SAgs, interacts with TCRs in a Vβ-restricted fashion, the binding of MAM to TCR is influenced by the CDR3 region of the TCR (Hodtsev *et al.*, 1998). This is in contrast to other SAgs, suggesting that MAM represents a new type of ligand for TCRs distinct from both conventional peptide antigens and other known SAgs. Here, we report the crystallization and preliminary crystallographic analysis of MAM complexed with its

human immunoreceptor, the class II MHC molecule, HLA-DR1, in the context of a haemagglutinin peptide (HA).

2. Material and methods

2.1. Expression and purification

Soluble MAM was overexpressed in *Escherichia coli* using the pGEX-6P-1 GST fusion protein expression system (Amersham Pharmacia) with a modified protocol (Langlois *et al.*, 2000). Briefly, *E. coli* BL21(DE3) cells transformed with the plasmid pWM.25B (Langlois *et al.*, 2000) were grown to an OD_{600} of 0.8 at 310 K with shaking in superbroth medium containing 0.1 mg ml^{-1} ampicillin. The temperature was reduced to 303 K and isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 m*M*. After 5 h induction, cells were harvested by centrifugation and resuspended in 100 ml phosphate-buffered saline (PBS) per litre of culture. After sonication of the resuspended cells, the lysate was clarified by centrifugation. The supernatant was loaded onto a glutathione Sepharose 4B (Amersham Pharmacia) column and purified according to the manufacturer's instructions. The purified MAM-GST fusion protein was digested overnight at 277 K with PreScission protease (Amersham Pharmacia) at a ratio of 2 U per milligram of fusion protein by dialysis against a protease buffer containing 50 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl, 1 m*M* EDTA and 1 m*M* DTT. The completeness of digestion was monitored by SDS–PAGE. MAM was separated from GST by passing twice through glutathione Sepharose 4B. The flowthrough was concentrated and applied to a Waters Protein-Pak 300SW column using a Waters 625 HPLC system (Waters Corporation) for further purification in a buffer containing 10 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl and 1 m*M* DTT. The purity and molecular mass were verified using electrospray mass spectrometry with a Finnigan-MAT TSQ-700 triple-quadrupole mass spectrometer equipped with an electrospray ionization source (San Jose, CA, USA) (Figs. 1*b* and 1*c*).

The HLA-DR1, loaded with a haemagglutinin peptide (HA) (PKYVKQNTLKLAT), was prepared from individually expressed inclusion bodies of the DR1 α and β chains using a previously described refolding protocol (Frayser *et al.*, 1999). To prepare the selenomethioine-substituted (SeMet) derivative, the plasmid pLM1-α encoding the α chain of HLA-DR1 was transformed into the methionine-auxotrophic *E. coli* strain B834 (Novagen). Cells were grown in minimal medium containing SeMet at 310 K. Purification of the inclusion body of the SeMet α chain was performed using the same method as that for the native α chain. Through use of the same protocol to produce the native protein, the purified inclusion body of the SeMet α chain was

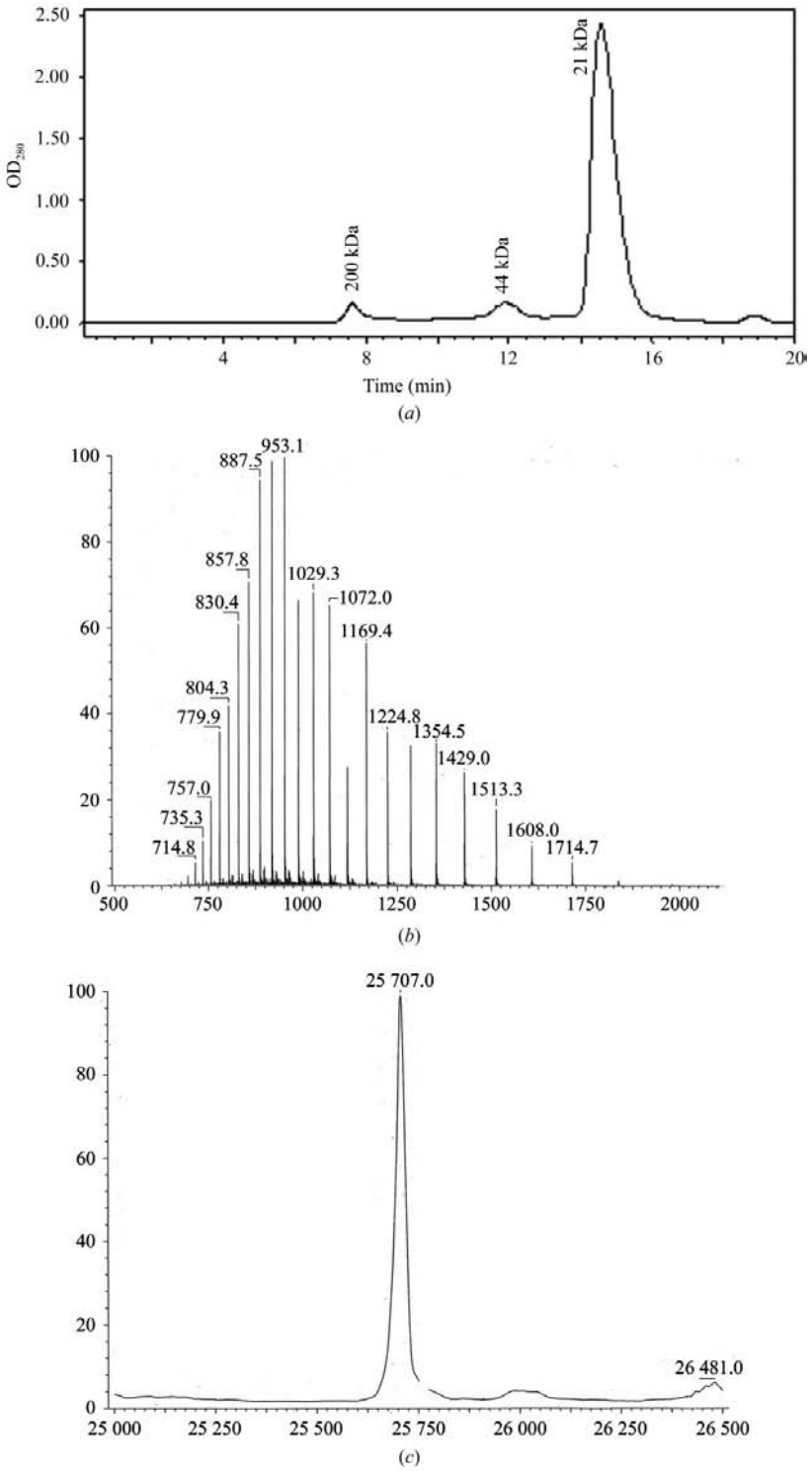

Figure 1
Profiles of chromatography and mass spectrometry of recombinant MAM. (*a*) Profile of gel-filtration chromatography of recombinant MAM on a Waters Protein Pak 300SW column. The positions were labeled for the Bio-Rad molecular-weight standard, which was run under the same conditions on the same column. (*b*) Electrospray ionization (ESI) mass spectra of recombinant MAM. (*c*) Deconvoluted ESI mass spectra of recombinant MAM, showing the molecular weight detected. The molecular weight calculated for the sequence is 25 704 Da.

refolded together with the native DR1 β chain in the presence of the HA peptide.

2.2. Crystallization

Purified MAM was concentrated to 8.8 mg ml^{-1} using a Millipore stirred cell. Native and SeMet HLA-DR1–HA complexes were buffer-exchanged to 20 m*M* HEPES pH 7.5, 100 m*M* NaCl and 1 m*M* DTT and concentrated to 5.4 mg ml^{-1}. For crystallization, 50 µl of the HLA-DR1–HA complex was mixed with 17 µl MAM and 9 µl 9 m*M* $Zn(OAc)_2$, resulting in an equimolar ratio of MAM and DR1. Prior to crystallization, the complex mixture was incubated for 30 min. Initial crystallization conditions were established using Hampton Research Crystal Screen I, a commercially available sparse-matrix screen. Crystals appeared in three conditions, Nos. 32, 35 and 36, in which sulfate, phosphate and PEG 8000 were used as the precipitant, respectively. These conditions were further optimized. Microseeding was used to produce large crystals of both the native and SeMet complexes. The best crystals were grown at room temperature in hanging drops by mixing 2 µl protein solution with an equal volume of reservoir solution containing 1.7 *M* potassium sodium phosphate and 0.1 *M* HEPES pH 7.5. After equilibration overnight, the drop was introduced with 0.5 µl of stock solution of crystal seeds produced by crushing small crystals and diluting to varying degrees in the crystallization buffer. New crystals of both native and SeMet complexes appeared within 3 d and continued to grow for 2–3 weeks.

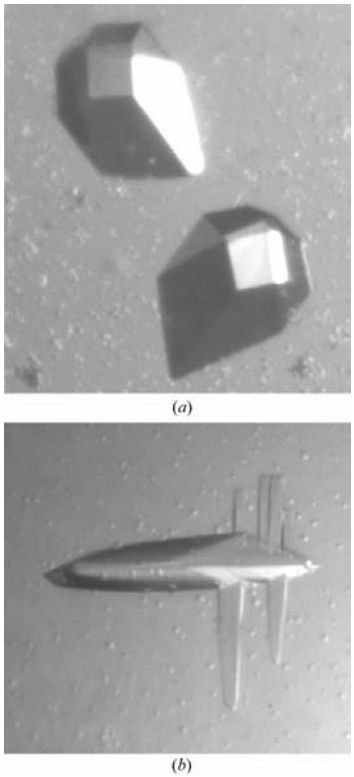

Figure 2
Native crystals of the MAM–HA–DR1 complex grown under (*a*) sodium potassium phosphate condition and (*b*) ammonium sulfate condition.

2.3. Data collection

Prior to data collection, all crystals were transferred to a cryoprotectant solution (mother liquor supplemented with 20% glycerol) and flash-cooled in a nitrogen-gas stream. Native data were collected to 3.3 Å resolution at 100 K using a CCD detector at beamline 19-BM of the Advanced Photon Source (APS), Argonne National Laboratory, USA. A set of single-wavelength anomalous diffraction (SAD) data (λ = 0.978 Å) for the SeMet complex crystal was collected to 3.4 Å resolution at 100 K using a Brandeis-B4 CCD detector at beamline X12C of the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory. All data were processed, scaled and reduced using the program *HKL*2000 (Otwinowski & Minor, 1997) and the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

2.4. Self-rotation functions

A self-rotation function was calculated for the SAD data with the program *GLRF* from the *REPLACE* suite of programs (Tong & Rossmann, 1990). Data in the resolution range 20–4 Å were used, with an integration radius of 25 Å.

2.5. Molecular replacement and SAD phasing

With the crystal structure of the HLA-DR1–HA complex (Stern *et al.*, 1994) as a search model, the cross-rotation function was calculated using the native data with the program *AMoRe* (Navaza, 1994). Molecular replacement (MR) generated clear solutions for the HLA-DR1 molecules. Using the program *SOLVE* (Terwilliger & Berendzen, 1999), we performed SAD phasing with the phase contribution from the DR1–HA complexes using the 3.4 Å SAD data. The positions for six Se atoms were clearly defined. With the combined MR and SAD phasing information, non-crystallographic symmetry (NCS) averaging using the program *RESOLVE* (Terwilliger, 2001) was able to produce an interpretable map.

Table 1
Data collection, molecular replacement, SAD phasing and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	Native	SeMet derivative
Wavelength (Å)	1.03	0.978
Resolution (Å)	3.3 (3.4–3.3)	3.4 (3.5–3.4)
Redundancy	2.1	4.7
Completeness (%)	81.3 (74.0)	99.5 (99.9)
Average $I/\sigma(I)$	7.4 (3.9)	12.7 (5.3)
R_{sym} (%)	11 (43.9)	13.9 (33.4)
SAD phasing		
Resolution range (Å)		20-3.4
R_{ano} (%)		6.8
Se sites		6
Overall FOM		
SAD alone		0.09
Combined†		0.62

† The overall FOM was calculated using combined phase information from the MR and SAD phasing.

3. Results and discussion

The MAM protein purified through our protocol is very pure, as judged by the single monomeric peak seen in gel-filtration chromatography and electrospray mass spectrometry (Figs. 1*a* and 1*c*).

Bar-shaped crystals of the MAM–HA–DR1 complex were first observed from condition Nos. 32 and 36 from Hampton Research Crystal Screen I with ammonium sulfate and PEG 8000 as the precipitant, respectively (Fig. 2*b*). Wedge-like crystals appeared under the phosphate condition (No. 35) almost six months later (Fig. 2*a*). The crystals obtained using the phosphate condition diffracted slightly better than those from the other two conditions. Therefore, the phosphate condition was further optimized. Even with microseeding, the majority of crystals appeared to be single when small but tended to cluster together when they grew larger. Using a hair-tool to remove small crystals attached to large crystals under a microscope, we were able to obtain single crystals with average dimensions of 0.15 × 0.1 × 0.1 mm. Although the crystals are still small, they diffracted to 3.3 and 3.4 Å resolution for the native and SeMet complex crystals, respectively, using synchrotron X-ray sources. Native and SAD diffraction data were then able to be collected (Table 1).

Surprisingly, although the crystals obtained from the three conditions display different morphologies, they all belong to the same space group *I*222, with the same unit-cell parameters: a = 137.4, b = 178.2, c = 179.6 Å. Therefore, it is possible that each asymmetric unit of the crystal lattice contains two, three or four MAM–HA–DR1 complexes. The Matthews coefficients (V_M) were accordingly calculated to be 4.0, 3.0 or 2.0 Å^3 Da^{-1}, respectively.

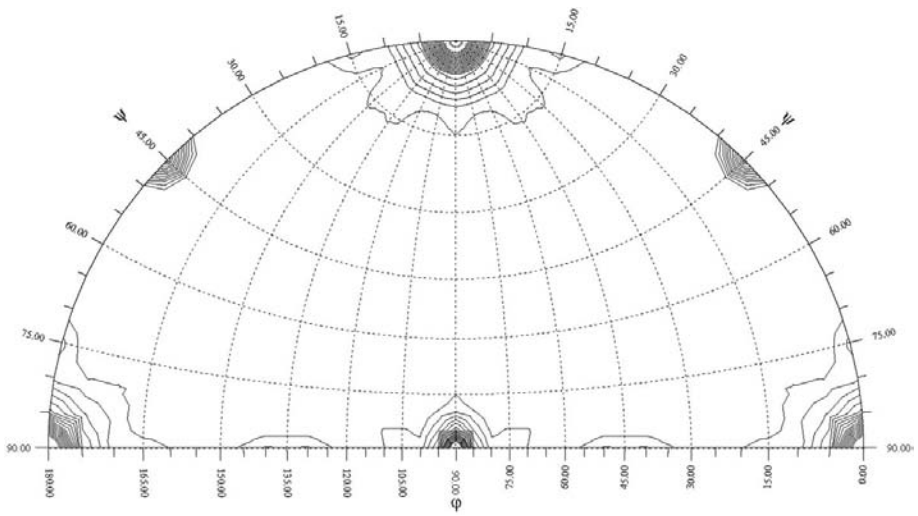

Figure 3
Projection of the $\kappa = 180^\circ$ section of the self-rotation function, showing the presence of a non-crystallographic twofold axis at spherical polar coordinates ($\varphi = 0$, $\psi = 45$, $\kappa = 180^\circ$).

The self-rotation function shows a peak of 6.4σ height with a signal-to-noise (S/N) ratio of 3.6 at the spherical polar coordinates ($\varphi = 0$, $\psi = 45$, $\kappa = 180^\circ$) (Fig. 3), indicating the presence of a twofold non-crystallographic symmetry (NCS) axis. This implies that either two or four complexes must be present in each asymmetric unit, resulting in solvent contents of 69 or 39%, respectively. Cross-rotational and translational function calculations further confirmed there to be two MAM–HA–DR1 complex molecules in an asymmetric unit. At a resolution range 20–4 Å, two solutions for the HLA-DR1 molecules could be obtained with correlation coefficients (CC) of 15.0 and 14.6%, respectively. The next highest noise peak showed a CC value of 11.2%. Using the solutions of the cross-rotation function, single-body translation-function calculation at 8–4 Å generated two solutions with CC values of about 21% for both solutions and of 10% for the highest noise peak. When one of the solutions was fixed, translation-function calculation produced a solution for the second HA–DR1 molecule with the same crystallographic origin as the first one, resulting in a dramatically increased CC value to 40% and an R factor decrease from 55 to 47%. The angular relationship between the two solutions was consistent with the result from the self-rotation function calculation.

The resultant electron-density maps showed some densities beyond the boundary of the DR1–HA molecules, but were insufficient to enable tracing of the MAM molecule. Combining the phase contribution from the MR solutions of the MHC molecules, we performed a phased SAD calculation using the 3.4 Å SAD data with the program *SOLVE* (Terwilliger & Berendzen, 1999). The positions for all six Se atoms were clearly defined. These positions were further verified on the expected sites on the α chain of the MHC molecules, which were independently determined using the molecular-replacement method. Although the SAD phasing only gave a figure-of-merit (FOM) of 0.09, the FOM of the combined MR and SAD phases was 0.62. With the combined MR and SAD phase information, non-crystallographic symmetry (NCS) averaging using the program *RESOLVE* (Terwilliger, 2001) was able to produce an interpretable map. Structure determination is currently in progress.

This research was supported by the National Institutes of Health (NIH) grant AI50628 (to HL). We thank L. J. Stern and W. Mourad for the gifts of HLA-DR1 and MAM expression plasmids, respectively, the Peptide Synthesis and Biological Mass Spectrometry Core facilities at the Wadsworth Center for the synthesis of the HA peptide, and R. G. Zhang at the APS and R. Sweet at the NSLS for assistance in X-ray data collection. Both APS and NSLS are supported by the Department of Energy and by grants from the NIH.

References

Abe, J., Kotzin, B. L., Jujo, K., Melish, M. E., Glode, M. P., Kohsaka, T. & Leung, D. Y. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 4066–4070.
Acha-Orbea, H. & MacDonald, H. R. (1995). *Annu. Rev. Immunol.* **13**, 459–486.
Choi, Y. W., Herman, A., DiGiusto, D., Wade, T., Marrack, P. & Kappler, J. (1990). *Nature (London)*, **346**, 471–473.
Cole, B. C. (1991). *Curr. Top. Microbiol. Immunol.* **174**, 107–119.
Cole, B. C. & Atkin, C. L. (1991). *Immunol. Today*, **12**, 271–276.
Cole, B. C. & Griffiths, M. M. (1993). *Arthritis Rheum.* **36**, 994–1002.
Cole, B. C., Knudtson, K. L., Oliphant, A., Sawitzke, A. D., Pole, A., Manohar, M., Benson, L. S., Ahmed, E. & Atkin, C. L. (1996). *J. Exp. Med.* **183**, 1105–1110.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Conrad, B., Weidmann, E., Trucco, G., Rudert, W. A., Behboo, R., Ricordi, C., Rodriquez-Rilo, H., Finegold, D. & Trucco, M. (1994). *Nature (London)*, **371**, 351–355.
Dellabona, P., Peccoud, J., Kappler, J., Marrack, P., Benoist, C. & Mathis, D. (1990). *Cell*, **62**, 1115–1121.
Fraser, J., Arcus, V., Kong, P., Baker, E. & Proft, T. (2000). *Mol. Med. Today*, **6**, 125–132.
Frayser, M., Sato, A. K., Xu, L. & Stern, L. J. (1999). *Protein Expr. Purif.* **15**, 105–114.
Hodtsev, A. S., Choi, Y., Spanopoulou, E. & Posnett, D. N. (1998). *J. Exp. Med.* **187**, 319–327.
Kappler, J. W., Staerz, U., White, J. & Marrack, P. C. (1988). *Nature (London)*, **332**, 35–40.
Kotzin, B. L., Leung, D. Y., Kappler, J. & Marrack, P. (1993). *Adv. Immunol.* **54**, 99–166.
Langlois, M. A., Etongue-Mayer, P., Ouellette, M. & Mourad, W. (2000). *Eur. J. Immunol.* **30**, 1748–1756.
Li, H., Llera, A., Malchiodi, E. L. & Mariuzza, R. A. (1999). *Annu. Rev. Immunol.* **17**, 435–466.
McCormick, J. K., Yarwood, J. M. & Schlievert, P. M. (2001). *Annu. Rev. Microbiol.* **55**, 77–104.
MacDonald, H. R., Lees, R. K., Baschieri, S., Herrmann, T. & Lussow, A. R. (1993). *Immunol. Rev.* **133**, 105–117.
Marrack, P. & Kappler, J. (1990). *Science*, **248**, 1066.
Mitchell, D. T., Levitt, D. G., Schlievert, P. M. & Ohlendorf, D. H. (2000). *J. Mol. Evol.* **51**, 520–531.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Papageorgiou, A. C. & Acharya, K. R. (2000). *Trends Microbiol.* **8**, 369–375.
Renno, T. & Acha-Orbea, H. (1996). *Immunol. Rev.* **154**, 175–191.
Stauffer, Y., Marguerat, S., Meylan, F., Ucla, C., Sutkowski, N., Huber, B., Pelet, T. & Conrad, B. (2001). *Immunity*, **15**, 591–601.
Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L. & Wiley, D. C. (1994). *Nature (London)*, **368**, 215–221.
Sundberg, E. J., Li, Y. & Mariuzza, R. A. (2002). *Curr. Opin. Immunol.* **14**, 36–44.
Sutkowski, N., Conrad, B., Thorley-Lawson, D. A. & Huber, B. T. (2001). *Immunity*, **15**, 579–589.
Swaminathan, S., Furey, W., Pletcher, J. & Sax, M. (1992). *Nature (London)*, **359**, 801–806.
Terwilliger, T. C. (2001). *Acta Cryst.* D**57**, 1755–1762.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Tong, L. A. & Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 783–792.
White, J., Herman, A., Pullen, A. M., Kubo, R., Kappler, J. W. & Marrack, P. (1989). *Cell*, **56**, 27–35.