Protein Science (1998), 7:1245–1249. Cambridge University Press. Printed in the USA.

FOR THE RECORD

Assembly and crystallization of the complex between the human T cell coreceptor CD8α homodimer and HLA-A2

GEORGE F. GAO,[1,4] ULRICH C. GERTH,[1,4] JESSICA R. WYER,[1] BENJAMIN E. WILLCOX,[1] CHRISTOPHER A. O'CALLAGHAN,[1] ZHIHONG ZHANG,[2] E. YVONNE JONES,[3] JOHN I. BELL,[1] AND BENT K. JAKOBSEN[1]

[1]Molecular Immunology Group, Institute of Molecular Medicine, Oxford University, John Radcliffe Hospital, Oxford OX3 9DS, United Kingdom
[2]Dyson Perrins Laboratory, South Parks Road, Oxford OX1 3QY, United Kingdom
[3]Laboratory of Molecular Biophysics, The Rex Richards Building, South Parks Road, Oxford OX1 3QU, United Kingdom

(RECEIVED September 10, 1997; ACCEPTED February 4, 1998)

Abstract: A strategy for overexpression in *Escherichia coli* of the extracellular immunoglobulin domain of human CD8α was devised using codon usage alterations in the 5′ region of the gene, designed so as to prevent the formation of secondary structures in the mRNA. A fragment of CD8α, comprising residues 1–120 of the mature protein, excluding the signal peptide and the membrane-proximal stalk region, was recovered from bacterial inclusion bodies and refolded to produce a single species of homodimeric, soluble receptor. HLA-A2 heavy chain, β2-microglobulin and a synthetic peptide antigen corresponding to the pol epitope from HIV-1 were also expressed in *E. coli*, refolded and purified. CD8α/HLA-A2 complexes were formed in solution and by co-crystallization with a stoichiometry of one CD8$\alpha\alpha$ dimer to one HLA-A2-peptide unit.

Keywords: CD8; HLA-A2; protein crystallization; protein folding; receptor-ligand complex

Major histocompatibility complex proteins (MHC or HLA in man) bind peptide antigens and present them on the cell surface where they can be recognized by T lymphocytes expressing a unique T cell receptor (TCR) matching the specific MHC-peptide combination. The transmembrane glycoproteins CD8 and CD4 are characteristic of distinct populations of T lymphocytes whose antigen responses are restricted by class I and class II MHC molecules, respectively. CD8 and CD4 play major roles both in the differentiation and selection of T cells during thymic development and in the activation of mature T lymphocytes in response to antigen presenting cells (Fung-Leung et al., 1991, 1993; The et al., 1991; Killeen et al., 1992; Zamoyska, 1994). CD8 and CD4 are, therefore, considered to be the main accessory molecules for T cell receptors. Although CD8 and CD4 are immunoglobulin superfamily proteins and determine antigen restriction by binding to MHC molecules the structural basis for their similar functions appears to be very different. Their sequence and structural differences are significant (Maddon et al., 1985), and whereas CD4 is expressed on the cell surface as a monomer CD8 is expressed as an $\alpha\alpha$ or an $\alpha\beta$ dimer (Norment & Littman, 1988).

Several lines of evidence have demonstrated that CD8 possesses binding specificity for class I MHC molecules independent of the peptide antigen carried in the complex (Norment et al., 1988). Mutational analyzes have indicated that the main recognition site for CD8 is the flexible loop of the α3 domain of MHC molecules, comprising residues 223–229 (Potter et al., 1989; Salter et al., 1990; Sun et al., 1995).

The analysis of the molecular basis underlying cellular immune responses has suffered from the lack of availability of soluble protein reagents with which to study structural and kinetic aspects of the interactions involved. Recently, however, a number of groups using expression systems ranging from *Escherichia coli* to mammalian cells have reported strategies for obtaining soluble TCR (Eilat et al., 1992; Hilyard et al., 1994; Ishii et al., 1995; Garboczi et al., 1996; Gregoire et al., 1996; Garcia et al., 1996a) and CD8 molecules (Leahy et al., 1992; Alcover et al., 1993; Garcia et al., 1996b). Here, we report a strategy for high-level expression of the extracellular immunoglobulin domain of human CD8α in *E. coli* and for the folding of the CD8$\alpha\alpha$ dimer as a MHC-binding receptor. Complex between CD8$\alpha\alpha$ and HLA-A2 associated with peptide antigen was assembled in solution by co-refolding and obtained by co-crystallization.

Results and discussion: To construct CD8α expression vectors for use in a range of expression systems a DNA fragment constituting the entire open reading frame was first generated by the polymerase chain reaction (PCR) using cDNA prepared from a cytotoxic T-lymphocyte cell line as template. A plasmid containing

Reprint requests to: Bent K. Jakobsen, Molecular Immunology Group, Institute of Molecular Medicine, Oxford University, John Radcliffe Hospital, Oxford OX3 9DS, United Kingdom; e-mail: jakobsen@pinnacle.jr2.ox.ac.uk.

[4]These authors contributed equally to this work.

this fragment was then used as template in a second generation of PCR reactions to synthesize five CD8α cDNA fragments which were inserted in plasmids containing the T7 promoter. These constructs all encode polypeptide chains in which an initiation codon is followed by the mature CD8α N-terminal region, comprising the entire immunoglobulin domain but truncated at different points C-terminally of this (Fig. 1). Constructs with equivalent C-terminal truncations but incorporating the CD8α signal sequence were tested for expression from a baculovirus promoter in insect cells.

Previously, expression of human CD8α in insect and CHO cells has indicated that disulfide bonding of the dimer through Cys143 is critical for the production of a secreted receptor (Leahy et al., 1992; Alcover et al., 1993). In contrast, Cys160, which is also involved in covalently linking the membrane-bound receptor, was found to prevent expression of soluble CD8$\alpha\alpha$ (Leahy et al., 1992). In accordance with this we found that in insect cells CD8α 1–146, 1-120/141–146, and 1–116/141–146 could be expressed as secreted dimers although at low levels (data not shown). The requirement for the Cys143 residue for secretion of CD8$\alpha\alpha$ dimer from insect cells indicates that this residue is essential either for the folding of the protein or for its stability during intracellular transport. These possibilities cannot at present be distinguished and since cysteine residues pose a potential problem when refolding protein in vitro constructs incorporating or omitting Cys143 were tested in bacteria (Fig. 1).

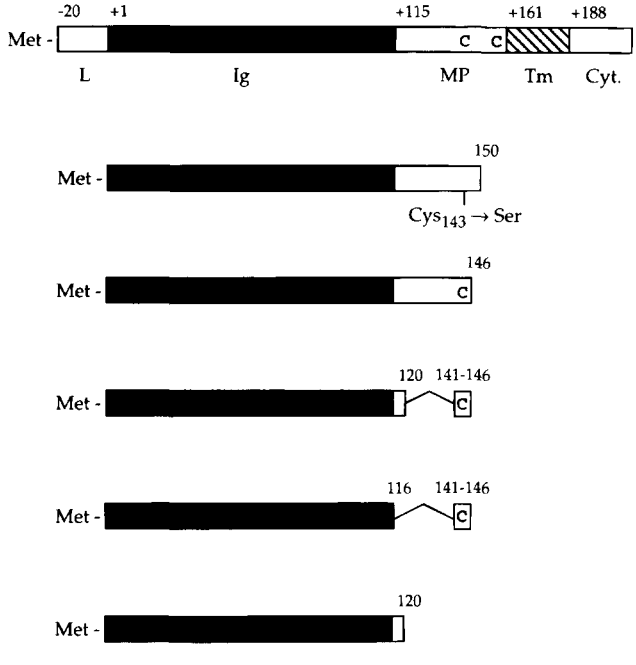


Fig. 1. CD8α and truncated proteins expressed in *E. coli*. Schematic presentation of CD8α illustrating domain organization (top) and the extent of the truncated proteins expressed in bacteria (below). L = signal sequence, Ig = immunoglobulin domain, MP = membrane proximal domain, Tm = transmembrane domain, Cyt. = cytoplasmic domain. Numbers indicate amino acid positions of domain boundaries in relation to the mature protein and the endpoints of proteins expressed in *E. coli*. "C" signifies the presence and positions of the cysteine residues at positions 143 and 160 involved in forming the interchain disulfide bonds in the CD8$\alpha\alpha$ dimer. Expression vectors containing the codon information for the illustrated truncated proteins were constructed both with the original codon usage for human CD8α and with codon usage changes in the 5′ end of the insert as described.

From plasmids containing inserts with the original human CD8α codon usage, no induction of expression could be detected in *E. coli* as assessed by comparison of whole cell extracts after gel electrophoresis and Coomassie staining. In contrast, expression was readily detectable when six codons in the CD8α gene were altered by silent mutations, using the primer 5′-GGAATTCCATAT GAGTCAATTTCGTGTATCACCGCTGGATCG-3′, to make the 5′ end of the gene more A/T rich (data not shown). It has previously been reported that such changes, as would be expected with a lower G/C content and therefore weaker base-pairing potential, decrease the tendency for hairpin formation in the 5′ end of the CD8α mRNA (de Smit & van Duin, 1990, 1994; Makrides, 1996). It appears likely, therefore, that secondary structure formation in the original sequence has a severe adverse effect on the potential for expression in *E. coli* (Garboczi et al., 1996). Expression from the mutated constructs are estimated to yield in excess of 100 mg/L, an increment of at least two orders of magnitude over the expression level obtained with the unmutated sequence.

Analysis of cellular fractions from lysed bacteria expressing CD8α showed that virtually all the protein is deposited in inclusion bodies as has previously been observed with bacterially expressed MHC heavy chain and β2m polypeptides (Garboczi et al., 1992). Purification of the CD8α inclusion bodies by washing yielded protein estimated to be approximately 90% pure (not shown). Similar inclusion body purity was obtained with HLA-A2 and β2m. CD8α could not be resolubilized in the pH 6.5 buffer used for the MHC chains (Garboczi et al., 1992) but adjustment of pH to 8.0, NaCl to 500 mM and the addition of 10% Glycerol made it readily soluble.

The resolubilized CD8α polypeptide (2 mL of 10 mg/mL) was incubated 30 min in reducing and denaturing buffer (15 mL of mM guanidine hydrochloride, 50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 mM NaCl, 10 mM DTT), and then diluted into conditions permitting formation of the native protein conformation (1 L of 200 mM Tris-HCl pH 8.0, 10 mM EDTA, 1 M L-arginine, 0.1 mM PMSF, 6.5 mM cysteamine, and 3.7 mM cystamine). After adding two more pulses of protein at 12 h intervals and allowing refolding to proceed for 48 h, soluble protein was concentrated for analysis and purification by gel filtration. Four of the polypeptides expressed and illustrated in Figure 1 produced mainly aggregate peaks of high molecular weight. In contrast, the elution profile of the 1–120 polypeptide showed one major peak of approximately 30 kDa as expected for the CD8$\alpha\alpha$ dimer. In this elution profile no indication of aggregates was evident (data not shown). The CD8$\alpha\alpha$ dimer was further purified by cation exchange on a Mono-S column eluting as a single peak at approximately 175 mM NaCl. Gel analysis of this peak showed only one band at the expected position for CD8α.

The identity of the purified CD8α was verified by two methods. Edman sequencing produced a run of ten amino acids corresponding to the mature N-terminus of CD8α, which indicated that the introduced methionine residue had been cleaved off in the bacteria. Electrospray mass spectrometry confirmed the molecular weight of CD8α (13,464.0) as very close to the theoretical value of 13,463.2 Da for residues 1–120, excluding the initiating methionine (Fig. 2). A smaller peak of slightly higher molecular weight is likely to correspond to a small fraction of protein which has retained the N-terminal methionine residue (Fig. 2).

To test whether CD8$\alpha\alpha$/HLA-A2 complexes could be formed in solution, a co-refolding experiment was performed. Gel filtration analysis of the concentrated protein reveals that the folding of

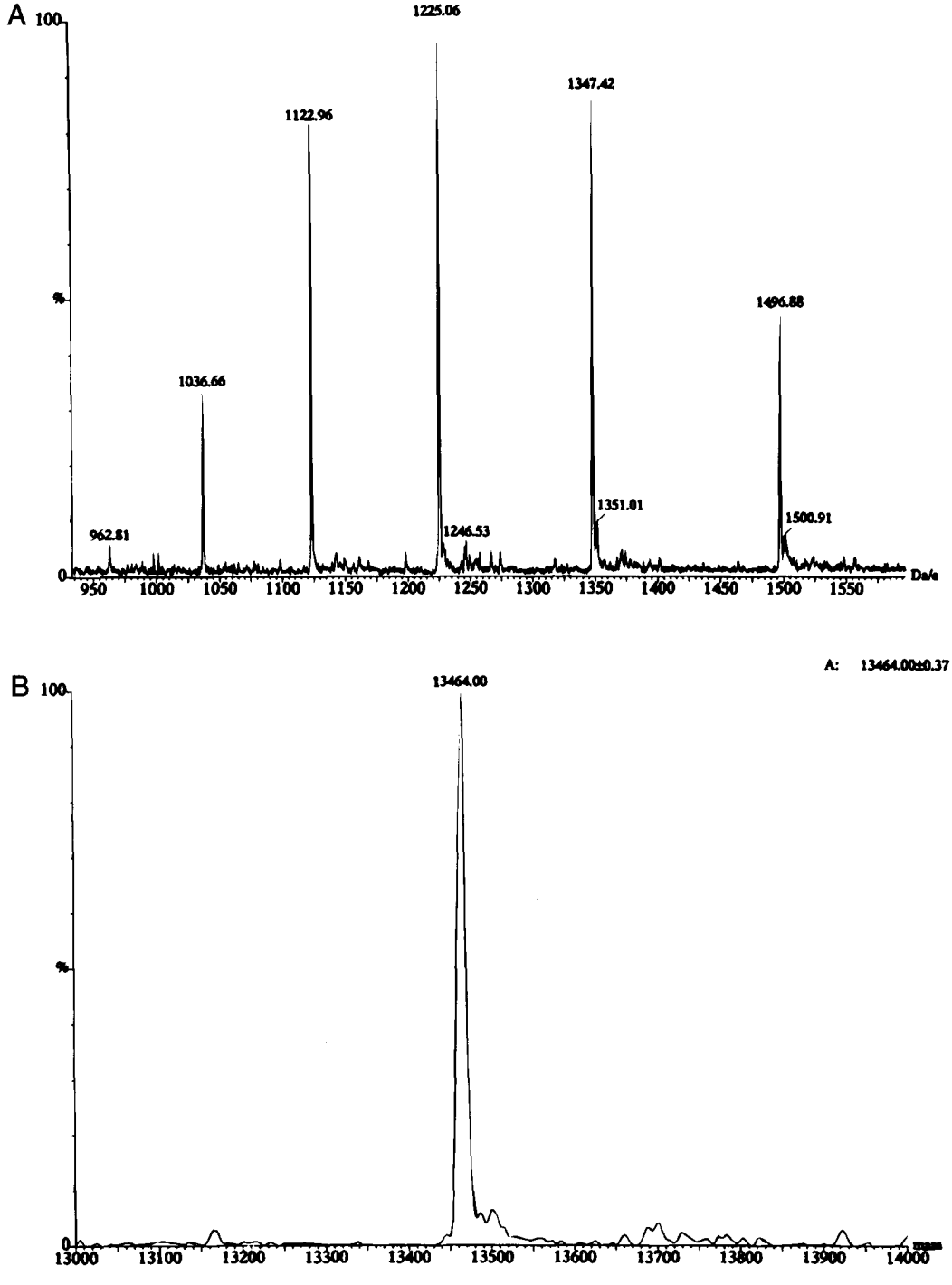
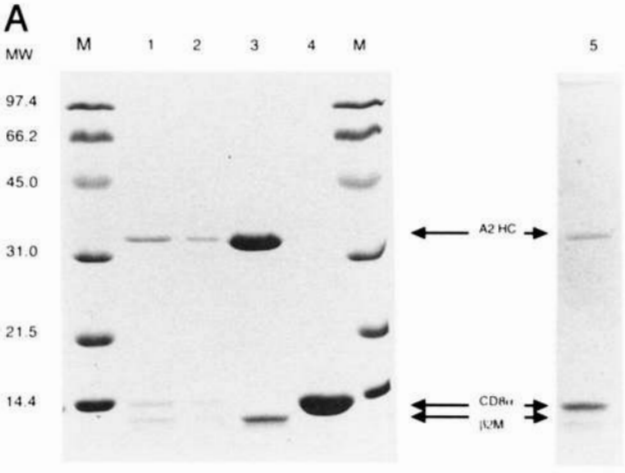


Fig. 2. Electrospray mass spectrometry analysis of soluble CD8$\alpha\alpha$ receptor. The mass spectrometer was calibrated using horse heart myoglobin (Mr 16951.48). **A:** Electrospray mass spectrum of CD8$\alpha\alpha$ showing mass-to-charge ratio peaks after proton bombardment. **B:** Deconvolution of the spectrum shown in A indicating the molecular weight of the single dominant peak.

both CD8$\alpha\alpha$ and HLA-A2 are impaired and that more aggregates are formed under these conditions than during individual refolding. A peak is detectable, however, at the expected position corresponding to approximately 75 kDa for the CD8$\alpha\alpha$/HLA-A2 complex, which was not present in either the CD8α refolding or in the HLA-A2 refolding (not shown). Analysis of the protein in this peak demonstrated that the components expected in the complex are present in amounts corresponding to one CD8$\alpha\alpha$ dimer to one HLA-A2 (Fig. 3A, lanes 1 and 5) indicating that stable complex formation can take place in solution.

Crystallization trials were performed at room temperature by vapor diffusion from sitting drops on microbridges with a 2:1



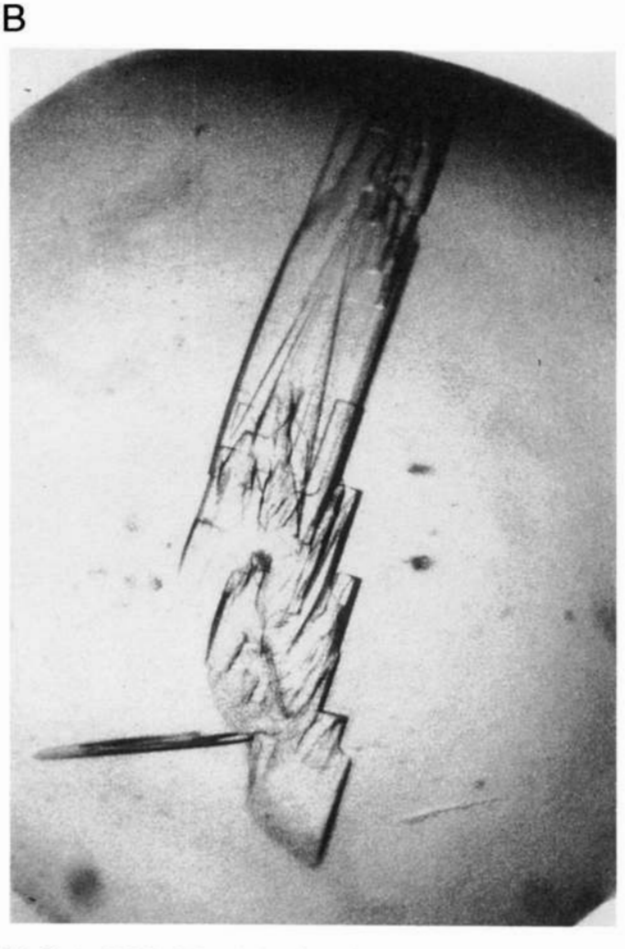


Fig. 3. A: SDS-PAGE analysis of co-refolded and co-crystallized CD8$\alpha\alpha$ and HLA-A2. Lane 1–4 and M were stained with Coomassie Blue, lane 5 was silver stained. Lanes 1 and 5: protein from the peak corresponding to complex formed from co-refolded CD8α, HLA-A2, β2m and peptide; lane 2: resolubilized crystal set up with CD8$\alpha\alpha$ and HLA-A2 complex; lane 3: HLA-A2; lane 4: CD8$\alpha\alpha$, M: protein size markers as indicated. The bands corresponding to A2 heavy chain (A2 HC), β2m and CD8α are indicated with arrows. **B:** Crystal of CD8$\alpha\alpha$-HLA-A2/pol peptide complex grown in 12% PEG 20,000, 100 mM MES pH 6.5.

mixture of separately refolded HLA-A2 and CD8$\alpha\alpha$. A 1 μL protein solution containing CD8$\alpha\alpha$ at 10 mg/mL and HLA-A2/pol at 20 mg/mL were mixed with 1 μL reservoir solution from Crystal Screen kits I and II (Hampton Research, California). Crystals appeared with two reservoir conditions, 12% PEG 20,000, 100 mM MES pH 6.5, and 15% PEG 6000, 50 mM MES pH 6.0 after approximately seven days (Fig. 3B). One crystal later diffracted at 2.7 Å resolution (Gao et al., 1997). To initially resolve the components constituting the crystal a fragment was washed, resolubilized, and analyzed by gel electrophoresis (Fig. 3A, lane 2). The protein bands present in this sample correspond exactly to those from the individual HLA-A2 and CD8α (Fig 3A, lanes 3 and 4, respectively). Furthermore, the ratios of the CD8α to the HLA-A2 heavy chain and β2m bands are identical to those found in the sample from the co-refolding peak assumed to be the receptor-ligand complex (Fig. 3A, lane 1).

The estimated molar ratios of the polypeptide bands representing HLA-A2 assembly and CD8$\alpha\alpha$ are identical in the complex formed in solution and in the crystal but the stoichiometry indicated by the band intensities is somewhat unexpected. Because CD8 has been detected on the T cell surface as a dimer, whereas CD4 is detected as a monomer, it has been speculated that CD8 would bind two MHC molecules and that a bivalent stoichiometry could be involved in T cell-target cell ligation and T cell activation (Giblin et al., 1994). However, the gel analysis of the complexes, taking into consideration that CD8α stain weaker with Coomassie than does β2m, are more indicative of a 1:1 HLA-A2:CD8$\alpha\alpha$ ratio and this has been confirmed by the crystal structure (Gao et al., 1997). The space-filling model of the complex based on the crystal structure furthermore rules out that more than one CD8$\alpha\alpha$ dimer can bind the MHC (Gao et al., 1997).

Acknowledgments: We are grateful to A. Willis for N-terminal sequencing of CD8, to Z. Rao for advice on crystal trials, K. Harlos for assistance and J. Tormo, D. Stuart, A. McMichael, D. Garboczi, and D. Wiley for discussion. The plasmid pHN1-β2m used for expression of the β2 microglobulin subunit was kindly provided by David N. Garboczi and Don C. Wiley, Harvard Medical School. This work was funded by the MRC. E.Y.J. was supported by the Royal Society, C.A.O. by the MRC.

References

Alcover A, Herve F, Boursier JP, Spagnoli G, Olive D, Mariuzza RA, Acuto O. 1993. A soluble form of the human CD8 alpha chain expressed in the baculovirus system: Biochemical characterization and binding to MHC class I. *Mol Immunol 30*:55–67.

de Smit MH, van Duin J. 1990. Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. *Proc Natl Acad Sci USA 87*:7668–7672.

de Smit MH, van Duin J. 1994. Control of translation by mRNA secondary structure in *Escherichia coli*. A quantitative analysis of literature data. *J Mol Biol 244*:144–150.

Eilat D, Kikuchi GE, Coligan JE, Shevach EM. 1992. Secretion of a soluble, chimeric gamma delta T-cell receptor-immunoglobulin heterodimer. *Proc Natl Acad Sci USA 89*:6871–6875.

Fung-Leung WP, Schilham MW, Rahemtulla A, Kundig TM, Vollenweider M, Potter J, van Ewijk W, Mak TW. 1991. CD8 is needed for development of cytotoxic T cells but not helper T cells. *Cell 65*:443.

Fung-Leung WP, Wallace VA, Gray D, Sha WC, Pircher H, The HS, Loh DY, Mak TW. 1993. CD8 is needed for positive selection but differentially required for negative selection of T cells during thymic ontogeny. *Eur J Immunol 23*:212–216.

Gao GF, Tormo J, Gerth UC, Wyer JR, McMichael AJ, Stuart DI, Bell JI, Jones EY, Jakobsen BK. 1997. Crystal structure of the complex between human CD8$\alpha\alpha$ and HLA-A2. *Nature 387*:630–634.

Garboczi DN, Hung DT, Wiley DC. 1992. HLA-A2-peptide complexes: Refolding and crystallization of molecules expressed in *Escherichia coli* and complexed with single antigenic peptides. *Proc Natl Acad Sci USA 89*:3429–3433.

Garboczi DN, Utz U, Ghosh P, Seth A, Kim J, VanTienhoven EA, Biddison WE, Wiley DC. 1996. Assembly, specific binding, and crystallization of a human TCR-alphabeta with an antigenic Tax peptide from human T lymphotropic virus type 1 and the class I MHC molecule HLA-A2. *J Immunol 157*:5403–5410.

Garcia KC, Degano M, Stanfield RL, Brunmark A, Jackson MR, Peterson PA, Teyton L, Wilson IA. 1996a. An alphabeta T cell receptor structure at 2.5 Å and its orientation in the TCR-MHC complex [see comments]. *Science 274*:209–219.

Garcia KC, Scott CA, Brunmark A, Carbone FR, Peterson PA, Wilson IA, Teyton L. 1996b. CD8 enhances formation of stable T-cell receptor/MHC class I molecule complexes [see comments]. *Nature 384*:577–581.

Giblin PA, Leahy DJ, Mennone J, Kavathas PB. 1994. The role of charge and multiple faces of the CD8 alpha/alpha homodimer in binding to major histocompatibility complex class I molecules: Support for a bivalent model. *Proc Natl Acad Sci USA 91*:1716–1720.

Gregoire C, Lin SY, Mazza G, Rebai N, Luescher IF, Malissen B. 1996. Covalent assembly of a soluble T cell receptor-peptide-major histocompatibility class I complex. *Proc Natl Acad Sci USA 93*:7184–7189.

Hilyard KL, Reyburn H, Chung S, Bell JI, Strominger JL. 1994. Binding of soluble natural ligands to a soluble human T-cell receptor fragment produced in *Escherichia coli*. *Proc Natl Acad Sci USA 91*:9057–9061.

Ishii Y, Nakano T, Honma N, Yuyama N, Yamada Y, Watarai H, Tomura T, Sato M, Tsumura H, Ozawa T. 1995. Preparation of soluble recombinant T cell receptor alpha chain by using a calmodulin fusion expression system. *J Immunol Methods 186*:27–36.

Killeen N, Moriarty A, The HS, Littman DR. 1992. Requirement for CD8-major histocompatibility complex class I interaction in positive and negative selection of developing T cells. *J Exp Med 176*:89–97.

Leahy DJ, Axel R, Hendrickson WA. 1992. Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2.6 Å resolution. *Cell 68*:1145–1162.

Maddon PJ, Littman DR, Godfrey M, Maddon DE, Chess L, Axel R. 1985. The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: A new member of the immunoglobulin gene family. *Cell 42*:93–104.

Makrides SC. 1996. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev 60*:512–538.

Norment AM, Littman DR. 1988. A second subunit of CD8 is expressed in human T cells. *Embo J 7*:3433–3439.

Norment AM, Salter RD, Parham P, Engelhard VH, Littman DR. 1988. Cell-cell adhesion mediated by CD8 and MHC class I molecules. *Nature 336*:79–81.

Potter TA, Rajan TV, Dick RFd, Bluestone JA. 1989. Substitution at residue 227 of H-2 class I molecules abrogates recognition by CD8-dependent, but not CD8-independent, cytotoxic T lymphocytes. *Nature 337*:73–75.

Salter RD, Benjamin RJ, Wesley PK, Buxton SE, Garrett TP, Clayberger C, Krensky AM, Norment AM, Littman DR, Parham P. 1990. A binding site for the T-cell co-receptor CD8 on the alpha 3 domain of HLA-A2. *Nature 345*:41–46.

Sun J, Leahy DJ, Kavathas PB. 1995. Interaction between CD8 and major histocompatibility complex (MHC) class I mediated by multiple contact surfaces that include the alpha 2 and alpha 3 domains of MHC class I. *J Exp Med 182*:1275–1280.

The HS, Garvin AM, Forbush KA, Carlow DA, Davis CB, Littman DR, Perlmutter RM. 1991. Participation of CD4 coreceptor molecules in T-cell repertoire selection. *Nature 349*:241–243.

Zamoyska R. 1994. The CD8 coreceptor revisited: One chain good, two chains better. *Immunity 1*:243–246.