The EMBO Journal vol.6 no.11 pp.3313–3316, 1987

Molecular cloning of two CD7 (T-cell leukemia antigen) cDNAs by a COS cell expression system

Alejandro Aruffo and Brian Seed

Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA

Communicated by E.Winnacker

The human CD7 antigen (gp40) is a cell surface glycoprotein found on thymocytes and mature T-cells. It is one of the earliest antigens to appear on cells of the T-lymphocyte lineage, and the most reliable clinical marker of T-cell acute lymphocytic leukemia. This report describes the isolation and nucleotide sequence of a full length CD7 cDNA, and of a cDNA for an unusual intron-bearing precursor. The DNA sequence of the clone predicts a highly glycosylated membrane protein with homology to members of the immunoglobulin superfamily, and no relationship to known oncogenes. Overexpression of CD7 RNA was observed in only one T-cell tumor line, and genomic DNA rearrangement was not observed in any lines. Prompted by a recent suggestion that CD7 plays a role in IgM binding, COS cells expressing CD7 were tested and found not to bind IgM or IgM immune complexes.

Key words: expression library/gp40/IgM receptor/surface antigen

Introduction

Antibodies in the CD7 cluster of differentiation (Palker *et al.*, 1985) recognize a 40-kd glycoprotein on the surface of peripheral blood T-cells and thymocytes. Like CD2 (the T-cell erythrocyte receptor), CD7 is one of the earliest antigens of the T-cell lineage, and can be found on rare cells in bone marrow (van Dongen *et al.*, 1985).

Early studies with CD7 antibodies showed that $CD7^+$ T-cells enhance immunoglobulin (Ig) synthesis by B-cells (Morishima *et al.*, 1982), suppress B-cell Ig synthesis when stimulated with Concanavalin A (Haynes *et al.*, 1979) and are the precursors of cytotoxic T-cells generated in mixed lymphocyte culture (Morishima *et al.*, 1982). Clinically, CD7 has been found to be the most reliable marker for the identification of T-cell acute lymphocytic leukemia (Link *et al.*, 1983). Cytotoxins coupled to CD7 antibody have been used to purge bone marrow prior to reinfusion to avoid early relapse in autologous bone marrow transplants or as prophylaxis against graft versus host disease in allogeneic bone marrow transplants (Ramakrishnan *et al.*, 1985). Similarly, CD7 antibodies have been administered to avert acute allograft rejection (Raftery *et al.*, 1985) in renal transplants.

At present the physiological role of CD7 is not understood. CD7 antibodies are not mitogenic, and do not block the T-cell response to PHA or tetanus toxoid (Palker *et al.*, 1985), although the CD7 antibody 7G5 significantly inhibits the primary mixed lymphocyte reaction (Lazarovits *et al.*, 1987). In this article we describe the nucleotide sequence and preliminary characterization of two cDNAs encoding the human CD7 antigen. Prompted by the recent suggestion that CD7 may be, or may be part of, the T-cell IgM receptor (Sandrin *et al.*, 1987) we evaluated the ability of COS cells expressing CD7 to bind IgM or IgM immune complexes. The results do not support the simple notion that CD7 itself is an IgM receptor.

Results and discussion

Isolation of a cDNA encoding the human CD7 antigen

CD7 cDNAs were isolated from a large plasmid library prepared from poly A^+ RNA isolated from the human T-cell leukemia line HPB-ALL as described previously (Aruffo and Seed, 1987). The expression library was introduced into COS cells by spheroplast fusion and allowed to replicate and express the inserted cDNA. Forty-eight to seventy-two hours after transfection the cells were harvested without trypsin, treated with monoclonal antibodies recognizing CD7 and other surface antigens and distributed in dishes coated with affinity purified anti-mouse immunoglobulin antibody (Seed and Aruffo, 1987). Under these conditions cells expressing surface antigen adhere and the remaining cells can be washed away. From the adherent cells, a Hirt (1967) episomal DNA fraction was prepared and the resulting DNA tranformed back into *Escherichia coli* for further rounds of fusion and selection. In the third round of selection a portion of the harvested cells was treated with two antibodies specific for CD7 (7G5 and Leu 9), allowed to attach to the anti-mouse immunoglobulin dishes, washed and subjected to the Hirt procedure. After transformation of the DNA into *E.coli*, eight colonies were picked and the plasmid DNA prepared from them by an alkaline miniprep procedure (Maniatis *et al.*, 1982). The DNA was transfected into eight individual COS cell cultures and, three days later, surface expression of the CD7 antigen was detected by indirect immunofluorescence in seven of eight transfected dishes. Restriction enzyme digestion of the corresponding plasmid DNAs revealed two species. One contained a 1.2-kb insert, and the other a 1.7-kb insert.

CD7 cDNA sequence analysis

Both isolates were sequenced by the dideoxynucleotide method, The 1.2-kb cDNA encodes a long open reading frame of 240 residues having the typical features of an integral membrane protein. Our initial assignment of the signal sequence cleavage site by the method of von Heijne (1986) was at the 18th residue, but while this work was in review, Alan Williams (personal communication) suggested that the homology with immunoglobulin variable regions (see below) would better predict the mature terminus at residue 26; this assignment would also correlate well with the position of the intron as discussed below and is shown in Figure 1. Removal of the predicted N-terminal signal sequence gives a mature protein of 215 residues with a predicted molecular mass of 23 kd. In the extracellular domain are two N-linked glycosylation sites (Asn-X-Ser/Thr), in agreement with the results of Sutherland *et al.* (1984), who also showed the presence of O-linked glycans and covalently associated palmitic acid on the mature protein. In the 27-amino-acid hydrophobic membrane spanning domain is a single cysteine residue which may be the site of fatty acylation (Rose *et al.*, 1984; Kaufman *et al.*, 1984). The length of the cytoplasmic domain, 39 residues, is in good agreement with the 30–40 amino acids predicted by protease

Fig. 1. Nucleotide sequence of the 1.7-kb cDNA. Nucleotide numbering is given in parentheses at right. The amino acid sequence is numbered from the projected cleavage site of the secretory signal sequence. Splice donor and acceptor sites indicated by (/). The location of the potential sites for addition of asparagine-linked carbohydrate (CHO) are shown, the potential fatty acid esterification site is denoted (*), and the predicted transmembrane domain (TM) is underlined. Nucleotide sequences potentially involved in hairpin formation are denoted by (·). The presumed polyadenylation signal is underlined.

digestion of the CD7 precursor in rough microsomal membrane fractions (Sutherland *et al.*, 1984).

Sequence analysis of the 1.7-kb clone (Figure 1) revealed the presence of an intron located 121 bp from the 5′ end. The 411-bp intron contains stop codons in all three reading frames and is located just downstream of the secretory signal sequence, as is frequently observed for secreted or surface proteins. Both the 5′ and 3′ ends of the intron conform to the splice donor/acceptor consensus AAG↓GTRAGA/ . . . /Y_{6-11}NYAG↓A (Mount, 1982). Because both the 1.2- and 1.7-kb clones express CD7 antigen equally well in COS cells, the intron must be excised in COS cells fairly efficiently.

Comparison of the amino acid sequence with the National Biomedical Research Foundation database revealed substantial homology with human and mouse immunoglobulin κ chain and T-cell receptor γ chain variable regions over almost the entire extracellular portion of the molecule (Figure 2). Two cysteine residues shared in approximately equal spacing by all three structures fall in the conserved sequences Ile-Thr-Cys and Tyr-X-Cys. In κ chain variable regions these cysteines form a disfulfide bridge. The presence of at least one intrastrand disulfide bond in the CD7 structure has previously been proposed by Sutherland *et al.* (1984) who noted that immunoprecipitation of CD7 gave rise to a band with an apparent molecular mass of 40 kd under reducing conditions and 38 kd under nonreducing conditions.

Based on the homology with immunoglobulin V-regions, we predict that CD7 contains a disulfide bond linking Cys 23 and Cys 89. A second disulfide bond, linking Cys 10 and Cys 117, has been proposed by Alan Williams, based on the structural similarity between CD7 and Thy-1 (A.F.Williams, personal communication). The extracellular domains of both Thy-1 and CD7 have four cysteine residues, in roughly homologous positions. The four cysteine residues of Thy-1 are joined in two internal disulfide bridges between Cys 9−111 and Cys 19−85 (Williams and Gagnon, 1982). In Thy-1 Cys 111 forms an amide bond with the ethanolamine moiety of a substituted phosphatidylinositol, and is thus the last residue of the mature molecule (Tse *et al.*, 1985). In CD7, Cys 117 is followed by four repeats of a sequence whose consensus is Xaa-Pro-Pro-Xaa-Ala-Ser-Ala-Leu-Pro, and which we propose plays the role of a stalk projecting the V-like domain away from the surface of the cell.

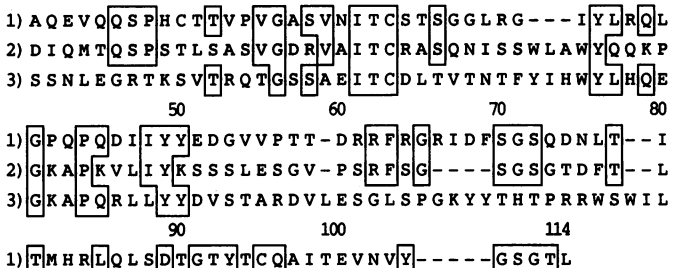

Fig. 2. Homology between the CD7 cDNA (**1**), human κ chain (**2**) and human T-cell receptor γ chain (**3**); variable regions are shown boxed.

In addition to the homologies shown in Figure 2 and mentioned

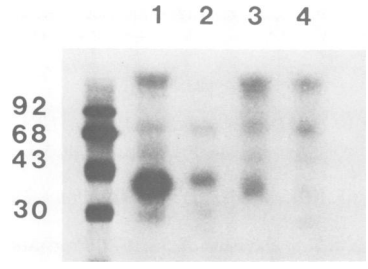

Fig. 3. Surface expression of CD7 antigen. **Lanes 1, 2, 3**: immunoprecipitates of transfected COS cells, HPB-ALL cells and activated T-cells with CD7 antibody. **Lane 4**, immunoprecipitate of CD7 transfected COS cells treated with anti-CD4 antibody.

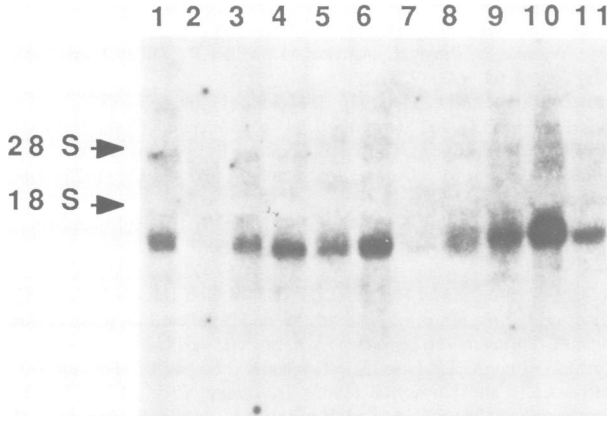

Fig. 4. RNA blot analysis of CD7 transcripts. RNA sources were: **lane 1**, thymocytes; **lane 2**, U937 (histiocytic leukemia); **lane 3**, HuT-78 (adult T-cell leukemia); **lane 4**, activated T-cells; **lane 5**, HPB-ALL (T-cell leukemia); **lane 6**, Jurkat J3R7 (T-cell leukemia); **lane 7**, Namalwa (Burkitt lymphoma); **lane 8**, MOLT4 (T-cell leukemia); **lane 9**, HSB-2 (T-cell leukemia); **lane 10**, PEER (T-cell leukemia); and **lane 11**, senescent (rested) T-cells. The locations of the 18S and 28S ribosomal RNA bands are indicated.

above, the extracellular domain of CD7 has significant homology with both chains of the rat CD8 heterodimer (Johnson and Williams, 1986), and the myelin P_0 protein (Lemke and Axel, 1985 (A.F.Williams, personal communication). These proteins, like CD7, are composed of a single Ig V-like domain followed by transmembrane and cytoplasmic domains (Williams, 1987).

CD7 directs the production of a 40 kd protein in transfected COS cells

Immunoprecipitation of CD7 antigen from transfected COS cells was carried out with monoclonal antibody 7G5 (Lazarovits *et al.*, 1987). The material obtained from COS cells migrated as a broad band of molecular mass 40 kd under reducing conditions (Figure 3). Parallel immunoprecipitation of the antigen on HPB-ALL and activated T-cells gave bands with molecular masses of 41 and 39 kd respectively. In both COS cell and HPB-ALL immunoprecipitations a faint band with a mol. wt of 30 kd was also observed, possibly corresponding to a partially glycosylated precursor (Sutherland *et al.*, 1984).

RNA blot analysis and origin of the 1.7-kb clone

Equal amounts of total RNA prepared from cell types expressing or lacking CD7 were subjected to blot analysis as described previously (Seed and Aruffo, 1987). A single 1.3-kb species was visible in lanes containing RNA from thymocytes, activated T-cells, resting T-cells and the T-cell leukemia line HuT-78, HPB-ALL, Jurkat J3R7, HSB-2 and PEER (Figure 4). With the exception of the PEER cell line, none of the T-cell tumors showed significant overexpression of CD7 transcripts. CD7 RNA was detected in all of the thymus-derived cells, but not in RNA from U937 (histiocytic leukemia) and Namalwa (Burkitt lympyhoma) cells. No band corresponding to the 1.7-kb cDNA could be detected, suggesting that this species is artificially enriched during the cloning or library amplification process.

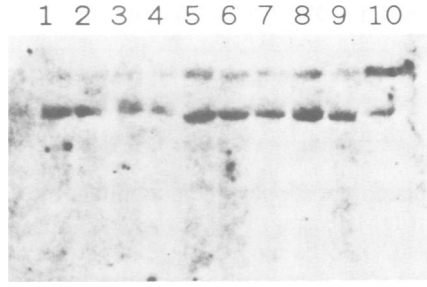

Fig. 5. Genomic DNA blot analysis. Twenty micrograms of each genomic DNA digested with *DraI* was electrophoresed through a 0.8% agarose gel, blotted and hybridized to a CD7 probe. DNA samples were prepared from: **lane 1**, HuT-78; **lane 2**, PEER; **lane 3**, MOLT4; **lane 4**, HSB-2; **lane 5**, placenta; **lane 6**, Jurkat J3R7; **lane 7**, total peripheral blood lymphocytes; **lane 8**, HPB-ALL; **lane 9**, activated T-cells; and **lane 10**, HeLa cells.

Enrichment during amplification seems unlikely because the 1.2-kb cDNA clone propagates as well in *E.coli* as the 1.7-kb clone. However, immediately upstream, and downstream of the site of insertion of the intron, are sequences that could form an interrupted stem and loop structure (Figure 1). Eight of the ten basepairs of the potential stem are GC pairs, perhaps giving the structure sufficient stability to interfere with elongation of the cDNA first strand. The presence of the intron greatly separates the two halves of the stem, potentially eliminating the structure via unfavorable loop entropy and allowing efficient first strand synthesis.

The CD7 gene is not rearranged

Genomic DNAs from placenta, peripheral blood lymphocytes, T-cells, HeLa cells or the tumor lines used in the RNA blot analysis above showed identical *DraI* digest patterns (Figure 5). Thus the CD7 gene is not grossly altered during development, and the high level of expression in the PEER cell line is not the consequence of a substantial genomic rearrangement.

COS cells expressing CD7 do not bind IgM

Human peripheral blood T-lymphocytes express receptors for IgM antibodies (FcRμ; Moretta *et al.*, 1975; McConnell and Hurd, 1976). Recently it has been reported that CD7 might play a role in IgM binding by T-cells (Sandrin *et al.*, 1987). L-cells, normally $CD7^-$ and $FcR\mu^-$, become $CD7^+$ and $FcR\mu^+$ when transfected with a 16-kb genomic fragment encoding the CD7 antigen (Sandrin *et al.*, 1987). Furthermore, IgM binding to CD7 positive cells can be blocked by the anti-CD7 monoclonal antibody Huly-m2 (Thurlow *et al.*, 1984), and IgM columns bind to a 37-kd protein from radiolabeled lysates of peripheral blood T-lymphocytes (Sandrin *et al.*, 1987).

Accordingly, COS cells expressing CD7 were tested for their ability to bind IgM. IgM receptor activity was assayed either by direct binding (Hardin *et al.*, 1979) or by a rosette assay with ox erythrocytes coated with an IgM fraction of rabbit anti-bovine red cell serum as described by Ercolani *et al.* (1981). Cells expressing CD7 neither bound human IgM nor formed rosettes with IgM coated erythrocytes. Under the same conditions COS cells

transfected with a cDNA encoding the human IgG receptor CDw32 (S.Stengelin and B.Seed, in preparation) bound IgG directly and formed rosettes with IgG coated erythrocytes. Erythrocytes coated with IgM or IgG antibodies also adhered to a fraction of peripheral blood lymphocytes as reported (Moretta *et al.*, 1975).

These results do not support the notion that the CD7 antigen is by itself an IgM receptor, although they do not exclude the possibility that COS cells suppress IgM binding activity in some manner, or that CD7 is part of, or modified to become, an IgM receptor. That CD7 is not by itself an IgM receptor is supported by the observation that a number of $CD7^+$ T-cell lines are $FcR\mu^-$ (Sandrin *et al.*, 1987).

Materials and methods

Preparation of cDNA library and recovery and characterization of CD7 clones

Preparation of an HPB-ALL cDNA library in the expression vector πH3 was described previously (Aruffo and Seed, 1987). Panning of the library was carried out as described by Seed and Aruffo (1987), using purified anti-CD7 antibody Leu9 (Becton Dickinson) and antibody 7G5 as ascites (Lazrovits *et al.*, 1987). The Leu9 antibody was used at a concentration of 1 μg/ml and the ascites fluid was diluted 1/1000. Methods for COS cell transfection, radioimmunoprecipitation, DNA and RNA blot hybridization and DNA sequencing were all as described (Seed and Aruffo, 1987).

IgM and IgG binding by COS cells transfected with CD7 and CDw32

Human IgM, IgG and IgA antibodies, affinity purified FITC conjugated goat anti-human immunoglobulins and antibodies [anti-Ig(G+M+A)], washed and preserved bovine red blood cells and IgG and IgM fractions of rabbit anti-bovine red blood cell antibodies were purchased from Cooper Biomedical (Malverne, PA). COS cells were transfected by the DEAE Dextran method (Seed and Aruffo, 1987) with cDNAs encoding the CD7, CDw32 and CD28 surface antigens. Forty-eight hours after transfection the cells were washed with PBS/0.5% BSA and incubated for 30 min at 4°C with FITC conjugated rabbit anti-human immunoglobulins. After washing, the cells were examined with a fluorescence microscope. The experiments were also performed in the presence of 0.1% azide with the same results.

Bovine erythrocytes for rosette assays were prepared as described by Ercolani *et al.* (1981). Briefly, a 2% suspension of bovine erythrocytes was washed with PBS/0.5% BSA and treated with subagglutinating amounts of either the IgG or the IgM fraction of rabbit anti-bovine erythrocyte antibodies at 4°C for 1 h. Erythrocytes were then washed twice with PBS/0.5% BSA and adjusted to a 2% solution. Two ml of antibody-coated erythrocytes were layered on 60-mm dishes containing COS cells which had been transfected 48 h earlier with either CD7, CD32 or CD28 by the DEAE Dextran method (Seed and Aruffo, 1987). The dishes were then centrifuged at 150 *g* at 4°C for 15 min. After an additional 45-min incubation at 4°C the plates were gently washed five times with 5 ml of PBS/0.5% BSA, and the COS cells were examined for rosette formation. These experiments were also performed in the presence of 0.1% sodium azide without alteration of the results.

Formation of T-cell rosettes with antibody-coated erythrocytes

Peripheral blood lymphocytes were obtained from heparinized blood by centrifugation at 4°C over a Ficoll-Hypaque gradient at 400 *g* for 30 min. Leukocytes at the interface were washed twice with PBS. The leukocytes were adjusted to 10^7 cells/ml in IMDM/10% Fetal Bovine Serum (FBS) and incubated in tissue culture dishes at 37°C for 30 min. Nonadherent cells were transferred to new dishes and PHA was added to stimulate proliferation of T-lymphocytes. On the next day the cells were washed with PBS and placed in fresh IMDM/10% FBS.

Rosette assays (Ercolani *et al.*, 1981) were performed three days later. Cells were washed with PBS/0.5% BSA and a 10-μl suspension of 2% Ig coated erythrocytes prepared as described above was added to 10 μl of PBS/0.5% BSA containing 5×10^6 cells/ml. The mixtures were placed in Falcon round bottom 96-well plates and centrifuged at 150 *g* for 15 min at 4°C. After an additional incubation of 45 min at 4°C, pellets were resuspended with 10 μl of PBS/0.5% BSA and the rosettes scored by phase contrast microscopy. The experiments were carried out in both the presence and absence of 0.1% sodium azide with no detectable difference.

Acknowledgements

We thank Andrew Lazarovits and David Camerini for their gift of 7G5 ascites, Alan Williams for advice on protein structure and homologies, Mauro Sandrin for preprints and discussion of his work prior to publication, Louis Ercolani for advice on rosette assays and Peter Heinrich for advice on sequencing. This work was supported by a grant to the Massachusetts General Hospital from Hoechst AG.

References

Aruffo,A. and Seed,B. (1987) *Proc. Natl. Acad. Sci. USA*, in press.
Ercolani,L., Novogrodsky,A., Suthanthiran,M., Rubin,A.L. and Stenzel,K.H. (1981) *J. Immunol.*, **127**, 2044–2051.
Hardin,J.A., Nakaoka,K.K. and Carboni,J.M. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 912–914.
Haynes,B.F., Eisenbarth,G.S. and Fauci,A.S. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 5829–5833.
Hirt,B. (1967) *J. Mol. Biol.*, **26**, 365–369.
Johnson,P. and Williams,A.F. (1986) *Nature*, **323**, 74–76.
Kaufman,J.F., Krangel,M.S. and Strominger,J.L. (1984) *J. Biol. Chem.*, **259**, 7230–7238.
Lazarovits,A.I., Colvin,R.B., Camerini,D., Karsh,J. and Kurnik,J.T. (1987) In McMichael,A.J. *et al.* (eds), *Leukocyte Typing III.* Oxford University Press, Oxford, in press.
Lemke,G. and Axel,R. (1985) *Cell*, **40**, 501–508.
Link,M., Warnke,R., Finlay,J., Amylon,M., Miller,R., Dilley,J. and Levy,R. (1983) *Blood*, **62**, 722–728.
McConnell,I. and Hurd,C.M. (1976) *Immunology*, **30**, 835–837.
Maniatis,T., Fritsch,E. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.
Moretta,L., Ferrarini,M., Durante,M.L. and Mingari,M.C. (1975) *Eur. J. Immunol.*, **5**, 565–569.
Morishima,Y., Kobayashi,M., Yang,S.Y., Collins,N.H., Hoffman,M.K. and Dupont,B. (1982) *J. Immunol.*, **129**, 1091–1098.
Mount,S.M. (1982) *Nucleic Acids Res.*, **10**, 459–472.
Palker,T.J., Scearce,R.M., Hensley,L.L., Ho,W. and Haynes,B.F. (1985) In Reinherz,E.L., Haynes,B.F., Nadler,L.M. and Bernstein,I.D. (eds), *Leukocyte Typing II.* Springer-Verlag, New York, pp. 303–313.
Raftery,M.J., Lang,C.J., Ivory,K., Tidman,N., Sweny,P., Fernando,O.N., Moorhead,J.F. and Janossy,G. (1985) *Transplant. Proc.*, **17**, 2737–2739.
Ramakrishnan,S., Uckun,F.M. and Houston,L.L. (1985) *J. Immunol.*, **135**, 3616–3622.
Rose,J.K., Adams,G.A. and Gallione,C. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2050–2054.
Sandrin,M.S., Vaughan,H.A., Lovering,K.E., Curnow,K., Deacon,N.J. and McKenzie,I.F.C. (1987) In McMichael,A.J. *et al.* (eds), *Leukocyte Typing III.* Oxford University Press, Oxford, in press.
Seed,B. and Aruffo,A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 3365–3369.
Shaw,S. (1987) *Immunology Today*, **8**, 1–3.
Sutherland,D.R., Rudd,C.E. and Greaves,M.F. (1984) *J. Immunol.*, **133**, 327–333.
Thurlow,P.J., Lovering,K.E. and McKenzie,I.F.C. (1984) *Transplantation*, **38**, 143–147.
Tse,A.G.D., Barclay,A.N., Watts,A. and Williams,A.F. (1985) *Science*, **230**, 1003–1008.
van Dongen,J.J.M., Hooijkaas,H., Comans-Bitter,M., Hählen,K., de Klein,A., van Zanen,G.E., van't Veer,M.B., Abels,J. and Benner,R. (1985) *J. Immunol.*, **135**, 3144–3150.
von Heijne,G. (1986) *Nucleic Acids Res.*, **14**, 4683–4690.
Williams,A.F. (1987) *Immunology Today*, in press.
Williams,A.F. and Gagnon,J. (1982) *Science*, **216**, 696–703.

Received on July 8, 1987; revised on August 21, 1987