Defining the repeating elements in the cysteine-rich region (CRR) of the CD18 integrin β2 subunit

Suet-Mien Tan[a], Susannah E. Walters[a], Elizabeth C. Mathew[a,1], Martyn K. Robinson[b], Karel Drbal[c], Jacqueline M. Shaw[a], S.K. Alex Law[a,*]

[a] *The Medical Research Council Immunochemistry Unit, Department of Biochemistry, University of Oxford, Oxford, UK*
[b] *Celltech plc, Slough, UK*
[c] *Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Prague, Czech Republic*

Received 16 May 2001; revised 2 August 2001; accepted 2 August 2001

First published online 14 August 2001

Edited by Veli-Pekka Lehto

Abstract The cysteine-rich region (CRR) of the integrin β subunits is organised into four repeating elements. By expression of a panel of truncated β2 subunits, and CRR segments fused to the C-terminal end of a CD4 soluble fragment, the segment required for the expression of two monoclonal antibody conformational epitopes was determined. This segment, E482–Q574, contains 16 cysteines representing two repeating units. We have thus defined the CRR unit motif of 'xC—C—C—CxCxxCxC—Cx', where 'x' represents a single residue, and '—' represents a stretch of four to 14 residues. © 2001 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.

Key words: Integrin; β Subunit; Cysteine-rich region

1. Introduction

The integrins are a family of cell adhesion molecules that mediate a wide repertoire of biological functions [1]. The integrins are obligate heterodimers, each composed of an α subunit and a β subunit. Both the α and β subunits are type I membrane proteins with relatively large extracellular domains (~700–1000 residues). Analyses of the primary structure of the α subunits showed seven repeating elements each of about 60 residues at the N-terminal. They have been predicted to assume the fold of a β propeller [2]. Nine of the 18 α subunits have an I-domain inserted between the second and third blade of the proposed β-propeller structure. The I-domains of the α2, αL, and αM subunits have been crystallised and shown to have a dinucleotide binding domain fold. In these integrins, the I-domains have an important role in binding to ligands [1]. The structure of the remaining extracellular domains of the α subunits was predicted to form four to six Ig-like domains [3]. The extracellular domains of the β subunits are organised linearly, from the N-terminal, a plexin–semaphorin–integrin domain [4], a spacer segment, a highly conserved region which has been predicted to assume an I-domain-like fold [5,6], a mid-region, and the major cysteine-rich region (CRR). Perhaps the most outstanding feature of the β subunit is the high level of cysteine residues arranged in a very conserved and characteristic pattern. The CRR has been postulated to contain three or four repeating elements, each containing eight cysteine residues [7,8].

The domain organisation of the CRR remains unclear. Primary structure analyses suggested that each repeating element in the CRR should contain eight cysteines. However, it is not possible to determine which sliding segment with eight cysteines should constitute a repeating unit. Thus, various assignments have been proposed (for example, see [7–12]). The gene of the β2 integrin contains 16 exons, with the CRR encoded in exons 11–14 [13]. However, the size of the exons are irregular, and the exon boundaries do not provide any information regarding the postulated repeating structures in the CRR. In this article, we describe the experiments leading to the determination of the repeating units. We make use of our earlier observation that the series of β2 C-terminal truncated variants can support LFA-1 and Mac-1 expression and function [14]. The shortest of these variants was β2-C445*, in which the codon for the cysteine residue at position 445 was converted to a stop codon. Thus, we employ the strategy of constructing truncated variants of the β2 of increasing length from β2-C445*. Since the cysteine residues hold the key in defining the domains in this region, the codons for the cysteine residues were systematically converted into stop codons. The cDNA constructs were transfected into COS-7 cells and the expression of the epitopes of two monoclonal antibodies (mAbs), KIM127 and MEM-48, was monitored. The epitopes of these mAbs had been mapped to the CRR of the integrin β2 subunit [11,12,15–17], and had been shown to be conformational. In particular, they fail to recognise the denatured β2 subunit under reducing conditions by Western blot analyses [15,16]. The rationale of this strategy is based on the argument that the expression of each of these epitopes is possible only when a complete repeating element is included in the construct, and only when the appropriate disulphide bonds are formed. After the C-terminal ends for the repeating units were identified, truncated constructs were made from the N-terminal. They were linked to the C-terminal of a construct encoding the Ig domains 3 and 4 of rat CD4. The fusion proteins were expressed, and the appearance of the epitopes for the two mAbs was monitored. The repeating units in the CRR were thus defined.

*Corresponding author. Fax: (44)-1865-275729.
E-mail address: alaw@molbiol.ox.ac.uk (S.K.A. Law).

[1] Present address: Microbiology and Immunology, University of California, San Francisco, CA, USA.

Abbreviations: mAb, monoclonal antibodies; CRR, cysteine-rich region; CRR1–4, cysteine-rich repeats 1–4

PII: S0014-5793(01)02778-8

2. Materials and methods

The mAbs MEM-148 [18], KIM127 [15] and MEM-48 [16] were described previously. MHM24 [19] was a gift from A.J. McMichael (John Radcliffe Hospital, Oxford, UK). KB43 [20] was a gift from K. Pulford (Leukaemia Research Fund Immunodiagnostic Unit, John Radcliffe Hospital, Oxford, UK). The mAb OX68 is against the Ig domains 3 and 4 of rat CD4, and is provided by M. Puklavac (Sir William Dunn School of Pathology, Oxford, UK). The cDNA construct for Ig domains 3 and 4 of rat CD4 [21] was a gift from M.H. Brown (Sir William Dunn School of Pathology, Oxford, UK).

The cDNAs of the integrin β2 subunit truncation variants were constructed with standard molecular biology techniques as previously described [14]. The cDNA for the CD4 domains was introduced into pcDNA3.1/Zeo(+) vector (Invitrogen, Groningen, The Netherlands). β2 Integrin fragments were synthesised by PCR and inserted into the *Eco*RI and *Bam*HI sites in the CD4 expression construct [21]. All constructs were verified by sequencing (DNA Sequencing Facility, Department of Biochemistry, University of Oxford, Oxford, UK).

The constructs were transfected into COS-7 cells as previously described [14]. To detect intracellular proteins, cells were harvested and washed once with phosphate-buffered saline (PBS). Cells were fixed with 3.7% paraformaldehyde (w/v) in PBS for 10 min at room temperature. Fixed cells were washed once with PBS and incubated in CSK buffer (100 mM NaCl, 300 mM sucrose, 10 mM PIPES, 3 mM $MgCl_2$, 1 mM EGTA, pH 6.8) containing 0.25% (v/v) Triton X-100, 0.5 mM phenylmethylsulfonyl fluoride, and 2.5 mM iodoacetamide for 2 min at room temperature. Permeabilised cells were stained with the mAbs of interest as described before [14].

Immunoprecipitation of metabolically labelled proteins from tissue culture supernatants was carried out as previously described [14].

3. Results

3.1. Progressive appearance of mAb epitopes in truncated β2 variants

Previously, we found that conversion of the codon for C445 into a stop codon [14] results in the expression of the shortest

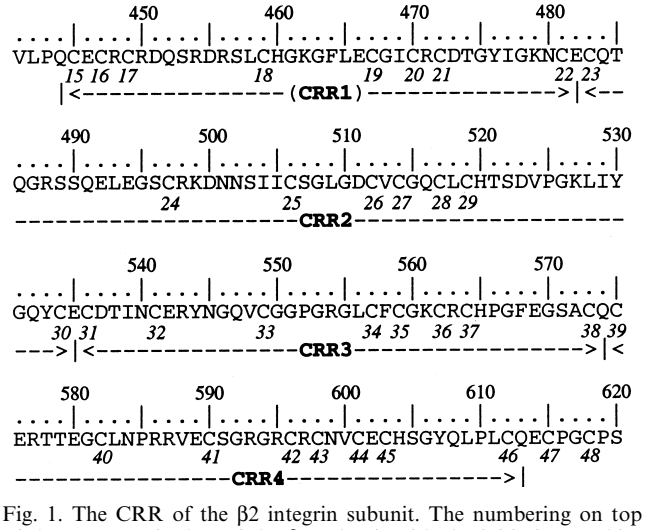

Fig. 1. The CRR of the β2 integrin subunit. The numbering on top of the sequence is that of the β2 subunit with the initiation methionine as '1'. The numbering below the sequence only counts the cysteines. Based on the experiments described in this article, the four repeats are shown as CRR1, CRR2, CRR3, and CRR4. The pseudo-repeat CRR1 is shown in brackets. Since the pattern of cysteine residues are very conserved and characteristic, we find it more convenient to refer to the extracellular cysteines *C1* to *C56* from the N-terminus [22], particular to those who work on integrin β subunits other than the one under study. The numbers are in *italics* under the main amino acid sequence.

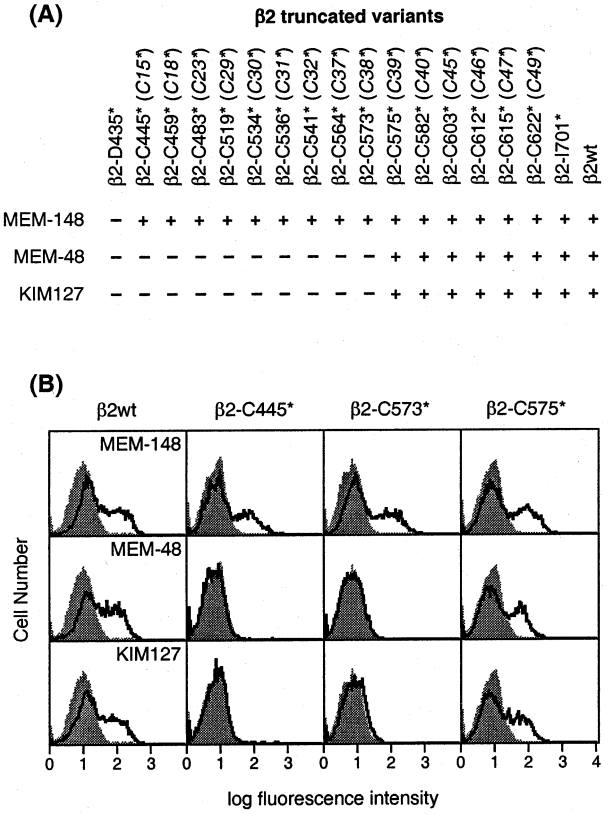

Fig. 2. The truncated variants of the β2 integrin subunit were transfected into COS-7 cells, and the intracellular expression of the epitopes of the mAbs MEM-148, KIM127 and MEM-48 were monitored by flow cytometry. A summary of the results is shown in (A). The residues at which stop codons were introduced are given both as the residue on the β2 sequence and, in the cases where the residues are cysteines, the numbering counting only cysteines are also included in brackets. The key histogram plots are shown in (B). Whereas the KIM127 and MEM-48 epitopes are not expressed in the transfectants of truncated variant β2-C573*, they are expressed in the transfectants of β2-C575*.

β2 truncated variant that can support surface expression of a functionally active LFA-1. In this variant, only the epitopes of mAb mapped to the mid-region, such as the mAbs H52 and MEM-148, were expressed. In this work, we introduced stop codons in the CRR to generate a series of β2 truncation variants of increasing length (for the β2 CRR sequence, see Fig. 1). The cDNA of these variants was transfected into COS-7 cells. Since these variants were not expressed on the cell surface without an integrin α subunit, we adopted the strategy of first fixing the cells, and then analysing for expression of the mAb epitopes in the intracellular pool by flow cytometry. The results are shown in Fig. 2A. The epitopes for KIM127 and MEM-48 appear on β2-C575* and longer variants, but not on β2-C573* and shorter variants. In addition, we have constructed the variant β2-D435*, which fails to express the MEM-148 epitope. The key histograms are shown in Fig. 2B.

3.2. Expression of discrete CRR domains as fusion proteins to CD4

The above results demarcate the C-terminal boundary, at C575 of the β2 subunit, required for the expression of the

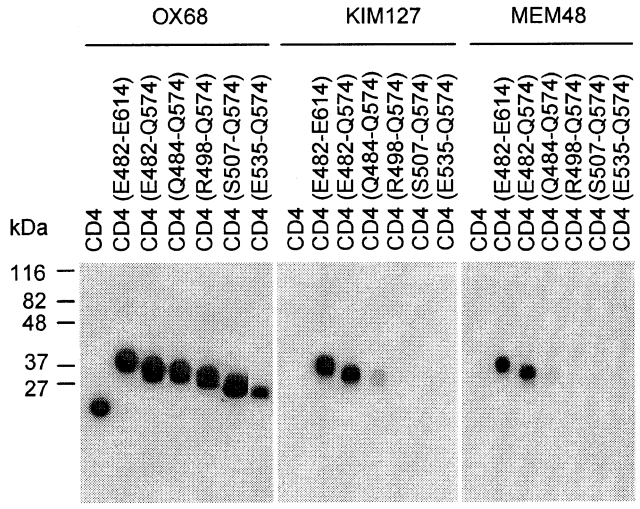

Fig. 3. The expression of CD4-CRR fusion proteins. COS-7 transfectants were metabolically labelled with [^{35}S]Met/Cys. Tissue culture supernatants were subjected to immunoprecipitation with OX68 (against rat CD4), and KIM127 and MEM-48 (against the CRR of the β2 integrin subunit).

KIM127 and MEM-48 epitopes. To determine the N-terminal boundary, we made a series of constructs, having different N-termini, but with the stop codon at C575, and ligated them to the C-terminal of the third and fourth Ig domain of rat CD4. These constructs were transfected into COS-7 cells and analysed, first by intracellular staining as above, followed by immunoprecipitation of secreted fusion proteins after metabolic labelling. Since the results are identical, only the immunoprecipitation results are shown. The supernatants were subjected to immunoprecipitation with OX68 against the CD4 domains, and the two mAbs KIM127 and MEM-48 (Fig. 3). OX68 precipitated the series of CD4-CRR fusion proteins, indicating that they are all secreted into the tissue culture medium. CD4(E482–Q574) and CD4(E482–E614) were also precipitated with KIM127 and MEM-48. However, if the two N-terminal amino acids, namely E482 and C483, were removed from CD4(E482–Q574) to give CD4(Q484–Q574), only weak bands were detected in the immunoprecipitate with KIM127 and MEM-48. Since a full band was seen with OX68, the inefficiency of precipitation with KIM127 and MEM-48 suggested the expression of defective or incomplete epitopes.

3.3. MEM-48 and KIM127 are two distinct mAbs

Although the mAbs MEM-48 and KIM127 give identical results in this series of analyses, they can be distinguished in the following experiment. Whereas the epitope of MEM-48 is expressed at 4°C on the LFA-1 of MOLT-4, that of KIM127 requires incubation at 37°C in the presence of 5 mM $MgCl_2$ and 1.5 mM EGTA (Fig. 4). This is in agreement with previous reports that the expression of the KIM127 epitope is temperature-sensitive and can be activated with the mAb KIM185 [23,24].

4. Discussion

Although it is documented that CRR of the integrin β subunits contains three or four repeating elements, the unit element remains undefined. In this study, we employed the strategy of monitoring the expression of the epitopes of two mAbs in segments of the CRR. A series of truncated β2 subunits were constructed. The epitopes recognised by the mAbs KIM127 and MEM-48 were expressed in the β2-C575* variants and those that were longer, but not in the β2-C573* and those that were shorter. Thus, a boundary has been determined to lie between the residues C573 and C575. A series of rat CD4 fusion proteins were constructed with the different lengths of β2 integrin segments ending at C575. We established that the segment inclusive of E482 and Q574 of the integrin β2 subunit is needed for the expression of the KIM127 and MEM-48 epitopes. The epitopes are also expressed on the extended segment to end at E614, but their expression is severely compromised if the segment starts at Q484 instead of E482.

The minimum construct that expresses the epitopes of KIM127 and MEM-48 is inclusive of residues E482 and Q574. The disulphide bonds in this segment must be formed correctly since both epitopes are lost upon reduction [15,16]. At the onset of the experiments, we argued that only when complete units are included in the construct would they fold correctly. By alignment analyses, each repeating unit should contain eight cysteines [7,8]. The experimental result is therefore consistent with the preset conditions because the segment E482–Q574 contains 16 cysteines, and therefore represents two complete repeating units. Accordingly, three boundaries are defined at E482, E535, and Q574. Each of the two units has the same arrangement of cysteines with the pattern: 'xC—C—C—CxCxxCxC—Cx', where '—' represent segments of variable lengths from four to 14 residues, and 'x' represents a single residue (it should be noted that the residue before the first cysteine and the residue after the last cysteine are always included in the expression constructs). Another repeating unit is found in the segment C-terminal to Q574, i.e. Q574–E614, which has the eight cysteines arranged in an identical pattern. We have previously determined the C-terminal residue of the mid-region of the β2 subunit at Q444 [12]. The segment C455–E482 is therefore bound by the mid-region at the N-terminal side, and the cysteine repeating unit E482–E535 on the C-terminal side. Eight cysteines are also found in this segment and they are arranged in a pattern similar, but not identical, to the other three repeating units. It is nonetheless convenient, at this point, to refer to it as cysteine-rich region repeat 1

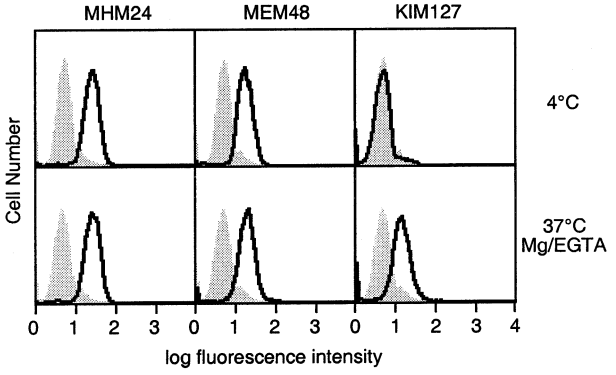

Fig. 4. The distinction between KIM127 and MEM-48. MOLT-4 cells were stained with the two mAbs in 10 mM HEPES, 150 mM NaCl, and 0.2% (w/v) bovine serum albumin, pH 7.5, at 4°C or at 37°C, supplemented with 5 mM $MgCl_2$ and 1.5 mM EGTA. The epitope of KIM127 is not expressed at 4°C. The expression of LFA-1 is shown with the mAb MHM24 [19]. The background is obtained with the mAb KB43 against the αXβ2 integrin [20].

(CRR1) spanning the region C445–E482. Thus, CRR2 will span the region E482–E535, CRR3 E535–Q574, and CRR4 Q574–E614 (see Fig. 1).

The CRR repeating elements, before being defined in this work, have been assumed to be epidermal growth factor (EGF)-like [1]. Indeed, a protein TIED has been described to have ten integrin eGF-like domains with no other features [25]. The boundaries of each repeat was offset by one cysteine as those assigned here, i.e. each repeat would start with the second cysteine and end with the first cysteine in the subsequent repeat. Curiously, the first cysteine in the entire sequence was ignored, and the final repeat has only seven cysteines. Thus, if the 10 units were assigned starting with the N-terminal of the protein, such that each unit would have eight cysteines, the pattern of cysteines in all 10 units would be identical to those in CRR2, CRR3 and CRR4.

Although the cysteine residues are highly conserved in the integrin β subunits, the conservation is not absolute. Two cysteines in the core of the CRR are absent in the human β4 and β8 subunits. They are equivalent to C536 and C559 of the human β2 subunit (see Fig. 1). It was proposed that when a pair of cysteines are missing in homologous proteins, they are likely, when present, to form a disulphide bond [26]. When applied to the CRR of the integrin β4 and β8 subunits, both missing cysteines would be found in CRR3 according to the proposed assignment herein. Thus, CRR3 of the integrin β4 and β8 subunits would remain as self-contained structures, each having six cysteines engaged in three disulphide bonds. Pursuing the same argument, the two cysteine residues, when present in an integrin subunit such as β2, should engage in a disulphide bond. The conjecture will require future experiments for verification.

The epitopes of the mAbs KIM127 and MEM-48 are not identical (Fig. 4). Using human–mouse chimeras of the β2 integrin subunit, the species-specific residues contributing to the KIM127 epitope have been identified to G526, L528 and Y530 in CRR2, and those to MEM-48 to L556, F558, and H565 in CRR3 [12]. However, the epitope of KIM127 is not expressed in the β2 truncated protein with a complete CRR2 (β2-C536*), a complete CRR3 (β2-C575*) is also required (Fig. 2). Conversely, the epitope of MEM-48 is not expressed in the CD4 fusion protein containing CRR3 only (E535–Q574), the inclusion of a complete CRR2 (E482–E535) is also required. Taken together, these results suggest that either (1) residues common to both human and mouse in CRR3 also contribute to the expression of the full epitope of KIM127, and those in CRR2 to the expression of MEM-48 or (2) the two CRR repeats require each other to fold correctly for the expression of the two mAb epitopes. These alternatives cannot be distinguished at present.

In conclusion, we have determined the boundaries of the CRR repeats in the integrin β2 subunit, and shown that the double repeats of CRR2–CRR3 can be expressed as CD4 fusion proteins independent of other integrin structures. This finding is likely to be applicable to other integrin β subunits, and will facilitate the design of future experiments in the analyses of the structure of the CRR repeating units.

Acknowledgements: We would like to thank Dr Marion Brown for providing the cDNA for Ig domains 3 and 4 of rat CD4 and the advice for constructing the fusion proteins, and Vaclav Horejsi for comments on the manuscript. This work was supported in part by Grants L0509 and L0524 from the Arthritis Research Campaign, UK, and a grant from the Sir Samuel Scott of Yews Trust, UK.

References

[1] Humphries, M.J. (2000) Biochem. Soc. Trans. 28, 311–339.
[2] Springer, T.A. (1997) Proc. Natl. Acad. Sci. USA 94, 65–72.
[3] Lu, C., Oxvig, C. and Springer, T.A. (1998) J. Biol. Chem. 273, 15138–15147.
[4] Bork, P., Doerks, T., Springer, T.A. and Snel, B. (1999) Trends Biochem. Sci. 24, 261–263.
[5] Tuckwell, D.S. and Humphries, M.J. (1997) FEBS Lett. 400, 297–303.
[6] Huang, C., Zang, Q., Takagi, J. and Springer, T.A. (2000) J. Biol. Chem. 275, 21514–21524.
[7] Kishimoto, T.K., O'Connor, K., Lee, A., Roberts, T.M. and Springer, T.A. (1987) Cell 48, 681–690.
[8] Law, S.K.A., Gagnon, J., Hildreth, J.E.K., Wells, C.E., Willis, A.C. and Wong, A.J. (1987) EMBO J. 6, 915–919.
[9] Tamkun, J.W., DeSimone, D.W., Fonda, D., Patel, R.S., Buck, C., Horwitz, A.F. and Hynes, R.O. (1986) Cell 46, 271–282.
[10] Calvete, J.J., Henschen, A. and Gonzalez-Rodriguez, J. (1991) Biochem. J. 274, 63–71.
[11] Zang, Q. and Springer, T.A. (2001) J. Biol. Chem. 275, 6922–6929.
[12] Lu, C., Ferzly, M., Takagi, J. and Springer, T.A. (2001) J. Immunol. 166, 5629–5637.
[13] Weitzman, J.B., Wells, C.E., Wright, A.H., Clark, P.A. and Law, S.K.A. (1991) FEBS Lett. 294, 97–103.
[14] Tan, S.M., Hyland, R.H., Al-Shamkhani, A., Douglass, W.A., Shaw, J.M. and Law, S.K.A. (2000) J. Immunol. 165, 2574–2581.
[15] Stephens, P., Romer, J.T., Spitali, M., Shock, A., Ortlepp, S., Figdor, C.G. and Robinson, M.K. (1995) Cell Adhes. Commun. 3, 375–384.
[16] Bazil, V., Stefanova, I., Hilgert, I., Kristofova, H., Vanck, S. and Horejsi, V. (1990) Folia Biol. Praha 36, 41–50.
[17] Huang, C., Lu, C. and Springer, T.A. (1997) Proc. Natl. Acad. Sci. USA 94, 3156–3161.
[18] Drbal, K., Angelisova, P., Cerny, K., Pavlistova, D., Cebecauer, M., Novak, P. and Horejsi, V. (2000) Biochem. Biophys. Res. Commun. 275, 295–299.
[19] Hildreth, J.E.K., Gotch, F.M., Hildreth, P.D.K. and McMichael, A.J. (1983) Eur. J. Immunol. 12, 202–208.
[20] Mathew, E.C., Shaw, J.M., Bonilla, F.A., Law, S.K.A. and Wright, D.A. (2000) Clin. Exp. Immunol. 121, 133–138.
[21] Brown, M.H. and Barclay, A.N. (1994) Protein Eng. 7, 515–521.
[22] Nolan, S.M., Mathew, E.C., Scarth, S.L., Al-Shamkhani, A. and Law, S.K.A. (2000) FEBS Lett. 486, 89–92.
[23] Robinson, M.K., Andrew, D., Rosen, H., Brown, D., Ortlepp, S., Stephens, P. and Butcher, E.C. (1992) J. Immunol. 148, 1080–1085.
[24] Andrew, D., Shock, A., Ball, E., Ortlepp, S., Bell, J. and Robinson, M.K. (1993) Eur. J. Immunol. 23, 2217–2222.
[25] Berg, R.W., Leung, E., Gough, S., Morris, C., Yao, W.P., Wang, S.X., Ni, J. and Krissansen, G.W. (1999) Genomics 56, 169–178.
[26] Kreisberg, R., Buchner, V. and Arad, D. (1995) Protein Sci. 4, 2405–2410.