

Molecular cloning and sequencing of the low-affinity IgE receptor (CD23) for horse and cattle

Johanna L. Watson^{a,*}, Kenneth A. Jackson^b,
Donald P. King^b, Jeffrey L. Stott^b

^a*Department of Medicine and Epidemiology, School of Veterinary Medicine,
University of California, Davis, CA, USA*

^b*Department of Pathology, Microbiology and Immunology, School of Veterinary Medicine,
University of California, Davis, CA, USA*

Received 1 October 1999; received in revised form 6 December 1999; accepted 16 December 1999

Abstract

Expression of the low-affinity IgE receptor (CD23) on the surface of mononuclear cells is a critical event in the development of IgE-mediated immunologic responses. Using PCR and cDNA library screening, a 2.7 kb cDNA encoding equine CD23 and a 627 bp PCR fragment of cattle CD23 were sequenced and characterized. The equine CD23 sequence encodes a complete and continuous open reading frame of 327 amino acids, while the shorter cattle fragment encodes 209 amino acids corresponding to nucleotides 325–1098 of the equine CD23 transcript. In addition to high identities in their nucleotides and translated amino acids with CD23 sequences published for other species, the translated equine CD23 protein also shares many of the structural features of this molecule described for human and rodents. Interestingly, an additional repetitive element of possible functional significance consisting of 18 amino acids, located between the transmembrane region and the carbohydrate-binding lectin domain of horse CD23, was also identified. Based on these sequences, molecular assays will be developed to measure CD23 mRNA in tissues and expression of recombinant proteins will be essential to the production of species-specific antibody reagents. These assays and reagents will be useful in future studies of allergic and lymphoproliferative diseases in horses and cattle. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: CD23; IgE receptor; IgE; Allergy; Cattle; Horse

* Corresponding author. Fax: +1-530-752-0414.

E-mail address: jlwatson@ucdavis.edu (J.L. Watson).

1. Introduction

CD23 (Fc ϵ RII), the low-affinity receptor for IgE, is expressed on the surface of B-cells prior to isotype switching, as well as on monocytes, eosinophils and platelets. This type II transmembrane glycoprotein is a member of the C-type lectin family. CD23 plays an important role in the differentiation and activation of B lymphocytes and is upregulated, or induced, by interleukin 4 (IL-4) (Conrad, 1990). Expression of CD23 is thought to play a central role in IgE-associated immune responses, including allergies (Yu et al., 1994; Fugiwara et al., 1994) and in B lymphocyte neoplasia (Fournier et al., 1994). CD23 can be induced on human alveolar macrophages and has been associated with induction of allergic asthma (Williams et al., 1992). Horses suffer from a number of allergic conditions in which CD23 has the potential to play a major role in the etiology, including anaphylaxis, recurrent urticaria and chronic obstructive pulmonary disease (COPD). The development of species-specific reagents for equine CD23 are necessary in order to better define the pathogenesis of these latter conditions. This paper describes the sequencing and characterization of the entire cDNA encoding equine CD23. An additional polymerase chain reaction (PCR) fragment of cattle CD23 was also cloned and sequenced and is included for comparative purposes.

2. Materials and methods

2.1. Polymerase chain reaction (PCR)

The cDNA encoding a fragment of horse and cow CD23 was amplified by PCR. Lambda DNA (Wizard lambda DNA purification system, Promega, Madison, WI) purified from either a horse mesenteric lymph node cDNA library (Watson et al., 1997) or a cattle thymic cDNA library (Custom Library, Stratagene, Carlsbad, CA) was used as a template for these PCRs. Fifty microliters PCR contained 50 mM KCl, 2 mM MgCl₂, 10 mM Tris-HCL pH 9.0, 0.2 mM of each dNTP, 40 pmol of each primer, and 0.5 units of Taq DNA polymerase (Promega, Madison, WI). A degenerate sense primer, 5' CTG CTT CT(CT) CTG TGG CAC T 3' and an antisense primer, 5' GCA TCA TCA CGC AGT CCT C 3', designed from conserved regions of the mouse, rat and human CD23 mRNA sequences, generated products of 775 bp and 629 bp for the horse and cow respectively. These PCR products were cloned into pGEM-T-easy (Promega, Madison, WI) and were subsequently transfected into DH5 α cells (Gibco BRL, Gaithersburg, MD). Plasmid was purified using QIAprep spin minikit (Qiagen, Valencia, CA) and presence of the inserts of appropriate size were verified by restriction digests. Three independent plasmid clones containing the PCR insert were sequenced.

2.2. Library screening

In order to obtain a full length transcript for the horse, the mesenteric lymph node cDNA library was subsequently screened using the equine CD23 PCR fragment. The probe was prepared from the pGEM-T-easy plasmid by EcoRI digestion and gel

purification, and was labeled with [α - ^{32} P]dCTP (Hexanucleotide labeling system, Boehringer Mannheim, Indianapolis, IN). Plaque hybridization was carried out for 18 h at 42°C in 5X SSPE, 5X Denhardt's, 1% sodium dodecyl sulfate (SDS), 50% formamide, and 400 μ g/ml denatured herring sperm DNA. Filters (Immobilon Ny+, Millipore, Bedford, MA) were washed twice with 2X SSPE 0.5%, once with 0.5X SSPE, and once with 0.1X SSPE. All wash solutions contained 0.5% SDS. Secondary and tertiary screenings were performed using identical conditions to select the positive plaques. Inserts were subcloned into the Bluescript phagemid vector by *in vivo* excision with the Ex Assist helper phage (Stratagene, Carlsbad, CA). Inserts corresponding to two independent primary plaque isolates were sequenced (in triplicate) by chain termination.

2.3. DNA sequencing and analysis

Cloned PCR products and isolated cDNA library clones were sequenced by the chain termination method (Sanger et al., 1977) using an automated sequencer with dye terminators (Prism, Perkin–Elmer, Norwalk, CT). DNA sequences were assembled and analyzed with the University of Wisconsin Genetics Group (GCG) program.

3. Results

The final 2.67 kb sequence of horse CD23 encoding an open reading frame of 327 amino acids is shown in Fig. 1a. The 629 bp cDNA fragment obtained for cattle, encoding 209 amino acids corresponding to nucleotides 325–1098 of the equine CD23 transcript, is shown in Fig. 1b. Comparisons calculated using the BEST FIT program of GCG (University of Wisconsin, Genetics Computer Group) demonstrate that the horse nucleotide sequence has identities of 78 and 76% with human and mouse CD23 sequences, respectively. The predicted amino acid sequence is 64 and 59% identical and 70 and 65% similar to the human and mouse translated protein, respectively. Similar analysis of the cattle CD23 shows that when compared with analogous regions of horse, human and mouse CD23, the sequence exhibits identities of 82, 76 and 72%, respectively, at the nucleotide level, and 75, 63 and 59%, respectively, at the amino acid level. There are two potential *N*-glycosylation sites (amino acid residues 101 and 194) in the equine CD23 sequence, both of these are positionally conserved with the cattle sequence, although only one is conserved with human CD23 (Fig. 2).

4. Discussion

Analysis of the predicted amino acid sequence for horse CD23 highlights a transmembrane region and a lectin-like domain in analogous positions to the human, mouse and rat CD23 protein (Fig. 2). The shorter cattle fragment spans much of the lectin-like domain. Interestingly, this lectin-like domain shares homology with other animal lectins, such as human and rat asialoglycoprotein receptors, rat and chicken cartilage proteoglycans, human and canine pulmonary surfactant apoproteins, and

(a)

```

1  ggcacgaggg  tttagacctgt  caccacggcc  tgtggcccct  gccggcgctg  cccaacggcc  ccaacagact
71  agggaggaga  cgacgcgctc  gaaactcata  tcccagatcc  cagccagagc  tgtgactgtg  ccagtgcaat
141 ggactgcctt  gacaggtggg  gagcactctg  tcactgggag  aatccaagca  gggctggcAT  GGAGGAACAT
                                     M   E   E   H
211 GCATCTCAG  AGTTCCCAAA  GTTTCACAG  AGGCGCGCG  AGTGTGCGCA  GCGGGGAGCG  CTGCTCGCAC
    A Y S      E F P K      F P R      R R R      Q C C Q      R G T      L L A
281 TGCTGGGACT  GGTGACCACC  GCGATGTGGG  CCGGGCTGCT  GACCTGCTCT  CTCTCTGGCG  ACTGGGAGAG
    L L G L      V T T      A M W      A G L L      T L L      L L W      H W E T
351 TGTGCAGAA  CTGAACAGCG  TGGAAAGAC  CGCGGCCAG  AACCACTTGG  AGACTGTGCA  GAAGCTGAAA
    V Q K      L K Q      L E K T      A A Q      N D L      E T V Q      K L K
421 CAGCTGGAAA  AGACCGCGCG  CAGAAGCAG  TTGGAGACTG  TGCAGAAAGCT  GAAACAGCTG  GAAAAGACCG
    Q L E      K T A      A Q N D      L E T      V Q K L      K Q L      E K T
491 CGGCCAGAAA  CGTCTCTCAG  GTTTCACAG  ACTTGGAAAA  ACACAAGGGG  GAGCAGGTGG  CCGAGGAATC
    A A Q N      V S Q      V S K      D L E K      H K G      E Q V      A Q E S
561 CCAGGCTGCC  AAGATGTCCG  AGAGCATGGA  GAAATCCAA  GATGACGAGA  ACAGAAATGA  ATCTCAGGAC
    Q A A      K M S      Q S M      K I Q      D E Q      N R M      K S Q D
631 TCGGAGCTCT  CCGCGAACTT  GAACCGACTT  CGCGAGGACC  TGAGCGACCT  CCAGTCCGAG  GGCTTGAATG
    S E L      S R L      N G L      R E D      L S D      L Q S      Q G L N
701 AGAGACGCGC  AGCCTTGAC  GCACTGGACA  GGCTCCAGGA  GGAAGTGGCG  AAGCTGTGGA  TAGAGTACAG
    E R R      A L H      A L D      R L Q      E V A      K L W      I E L R
771 CGTGTCCAAC  GGCTCCACGT  GTAACACGTG  CCCGACGAC  TGGGTCCATT  TCCAAAAGAA  GTGCTACTAC
    V S N      G S T      C N T      C P D      D W V      H F Q      K K C Y Y
841 TTCGGGGAGG  GCGCCCAAGC  GTGGATCCAG  GCGCGGTTCG  CTGCGACGAA  GTGCGACGGG  CGGCTGGTCA
    F G E      G P K      R W I      Q A R      F A C      S K      L H G      R L V
911 GCATCCACAG  CCAGGAGGAG  CAGGACTTCC  TGACCCGCGA  CGCCACSAK  AAAGGCTCCT  GGATTCGCCCT
    S I H      S Q E      E Q D      F L T      R H      A N K      K G S      W I G L
981 TCGGGACCTG  AACATAGAAG  GGGAGTTTGT  CTGGATGGAC  CAGAAACCCC  TGGACTACAG  CAACTGGCAG
    R D L      N I E      G E F      V W M      D Q N      P L D      Y S N W Q
1051 CCGGGGAGCG  CCAACAACGG  GGGCCAAGCG  GAGGACTGCG  TCATGTGTGCG  GGCCTCCGGG  CACTGGAAGC
    P G E      P N E      G Q G      E D C      V M M      R A S      G G      H W N
1121 ACGCCTTCTG  TGGCAGCTAC  CTGGACGGCT  GGGTGTGTGA  CCGGCTGGTC  ACGTGTGTA  tgccccaccg
    D A F      C G S      Y L D      G W V      C D      R L V      T C *
1191 ccacctcggc  cccctcgggt  ccacgggggt  ttgtcctgag  tcgtggaccc  tgacagcccc  ctgcccggcc
1261 accagagagc  cccctctgtg  tcaactcccc  atctcaaaaa  cgggaacacg  aaggccccatg  gcaggacctc
1331 gaggaccccc  gccaagggcg  tggacacctc  ttgtgtgggt  gaaggggtcct  ggtgatactt  cctacacccc
1401 aaatggaagc  agctgatacca  tccctcggctc  ctgccaaacgc  cccggaggggc  ttgtccctgt  ttctcaggcg
1471 caagctggcc  attgccaact  cccctgacttt  aaaggccttga  ggtatccccc  tcgagaaccc  ctgcccacga
1541 gcttgaagcg  tagcaccccc  ctctcctactg  gggcacagct  cctggcagcc  cagcgcgtca  gggagcgccc
1611 ccgaccctgt  tgttctctgt  cgccccgacc  tggcgtctgt  ctggagtcag  gaagtgaacc  ccaccctgac
1681 cctccccaga  cgccgcgcgt  gtccggggcac  cacacgacag  agtccccccg  ctaggagccc  agtctctccc
1751 acccgctctg  catcagtaag  catcagtaag  acagggcgac  cagccttgcc  cagcctctag  ggccctgacg
1821 gtgtggaagg  tgggagctgg  gaggccacag  gccccagcag  ggggtggggac  aggatgtggc  cctctggagc
1891 aaagatgggc  tggactccag  gtcacactgt  ctctctattc  actgcccacc  cctagatgtc  tgtgtcccc
1961 agctccccac  cccggaaact  ccggccagcg  caaaaactgcc  tacctagacc  ctccacactg  tattcgaggc
2031 tccccctetac  cagtcactca  gggcccgact  gggggcaccc  tggccacctc  gtccctcact  ttccccctcc
2101 aaaccataaa  ggagagccat  gggccgtggt  gtccagacac  attcagaatc  caccacactg  cttggcgtgg
2171 ccccatcctc  cctctccctc  ggtctccctc  ttcaactgcc  ccaaagtttc  tagcaatttc  ctttcaatat
2241 agaggaaaaa  ccaaagcctt  acagggtcct  ttgcgatctc  agaacccttc  cccagtgcc  tatctcccca
2311 acttcacttt  ctcttctctc  tccccgcacc  cggcacccac  cccagccacc  cctctcagta  tttaacattt
2381 aggctcactc  ccacccagag  cctcttgca  atgctgcctc  cctctgtagc  tctctcagta  ggcctatcct
2451 gagctctccc  tccctcccat  ttaaaatgac  agccctgact  ccaagtcctg  tttactcttc  tgtactctcc
2521 caccacggca  cccctctcct  tccgacaaga  tgtgtaatct  actgatctgt  tcaactgctc  tgtacacggg
2591 ttagaatgca  agcttcagga  gggcaggtgc  ctgcctttgt  ttggtttcta  gctgtcccca  gcgcctggca
2661 cacagtagct  cgtgccc

```

(b)

```

1  GGGAAATATGC  ACGAAATCTA  AAACAGCTGG  AGGAGACCGC  TGCCCTGAAT  GTCTCTCAGG  TGTCAAAGA
    E N A      R N L      K Q L      E E T      A A L      N V S      Q V S K D
71  CTTGGAAGA   CAAAAGGGTG  ACCAGATGGC  CCAGAAATCC  CAGGCTGCCC  AGATGATGCA  GGACATGGAA
    L E R      Q K G      D Q M      A Q K      S Q A      A G M      M Q D M E
141 CGAATCAAAA  CTGAACAGAA  GAGAATGGAA  TCTCAGGAGT  CTGAGCTCTC  CTGGAACCTG  GATGGTCTTC
    R I Q      T E Q      K R M      E S Q      E S E      L S W      N L D G G L R
211 GAGCTGACCT  GAGTGGACCT  AAGTCCCGCG  GCTTGAATGA  GATCGTCGAC  GCGTGTGATT  CACTGGGAAG
    A D L      S D L      K S R      G L N      E M R      Q A L      D S L G R
281 ACTCCAAGAG  GAGGTGGGAA  AGTTGTGGAT  CGAGCTACGC  GCTGCCAACG  GCTCTGTGTG  CAACACGTGC
    L Q E      E V G      K L W      I E L      R A A      N G S V C      N T C
351 CCCGAGGCAT  GGATCTATTT  CCAAAAGAAG  TGCTACTACT  TCGGGGAGCG  CGCCAAGAAA  TGGATCCAGG
    P E A      W I Y      F Q K      K C Y      Y F G      E G A K      K W I      Q A
421 CCCGGTACGC  CTGTGAAAAA  CTGCACGGGC  GGCTGGTTAG  CATCCACAGC  CCAGAGGAGC  AGGACTTCCT
    R Y A      C E N      L H G      R L V      S I H      S P E      E Q D F L
491 GACCAACGCT  GCCAATCGGA  GGGGCTCCTG  GATTGGCCTT  CGGACCTGCG  ACATTGAAAG  GGAGTTTATC
    T K R      A N W      R G S      W I G      L R D      L D I E      G E F I
561 TGGATGGACA  ACCAGCCCTT  GGACTATAGC  AACTGGCAGC  CAGGGGAGCC  CAACGATGCA  GGCCAGGCT
    W M D      N Q P      L D Y      S N W      Q P G      E P N      D A G      Q G

```

Fig. 1. (a) The complete nucleotide sequence of horse CD23 (GenBank accession No. AF141931). Predicted amino acid residues for the translated protein are shown and the stop codon (TGA) is underlined. The PCR fragment used as a probe to screen the horse library is shaded. (b) Nucleotide sequence and predicted amino acid sequence of a fragment of cattle CD23 (GenBank accession No. AF143722).

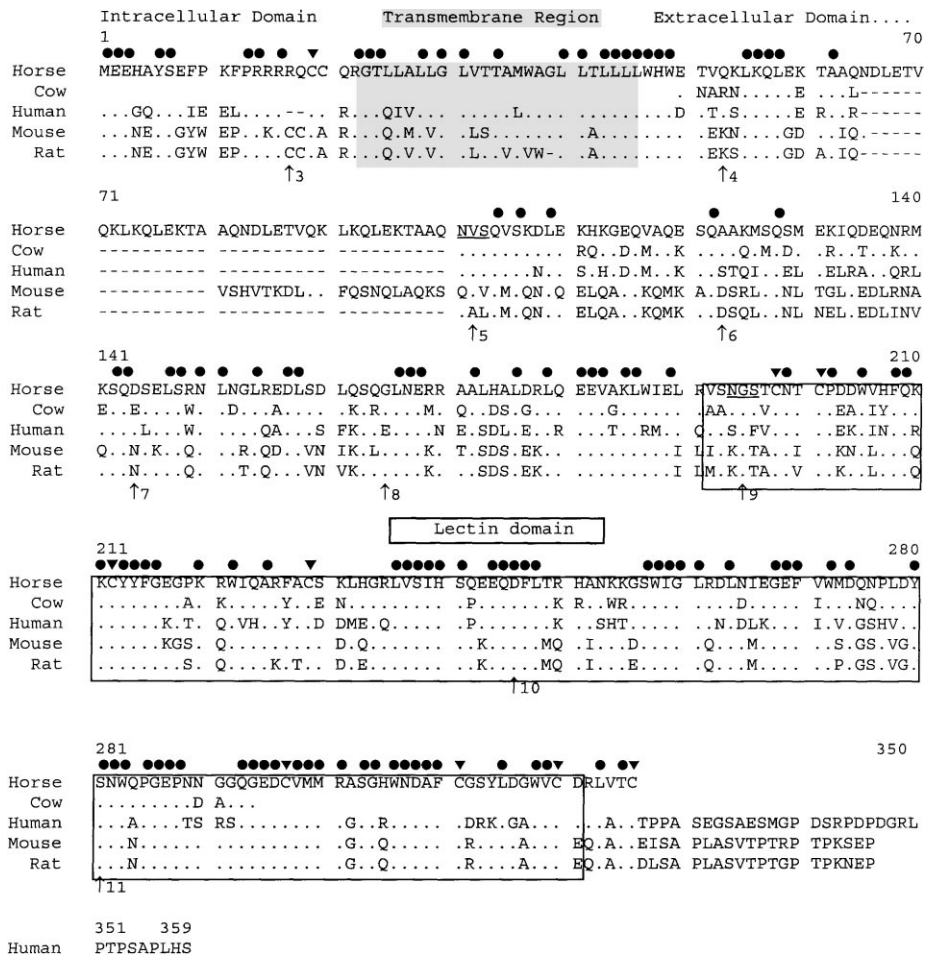


Fig. 2. Alignment of deduced amino acid sequences of horse, cattle, human (GenBank accession No. M15059, Ikuta et al., 1987), mouse (GenBank accession No. M99371, Bettler et al., 1989) and rat (GenBank accession No. X73579) CD23. Dots (·) represent amino acid residues that are conserved with the horse sequence. Filled triangles (▼) indicate conserved cysteine residues and filled circles (●) other invariant residues across all five species shown. Gaps were inserted to maximize alignment. Potential *N*-glycosylation sites present on the horse sequence are underlined. The lectin domain and the transmembrane region are identified as boxed and shaded text, respectively. The numbered arrows shown indicate exon arrangement for the human CD23 gene.

corresponds to the carbohydrate-binding domain of these proteins (Conrad, 1990; Suter et al., 1987). Analysis of the amino acid residues of CD23 for this region shows that 68/130 (52%) are invariant across all five species investigated to date (Fig. 2). In comparison, only 59/161 (37%) of the residues outside this lectin-like domain are conserved (unique repetitive elements in the equine and mouse CD23 sequence excluded for this analysis). This suggests that there is evolutionary pressure to maintain the structure of this biologically important region.

There are two isoforms of CD23 in the human; CD23a (constitutively expressed) and CD23b (IL-4 inducible). The presence of a conserved tyrosine residue at position 6 in the amino acid sequence of horse CD23 would appear to identify it as a homolog of human CD23a (Gordon, 1994). Human CD23 contains a repeat structure which makes up exons 5, 6 and 7. Each of these repeats are 21 amino acids long and terminate in SQ, and may have arisen from exon duplication (Suter et al., 1987). The mouse sequence appears to contain four such repeats. While the gene structure of horse CD23 has not been determined, it appears to be more similar to the human sequence containing three of these highly conserved repeats. Interestingly, the horse sequence contains two exact duplications of an 18 amino acid segment prior to these repeats which corresponds to exon 4 in the human sequence. This sequence repeat in the horse duplicates a highly conserved motif, -LKQL-, and appears to be between the transmembrane region and the lectin domain. This increases the number of heptadic leucine repeats available to function as a leucine-zipper and increases the length of this peptide which could form an α -helix coiled-coil. As a result, this segment may be functionally significant. The horse cDNA for CD23 terminates earlier than the human, mouse and rat sequences, and extends just five amino acid residues past the lectin domain.

In summary, using PCR and plaque screening, the entire cDNA for equine CD23 and a cDNA PCR fragment of cattle CD23 have been identified and characterized. In addition to their utility in the identification of cellular expression of CD23 by reverse-transcriptase PCR, these cDNA sequences will allow future recombinant protein expression approaches to be employed to raise CD23-species-specific antibody reagents. These reagents will be useful in future studies of allergic and lymphoproliferative diseases in horses and cattle.

Acknowledgements

This project was supported by the California Center for Equine Health and Performance with funds provided by the Oak Tree Racing Association, California satellite wagering fund, and contributions by private donors.

References

- Bettler, B., Hofstettler, H., Rao, M., Yokoyama, W.H., Kilchherr, F., Conrad, D.H., 1989. Molecular structure and expression of the murine lymphocyte low-affinity receptor for IgE (Fc epsilon RII). *Proc. Natl. Acad. Sci. USA* 86, 7566–7570.
- Conrad, D.H., 1990. Fc ϵ RII/CD23: the low affinity receptor for IgE. *Annu. Rev. Immunol.* 8, 623–645.
- Fournier, S., Rubio, M., Delespesse, G., Sarfati, M., 1994. Role for low-affinity receptor for IgE (CD23) in normal and leukemic B-Cell proliferation. *Blood* 84, 1881–1886.
- Fugiwara, H., Kikutani, K., Suematsu, S., Naka, T., Yoshida, K., Tanaka, T., Suemura, M., Matsumoto, N., Kojima, S., Kishimoto, T., Yoshida, N., 1994. The absence of IgE antibody-mediated augmentation of immune responses in CD23-deficient mice. *Proc. Natl. Acad. Sci. USA* 91, 6835–6839.
- Gordon, J., 1994. B-cell signalling via the C-type lectins CD23 and CD72. *Immunol. Today* 15, 411–417.
- Ikuta, K., Takami, M., Kim, C.W., Honjo, T., Miyoshi, T., Tagaya, Y., Kawabe, T., Yodoi, J., 1987. Human lymphocyte Fc receptor for IgE: sequence homology of its cloned cDNA with animal lectins. *Proc. Natl. Acad. Sci. USA* 84, 819–823.

- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- Suter, U., Bastos, R., Hofstetter, H., 1987. Molecular structure of the gene and the 5'-flanking region of the human lymphocyte immunoglobulin E receptor. *Nucleic Acids Res.* 15, 7295–7308.
- Watson, J.L., Pettigrew, H.D., Wilson, L.K., Gershwin, L.J., 1997. Isolation and partial expression of a cDNA encoding equine IgE heavy chain and production of polyclonal antibodies to the recombinant protein. *J. Vet. Allergy Clin. Immunol.* 5, 135–142.
- Williams, J., Johnson, S., Mascali, J.J., Smith, H., Rosenwasser, L.J., Borish, L., 1992. Regulation of low affinity IgE receptor (CD23) expression on mononuclear phagocytes in normal and asthmatic subjects. *J. Immunol.* 149, 2823–2829.
- Yu, P., Kosco-Vilbois, M., Richards, M., Kohler, G., Lamers, M., 1994. Negative feedback regulation of IgE synthesis by murine CD23. *Nature* 369, 753–756.