Two vaccinia virus proteins structurally related to the interleukin-1 receptor and the immunoglobulin superfamily

Geoffrey L. Smith1* and Y. Sang Chan2

1Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE and
2Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, U.K.

The structures of two vaccinia virus genes (B15R and B18R) from near the right inverted terminal repeat are described. These genes encode proteins of 36.5K and 40.7K, respectively, which have an N-terminal hydrophobic sequence, possible sites for attachment of N-linked carbohydrate and a short string of hydrophobic residues near the C terminus. These properties are consistent with the mature proteins being either virion, cell surface or secretory glycoproteins. Protein sequence comparisons established that the two gene products are related to each other (20% identity) and to the immunoglobulin (Ig) superfamily. Intriguingly, the nearest homologues of these proteins in the SWISS-PROT (version 14) database are the human and murine interleukin-1 receptors, although both proteins are related to a wide range of Ig superfamily members, including the interleukin-6 receptor. The product of one of these genes is known to be expressed on the cell surface early during infection and immunity directed against it confer resistance to virus infection without directly neutralizing virus infectivity. We propose a novel method for virus immune evasion in which the product of one or both of these proteins may bind interleukin-1 and/or interleukin-6 and prevent these cytokines reaching their natural receptors. In consequence the inflammatory response would be diminished and virus replication enhanced.

Introduction

Vaccinia virus is a member of the poxvirus group of large DNA viruses which replicate in the cytoplasm of infected cells and encode many enzymes for transcription and DNA replication (Moss, 1990). The virus genome is a dsDNA molecule of 186 kbp with inverted terminal repeats (ITR) of 10 kbp. The sizes of the ITR and proximal regions show heterogeneity in different strains of vaccinia virus and among other orthopoxviruses, whereas the central region of the genome is conserved (Mackett & Archard, 1979); many essential virus structural proteins and enzymes are encoded here (Condit & Niles, 1990). Genes located near the left ITR are not essential for virus replication (Perkus et al., 1986; Kotwal & Moss, 1988a).

To characterize the right end of the vaccinia virus genome and with the hope of finding genes that might influence virus virulence, the nucleotide sequence of a 42 kbp region of the virus has been determined. Comparisons of the predicted protein sequences encoded by these genes against protein databases established a surprising number of homologies, some of which have been reported (Kotwal & Moss, 1989; Smith et al., 1989a, b, c; Howard & Smith, 1989; Traktman et al., 1989; Rempel et al., 1990). Here, the sequences of two more genes are described. These genes are both predicted to encode glycoproteins that may be located on the surface of the infected cell or virus particle or may be secreted from the cell. These proteins are related to each other and to members of the immunoglobulin (Ig) superfamily. Members of this large family contain varying numbers of structurally similar domains (Ig domains) and perform diverse functions, although a unifying theme is the involvement in cell surface interactions, either between cells or by binding cytokines (Williams & Barclay, 1988). The members of the Ig superfamily to which the vaccinia virus genes B15R and B18R are most closely related are the human (Sims et al., 1989; Chua & Gubler, 1989) and murine (Sims et al., 1988) interleukin-1 receptors (IL-1R). Another vaccinia virus glycoprotein, the haemagglutinin, is also a member of this superfamily (Jin et al., 1989).

The sequence of one of these genes (B18R) has been reported recently (Ueda et al., 1990) but its relationship to the Ig superfamily or to IL-1R was not described. The B18R protein was shown to be expressed on the cell surface early after infection with vaccinia virus and is the previously described soluble antigen (Ueda et al., 1969; Ueda & Tagaya, 1973; Ikuta et al., 1980). Antibody
against this antigen does not neutralize virus infectivity but does induce resistance to virus inoculation in rabbits (Ueda & Tagaya, 1973). The present report shows that the proteins predicted to be encoded by B15R and B18R are related to the Ig superfamily, and in particular to IL-1R, and suggests that one or both proteins may interfere with the induction of host immunity by binding IL-1 and/or IL-6, thereby reducing the inflammatory response. This constitutes a novel mechanism of virus immune evasion.

### Methods

**DNA sequencing.** The sequence of the SalI I restriction fragment of the WR strain of vaccinia virus was determined as previously described (Smith et al., 1989a). Briefly, the SalI I fragment was purified, circularized and sonicated to generate random fragments of > 300 bp in length. These were cloned into M13 vectors and sequenced using dideoxynucleoside triphosphates, [35S]dATP and the Klenow fragment of DNA polymerase I.

**Computing.** Random DNA sequences were read using a sonic digitizer and assembled into contiguous sequence using the computer programs DBUTIL and DBAUTO (Staden, 1982). The sequences were shown to read from both DNA strands except for 81 nucleotides from gene B15R that were read from four to six independent clones of the same polarity. Open reading frames (ORFs) were identified using the ORFFILE program and files for individual protein sequences were created using DELIB (both kindly provided by M. E. G. Boursnell, Institute for Animal Health, Houghton, U.K.). Protein sequences were compared against the SWISSPROT protein database (version 14) using the FASTP program (Lipman & Pearson, 1985). Multiple protein alignments were computed with the MULTALIGN program (Barton & Sternberg, 1987) and sequence similarities were evaluated using the ALIGN program (Dayhoff et al., 1983).

### Results

The complete nucleotide sequence of the vaccinia virus (strain WR) SalI I fragment was determined by computer-aided compilation of sequences from randomly generated DNA fragments. Within this region, ORFs encoding 66 amino acids or more that did not extensively overlap other ORFs were selected for analysis. Of these ORFs, two, designated B15R and B18R, are the subject of this paper. This nomenclature indicates that the genes are the fifteenth and eighteenth ORFs from the left-hand end of the HindIII B restriction fragment and that they are transcribed rightwards towards the genomic terminus. The position of the 9.8 kb SalI I fragment within the virus genome and the location and direction of transcription of genes B15R and B18R are shown in Fig. 1.

Fig. 2 and 3 show the nucleotide sequence and deduced amino acid sequence of genes B15R and B18R, respectively; potential transcriptional regulatory signals are underlined. Gene B15R contains the vaccinia virus early transcriptional termination signal T₅NT (Yuen & Moss, 1987) 18 to 24 nucleotides upstream and 33 to 41 and 1023 to 1029 nucleotides downstream of the A of the presumed initiating methionine codon; early transcripts would be expected to terminate approximately 90 nucleotides into the coding region. B15R lacks the late promoter consensus sequence TAAAT (Davison & Moss, 1989) within 200 nucleotides upstream of the ORF. Taken together these data suggest that high level expression of this gene is unlikely. Gene B18R contains T₅NT motifs starting 66 nucleotides upstream and 13 nucleotides downstream of the ORF. There is also a TAAAT sequence at positions -27 to -31 relative to the initiating methionine codon. These signals might allow gene B18R to be transcribed constitutively, a prediction consistent with the recent demonstration that the gene is transcribed and translated at least early during infection (Ueda et al., 1990).

B15R was predicted to encode a protein of 326 amino acids with an Mᵦ of 36-5K. The protein contains a potential N-terminal signal peptide of 18 amino acids and five potential sites for the addition of N-linked sugar residues. Near the C terminus there is another string of uncharged residues; however, structural considerations (below) suggested that this does not function as a membrane anchor domain.

B18R encodes a protein predicted to consist of 351 amino acids with an Mᵦ of 40-7K. This differs from the published amino acid sequence from the IHD strain of vaccinia virus (Ueda et al., 1990) at only three positions (Fig. 3). Like B15R, it contains a potential hydrophobic signal peptide (11 amino acids) and five sites for the attachment of N-linked carbohydrate. At the C terminus there is a short run of hydrophobic amino acids (eight residues) followed by a single charged residue. This structure is too short to be a conventional anchor sequence and (as for the comparable region of B15R) is predicted to form a β-strand within an Ig domain (Williams & Barclay, 1988) and to be unavailable to function as an anchor. Nonetheless, at least some of this...
protein is located on the cell surface early during infection (Ueda et al., 1990).

Comparisons of the predicted protein sequence of these two genes with the SWISSPROT protein database (version 14) and a library of vaccinia virus proteins using the FASTP program established that the encoded proteins are related to each other (20% identity) and to members of the Ig superfamily. The nearest homologues

in the protein database were IL-1R precursors. B15R had an optimized FASTP score of 214 against human IL-1R (Sims et al., 1989; Chua & Gubler, 1989), with 21~ identity over 323 amino acids; B18R had a FASTP score of 240 against this protein. Some other members of the Ig superfamily that were identified with a FASTP search were the neural cell adhesion proteins (NCAM) from mouse (Moos et al., 1988) and chicken (Hemperly et al., 1989) with 166 against this protein. These results indicate that the IL-1R and B15R/B18R genes are members of a new family of receptor homologues in the vaccinia virus genome.

Fig. 2. Nucleotide sequence and deduced amino acid sequence of gene B18R (EMBL database accession number X56122). Potential transcription control signals are underlined and a possible signal peptide at the N terminus is boxed. Sites for the addition of N-linked carbohydrate (NXS/T except where X is P) are boxed and the cysteine residues likely to form disulphide bonds within Ig domains are stippled.

Fig. 3. Nucleotide sequence and deduced amino acid sequence of gene B18R (EMBL database accession number X56122). The positions at which the sequence differs from the published sequence of this gene from vaccinia virus strain IHD (Ueda et al., 1990) are shown by the inclusion of the different amino acid from IHD above the line. Other features are marked as in Fig. 2.
Fig. 4. Amino acid alignment of the Ig domains from B15R and B18R with the Ig domains of the human and murine IL-1Rs (H-IL-1R and M-IL-1R respectively) (Sims et al., 1988, 1989; Chua & Gubier, 1989), human IL-6R (H-IL-6R) (Yamasaki et al., 1988), vaccinia virus haemagglutinin (VV HA) (Jin et al., 1989), the Ig domain of Ig κ (Hieter et al., 1980), domain 1 of fasciclin II (FAS 1) (Harrelson & Goodman, 1988), domain 3 of myelin-associated glycoprotein (MAG) (Salzer et al., 1987) and domain 1 of chicken NCAM (CHNCAM 1). The regions predicted to form β-strand structures of Ig domains are indicated above the sequences. Residues identical in six or more sequences are boxed. A few residues between β-strands B and C have been omitted. Also omitted for brevity are E-strands D and, where appropriate, C' and C". Higher numbers of residues (about 30 or more) between strands C and E are indicative of the V domains (Williams & Barclay, 1988).

An assessment of the degree of sequence similarity of each domain with individual domains of members of the Ig superfamily was made using the ALIGN program (Dayhoff et al., 1983). This compares the best alignment score for two sequences with the mean score of multiple (100 or more) alignments of randomly scrambled sequences. The scores for the best alignment of the real sequences are presented as the number of standard deviations from the mean score of the randomized sequences. Values of greater than 3.1 are significant (P \(10^{-3}\)) and values of 4.8, 6.0 and 7.9 indicate probabilities of \(10^{-8}\), \(10^{-9}\) and \(10^{-15}\), respectively. In Table 1 the alignment scores of the six vaccinia virus Ig-like domains against selected domains of Ig superfamily members are presented. With B15R the highest individual scores were found against the human and murine IL-1R domains. Domain 2 of B18R also scored well against the IL-1R domains and overall there were highly significant scores against a wide range of Ig domains, supporting the contention that both these proteins belong to the Ig superfamily.

The relationship between B15R and the murine and human IL-1Rs was further explored by an amino acid sequence alignment of the whole of the virus protein with the external regions of the IL-1Rs, which contain the three Ig domains up to the transmembrane anchor, and the signal sequence and single C2 domain of IL-6R (Fig. 5). This comparison and the alignment of individual domains (Fig. 4) indicate a closer relationship between IL-1Rs and IL-6R and the vaccinia virus proteins than between the vaccinia virus proteins and other Ig...
superfamily members. For example, (i) B15R and the external region of the IL-1Rs have a very similar length. (ii) There are additional conserved cysteines in B15R, B18R and the IL-1Rs located near the beginning of β-strands A and G in domain 1 and at similar positions in domain 2 of the IL-1Rs and B18R. These cysteines lie within the three-dimensional structure of an Ig C domain in positions probably allowing the formation of another intradomain disulphide bond. (iii) In B15R, both IL-1Rs and IL-6R there is a proline following the invariant cysteine in β-strand B of domains 1 and 2, an unusual residue in this position for most Ig domains. B18R domain 1 also contains proline at this position. (iv) In β-strand F of domain 3 of B15R and both IL-1R sequences, the glycine typical of other Ig domains is absent. Moreover, the otherwise invariant tyrosine is replaced in both IL-1Rs and in B15R with phenylalanine. (v) A glycosylation site is conserved in domain 1 β-strand F of both IL-1Rs and B15R despite divergence of amino acid sequence. (vi) Domain 3 does not contain additional cysteines and is longer than domains 1 and 2 in B15R, B18R and both the IL-1Rs.

Discussion

The primary structure of two genes from within the Sall I fragment near the end of the vaccinia virus genome are presented. Both genes are predicted to encode proteins that are related to each other, to members of the Ig superfamily (Williams & Barclay, 1988) and to interleukin receptors (Sims et al., 1988, 1989; Chua & Gubler, 1989). This relationship was initially established by a FASTP search of protein databases and was confirmed by statistical computational analysis using the ALIGN program (Dayhoff et al., 1983). These analyses established that both proteins have three Ig domains that are related to Ig domains from a broad range of members of this superfamily. For B15R, the highest scores with both FASTP and ALIGN were against the human and murine IL-1Rs. B18R is also related to IL-1R, and both B15R and B18R are related to the single Ig domain of the IL-6R.

The B18R protein is located on the cell surface early during infection (Ueda et al., 1990) despite the C-terminal hydrophobic residues being predicted to form the final β-strand of an Ig domain and therefore unlikely to function as a transmembrane anchor. This localization may be mediated by protein–protein interactions possibly involving intermolecular disulphide bonds between additional cysteines (as occurs with other members of the Ig superfamily, e.g. the immunoglobulins). Alternatively, these cysteines may form intradomain disulphide bridges between the A and G β-strands. Similar arguments apply for B15R but in addition it is possible that the N-terminal hydrophobic sequence of this protein might act as both a signal and anchor sequence, as does the corresponding region of influenza neuraminidase which is of similar size. Irrespective of whether the proteins remain on the cell surface or are released, a plausible function for either protein is to bind and sequester interleukin-1 (IL-1) or IL-6 at the site of tissue

Table 1. Similarity scores for the Ig domains of vaccinia virus proteins B15R and B18R against selected domains from other Ig superfamily members computed using the ALIGN program

<table>
<thead>
<tr>
<th></th>
<th>B15R*</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>B15R*</td>
<td>4.13</td>
<td>2.12</td>
<td>0.62</td>
<td>4.13</td>
<td>2.12</td>
<td>0.62</td>
<td>4.13</td>
<td>2.12</td>
<td>0.62</td>
<td>4.13</td>
<td>2.12</td>
<td>0.62</td>
<td>4.13</td>
</tr>
<tr>
<td>B18R†</td>
<td>4.13</td>
<td>2.12</td>
<td>0.62</td>
<td>4.13</td>
<td>2.12</td>
<td>0.62</td>
<td>4.13</td>
<td>2.12</td>
<td>0.62</td>
<td>4.13</td>
<td>2.12</td>
<td>0.62</td>
<td>4.13</td>
</tr>
</tbody>
</table>

* B15R domains are amino acids 28 to 119 (1), 121 to 214 (2) and 222 to end (3).
† B18R domains are amino acids 53 to 149 (1), 152 to 241 (2) and 252 to end (3).
‡ Murine IL-1R domains are amino acids 26 to 119 (1), 125 to 219 (2) and 231 to 335 (3) (Sims et al., 1988).
§ Human IL-1R domains are amino acids 24 to 116 (1), 122 to 216 (2) and 228 to 332 (3) (Sims et al., 1989).
‖ The human IL-6R domain is amino acids 27 to 116 (Yamasaki et al., 1988).
§§ Rabbit poly(Ig) receptor domain 5, amino acids 458 to 558 (Mostov et al., 1984).

The B18R protein is located on the cell surface early during infection (Ueda et al., 1990) despite the C-terminal hydrophobic residues being predicted to form the final β-strand of an Ig domain and therefore unlikely to function as a transmembrane anchor. This localization may be mediated by protein–protein interactions possibly involving intermolecular disulphide bonds between additional cysteines (as occurs with other members of the Ig superfamily, e.g. the immunoglobulins). Alternatively, these cysteines may form intradomain disulphide bridges between the A and G β-strands. Similar arguments apply for B15R but in addition it is possible that the N-terminal hydrophobic sequence of this protein might act as both a signal and anchor sequence, as does the corresponding region of influenza neuraminidase which is of similar size. Irrespective of whether the proteins remain on the cell surface or are released, a plausible function for either protein is to bind and sequester interleukin-1 (IL-1) or IL-6 at the site of tissue
damage during virus infection. These cytokines would be unable to reach their natural receptors, resulting in a diminished inflammatory response and an increased ability of the virus to spread and cause disease in the mammalian host. The B18R gene product is not essential for virus replication in vitro because viable mutants of cowpox virus exist which lack this gene and immunological analyses show that the Lister strain of vaccinia virus does not express this protein (Ueda et al., 1969; Ueda & Tagaya, 1973; Ikuta et al., 1980).

Another possible function for the B15R and B18R proteins is to facilitate virus spread by mediating the interaction of virus particles or infected cells with other cells. There is ample precedent for such a cell–cell interaction with members of the Ig superfamily, e.g. the major histocompatibility complex (MHC) antigens, NCAM, fasciclin II (Harrelson & Goodman, 1988) and the intercellular protein amalgam (Seeger et al., 1988). However, the closer homology to the IL-1R and IL-6R make the binding of cytokines the more likely function. Indeed, antibodies to the B18R gene product are known to restrict virus replication without directly neutralizing virus infectivity (Ueda & Tagaya, 1973). These observations are consistent with the above model in which the B18R protein would bind cytokines (IL-1 or IL-6) and prevent them mediating their normal function in cells bearing their receptors. Antibody to B18R would block the sequestering of cytokines in this way so that a normal inflammatory response would ensue and virus replication would be restricted. It was recently demonstrated that blockade of the IL-1R directly attenuates the host inflammatory response (Gershenvald et al., 1990).

Such a mechanism would constitute another novel defence mechanism employed by vaccinia virus against the host immune response. Other mechanisms, proposed or proven, include interference with the complement cascade by the expression of a secretory homologue of C4b-binding protein (Kotwal & Moss, 1988b) and a membrane-associated homologue of factor H (S. T. Howard & G. L. Smith, unpublished results), the expression of three serine protease inhibitors which may block the presentation of peptides to cytotoxic T cells (Boursnell et al., 1988; Kotwal & Moss, 1989; Smith et al., 1989a) and restrict the infiltration of infected lesions with leukocytes (Palumbo et al., 1989; Chua et al., 1990), and the possession of a homologue of the tumour necrosis factor (TNF) receptor (Howard et al., 1991). The possession of receptors that might bind and sequester one or more of IL-1, IL-6 and TNF would constitute a major advantage to the virus in the suppression of the inflammatory response. Other viruses encode protein(s) related to the Ig superfamily, e.g. human cytomegalovirus (Beck & Barrell, 1988) and Epstein–Barr virus (Baer et al., 1984), and in the former case the encoded protein is also proposed to aid immune evasion. This protein is related to class I MHC antigens, binds β-2 microglobulin (Browne et al., 1990) and prevents host MHC antigens binding this antigen, a process required for MHC transport to the cell surface. Thus cytokotic T cells directed against virus peptides are less able to detect and destroy virus-infected cells. The mechanisms devised by viruses to evade the host immune system are varied and fascinating.

We are most grateful to A. Neil Barclay and Alan F. Williams for computing assistance and helpful discussions. G. L. S. is a Lister Institute-Jenner Research Fellow. This work was supported by grants from the Medical Research Council.
References


a polypeptide with extensive homology to DNA ligases. *Nucleic Acids Research* 17, 9051–9062.


(Received 9 October 1990; Accepted 13 November 1990)