Sequence and expression of a baculovirus protein with antigenic similarity to telokin

Deborah A. Raynes, David J. Hartshorne and Vince Guerriero, Jr*

Department of Animal Sciences, University of Arizona, Tucson, Arizona 85721, U.S.A.

A protein from baculovirus-infected cells reacted with an antibody against the smooth muscle protein telokin. Because of this unusual similarity, the protein, termed telokin-like protein-20 (TLP20), was isolated and characterized. Its $M_r$ on denaturing polyacrylamide gels was 28K and the protein contained a high proportion of $\beta$ structure. A cDNA for TLP20 was isolated and sequenced. The 3’ non-coding sequence contained a region of high identity with the 5’ end of two other baculovirus genes. The 5’ non-coding region contains several baculovirus regulatory elements. Surprisingly, the derived amino acid sequence showed no homologies to telokin. The cDNA was cloned into a bacterial expression vector and the subsequently expressed protein had a slightly lower $M_r$ than the native protein, but cross-reacted with telokin antibody. This paper reports the characterization of a new baculovirus protein that shares some antigenic similarities to the smooth muscle protein telokin.

Myosin light chain kinase (MLCK) is a key enzyme involved in the regulation of smooth muscle contraction (Hartshorne, 1987). A protein called telokin is identical to the C-terminal 157 amino acids of MLCK (Ito et al., 1989) but the function of this protein or the corresponding region of MLCK is unknown. An antibody against smooth muscle telokin was found to bind to a protein in cell extracts from Autographa californica nuclear polyhedrosis virus (AcMNPV)-infected SF9 cells and from Manduca sexta larvae that had been infected with AcMNPV. This protein has been named telokin-like protein-20 (TLP20) because the calculated $M_r$, based on amino acid composition, is 20K. TLP20 was purified from AcMNPV-infected SF9 cells using anion exchange and phenyl-Sepharose chromatography. The $M_r$ of TLP20 (28K) was slightly higher than that of telokin (24K), as demonstrated by SDS-PAGE (Fig. 1 and Table 1), and the antibody against telokin recognized the purified protein. Isoelectric points of 4.5 for telokin and 4.7 for TP20 were determined by two-dimensional gel electrophoresis (Table 1).

TLP20 was linked to agarose beads to make an affinity column. Affinity-purified anti-telokin antibody was applied to the column and about 3% of the antibody bound and could be eluted at low pH. The observation that only a small portion of the polyclonal antibody reacted to TLP20 indicated that only one or a few epitopes were antigenically related. Western blot analysis was used to determine the specificity of this cross-reaction. The antibody reacted to TLP20, telokin and MLCK but, as with the whole anti-telokin antibody, did not react to MLCK with the telokin portion removed by proteolysis. Therefore, the antibody that was purified over the TLP20 column was specific for telokin and the telokin-related portion of MLCK.

A cDNA library was constructed from baculovirus-infected SF9 cells and screened with an antibody against TLP20. Clones 3-3 and 28-4 were isolated and the 5’ end of 3-3 was used to isolate clone 7-1. These overlapping clones were sequenced and their relationships to each other were established as indicated in Fig. 2. An open reading frame was found that coded for a protein of 180 amino acids. The amino acid and nucleotide sequences were compared to telokin using the program FASTA (Genetics Computer Group Sequence Analysis software package version 7.0; Devereux et al., 1984) and no homology could be found. The amino acid sequence was used to search the peptide sequence databases at the National Center for Biotechnology Information (NCBI) using the BLAST network service and no related proteins could be found. The nucleotide sequence was also used to search GenBank at the NCBI using the BLAST network service and a region of identity was found with the gene encoding gp41, a structural glycoprotein of AcMNPV (Whitford & Falkner, 1992). There is more than 98% identity between TLP20 and gp41 in a region that extends for 364 nucleotides, from positions 1082 to 1446 of the TLP20 sequence. This region of identity
covers the entire 5’ non-coding region of gp41 and extends over the first eight amino acids of the protein. Another match was found with the gene for p40, a gene isolated from *Bombyx mori* NPV (Nagamine *et al.*, 1991). This region of identity starts at nucleotide 1328 and ends at 1446. Again, this included the non-coding 5’ end and the first eight amino acids of the p40 protein. Interestingly, both of these matching sequences are in the 3’ end of the non-coding region of TLP20 gene.

Analysis of the sequence reveals that there are regulatory elements consistent with baculovirus genomic DNA. The sequence A/GTAAG is a late promoter element (Blissard & Rohrmann, 1990) and is found twice in the 5’ non-coding region (from nucleotides 168 and 331), once in the coding region (from nucleotide 674) and twice in the 3’ non-coding region (from nucleotides 1372 and 1403). The sequence CAGT is thought to be an initiation site for baculovirus early genes (Blissard & Rohrmann, 1986). Starting at nucleotide 134, the sequence TGGAAG matches the core element of the simian virus 40 enhancer (Weiher *et al.*, 1983). Thus, the 900 nucleotides in bold are discussed in the text. The underlined sequence has a close similarity to the genes gp41 and p40 as discussed in the text.
indicated that TLP20 is encoded in the viral genome. The promoter and enhancer genomic elements are not a recognition site for polyadenylation (AATAAA). The promoter and enhancer genomic elements are not consistent with a cDNA library. Therefore it is possible that the poly(A) RNA isolated for the construction of the library included viral DNA and this was ligated into the vector and was part of the library. A probe made from the 5' end of clone 3-3 (nucleotides 181 to 482) hybridized to baculovirus DNA (data not shown), which indicated that TLP20 is encoded in the viral genome.

An explanation for a lack of amino acid similarity between TLP20 and telokin may be that the isolated clones do not represent the protein that the telokin antibody reacted with in the cell homogenates. Verification that the open reading frame in Fig. 2 coded for the correct protein was achieved by inserting this sequence into the expression vector pET-3d (Studier et al., 1990) and purifying the protein from bacteria. The protein expressed in bacteria had a slightly lower Mr when compared to the cell-isolated protein (Fig. 1). However, the anti-telokin antibody recognized the bacterial protein (Fig. 1). This size difference may be due to a lack of post-translational modification in bacteria, since a number of the baculovirus proteins are known to be glycosylated.

A lack of amino acid sequence similarities between telokin and TLP20 indicated that they have few antigenic sites in common. These sites could consist of short stretches of amino acids or a region of identical structure. The data from circular dichroism spectra were used to estimate secondary structure with the program PROSEC (Chang et al., 1992), to determine whether there was a structural similarity. The three-dimensional structure of telokin has recently been determined (Holden et al., 1992) and it consists of seven strands of antiparallel β-pleated sheets that form a barrel. Table 1 shows a comparison of these two proteins, including the circular dichroism estimates of secondary structure. As predicted from the three-dimensional structure, telokin has few regions of alpha helix (4%). This is similar to TLP20 (6%). The estimates for the other structures including β-pleated sheets are all similar for both proteins.

This paper reports the identification of a baculovirus protein (TLP20) that cross-reacts with an antibody against telokin. The nucleotide sequence of the cDNA revealed known baculovirus regulatory regions in the 5' non-coding region. Future studies to determine the three-dimensional structure of TLP20 may help to predict the function, as yet unknown, of this protein.

This work was supported by NIH grant HL43651.

Note added in proof. We have partially completed structural analysis of TLP20 using X-ray crystallography, in collaboration with Dr I. Rayment (University of Wisconsin, Wis., U.S.A.). The very high percentage of β structure indicated by this analysis is in agreement with the predictions made here using circular dichroism.

References


(Received 5 October 1993; Accepted 5 April 1994)