Short paper

Andrew M. Knight,
Georgina B. Harrison●,
Richard J. Pease●,
Peter J. Robinson and
P. Julian Dyson

Transplantation Biology Section
and Division of Molecular
Medicine●, Clinical Research
Centre, Harrow

Biochemical analysis of the mouse mammary tumor virus long terminal repeat product. Evidence for the molecular structure of an endogenous superantigen

Recent reports have shown that both exogenous and endogenous mouse mammary tumor viruses (MMTV) can encode superantigens. Transfection and transgenic studies have identified the open reading frame (ORF) present in the 3' long terminal repeat (LTR) as encoding superantigen function. In this study, we have used an *in vitro* translation system in an attempt to characterize the molecular nature of the protein encoded by the 3' ORF of Mtv-8. Using various constructs encoding full-length and truncated versions of the ORF product, we report that the hydrophobic region close to the amino terminus of the 36-kDa protein can function as a transmembrane domain. Protease digestion experiments also demonstrate that the protein has a type-II transmembrane conformation with an extra-cytoplasmic carboxy terminus. Since this hydrophobic region is conserved between all known MMTV, we speculate that LTR ORF, including those proposed to encode the minor lymphocyte stimulatory antigens, are also capable of encoding type-II transmembrane glycoproteins. The polymorphism between MMTV LTR ORF products, which correlates with deletion phenotypes, is predominantly in the carboxy-terminal extracellular region, consistent with a major role in interaction with the T cell receptor.

1 Introduction

Conventional antigen recognition by T cells involves a processed form of the antigen being presented in the antigen-binding pocket of a major histocompatibility molecule [1]. In conventional antigen recognition, all variable germ-line segments (V_α, J_α, V_β, D_β and J_β), as well as N-region additions, contribute to the specificity of the TcR [2]. In contrast to this, T cells stimulated by certain bacterial toxins [2] and the recently identified endogenous "superantigens" encoded by germ-line copies of mouse mammary tumor virus (MMTV; [3–6]) have a particular TcR V_β usage. The major effect of endogenous superantigens is seen as thymic deletion of T cells expressing particular V_β regions. In contrast, in the periphery, superantigen interaction leads to polyclonal activation of T cells expressing such TcR V_β regions with limited contribution from other TcR components [2].

Evidence suggests that superantigens are quite distinct from peptides in their interaction with both MHC and TcR molecules. First, exogenous superantigens do not require cellular processing for presentation to T cells [7] and, second, mutation studies of class II molecules have sug-

gested interaction of toxins with a region distinct from the peptide binding pocket [8]. Additionally, the cellular response of T cells can differ between "superantigen" and conventional antigen stimulation [9].

While bacterial toxins have been characterized at the molecular level, no such information is available for endogenous superantigens. Following recent reports identifying the MMTV 3' long terminal repeat (LTR) as the region required for superantigen function [10, 11], we have investigated the biochemical properties of the protein encoded by the 3' LTR open reading frame (ORF) of Mtv-8 (Mtv-8 ORF) which is responsible for partial thymic deletion of $V_\beta11^+$ T cells [6]. Previous *in vitro* translation studies have shown that the 3' LTR ORF encodes four overlapping polypeptides of 36, 24, 21 and 18 kDa by the use of alternative initiation sites [12, 13]. *In vivo* identification of MMTV ORF has been limited to the product of the partially deleted LTR present in the lymphoma cell line EL4 [14]. Using a number of truncated constructs from Mtv-8 ORF, we have translated *in vitro* the resulting transcripts and examined their protein products.

2 Materials and methods

2.1 Construction of Mtv-8 ORF plasmids

Mtv-8 ORF inserts for translation studies were constructed as follows (see Fig. 1): using the polymerase chain reaction (PCR), Mtv-8 3' Eco RI fragment [15] DNA (a kind gift from Dr. Gordon Peters, ICRF, London, GB) was amplified with the following oligonucleotide primers (see Fig. 1A). Construct 1 [see Fig. 1B and G; full-length Mtv-8 ORF amino acid (aa) no. 1-322] 5' primer: AAT ACG ACT

CAC TAT AGG GCG AAT TGC GCC ATG GCG CG and 3' primer: GCG TCG ACT TAT CCA TTA TAG CTC ATG CCA. Construct 2 (see Fig. 1C and H; MTV-8 ORF truncated at the carboxy terminus, aa 1-275) 5' primer: same as primer used for construct 1, and 3' primer: ATA AGT CGA CTT ATT TTA TGT CGT CTT TTT C. Construct 3 (see Fig. 1D and I; Mtv-8 ORF truncated at the carboxy terminus, aa 1-316) 5' primer: same as primers used for construct 1 and 2, and 3' primer: ATA AGT CGA CTT AAA AAG TAT CGT CAG AAA T. Construct 4 (see Fig. 1E and J; Mtv-8 ORF truncated at the amino terminus, so as to exclude region from aa nos. 1-122) 5' primer: AAT ACG ACT CAC TAT AGG GCG AAT TGC GCC ATG GTT and 3' primer as used for construct 1. All 5' primers are mismatched to introduce an Nco I restriction site at the first ATG start codon. All 3' primers include an in-frame stop codon and a Sal I restriction site. PCR conditions were as described [16] with 25 cycles each consisting of 1 min 94°C, 1 min 55°C, 30 s 72°C followed by an additional 9 min at 72°C, using 2 mM Mg^{2+}. PCR products were gel purified (1% low-melting point agarose, BRL, Uxbridge, GB), digested with Nco I and Sal I according to manufacturer's conditions (C.P. Biolabs, Bishops Stortford, GB) and ligated into the Nco I, Sal I linearized transcription vector (a kind gift from Drs. Mike Howell and Tim Hunt, University of Cambridge). Efficient cap-independent translation is obtained using this vector as it contains nucleotides 259-837 of the 5' untranslated region (UTR) from mouse encephalomyocarditis virus (EMCV) upstream of its cloning site [17]. Following linearization with Sal I, plasmids containing constructs 1,2,3 and 4 were transcribed using T7 RNA polymerase (Boehringer-Mannheim, Mannheim, FRG). Transcripts encode Met-Ala followed by residues encoded by the ORF.

2.2 In vitro translation studies

Transcripts from constructs 1,2,3 and 4 (see Fig. 1B–E) were translated in unfractionated reticulocyte lysates (Amersham, Int., Amersham, GB) with [35S] methionine (Amersham), either in the presence or absence of canine pancreatic microsomes (Promega, Madison, WI).

Protease digestions of translation mixtures were performed using 10 µg TPCK trypsin (Cambridge Bioscience, Cambridge, MA) for 90 min at 0°C in the presence of 20 mM Tris HCl, pH 7.5 and 2 mM tetracaine HCl (Sigma, Poole, GB) [18]. Digestion, was terminated with soybean trypsin inhibitor (Sigma) at a final concentration of 6 mg/ml. Glycosylation inhibition was performed using the competitive N-linked glycosylation inhibitor N-acetyl-Asn-Tyr-Thr-carboxyamide at a final concentration of 100 µM [19]. For carbonate treatment, post-translational rat liver microsomes were added to mixtures and ultracentrifuged in a Beckman airfuge for 15 min at 30 psi. Membrane fractions were resuspended in sodium carbonate (pH 11.5) for 30 min at 0°C. After another centrifugation for 15 min at 30 psi. the protein content of membrane fractions and supernatant were analyzed. All proteins were analyzed by SDS-PAGE on 8% to 15% gradient gels.

3 Results and discussion

Previous reports have identified a hydrophobic region close to the amino terminus of an MMTV ORF protein sequence [13] and suggested this may encode a membrane transfer sequence (signal sequence) [10]. Hydrophobicity plot analysis of the predicted amino acid sequence of Mtv-8 ORF did not reveal any other significant non-polar regions (see Fig. 1F). To determine functionally important domains of Mtv-8 ORF, various deletions were made within the DNA sequence (Fig. 1B–E) and the resulting constructs expressed and translated in vitro. Protein products could then be examined biochemically for potential signal/trans-membrane domains. Initial in vitro studies using the endogenous Mtv-8 ORF 5' UTR lead to translation products initiated at internal sites as well as the full-length polypeptide (data not shown), as previously reported
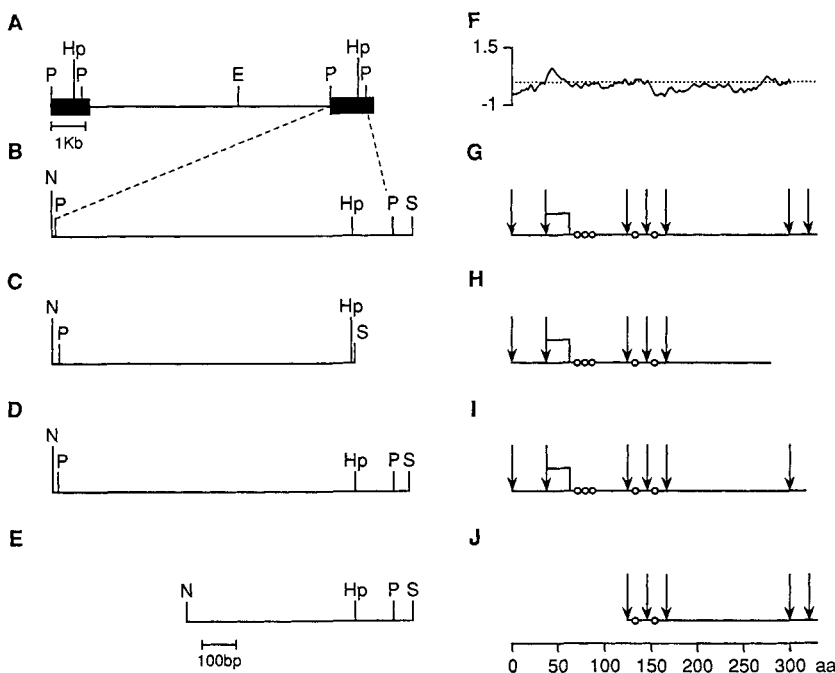

Figure 1. Schematic representation of in vitro constructs. All DNA inserts were PCR products cloned into the transcription vector, as described in Sect. 2.1 (A) Diagram of endogenous MMTV (Mtv-8). Shaded regions indicate LTR. (B) Construct 1, indicating region of LTR containing the ORF. (C, D, E) Constructs 2, 3 and 4, respectively. (N = Nco I, P = Pst I, Hp = Hpa I, E = EcoR I, S = Sal I. (F) Hydrophobicity plot of Mtv-8 ORF. (G, H, I, J) representation of anticipated translation products from constructs 1, 2, 3 and 4, respectively, ↓ = ATG initiation codons. Open box = hydrophobic region, Open circles = potential N-linked glycosylation sites.
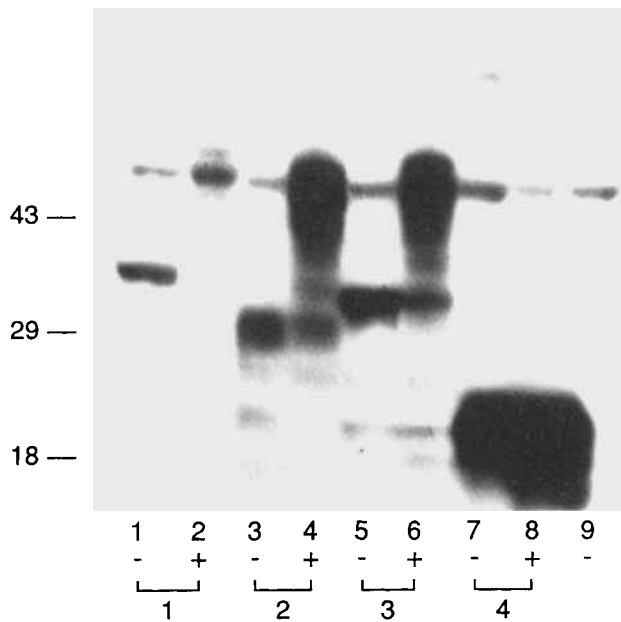
Figure 2. SDS polyacrylamide gradient (8 %–15 %) gel resolving proteins from Mtv-8 ORF constructs. Lanes 1 and 2, 3 and 4, 5 and 6, 7 and 8 are products from constructs 1, 2,3 and 4, respectively. Lane 9 = no RNA control. + = microsomes present, – = microsomes absent. Nos. indicate standard protein marker sizes in kDa. (It should be noted that the translation efficiency of construct 4, lanes 7 and 8, was significantly greater than that of the other constructs, probably due to the altered initiation site).

[12, 13]. To improve the efficiency of initiation at the first methionine residue, MMTV fragments were cloned into a transcription vector (see Sect. 2.1) containing the 5' UTR from EMCV upstream of its cloning site [17]. Fig. 2 shows that constructs 1,2,3 and 4 are all capable of producing major protein products of predicted size, 36, 31,36 and 22
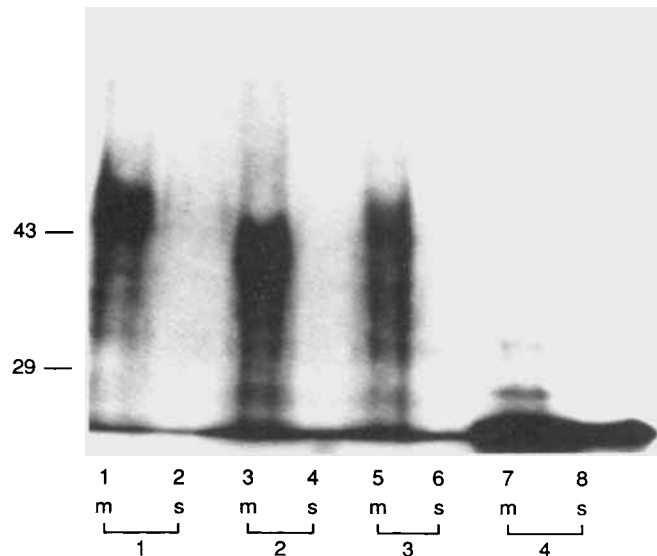
kDa, respectively (lanes 1, 3,5 and 7). On addition of canine microsomes to the translation mixtures, the molecular weight of the protein products from constructs 1, 2 and 3 increased by approximately 10 kDa (Fig. 2 lanes 2, 4 and 6). This increase in size would be accounted for by the post-translational addition of carbohydrate side chains to the protein backbone, in agreement with the usage of the consensus N-linked glycosylation sites at Asn 80, 90, 94, 132 and 147. It is clear, however, that the protein product from construct 4 does not change in molecular weight in the presence of microsomal membranes (Fig. 2, lanes 7, 8; even though the truncated sequence retains two potential glycosylation sites at Asn 132 and 147, see Fig. 1J). These results indicate that the domain directing membrane translocation would be expected to be located within residues 1–122.

To determine whether the products of the Mtv-8 ORF constructs are membrane anchored, we used alkaline sodium carbonate on microsomal translations to rupture the membranes [20]. Using this treatment, we have confirmed absolute discrimination between *in vitro* translated soluble proteins (yeast α mating factor, *E.coli* β lactamase) and a transmembrane protein (coronavirus E1 glycoprotein; data not shown). Fig. 3 shows the distribution of protein products after carbonate treatment of the translation mixtures. For the *in vitro* translation products of constructs 1, 2 and 3, exclusive association with the membrane fraction is seen (lanes 1–6) consistent with their being membrane-associated glycoproteins. In contrast, the product of construct 4 is clearly not tightly associated with the membrane fraction, as it is seen in both membrane and supernatant fractions (lanes 7 and 8). The weak membrane association observed is not understood but may reflect limited solubility.

These results are consistant with the utilization of the hydrophobic region from residues 38 to 63 (See Fig. 1F) as a
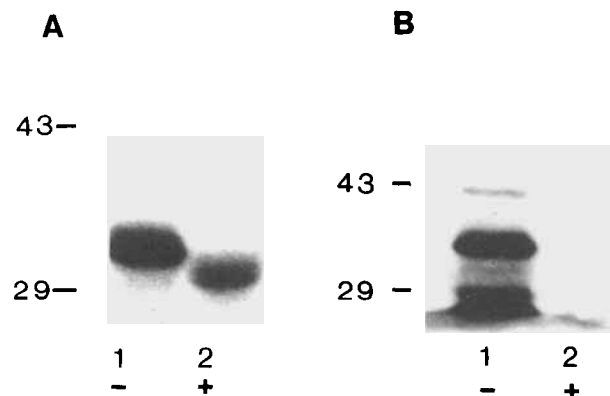

Figure 3. SDS polyacrylamide gradient (8 %–15 %) gel of protein products from Mtv-8 ORF constructs 1–4 following carbonate treatment. Numbers indicate the construct no. m = membrane fraction, s = supernatant fraction. Sizes are standard molecular mass markers in kDa.

Figure 4. SDS polyacrylamide gradient gels (8 %–15 %) showing the protein product from Mtv-8 ORF construct 1. (A) Lanes 1 and 2 show protein products in the presence of microsomal membranes (note in both lanes the N-glycosylation inhibitor N-acetyl-Asn-Tyr-Thr-carboxyamide is included.) (B) Lanes 1 and 2 show protein products in the absence of micosomal membranes. + = treatment of translation mixture with 10 μg TCPK trypsin as described in Sect. 2.2. - = treatment absent. Sizes are standard molecular mass markers in kDa.

transmembrane domain. Trypsin treatment in the presence of microsomes was, therefore, used to define, by the degree of protection, the extent to which the protein has translocated across the microsomal membrane, and hence its membrane orientation. To simplify the interpretation of these experiments, the competitive N-glycosylation inhibitor N-acetyl-Asn-Tyr-Thr-carboxyamide was included. Fig. 4A illustrates that for the product of constuct 1, a significant decrease of approximately 3 to 5 kDa is seen after trypsin treatment (lanes 1 and 2), suggesting that the majority of the protein is protected by the microsomal membranes. (B) shows that in the absence of microsomes no such protection is observed. These results are consistant with the hydrophobic domain (residues 38 to 63) anchoring the protein in a type-II membrane configuration. The product of construct 4 does not show evidence of microsomal protection from trypsin cleavage (data not shown) consistant with its lack of a membrane translocation domain. The product from construct 1, (Fig. 4A, lane 1) in the presence of the glycosylation inhibitor is the same size as that seen in the absence of microsomes (Fig. 2, lane 1 and Fig. 4B, lane 1), demonstrating that signal sequence cleavage is not occurring.

4 Concluding remarks

Since the identification of numerous MMTV as encoding superantigens [3–6], presumably through expression of 3' LTR ORF products, many questions about their mode of action can be answered. One important aspect of endogenous superantigens is whether they are presented as conventional peptides in association with MHC molecules or independently of such processing events, as are bacterial toxins. This study provides information concerning the molecular characteristics of the Mtv-8 ORF protein. Our *in vitro* analysis has shown that the 3' ORF of Mtv-8 is capable of encoding a type-II 45-kDa transmembrane glycoprotein. Sequence analysis of the ORF from various MMTV reveals a high degree of similarity [10]. The sequence of the ORF from Mtv-8, 9 and 11, for example, which we have previously shown to be involved in TcR $V_\beta 11^+$ T cell deletion [6], are very similar to each other, especially in the carboxy terminal region. Further, MMTV responsible for deletion of other classes of V_β^+ T cells have a carboxy terminal region distinct from that of Mtv-8, 9 and 11 (for example Mtv-1,$V_\beta 3$ deletion; Mtv-C3H,$V_\beta 14/15$ deletion [10]). The correlation between carboxy sequence and V_β deletion phenotype has suggested a role for the carboxy terminus in determination of V_β deletion specificity; for a type-II membrane protein this region would be able to interact with

the TcR without requirement for cellular processing. One possibility, currently under investigation, is that the type-II glycoproteins encoded by the 3'ORF of MMTV may bind either intracellularly or at the plasma membrane with class II and that this complex is what is recognized by endogenous superantigen-responsive T cells either thymically (leading to deletion) or peripherally (leading to expansion).

5 References

1 Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L. and Wiley, D. C., *Nature* 1987. *329:* 506.
2 Marrack, P. and Kappler, J., *Science* 1990. *248:* 705.
3 Marrack, P., Kushnir, E. and Kappler, J., *Nature* 1991. *349:* 524.
4 Frankel, W. N., Rudy, C., Coffin, J. M. and Huber, B. T., *Nature* 1991. *349:* 526.
5 Woodland, D. L., Happ, M. P., Gollob, K. J. and Palmer, E., *Nature* 1991. *349:* 529.
6 Dyson, P. J., Knight, A. M., Fairchild, S., Simpson, E. and Tomonari, K., *Nature* 1991. *349:* 531.
7 Dellabona, P., Peccoud, J., Benoist, C. and Mathis, D., *Cold Spring Harbor Symp. Quant. Biol.* 1989. *54:* 375.
8 Dellabona, P., Peccoud, J., Kappler, J., Marrack, P., Benoist, C. and Mathis, D., *Cell* 1990. *62:* 1115.
9 Patarca, R., Wei, F.-Y., Iregui, M. V. and Cantor, H., *Proc. Natl. Acad. Sci. USA* 1991. *88:* 2736.
10 Choi, Y., Kappler, J. W. and Marrack, P., *Nature* 1991. *350:* 203.
11 Acha-Orbea, H., Shakhov, A. N., Scarpellino, L., Kolb, E., Muller, V., Vessaz-Shaw, A., Fuchs, R., Blochlinger, K., Rollini, P., Billotte, J., Sarafidou, M., MacDonald, H. R. and Diggelmann, H. *Nature* 1991. *350:* 207.
12 Dickson, C. and Peters, G., *J. Virol.* 1981. *37:* 36.
13 Racevskis, J. and Prakash, O. M., *J. Virol.* 1984. *51:* 604.
14 Racevskis, J., *J. Virol.* 1986. *58:* 441.
15 Peters, G., Placzek, M., Brookes, S., Kozak, C., Smith, R. and Dickson, C., *J. Virol.* 1986. *59:* 535.
16 Love, J. M., Knight, A. M., McAleer, M. A. and Todd, J. A., *Nucl. Acids Res.* 1990. *18:* 4123.
17 Kaminski, A., Howell, M. T. and Jackson, R. J., *EMBO J.* 1990. *9:* 3753.
18 Scheele, G., *Methods Enzymol.* 1983. *96:* 94.
19 Kassenbrock, C. K., Garcia, P. D., Walter, P. and Kelly, R. B., *Nature* 1988. *333:* 90.
20 Howell, K. E. and Palade, G. E., *J. Cell Biol.* 1982. *92:* 822.