BACKGROUND PAPER

MAPPING EPITOPES ON CORONAVIRUS GLYCOPROTEINS

Hubert Laude

Institut National de la Recherche Agronomique
Station de Recherches de Virologie et d'Immunologie Moléculaires
78350 Jouy-en-Josas
France

INTRODUCTION

The purpose of this paper is to summarise the state of coronavirus epitope research, with an emphasis laid on the most recent achievements, i.e. the correlation between antigenicity and polypeptide primary structure. The antigenic properties of coronaviruses reside mainly in their envelope glycoproteins: the S (spike) protein, a high molecular weight polypeptide which constitutes the petal-shaped peplomers and the M (membrane) protein, a smaller integral membrane polypeptide. Some coronaviruses possess an additional surface glycoprotein, the haemagglutinin-esterase (HE) protein.

PEPLOMER PROTEIN S

Analysis of the predicted amino acid sequences now established for 5 coronaviruses (IBV, MHV, TGEV, FIPV and BCV) has led to a model in which the protein is divided into 2 domains: the amino-terminal half, which is assumed to form the globular part of the peplomer, and the carboxy-terminal half, which is assumed to form an elongated structure, the stalk of the peplomer, and also contains a putative transmembrane anchoring region near the carboxy terminus. Both parts bear numerous asparagine-linked complex glycan sidechains, which may modulate their antigenicity. The sequences of the amino-terminal halves exhibit very little conservation, whereas homologies averaging 30% are observed in the carboxy-terminal halves. Several coronaviruses, including IBV, MHV and BCV have a S protein which is cleaved approximately in the middle, thus defining an amino-terminal subunit (S1) and a carboxy-terminal subunit (S2). Several major antigenic sites have been identified on the S proteins of MHV, TGEV and IBV by competitive binding experiments; a number of these sites could be oriented, if not precisely localised, on the primary structure of the molecule.

MHV S Protein

Recombinant viruses generated *in vivo* by RNA-RNA recombination which contain approximately the 5' two-thirds of the MHV-JHM peplomer gene and the 3' one-third of the MHV-A59 peplomer gene have lost their reactivity to monoclonal antibodies (MAbs) defining the sites designated A and B on the JHM S protein. It was inferred that both sites are probably contained within the carboxy-terminal one-third of the protein (Makino *et al.*, 1987). However the positions of the crossover sites have only been mapped by ologonucleotide finger-printing and require confirmation by nucleotide sequencing. Site A

Coronaviruses and Their Diseases
Edited by D. Cavanagh and T.D.K. Brown
Plenum Press, New York, 1990

was defined by neutralising or non-neutralising MAbs and site B by neutralising MAbs; site A MAbs also inhibited cell fusion. Both sites contained epitopes which exhibited a natural variation from strain to strain and allowed the generation of escape mutants. In addition site B escape mutants showed markedly altered neurovirulence when inoculated into mice. Reactivity analysis of selected or spontaneous variants suggested that sites A and and B are topologically related (Fleming *et al.*, 1986; Taguchi and Fleming, 1989). Three antigenic sites delineated by other groups on the JHM S protein, designated B and C (Talbot *et al.*, 1984) and Ba (Wege *et al.*, 1984), apparently possess the same characteristics with respect to neutralisation, fusion inhibition, variation and generation of attenuated mutants (Dalziel *et al.*, 1986; Wege *et al.*, 1988). It is therefore tempting to speculate that sites B, B-C and Ba are topographically related to the above-mentioned site A. At least part of the latter sites might be conformation-dependent, as was the case for sites B and C (Talbot *et al.*, 1984); this would explain why they could not be mapped using a bacterial expression system.

Expression of fragments of a protein as a fusion polypeptide using a procaryotic vector such as pEX is a classical approach to the mapping of antigenic determinants. Most probably, however, this method mainly detects epitopes that are conformation-independent. Fine mapping of a linear epitope on JHM S protein has been achieved by combining pEX expression and peptide scanning (Lutyes *et al.*, 1989). This epitope belongs to the site designated A (Talbot *et al.*, 1984); this site is: i) recognised by strongly neutralising and fusion-inhibiting MAbs, ii) resistant to denaturation, iii) highly conserved among MHV strains, and iv)possibly crucial for virus infectivity (failure to select escape mutants). The antibody binding site stretched from residue 848 to 856, i.e. about 130 residues from the N-terminus of subunit S2. This 9 amino acid sequence is markedly hydrophobic and located immediately upstream of a region largely conserved in the coronavirus peplomer protein.

Finally, a synthetic decapeptide corresponding to the sequence 993-1002 of the JHM S protein , which was delineated using a surface probability algorithm, elicited a high level of neutralising antibodies and protected mice against a lethal challenge (Talbot *et al.*, 1989). Its position is near the middle of the S2 subunit, between two potentially alpha helical regions which have been suggested to frame the peplomer stalk. Whether this region corresponds to a natural antigenic site remains to be determined.

The above observations support the proposal that the carboxy-terminal S2 subunit of MHV i) contains distinct major antigenic sites, ii) bears neutralisation epitopes, strain-specific epitopes and essential virulence determinants. Surprisingly, there is date, no published evidence of MAb binding to the S1 subunit of MHV. However studies based on the localisation of point mutants or deletions conferring neutralisation-resistance have clearly established the existence of major neutralisation sites on S1 (see articles by Gallagher *et al* and Parker *et al* in Chapter of this volume).

IBV S Protein

In striking contrast, consistent evidence has been obtained that the S1 subunit of IBV is both the major inducer of neutralising MAbs and the major site of antigenic variation. Isolated S1 reacted with strongly neutralising MAbs (Mockett *et al.*, 1984) and elicited neutralising antisera (Cavanagh *et al.*, 1986). By comparing the amino acid sequences of different IBV strains it was found that most substitutions had occurred in the S1 subunit; S1 proteins can differ in upto 50% of their amino acids (Niesters *et al.*, 1986, Binns *et al.*, 1986, Kusters *et al.*, 1989). In particular, sequence alignments allowed the identification in Massachusetts serotype strains of two regions of clustered substitutions, the hypervariable regions, HVR1 (56-69) and HVR2 (117-133), which were suggested to contain neutralisation epitopes (Niesters *et al.*, 1986). Indeed a mutation which prevented neutralisation by two MAbs was localised in HVR1 by direct sequencing of the genomic RNA of the relevant escape mutants (Cavanagh *et al.*, 1988).

Whereas most neutralising MAbs investigated so far are reported to be directed against S1, the S2 subunit has been shown to react with MAbs having a weak neutralising activity (Koch *et al.*, 1986) Several overlapping conformation-independent epitopes recognised by such MAbs have been precisely localised in the 30 N-terminal residues of S2 through expression of random fragments of S using the pEX system (Lenstra *et al.*, 1989). At least

some of these epitopes are conserved in several serotypes as judged by reactivity of the fusion protein with different antisera. In contrast none of the pEX expression products containing fragments of S1 reacted with MAbs suggesting that this subunit contains conformation-dependent epitopes.

The above findings lend support to the view that the major neutralisation domain of IBV resides on the amino-terminal S1 subunit and iś composed mainly of serotype specific and conformation-dependent epitopes. Whether distinct regions of S1 are involved remains to be established. In addition at least one immunodominant site, defined by weakly-neutralising and broadly reactive MAbs, is present on the carboxy-terminal S2 subunit.

TGEV S Protein

Detailed epitope maps have defined 4 to 5 major antigenic sites on the TGEV S protein (Delmas *et al.*., 1986, Correa *et al.*, 1988). Most of the determinants critical for neutralisation were highly conserved among the strains and susceptible to denaturation (Laude *et al.*, 1986, Jimenez *et al.*, 1986). Recent data indicate that all of these sites are located in the amino-terminal half of the protein (see articles by Delmas *et al.* and Enjuanes *et al.* in this volume).

BCV S Protein

In the case of the BCV S protein all the neutralising antibodies characterised so far have been directed against gp100 which has recently been found to correspond to the S1 subunit of the S protein (Deregt & Babiuk, 1987, Vautherot *et al.*, this volume). These data are consistent with recent observations on the role of S1 in the neutralisation of MHV (see above) to which BCV is closely related (see Chapter 10 of this volume).

OTHER GLYCOPROTEINS

M Protein

The M glycoprotein is largely buried within the virus membrane or closely associated with its inner surface. Two adjacent epitopes recognised by polyclonal antibodies are located in the 15 carboxy-terminal residues of the MHV M protein as revealed by studies of fragments expressed in the pEX system. Proteolytic digestion abolished the binding of one MAb, presumably directed against the C-terminal region (Tooze & Stanley, 1986). Three overlapping epitopes have been mapped in the first 30 N-terminal residues of TGEV M protein, through localisation of amino acid substitutions in mutants resistant to complement-mediated neutralisation (Laude *et al.*, to be published). Thus in agreement with its predicted relationship with the virus envelope, the major antigenic sites of the M protein correspond to the short protruding hydrophilic domains located at each extremity of the molecule.

HE Protein

The HE protein, associated with a coronavirus subgroup including BCV, HEV and HCV-OC43, is responsible for the strong haemagglutinating activity of these viruses and is also able to elicit neutralising antibodies. Competition experiments between anti-BCV MAbs have defined at at least one major neutralisation site, not yet correlated with HE primary structure (Deregt & Babiuk, 1987).

CONCLUDING REMARKS

Despite significant advances in the physical mapping of B cell epitopes on coronavirus glycoproteins, the emerging picture is still incomplete. The published data mainly concern the peplomer proteins of MHV and IBV. However substantial information about the TGEV peplomer protein has been presented during the symposium. Moreover additional data reported on the MHV peplomer protein, which have established the presence of neutralisation sites on the S1 subunit, have reconciled the apparent discrepancy between MHV and other coronavirues. The difficulties inherent in the elucidation of antigenic structures on a polypeptide which is both very large and highly glycosylated such as the coronavirus peplomer protein should not be underestimated. Clearly, a single linear epitope is less

difficult to map than a highly conformation-dependent antigenic site. In this context, methods other than bacterial expression of antigen fragments need to be employed. Moreover, the use of complementary approaches should be preferred whenever possible, since each of them contains its own pitfall. Finally, it should be noted that the role of carbohydrate sidechains in the expression of antigenicity has received very little attention so far.

Continuing investigation in this area is of obvious importance for coronavirologists. First, the resulting information is most helpful for the development of recombinant or synthetic vaccines. In this respect, T cell-recognised structures deserve increased attention in future research. Second, as epitopes often coincide with functional determinants, these studies provide a unique opportunity to explain in terms of molecular structure, fundamental virus processes such as neutralisation, cell receptor recognition and antigenic variation.

SELECTED LITERATURE

Cavanagh, D., Davis, P.J. and Mockett, A.P.A. (1988) Amino acids within hypervariable region 1 of avian coronavirus IBV (Massachusetts serotype) spike glycoprotein are associated with neutralisation epitopes. Virus Res., 11: 141-150.

Lenstra, J.A., Kusters, J.G., Koch, G. and van der Zeijst, B.A.M. (1989) Antigenicity of the peplomer protein of infectious bronchitis virus. Mol. Immunol., 26: 7-15.

Lutyes, W., Geerts, D., Posthumus, W., Meloen, R. and Spaan, W. (1989) Amino acid sequence of a conserved neutralising epitope of murine coronaviruses. J. Virol., 63: 1408-1412.

Makino, S., Fleming, J.O., Keck, J.G., Stohlman, S.A. and Lai, M.M.C. (1987) RNA recombination of coronaviruses: localisation of neutralising epitopes and neuropathogenic determinants on the carboxy terminus of peplomers. Proc. Natl. Acad. Sci. USA. 84: 6567-6571.

Talbot, P.J., Dionne, G. and Lacroix, M. (1988) Vaccination against lethal coronavirus-induced encephalitis with a synthetic decapeptide homologous to a domain in the predicted peplomer stalk. J. Virol., 62: 3032-3036.

Tooze, S.A., and Stanley, K.K. (1986) Identification of two epitopes in the carboxyterminal 15 amino acids of the E1 glycoprotein of mouse hepatitis virus A%9 by using hybrid proteins. J. Virol. 60: 928-934.