Selection of peptide ligands for the antimucin core antibody C595 using phage display technology: definition of candidate epitopes for a cancer vaccine

P Laing, P Tighe, E Kwiatkowski, J Milligan, M Price, H Sewell

Abstract

Aims—To further define the specificity of the antimucin core antibody C595 by fitting it with a family of hexapeptide ligands by immunoselection of filamentous bacteriophage from a gene III display library of approximately 6•4 x 10^7 random hexapeptides.

Methods—Three rounds of immunoselection were used to enrich for C595 binding phage. DNA sequencing revealed the hexapeptides expressed. Bacteriophage and corresponding synthetic hexapeptides were used in ELISA assay to determine binding affinities.

Results—Twenty nine clones from this selected population were analysed. Seven contained the natural epitope RPAP, encoded by two different DNA sequences; 17/29 contained the motif RLPP. In all, 28/29 clones contained the motif RXXP and one clone (RVRPAP) contained the motif RXXP in two peptidic registers; 24/28 clones (6/8 DNA sequences) contained a hydrophobic residue (V or I) at position 1 relative to the RXXP motif. In addition the proximity of RXXP to glycine (position 5) suggests that this contributes in the natural epitope to antibody/antigen binding, which was not detected by chemical synthetic methods. One clone, KSKAGV, bears no obvious relationship to the natural epitope and therefore qualifies as a weakly binding mimotope.

Conclusions—This approach has rapidly defined the specificity of this antibody in an unprecedented detail, and provides a more comprehensive molecular basis for exploring the immune recognition of the MUC1 mucin by the C595 antibody. Importantly, the novel but related epitopes seen provide peptide specificities and a strategy which may prove useful in generating cancer vaccine candidates.

Keywords: MUC1 mucin, epitope library, peptide.

Human polymorphic epithelial mucins (PEM; MUC1 mucin; carcinoma associated epithelial mucin) are complex high molecular mass glycoproteins with lubricant and anti-infective properties in normal epithelium, but which are also present in altered forms in human breast carcinoma and other carcinomas.1 The PEM show genetic polymorphism evident at the DNA and protein level.2,3 They comprise a single chain polypeptide core, which consists predominantly of a highly conserved tandem repeat of a 20 amino acid sequence.4 Oligosaccharide side chains are attached in O-glycosidic linkage to serine and threonine residues of the core, and differences in the glycosylation status of the mucins in neoplasia have been implicated in their recognition by tumour selective monoclonal antibodies.1 Recently Price et al5 by synthesising multiple overlapping heptapeptides, have identified a tumour specific PEM epitope RPAP (arginine–proline–alanine–proline) recognised by a murine monoclonal antibody C595. Price et al5 have also delineated the fine specificity of this antibody by comprehensive replacement analysis of individual amino acid residues which comprise the epitope. These studies, using multiple peptide synthesis methodology,6 have provided detailed information about the nature of the epitope recognised by C595, and established the important contribution of R and P residues at the epitope boundary as determinants of binding. All of the peptides identified were analogues of the natural epitope which contained the critical RXXP motif.

As an independent means of assessing the epitope specificity of C595 we have applied the approach of Scott and Smith.7 Single random hexapeptides are encoded by, and expressed on, the surface of a filamentous bacteriophage “epitope library”. The peptide being expressed as part of the gene III (pIII) attachment protein is thus limited in expression to approximately five molecules per virion. Immunoselection of phage via the binding affinity of the expressed peptide allows purification of peptides which bind tightly to the antibody, the peptide sequence being determined after propagation and appropriate sequencing of the bacteriophage gene III coding region. This technique has the advantage that essentially all 64 million possible hexapeptides can be surveyed simultaneously for antibody binding. Also, it is not subject to the uncertainties of quality control of peptide synthesis which are inherent in multipeptide synthetic paradigms. In principle this allows a more comprehensive assessment of the peptide specificity of an antibody, and makes it possible to carry out multiple trials of diverse permutations of amino acid sequences which are beyond the capacity of conventional multiple
Peptide ligands for antimucin core antibody

Peptide synthesis methodology. The technique offers an entirely different methodological approach to the identification of epitopes, with advantages and disadvantages which are complementary to the chemical synthetic approach.

By comparing the multiple synthetic peptide route and "epitope display library" methods using a well characterised antibody (C595) it should therefore prove feasible to assess the accuracy of the phage display technique for identifying a known epitope, its high affinity substitution variants, and potential mimotope sequences. This will help evaluate the potential for routine use of epitope libraries for the identification of essential contact residues in the epitope-paratope interaction. In addition it is possible to evaluate whether the phage library technique has the capacity to identify common residues flanking the epitope motif, which might favour epitope presentation and optimal expression of antibody binding activity.

We now report the identification of peptide ligands for C595 by the phage display technique. We find that this method allows delineation of the fine specificity of this antibody in unprecedented detail, and extends the chemical data on the epitope specificity of C595.

Methods

**EPITOPE LIBRARY**

The bacteriophage FUSE5 hexapeptide display library and *Escherichia coli* K91 Kan were kindly provided by George P Smith. This library is a bacteriophage construct derived from fd-tet. Single random hexapeptides are expressed as part of the bacteriophage pIII attachment protein of which there are five copies at one tip of the virion. The sequences originally inserted to generate the library included a degenerate region (NNK)6, where N = any nucleotide and K = G or T deoxynucleotides only. The 6 NNK triplets can thus code for all 20 amino acids, but one stop codon (amber) only. Each recombinant thus encodes one hexamer of essentially random origin.

**MONOCLONAL ANTIBODY C595**

C595 is a murine IgG3 subclass monoclonal antibody which was raised by immunisation of BALB/c mice with human urinary mucin obtained by affinity purification using the monoclonal murine IgM antibody NCRC-11. The epitope recognised by C595 is present in polymorphic epithelial mucins expressed by breast and ovarian carcinomas. C595 has been used for tumour imaging and for evaluation of the prognostic value of mucins of tumour origin shed into the circulation of cancer patients.

**IMMUNOSELECTION OF PHAGE EXPRESSING ANTIGENIC PEPTIDES**

C595, 500 μl, was prepared at a concentration of 200 μg/ml in Tris buffered saline (TBS), pH 7.7, placed in a Nunc 1.5 ml cryopreservation tube (Life Technologies) containing a 6.5 mm etched polystyrene ball of 1 μg IgG binding capacity (Northumbria Biologicals), and left for two hours at room temperature. Blocking of potential non-specific binding sites was carried out by immersing the ball in TBS + 5% bovine serum albumin (BSA) for 20 minutes. The ball was placed in 300 μl of TBS + 0.05% Tween 20 + 0.1% BSA (TBS-Tween); 5 μl of the hexapeptide library (5 x 10⁶ phage) was then added, the tube capped and the mixture incubated at 37°C for three hours with rotation on a blood mixer, to facilitate phage/antibody interaction. The ball was then extensively washed with TBS-Tween and transferred to a sterile Nunc tube. Bound phage was eluted using 500 μl of elution buffer (0.1 M HCl, pH adjusted to 2-2 with glycine, 1 mg/ml BSA, 0.1 mg/ml phenol red). The ball and buffer were inverted five times in the capped tube, and the buffer then neutralised by the addition of a few microlitres of 0.05 M Tris HCl, pH 8-8. This material containing eluted phage was then concentrated and desalted using Millipore-Ultra MC filters (30 000 Da nominal molecular weight limit), by centrifugation in a microcentrifuge at 8000 g for 30 minutes. Phage retained by the membrane was taken up in 200 μl TBS-1% gelatine.

**ENUMERATION OF VIABLE PHAGE**

Since the phage gives tiny plaques, it is more practical to titrate it as tetracycline transducing units (TU). *E Coli* are infected with the phage, and after a period of gene expression induction, are spread on plates containing 30 μg/ml tetracycline, thus positively selecting transductants.

**AMPLIFICATION OF C595 SELECTED BACTERIOPHAGE**

Following elution and concentration, the phage preparation (200 μl) was mixed with an equal volume of K91 Kan in log phase (grown in LB medium, Sigma) in a sterile 1-5 ml Eppendorf tube and left for 30 minutes at room temperature. The infected *E Coli* were then transferred into 5 ml LB medium prewarmed to 37°C containing 0-2 μg/ml (inducer concentration) of tetracycline, and shaken for 50 minutes at 37°C. Further tetracycline was then added to this culture to a concentration of 20 μg/ml and shaken overnight at 37°C.

**PURIFICATION OF PHAGE**

Following overnight growth of the culture, phage particles were purified by standard PEG/NaCl precipitation. Resultant phage pellets were resuspended in sterile water. Average yields were 10⁵² TU of phage, stored at 4°C.

**ssDNA PREPARATION AND SEQUENCING**

Isolated tet* K91 kan colonies were picked from a culture plate and used to inoculate 5 ml of LB broth containing 50 μg/ml of tetracycline. After overnight growth at 37°C single stranded phage DNA was isolated from culture supernatant by standard methods.
In order to minimise requirements for virion ssDNA and avoid the problems of variable yields of DNA from fUse5, cycle sequencing was carried out using a commercially available kit (USB), with a primer, 5'-ACACTGAGTTTCGTCACCAGT-3' (Oswel DNA Service), to target the 3' end of the hexamer encoding region of pH1. This primer was set so as to incorporate seven molecules of (α-35S)dATP during the extension stage of the reaction, in order to enhance later reading of the sequence.

Cycle sequencing was carried out on a Hybaid "Omnigene" thermal cycler. The programme was set to act as follows: labelling reaction: 50 cycles of 95°C, 15 s; 55°C, 30 s. This was followed by a termination reaction comprising 50 cycles of 95°C, 30 s; 70°C, 60 s. All other procedures followed the manufacturer's instructions. Sequencing gels were performed using Bio-Rad equipment.

PHAGE CAPTURE ENZYME LINKED IMMUNOSORBENT ASSAY (ELISA)
To confirm specific binding of C595 to the clones eluted from the coated polystyrene ball, representative phages containing the sequences seen in the table were propagated in K91 Kan, as was a control phage known to contain the irrelevant insert sequence KFRFVFG. Following purification, the pellets were suspended in water and protein concentration was determined using a bicinchoninic acid kit (Sigma). Phage concentrations were adjusted using phosphate buffered saline (PBS), 1% BSA, 0-05% Tween 20 as diluent by appropriate serial dilutions in the range of 0-05 to 100 μg/ml.

A microtitre plate (Nunc) was coated with C595 at a concentration of 10 μg/ml in PBS. A control plate was coated with a murine IgG monoclonal antibody 791T/36, with colorectal tumour membrane glycoprotein specificity, which was not expected to recognise C595 reactive phage. Both plates were incubated overnight at room temperature, washed in PBS, and blocked for one hour with PBS-1% BSA, followed by washing with PBS-BSA-Tween 20; 50 μl of phage at each dilution were then applied to the plates, leaving a control well free of phage, followed by incubation for one hour at room temperature. The plates were then washed, and 50 μl of sheep anti-bacteriophage M13 antiserum (CP Laboratories) at a dilution of 1 in 500 added to all wells. Incubation was for one hour, followed by PBS-BSA-Tween washing, and addition of donkey anti-sheep alkaline phosphatase conjugate (Sigma) at a dilution of 1 in 1000 and incubation for one hour. The plates were then washed, and disodium p-nitrophenyl phosphate substrate added. Optical densities were read at 405 nm.

BINDING INHIBITION ASSAY
The following peptides were synthesised (Severn Biotech) and investigated for their ability to inhibit binding of C595 to antigen containing the immunodominant RPAP sequence: KSKAGV; DTRPAP (natural antigen sequence); VRYPRP; and RVRPAP. A Nunc Immuno Maxisorb microtitre plate (Life Technologies) was coated for 18 hours at room temperature with an artificial antigen containing the sequence RPAP. The antigen is termed AK-CG, and is composed of the amino acid sequence CAPDTRPAG conjugated to a synthetic branched chain peptide polymeric Carrier with a polylsine backbone (a gift from F Hudecz, Hungary). The coating of AK-CG was applied at a concentration of 8 μg/ml, 50 μl per well. Following plate coating and washing three times with PBS pH 7-2, PBS-1% BSA (100 μl per well) was used to block non-specific binding sites for one hour at room temperature.

Peptides for inhibition of C595 binding were diluted in 10-fold steps to give a range of 10^{-3} to 10^{-7}M in PBS-BSA-Tween. The plate was then washed four times with PBS-BSA-Tween, and duplicate 50 μl values of peptide at each concentration applied to the plate; 50 μl of C595 at a concentration of 2 μg/ml was then applied to all wells containing peptides.

To obtain a value for maximum C595/AK-CG binding, wells containing no peptide were also present on the microtitre plate, as were blank wells containing no peptide on C595, but coated with AK-CG only.

Following applications of peptide and C595, the plate was vigorously shaken for two minutes to ensure uniform mixing of C595 and peptides.

After one hour at room temperature, the plate was washed four times with PBS-BSA-Tween; 50 μl of sheep anti-mouse IgG conjugated to alkaline phosphatase (Sigma) was then added and incubated for one hour. After

---

**DNA sequence and encoded hexapeptide inserts from independent clones selected by monoclonal antibody C595**

<table>
<thead>
<tr>
<th>DNA sequence of insert (5'-3')</th>
<th>Amino acid sequence of insert</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGT GAT GCT CAG CGG GCT CTT</td>
<td>g a r V a P a g a a</td>
<td>5</td>
</tr>
<tr>
<td>TGG GAG AGG CGG CTT GCT GCT</td>
<td>g a l E a P a g a a</td>
<td>2</td>
</tr>
<tr>
<td>CTT CGT CTG CCG CCT AAAT</td>
<td>d a v l N a L P N g a</td>
<td>9</td>
</tr>
<tr>
<td>AAT ATT CGT TGG CCT CTT</td>
<td>g a N I a L P g a a</td>
<td>7</td>
</tr>
<tr>
<td>AAG ATT CGT TGG CCT CTT</td>
<td>g a K I a L P g a a</td>
<td>1</td>
</tr>
<tr>
<td>GAG CGT CCG GCT CCT CGG</td>
<td>d a e E a R A a G g a</td>
<td>2</td>
</tr>
<tr>
<td>GAG ATT CGT AGG CGG CCT</td>
<td>g a a E a R A a G g a</td>
<td>1</td>
</tr>
<tr>
<td>GTG CCG ATT CCT CCT CGT</td>
<td>d a v K S K A G V a a</td>
<td>1</td>
</tr>
</tbody>
</table>

Total = 29

1Invariant amino acid sequences of vector regions flanking the variable inserts are in lower case letters.
2The RXXP motif in the hexapeptides is aligned and shaded for comparative purposes.
Peptide ligands for antimucin core antibody

Results
PEPTIDE EPITOPE LIBRARY PANNING
Antibody panning with C595 was performed for three rounds, with propagation of affinity selected recombinant bacteriophage between each round. Yields of phage from each round of panning were $4 \times 10^5$, $1.5 \times 10^6$, and $6.4 \times 10^7$ for rounds 1, 2, and 3 respectively. Affinity selected phages obtained after the third round of panning were maintained as 'K91 Kan transformants and 29 were picked at random for further analysis.

DNA SEQUENCING AND EXPRESSED PEPTIDE ANALYSIS
The DNA sequence of the gene III insert of each phage clone was determined. From this the codon usage and amino acid sequence of the displayed hexamer peptide inserts could be deduced. These are presented in the table.

It is striking that 28 of 29 clones examined contained an RXXP motif, and this is indicated in the table. These sequences can be further subdivided by similarity at both the DNA and protein level. The 29 clones represent nine different DNA sequences, and thus nine unique members of a hexamer peptide display library.

Seven of 29 clones encoded peptide hexamers containing the natural motif RPAP. These originate from two separate DNA sequences. The hexamer VRVPRAP actually presents the RXXP motif in two separate "registers" (VRVPRAP and RVPRAP). A non-natural motif was most commonly seen, with 17/29 clones (three DNA sequences) expressing peptides with a core motif of RLPP. The most common sequence VRLPPN (9/29) has the RXXP motif in a different "register" (−1 relative to the majority of the hexamers observed) which is seen in only three other clones (two DNA sequences). Only 1/29 clones expressed a completely different peptide sequence (KSKAGV).

ANTIBODY BINDING AFFINITY OF BACTERIOPHAGE CLONES
Bacteriophages were prepared and isolated from each transformed K91 kan host. Titration of purified phage against immobilised C595 antibody in ELISA revealed a range of binding affinities (fig 1A). Of the phages tested, only KSKAGV and the control negative phage (expressing KFRFVG) showed very low binding affinity to C595. In titration with an irrelevant antibody (791T/36) the specific nature of the affinity of the selected phage displayed peptides is revealed, none binding significantly to this antibody (fig 1B).

ANTIBODY BINDING INHIBITION ASSAYS
In order to study the specificity of these displayed peptides in more detail, the hexamer sequences expressed by three phages were chemically synthesised in order to test their ability to inhibit binding of C595 to a known antigen (AK-CG) containing the C595 target motif RPAP. In addition a hexamer sequence from the native antigen (DTTRPAP) was also synthesised. The inhibition curves obtained are shown in fig 2. A negative control peptide APTPGRR which was not expected to bind to C595 shows, as expected, no inhibition until high concentrations are used. By contrast, the peptides containing phage derived sequence show inhibitory activity, with 50% inhibition...
Figure 2  Competitive inhibition of IgG monoclonal antibody C595 binding to solid phase synthetic antigen AK-CG by synthetic peptides of identical sequence to the hexamers expressed on bio-panned bacteriophage. All peptides are diluted in the range 10⁻² to 10⁻⁴ M. Peptide DTRPAP (●) represents the natural epitope sequence and APTPGR (●) is used as an irrelevant peptide control.

concentrations (Cl₅₀) from approximately 10⁻² molar (RVRPAP) up to 10⁻⁴ molar (KSKAGV). Peptide RVRPAP inhibits antibody binding significantly better than the natural epitope DTRPAP (Cl₅₀ ≈ 2 x 10⁻⁷ M). The fact that KSKAGV does inhibit C595 binding to AK-CG suggests it is a true peptide ligand, but of low affinity.

Discussion
The sequence of the 20 amino acid tandem repeat core of the MuC-1 PEM containing the C595 reactive epitope is PDTRPAPGS-TAPAHGVTSA. The importance of the arginine and proline residues in the C595 epitope, as found by Price et al using synthetic methods, is fully confirmed in the present study.

Twenty nine characterised individual phage clones, resolved to eight independent hexamer sequences at the DNA level, were found to encode the RXXP motif (table). However the frequency of isolation of the clones was different, which may be indicative of both binding affinity, that is, competition for antibody between the phages, and propagation characteristics of individual phage in the hot cell. The demonstration of considerable differences in binding affinity (fig 2) between individual peptides suggests that this may be a significant factor. Of the clones obtained, 18/29 (six different DNA sequences) contain glycine following the terminal proline of this motif, as seen in the natural MuC 1 epitope. In five of these six, glycine is part of the invariant sequence which follows the insert. In the remaining instance, glycine may contribute as a determinant of 595 antibody binding. This is in contrast to the chemical synthetic data reported previously by Price et al, which found no measurable contribution of glycine to antigenicity. However, the constraints upon conformation imposed by pin display and phage display may differ, and perhaps the conformational flexibility afforded by glycine is more important in the phage construct, in allowing the epitope to assume its appropriate conformation.

A further feature of the clones analysed is the presence of either valine (V) or isoleucine (I) at position 1 relative to the RXXP motif. This is seen in 24/28 clones, corresponding to 6/8 unique DNA sequences. These residues are more hydrophobic than the native threonine (T) found at that position; however, all three amino acids possess a methyl (−CH₃) group attached to the first carbon of the side chain. This suggests that a methyl group presented at this position may also play a part in determining the binding characteristics of C595. It is interesting that in the phage clones examined the motifs RAAP and RPPP, which were recognised as strongly by C595 as was the native RPAP motif when examined by tetramer synthetic peptide means, were not found. This may be a statistical phenomenon, reflecting the number of clones examined, but may equally be further evidence that the flanking amino acids play an important role in presentation of the RXXP motif to the antibody. Wreschner et al indicate that the RPPP sequence may be encoded in place of RPAP in the PEM by alternative splicing of gene transcripts. Our data suggest, purely by the absence of clones, that C595 does not strongly bind the motif RPPP when presented in the context of other, flanking, residues.

A ninth independent phage clone displayed the peptide KSKAGV. This sequence bears no strong resemblance to either the “natural” MUC 1 peptide or any of the other peptide ligands identified. KSKAGV may, however, be somewhat analogous to the natural epitope, this being inferred by hydrophobicity patterns in DTRPAP and KSKAGV. DTR is comprised of three hydrophobic residues; similarly KSK is hydrophilic, whereas PAP and AGV are comprised of neutral or moderately hydrophobic residues.

The antigenicity of the human polymorphic epithelial mucins is of interest from the point of view of vaccine development. Thus synthetic carbohydrate and peptide epitopes of this mucin have been used to provoke antitumour immunity in patients with breast cancer. However a fundamental problem with the elicitation of immune responses against tumour antigens is that these antigens are protected by immunological tolerance of self components. Since in the present study we have identified a series of analogues which are antigenically equivalent to the natural mucin epitope recognised by antibody C595, and also a mimotope sequence which is structurally different from the natural epitope, this knowledge may be applied to enhance the immunogenicity of mucin vaccine immunogens. It is possible that the approach described here could be used to generate a library of analogues and mimotopes of mucin epitopes which could be used in the
Peptide ligands for antimucin core antibody

We would like to thank Prof George Smith and Dr Jamie Scott for providing the epitope library and helpful discussion at the onset of this work.


