

# Antibodies Immunoreactive With Formalin-Fixed Tissue Antigens Recognize Linear Protein Epitopes

Seshi R. Sompuram, PhD,<sup>1,2</sup> Kodela Vani, MS,<sup>1</sup> Laurie J. Hafer, PhD,<sup>1</sup>  
and Steven A. Bogen, MD, PhD<sup>1,2</sup>

**Key Words:** Formalin fixation; Antigen retrieval; Monoclonal antibody; Epitope; Protein; Immunohistochemistry

DOI: 10.1309/6H0ARQF7K3Y608EH

## Abstract

*It is not clearly understood why some monoclonal antibodies bind to their antigens in formalin-fixed, paraffin-embedded tissue sections but others do not. To address this question, we analyzed the protein epitopes of 9 monoclonal antibodies that are immunoreactive after formalin fixation and antigen retrieval. We identified the antibody contact sites by using phage display and synthesized corresponding peptides derived from the GenBank database sequence that contain the predicted antibody binding sites. Our data indicate that all 9 antibodies bind to linear epitopes, ie, composed of contiguous amino acids. In addition, the amino acids proline, tyrosine, glutamine, and leucine are highly represented in these antibody contact sites. The epitopes tend to be mildly to moderately hydrophilic. These findings are the first detailed studies of antibody epitopes associated with antigen retrieval and suggest that antibodies must recognize linear sequences to bind after formalin fixation and antigen retrieval.*

Restoration of immunoreactivity to tissue antigens following antigen retrieval is a cornerstone technology for modern histopathologic diagnosis. Although the antigen-retrieval technique has been practiced for slightly more than a decade, the protein chemistry explaining it is poorly understood. A variety of potential mechanisms have been forwarded as explanations for the restoration of immunoreactivity after antigen retrieval. These include extraction of diffusible blocking proteins, precipitation of proteins, and rehydration of the tissue section, thereby allowing better penetration of antibody<sup>1</sup>; removal of cage-like calcium complexes<sup>2</sup>; and heat mobilization of trace remaining amounts of paraffin.<sup>3</sup> Based on the chemistry of formalin reactions, recently reviewed by Shi et al,<sup>4</sup> the most likely mechanisms are associated with protein cross-linking. Recently, formalin fixation and antigen retrieval were analyzed using model systems by biophysical techniques.<sup>5,6</sup> Those studies suggested that the formaldehyde reaction creates intramolecular and intermolecular cross-links, many of which are broken on antigen retrieval. Our own recent study using peptides that mimic antibody binding sites further suggested that these cross-links may be formed, at least in part, by the Mannich reaction.<sup>7</sup>

A variety of monoclonal antibodies (mAbs) have gained popularity because of their usefulness in conjunction with formalin-fixed tissue samples after antigen retrieval. These mAbs represent a select subset, based on their clinical usefulness and ability to stain formalin-fixed, paraffin-embedded tissue sections after antigen retrieval. Antibodies that are not immunoreactive after antigen retrieval generally are not used for routine clinical immunopathology. It is not known why some antibodies are immunoreactive after antigen retrieval and others are not. The ability of an antibody to stain tissue proteins after

formalin fixation and antigen retrieval is widely believed to be due to an intrinsic characteristic of the antibody's epitope to which it binds. According to this hypothesis, protein epitopes become inaccessible or conformationally altered by formalin fixation. Boiling the tissue section, such as in antigen retrieval, restores immunoreactivity by reversing the denaturing change(s) caused by formalin fixation. Although this hypothesis is widely believed to be true, there has not been a detailed analysis of antibody epitopes to identify characteristics that make some epitopes suitable for antigen retrieval and others not. To further study this question, we characterized the protein epitopes to which many popular, clinically useful antibodies bind.

Because these antibodies were selected over the years as useful in conjunction with antigen retrieval, we wondered whether their epitopes might share common characteristics associated with restoration of immunoreactivity after antigen retrieval. By using phage display of peptide combinatorial libraries, we identified the specific antibody contact sites (epitopes) for 8 clinically useful mAbs to estrogen and progesterone receptors (ER, PR, respectively), HER-2, and p53. The epitope of a ninth, the MIB-1 antibody to Ki-67, was identified and characterized previously<sup>8,9</sup> and incorporated into these studies. These antibodies are immunoreactive after formalin fixation and antigen retrieval. We then synthesized linear peptides containing the predicted 5 to 7 amino acid contact sites and flanking amino acids on both sides to confirm the binding specificity of each antibody. We report our finding, that the epitopes of each of the 9 mAbs, all of which are useful in the context of formalin-fixed tissue samples (after antigen retrieval), share an important characteristic.

## Materials and Methods

### Peptides and Antibody Reagents

ER, PR, and Ki-67 protein-specific mAb clones ER1D5, ER2-123, PR636, PR1294, and MIB-1, as well as anti-HER-2 rabbit polyclonal antibody (HercepTest), were obtained from DakoCytomation, Carpinteria, CA; clones ER6F11, PR1A6, and CB11 (HER-2) were obtained from Vision Biosystems, Norwell, MA. Mouse polyclonal IgG was obtained from Sigma Chemical, St Louis, MO. Immunohistochemical detection reagents were purchased from DakoCytomation. Synthetic peptides were custom manufactured to our specifications by Synpep, Dublin, CA.

### Immunohistochemical Staining

Peptides that were coupled to isocyanate-activated slides were subjected to a standard immunohistochemical staining procedure, as described earlier.<sup>10</sup> All immunohistochemical

stains were performed on formalin-fixed and paraffin-embedded human tissue samples. Serial tissue sections were deparaffinized in xylene and rehydrated in decreasing concentrations of ethyl alcohol to distilled water. Antigen retrieval was performed by incubating the slides in a pressure cooker (Nordic Ware, Minneapolis, MN) for 30 minutes in 0.1 mol/L of citrate buffer (pH 6.0).

Immunohistochemical staining was performed on the Artisan Staining System (DakoCytomation). Primary antibodies (to ER, PR, HER-2, or Ki-67) were incubated with the tissue sections at approximately 3 µg/mL at room temperature for ER1D5 mAb and at 37°C for others. Incubations were for 40 minutes for the ER1D5 mAb and 20 minutes for others. After the primary antibody, the slides were rinsed and incubated with biotin-conjugated horse antimouse IgG (heavy and light chain-specific) secondary antibody for 20 minutes at room temperature for the ER1D5 mAb and 35°C for others. The slides were rinsed again and incubated with horseradish peroxidase-conjugated streptavidin for 20 minutes at room temperature for ER1D5 and 35°C for others. The color was produced with liquid stable diaminobenzidine (3,3'-diaminobenzidine tetrahydrochloride)/hydrogen peroxide for 20 minutes at room temperature for ER1D5 and 35°C for others and enhanced with 5% (wt/vol) copper enhancer (cupric sulfate pentahydrate) for 20 minutes at room temperature for ER1D5 and 35°C for others.

### Peptide Coupling to Glass Slides

Synthetic linear peptides were coupled covalently to the protected isocyanate (PI)-derivatized glass surface of microscope slides, as previously described.<sup>10,11</sup> For coupling to isocyanate-coated microscope slides, appropriate working dilutions of peptides were made in potassium phosphate (0.5 mol/L, pH 8.9) buffer. Briefly, various peptides (1 µL) were spotted onto activated, isocyanate-derivatized microscope glass slides. The peptides were permitted to covalently couple to the glass for 15 minutes at 42°C. The slides then were rinsed, and the remaining reactive isocyanate groups were quenched with a protein cocktail containing bovine gamma globulins (0.025%, Sigma Chemical).

### Peptide Specificity

ER, PR, and Ki-67 peptides were tested for binding to different mouse mAbs or to a mouse polyclonal antibody. For each peptide, a row of 10 spots was applied to a PI-activated microscope slide. Each spot was composed of approximately 20 pmol in 1.5 µL potassium phosphate (0.5 mol/L, pH 8.9) buffer. Various antibodies then were reacted with the spots: ER clones 1D5, 6F11, and ER2-123; PR clones 636, 1294, and 1A6; HER-2 clone CB11; Ki-67 clone MIB1; rabbit polyclonal HercepTest antibody; and mouse polyclonal IgG. Each spot was separated from the others on the slide by drawing

hydrophobic barriers with a Pap pen (Research Products International, Mount Prospect, IL). After incubation with the primary antibodies, all spots were developed with the Artisan DAB Detection System (DakoCytomation).

### Inhibition of mAb Binding by Peptide

We tested the ability of the peptides to inhibit the binding of antigen-specific mAbs to native antigen. As a source of native antigen, we used tissue biopsy sections containing cells that strongly express the antigen in question. If a peptide inhibits the binding of a mAb to native antigen, then that is detectable as inhibition of immunohistochemical staining on the tissue section. Aliquots of mAbs were incubated at 37°C for 45 minutes with various concentrations of peptides. The mAb-peptide combination then was used as a primary antibody in an immunohistochemical assay. The immunohistochemical assay was performed as described in the preceding text. Immunohistochemical staining was quantified on a 5-point scale: 0, no staining; 1, faint staining; 2, weak but detectable staining; 3, moderate staining intensity; 4, intensely strong staining.

### Protein Epitope Mapping

Antibody epitopes were mapped by analyzing the consensus peptide sequence inserts obtained after successive rounds of biopanning. Homologous areas to the consensus sequence were identified based on the GenBank sequence for that particular protein.

We used phage libraries that contained rationally designed combinatorial libraries of peptide sequences inserted into the NH<sub>2</sub> terminus of the pIII minor coat protein of the M13 bacteriophage. The libraries were prepared and supplied by Dyax, Cambridge, MA. The libraries (termed *TN6*, *TN8*, *TN9*, *TN10*, and *TN12*) contained 2 conserved cysteine residues separated by 4, 6, 7, 8, and 10 amino acids, respectively. The cysteines formed a disulfide bridge, creating a conformationally constrained ring.<sup>12</sup> The amino acids within the ring and the 3 amino acids on either side of the ring were diversified to allow all amino acid types (except cysteine) with equal frequency, using trinucleotide-mutagenesis technology.<sup>13</sup> Trinucleotide-mutagenesis technology involves controlled polymerization of preformed trinucleotides.

The libraries were screened by biopanning using standard methods with a few modifications.<sup>14-18</sup> Briefly, paramagnetic beads coated with anti-mouse IgG (Dynabeads, Dynal, Lake Success, NY) were prepared by mixing with the primary mouse mAb (for positive enrichment) or the polyclonal mouse IgG (for negative depletion) and incubating overnight at 4°C on a rotator. Antibody-adsorbed Dynabeads were washed 4 or 5 times with phosphate-buffered saline (PBS) containing 0.05% polysorbate (Tween)-20 (PBST) and twice with PBS before use in biopanning of phage libraries.

A phage library containing 10<sup>11</sup> to 10<sup>12</sup> plaque-forming units was negatively depleted by incubation with Dynabeads (100 µL) coated with polyclonal mouse IgG for 1 hour at room temperature on a rotator. With this negative depletion step, phages that were bound to common, shared regions of mouse IgG were depleted from the library by attaching to the magnetic beads. The unbound phages (supernatant) then were positively selected on the ER, PR, or HER-2-specific target mAb-adsorbed Dynabeads and incubated for 2 to 3 hours on a rotator. The beads were washed 10 times with PBST and 3 times with PBS to remove nonspecifically bound phage.

Phages that were bound to the mAb-coated beads were eluted twice to ensure that high-affinity phages were completely eluted. The beads were eluted first with a buffer containing 0.1 mol/L of glycine hydrochloride (pH 2.2) and then by 0.1 mol/L of citrate buffer. After elution, the acidic buffers were neutralized with 2 mol/L of tris(hydroxymethyl)aminomethane hydrochloride (pH 9.0), and the supernatants were pooled. The eluted phages then were amplified and used in a second round of biopanning. The second round panned and eluted phages were diluted to infect log phase XL-1 Blue *Escherichia coli* (Stratagene, La Jolla, CA) and grown on an LB-agar plate covered with 8% agarose to isolate individual phage clones.

Phage clones that were highly immunoreactive with the primary antibody to which they were selected were submitted for further analysis by sequencing the nucleotide inserts coding for the combinatorial peptides. The sequencing template was prepared by polymerase chain reaction (PCR) amplification from an overnight phage culture. The primers used for sequencing were 5'-CGGCGCAACTATCGGTATCAAGCTG-3' and 5'-CATGTACCGTAACACTGAGTTTCGTC-3'.

We performed 30 rounds of PCR on an MJ Research Tetrad thermocycler (MJ Research, Waltham, MA). The PCR product was diluted 1:20 with distilled water. Sequencing was performed in the forward and reverse directions with the following primers: 5'-GATAAACCGATACAATTAAGGCTCC-3' and 5'-GTTTTGTCGTCTTCCAGACGTTAG-3'. ABI Big Dye (version 1.0) (Perkin Elmer Applied Biosystems, Foster City, CA) was used to perform a 5-µL sequencing reaction (2 µL of Big Dye, 1 µL of distilled water, 0.5 µL of primer [at 3 pmol/µL], and 1.5 µL of diluted PCR product). The samples were cycled for 45 rounds on an MJ Research Tetrad thermocycler. After cycling, 2.5 volumes of absolute ethanol were added, and the mixture was centrifuged at 1,850g for 30 minutes. The tubes were inverted over paper towels and then centrifuged at 100g for 30 minutes. The samples were resuspended in 5 µL of distilled water and detected on an ABI 3700 DNA Analyzer (Perkin Elmer Applied Biosystems). The results were analyzed using Phred and Phrap to generate a consensus from the forward and reverse readings.<sup>19,20</sup>

## Results

### Epitope Mapping mAb Contact Sites

By using phage display, we mapped the precise binding sites for 8 of the 9 mAbs used for this study: ER antibodies 1D5, ER2-123, and 6F11; PR antibodies PR636, PR1294, and 1A6; HER-2 antibody CB11; and p53 clone DO-7 (DakoCytomation). The MIB-1 binding site already was described,<sup>8,9</sup> allowing us to synthesize the corresponding peptide without mapping. A representative example of epitope mapping by phage display is shown in **Table 1**, which illustrates a list of aligned peptides that were isolated as binders to the ER1D5 mAb.

Each amino acid is represented using the single letter amino acid code. At the bottom of Table 1, we identify the deduced amino acid code that is predicted by the phage display–derived sequence data. Namely, the sequence data predicted the contact site to be Q-A-P-Y(Y or F), corresponding

**Table 1**  
Epitope Mapping by Phage Display for Estrogen Receptor Antibody Clone 1D5\*

| Native Sequence   |   |          |          |          |          |            |
|-------------------|---|----------|----------|----------|----------|------------|
| G                 | Q | Q        | V        | P        | Y        | Y          |
| Raw Data          |   |          |          |          |          |            |
| –                 | T | <b>Q</b> | T        | <b>P</b> | <b>Y</b> | H          |
| S                 | H | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | P          |
| <b>G</b>          | – | <b>Q</b> | A        | <b>P</b> | <i>F</i> | <b>Y</b>   |
| L                 | D | <b>Q</b> | T        | <b>P</b> | <b>Y</b> | –          |
| N                 | S | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | <b>Y</b>   |
| –                 | T | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | W          |
| R                 | C | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | <b>Y</b>   |
| –                 | D | <b>Q</b> | <b>V</b> | <b>P</b> | <b>Y</b> | L          |
| –                 | D | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | <b>Y</b>   |
| <b>G</b>          | Y | <b>Q</b> | W        | <b>P</b> | <b>Y</b> | –          |
| N                 | Y | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | I          |
| –                 | F | <b>Q</b> | K        | <b>P</b> | <b>Y</b> | W          |
| N                 | – | <b>Q</b> | T        | <b>P</b> | <b>Y</b> | <b>Y</b>   |
| –                 | S | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | <i>F</i>   |
| –                 | S | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | <i>F</i>   |
| D                 | R | <b>Q</b> | A        | <b>P</b> | W        | <b>Y</b>   |
| E                 | N | <b>Q</b> | Q        | <b>P</b> | <b>Y</b> | <i>F</i>   |
| P                 | N | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | –          |
| N                 | G | <b>Q</b> | R        | <b>P</b> | <b>Y</b> | L          |
| <b>G</b>          | W | <b>Q</b> | A        | <b>P</b> | <b>Y</b> |            |
| K                 | P | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | V          |
| D                 | P | <b>Q</b> | A        | <b>P</b> | <b>Y</b> | <i>F</i>   |
| –                 | T | <b>Q</b> | T        | <b>P</b> | <b>Y</b> | H          |
| D                 | W | F        | A        | <b>P</b> | <b>Y</b> | V          |
| D                 | P |          | Q        | <b>P</b> | <b>Y</b> | K          |
| Deduced sequence: |   | <b>Q</b> | <b>A</b> | <b>P</b> | <b>Y</b> | <b>Y/F</b> |

\* The native estrogen receptor sequence from the GenBank database is listed at the top, starting with position 125 (glycine, G). Dashes represent invariant amino acids in the phage vector, which are not helpful in identifying the amino acid. Amino acids whose identities were unclear from the sequencing data are represented with an empty space. Amino acids that exactly match the native sequence are in bold type. Amino acids that do not exactly match but are conserved substitutions are italicized. Amino acids are represented by the single letter amino acid code.

to glutamine (Q), alanine (A), proline (P), tyrosine (Y), and tyrosine or phenylalanine (F). Based on the sequencing data, the last position would be predicted to be tyrosine or phenylalanine, both of which have aromatic side chains. These amino acids align with a sequence starting at position 127 of the native ER protein, shown at the top of Table 1. The only divergence of the predicted sequence from the actual native ER protein sequence is at position 128, where we would predict an alanine instead of the valine that actually is present. Alanine represents a conserved substitution for valine. In this way, we were able to identify the contact sites for the mAbs that we studied.

### Confirmation of Antibody Specificity Using Synthetic Linear Peptides

To verify that the antibodies bind to the predicted linear sequences, we tested the immunoreactivity of the antibodies to the synthetic peptides. A checkerboard approach was used in which the peptides were tested for immunoreactivity with the mAbs. Because peptides do not attach readily by passive adsorption to plastic microtiter trays, we attached the peptides to microscope slides using the PI-activated glass surface coating that we developed.<sup>10,11</sup> **Figure 1** illustrates the immunoreactivity of various antibodies with their cognate peptides containing the antibody binding site (epitope). The slides were developed by standard immunohistochemical analysis, in which each peptide was spotted onto glass microscope slides. Thus, the spots shown in Figure 1 represent a positive (immunoreactive) result in this dot-blot format. Figure 1

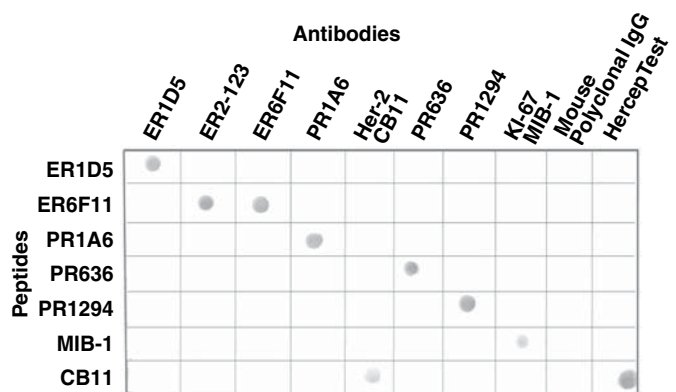


**Figure 1** Immunoreactivity of various monoclonal antibodies with peptides derived from the native sequences of estrogen receptor (ER), progesterone receptor (PR), Ki-67 (MIB-1), and HER-2 (HercepTest). The peptide sequence was chosen to contain the deduced antibody contact site (see Table 1). The figure represents a composite image from 3 glass slides because the spots and hydrophobic borders would not fit on 1 slide. The grid lines were superimposed after scanning the image to enhance clarity.

reveals that each peptide is recognized in an antigen-specific manner. The sequencing data revealed that the ER6F11 and ER2-123 mAbs bind to exactly the same epitope. As expected, both mAbs were immunoreactive with the same linear peptide. Murine polyclonal immunoglobulins did not bind to any of the peptides. The DakoCytomation HercepTest polyclonal anti-HER-2 antibody bound to the same peptide as the CB11 mAb. Both were raised to a carboxy-terminal peptide.

To further confirm the specificity of the linear peptides for the respective antibodies, we performed inhibition experiments. If the peptide inhibits antibody binding to native antigen, then we can infer that the peptide binds to the antibody at or near the antigen binding site. As a source of native ER, PR, or HER-2, we tested binding by immunohistochemical analysis with breast carcinoma samples known to express the protein in question. For these experiments **Image 1**, **Figure 2**, and **Figure 3**, the antibody was first allowed to react with varying doses of peptide. After a 45-minute preincubation of the antibody with the peptide, the antibody-peptide cocktail was used as a primary antibody for immunohistochemical staining. Image 1 and Figure 2 demonstrate representative staining data from experiments with the peptide to which the PR636 binds. Figure 2 depicts the quantitative assessment of staining. At a 16- $\mu\text{mol/L}$  concentration of linear peptide corresponding to the site at which the PR636 antibody binds, there was nearly complete inhibition of staining. Inhibition tapered off with decreasing peptide concentrations. Photomicrographs corresponding to the 16- $\mu\text{mol/L}$  concentration and irrelevant peptide groups are shown in Image 1.

A summary of the inhibition experiments for all of the ER, PR, and HER-2 antibodies and corresponding peptides is shown in Figure 3. In each, there is essentially complete inhibition of immunohistochemical staining at the 1.6- $\mu\text{mol/L}$  concentration for cognate peptide. A 10-fold higher concentration of an irrelevant peptide resulted in no detectable staining inhibition.

### Amino Acid Analysis of Antibody Contact Sites

An analysis of the amino acids comprising the antibody contact sites from these 9 mAbs revealed that certain amino acids tend to be disproportionately represented. **Figure 4A** shows the number of times each amino acid is represented in the antibody contact sites associated with all 9 antibodies. Each contact site typically comprises 5 to 7 amino acids and is based on the data obtained from the epitope mapping effort, such as shown in Figure 1. From all 9 antibody contact sites, there were 53 amino acids. We found that prolines (P) were the most preponderant. Approximately 13% (7/53) of the amino acids were prolines. Leucine (L), glutamine (Q), and tyrosine (Y) also were heavily represented. **Figure 4B** distills the data from Figure 4A to show only amino acids known to react chemically in one of the formaldehyde reactions.

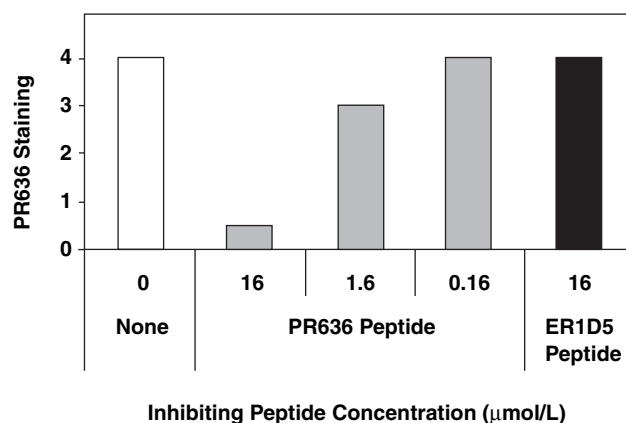


**Figure 2** The progesterone (PR)636 peptide inhibits the immunoreaction of the PR636 antibody with native antigen (in tissue sections) in a dose-dependent fashion. As a specificity check, an irrelevant peptide specific for the estrogen receptor (ER)1D5 antibody, at the highest concentration (16  $\mu\text{mol/L}$ ), yields no inhibition (far right). Immunohistochemical tissue staining is quantified on a 5-point scale (see the “Materials and Methods” section). The images from the PR636 16-  $\mu\text{mol/L}$  inhibiting peptide group and the irrelevant ER1D5 16- $\mu\text{mol/L}$  inhibiting peptide are shown in Image 1.

Tyrosine, glutamine, and serine were the most prevalent. Although lysines have long been suspected as important residues in formalin reactions, only 2 of the 9 antibody contact sites had lysines.

We also analyzed the antibody contact sites for their degree of hydrophilicity and hydrophobicity, using Kyte-Doolittle charts.<sup>21</sup> Hydrophilic stretches of the proteins are

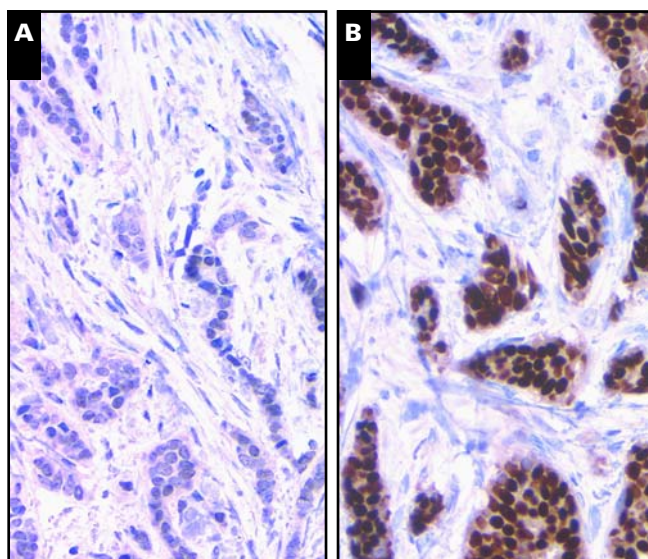


**Image 1** Images from immunohistochemical staining of a breast carcinoma, stained with the PR636 monoclonal antibody, inhibited either with the PR636 (**A**, 16  $\mu\text{mol/L}$ ) or ER1D5 (**B**, 16  $\mu\text{mol/L}$ ) peptide (**A** and **B**,  $\times 400$ ).

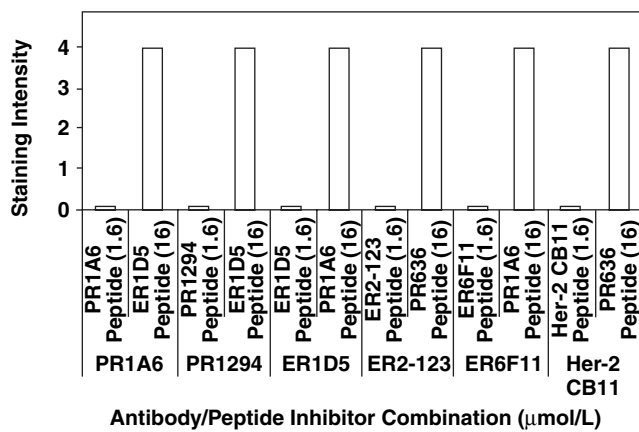


**Figure 3** Each peptide is able to inhibit its cognate antibody from binding to native antigen in tissue sections. For each, an irrelevant peptide at an even higher concentration demonstrates no inhibition. Immunohistochemical tissue staining intensity is quantified on a 5-point scale (see the “Materials and Methods” section). ER, estrogen receptor; PR, progesterone receptor.

represented below the y-axis, whereas hydrophobic regions are above. The more hydrophilic or hydrophobic a region, the greater the value below or above the line. **Figure 5** shows the analysis for each of the peptides, corresponding to a 17 to 25 amino acid stretch derived from the native protein. Within each peptide, the 5 to 7 amino acid contact site, where the antibody actually binds, is bolded in black. Figure 5 shows that the antibody contact sites generally are mildly to moderately hydrophilic. For the sake of clarity, we segregated the ER and

PR peptides into different graphs. Figure 5C depicts the hydrophobic characteristics of peptides from 3 proteins, HER-2, Ki-67, and p53. We also illustrate the Kyte-Doolittle hydrophobic charts for the entire ER and PR proteins for comparison.

## Discussion

To date, no one has undertaken a systematic analysis of the precise epitopes for antibodies that are effective in conjunction with formalin fixation and antigen retrieval. Because these antibodies were selected over the years as useful in conjunction with antigen retrieval, we wondered whether their epitopes might share common characteristics associated with restoration of immunoreactivity after antigen retrieval. By using phage display of peptide combinatorial libraries, we identified the specific antibody contact sites (epitopes) for 8 clinically useful mAbs to ER, PR, HER-2, p53, and Ki-67. The epitope of clone MIB-1 already had been described using deletional mapping techniques.<sup>8</sup> We confirmed the epitope mapping data by synthesizing the linear peptides containing the antibody contact sites and measuring immunoreactivity in a dot-blot type of assay (Figure 1). We further confirmed the specificity of the peptides by testing against other antibodies (Figure 1) and inhibition experiments (Image 1 and Figure 2). These data confirm that the antibody binds to contiguous amino acids and that the binding of each mAb is specific to the predicted sequence.

It is well known that formaldehyde is capable of chemically reacting with numerous amino acids, causing them to cross-link and thereby denature the protein. Thus, there are several possible chemical mechanisms to explain the fact that

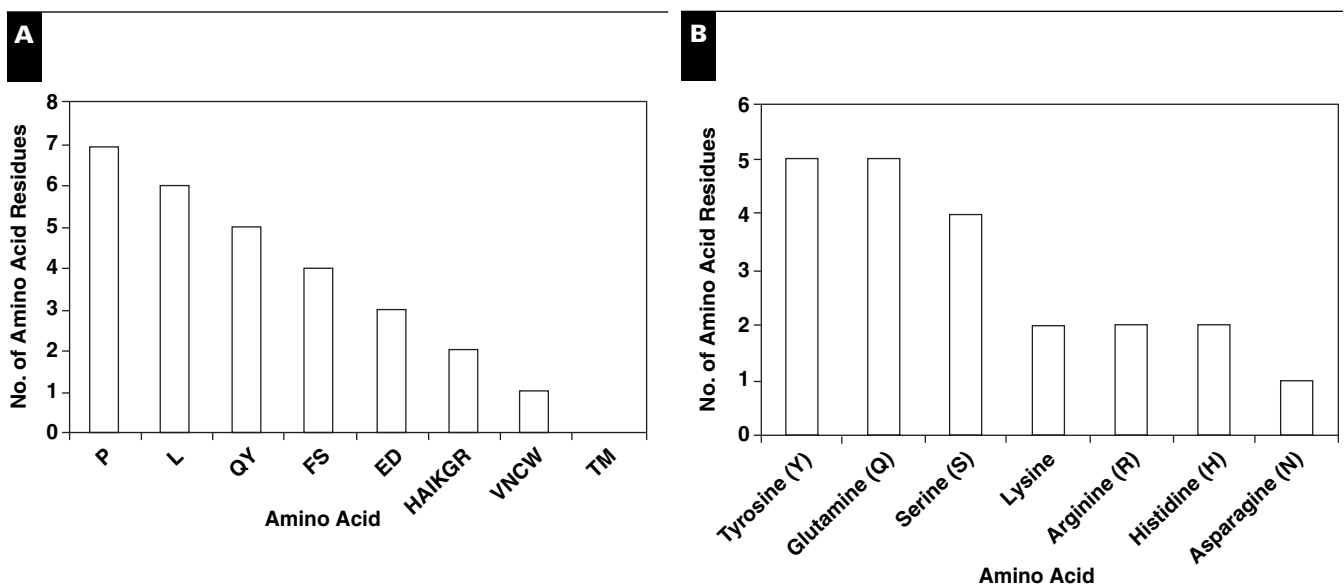


**Figure 4** Amino acid composition of antibody epitopes. **A**, The amino acids comprising the 9 antibody epitopes are listed using the single letter amino acid code. **B**, From the complete list of all amino acids, the formalin-reactive amino acids are listed.

many antibodies are not effective for immunohistochemical staining of formalin-fixed tissue samples.<sup>4</sup> In contrast, it is somewhat surprising and counterintuitive that boiling tissue sections reverses the effect. Since the first description of antigen retrieval,<sup>22</sup> it has remained a mystery why it is possible to recover immunoreactivity after boiling, a procedure that itself is also denaturing. If 3-dimensional (conformational) epitopes are not obliterated by formalin fixation, one would surely expect them to be denatured after boiling.

The data presented in this report shed light on this puzzle. Antibody epitopes can be classified as linear or conformational. The linear epitopes are described as a stretch of 5 to 7 amino acids that are contiguous in sequence. Although other amino acids may lend additional strength to the noncovalent interactions associated with antibody binding, these 5 to 7 amino acids are responsible for the majority of the binding energy. Conformational (or “discontinuous”) epitopes, on the other hand, are formed as a result of distantly and spatially distributed amino acids brought together by conformational folding or bonding. Most protein epitopes are conformational.<sup>23,24</sup> Therefore, it was surprising that each of the mAbs we studied recognized linear protein epitopes. Based on our study of 9 mAbs, the data strongly suggest that to be capable of binding

tissue antigens after formalin fixation and antigen retrieval, a mAb must recognize amino acids that are contiguous in the native protein.

These findings suggest a model to explain why antigen retrieval recovers immunoreactivity. It is known that formalin fixation induces protein cross-links.<sup>5,6</sup> Based on these data, it seems likely that formalin fixation cross-links the protein epitope to another nearby protein. The adjacent protein, now cross-linked to the epitope, blocks antibody binding. On boiling the tissue section, the formaldehyde-induced cross-linking reaction is reversed, reexposing the linear peptide epitope. A high temperature (ie, boiling water) is required for chemical reversal owing to the high activation energy of the reverse reaction. Although heat is also denaturing, 100°C is not capable of breaking peptide bonds and altering the primary amino acid sequence. Only higher order conformational aspects, such as secondary structure, will likely be denatured by heat. Thus, antigen retrieval needs only to regenerate and expose the primary amino acid sequence to restore immunoreactivity. This minimal requirement for antibody recognition is in accord with the denaturing conditions generally associated with antigen retrieval because the primary sequence will be the most resistant to denaturation.

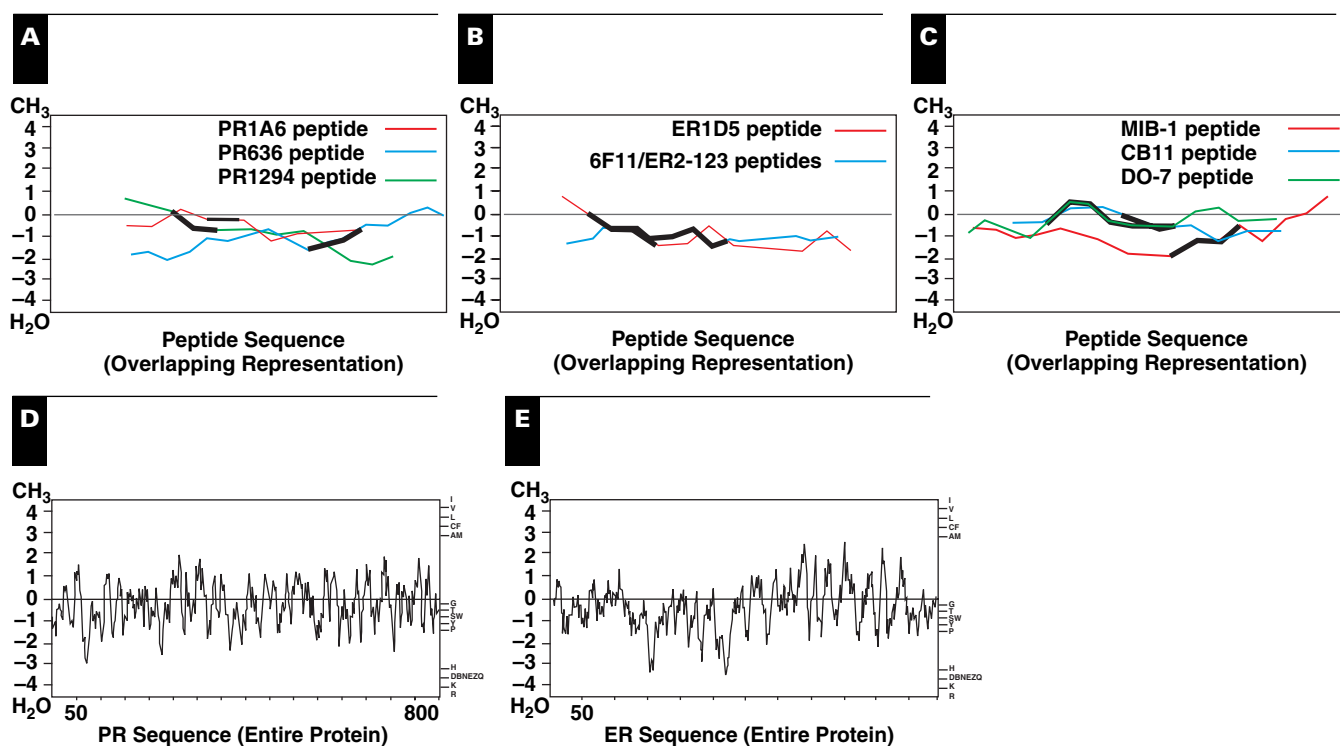


**Figure 5** Epitopes are mildly to moderately hydrophilic. **A, B, and C**, Each of the 9 antibody epitopes and surrounding amino acids were mapped in a Kyte-Doolittle plot. The antibody contact residues are bolded in black. Hydrophilic protein stretches are below the line, whereas hydrophobic regions are above the line. For comparison, the Kyte-Doolittle plot for the entire progesterone (**D**) and estrogen (**E**) receptor proteins are shown as well. The analysis window was set at 9 for each of these graphs. ER, estrogen receptor; PR, progesterone receptor.

Besides explaining the mechanism, the model also makes a prediction. If adjacent proteins block binding of mAbs to their respective epitopes, then the model predicts that mAbs will be blocked from binding to peptides *in vitro* by formaldehyde-induced cross-linking peptides to an irrelevant protein. A large, irrelevant protein that is cross-linked to a peptide will be expected to block antibody from binding to its peptide. Moreover, the model predicts that boiling will restore immunoreactivity, which causes dissociation of the irrelevant protein from the peptide. We recently tested this prediction and obtained the predicted results.<sup>25</sup>

The amino acid analysis (Figure 4) revealed that proline was the most predominant amino acid in the 9 antibody contact sites. Prolines disrupt secondary structure, forcing a linear stretch of contiguous amino acids. The fact that prolines constitute 7 of 53 amino acids (from the 9 antibody contact sites) is, therefore, in accord with the concept that antibodies bind to linear rather than conformational sequences. It is not clear what significance can be ascribed to the prevalence of leucine residues (6 of 53). The next most prevalent amino acids, glutamine (Q) and tyrosine (Y), are both capable of chemically reacting with certain other amino acids in the presence of formaldehyde. Tyrosine, for example, reacts with arginine in the presence of formaldehyde. In fact, the relatively high percentage of tyrosines is in agreement with previously published findings implicating the Mannich reaction in formalin fixation and antigen retrieval.

Of the 9 antibodies, 2 (CB11 and 1A6) were generated to synthetic peptides (product insert for CB11<sup>26</sup> and Bevitt et al<sup>27</sup> for 1A6). Therefore, it is not surprising that those 2 antibodies bind to linear epitopes. Nevertheless, the fact that these 2 antibodies bind to their targets after formalin fixation and antigen retrieval supports the hypothesis that linear epitopes are optimal for this clinical context. The other mAbs were generated by immunization with recombinant ER for clones 1D5<sup>28</sup> and 6F11<sup>29</sup>; recombinant complementary DNA fragment (MIB-1)<sup>30</sup>; recombinant p53 (DO-7),<sup>31</sup> and recombinant PR for clones PR1294 and PR636.<sup>32</sup>

Our data suggest a few guidelines in preparing mAbs to new, clinically useful protein targets so that the antibodies will bind after formalin fixation. Clearly, our data support the view that antibodies to formalin-fixed, paraffin-embedded tissue samples must recognize linear determinants. Therefore, synthetic peptides (conjugated to carrier proteins) may be suitable. This approach has been tested by Bevitt et al,<sup>27</sup> who found that immunization with linear peptides (conjugated to carrier proteins) sometimes, but not always, yielded antibodies suitable for formalin-fixed, paraffin-embedded tissue samples. Other times, the antibodies bind only to their protein targets in frozen sections, without formalin fixation. Therefore, antibody binding to a linear epitope may be required but not sufficient for binding after formalin fixation and antigen retrieval.

Our data additionally suggest that the antibody binding site should be mildly to moderately hydrophilic. This is not surprising because hydrophilic protein stretches tend to be exposed on the surface of a protein, making them accessible to antibody, rather than buried within a protein fold.

Last, our data suggest that peptides containing tyrosine, proline, leucine, and/or glutamine are more likely to be effective epitopes for mAbs that will bind to tissue sections after formalin fixation and antigen retrieval. The presence of proline increases the likelihood that the protein stretch is linear in the native configuration. Tyrosines and glutamines are likely to create reversible cross-links with other adjacent proteins on formalin fixation.

These data seem to be the first describing a common characteristic of antibodies that bind to their protein antigens in tissue after formalin fixation. Namely, mAbs chosen for their immunoreactivity with formalin-fixed tissue samples all recognize epitopes composed of contiguous amino acids. These findings are at odds with what would be predicted by random chance because most protein epitopes on large globular or transmembrane proteins such as these are conformational rather than linear.<sup>23,24</sup> By contrast, epitopes of mAbs that do not bind to their antigens after formalin fixation and antigen retrieval would be expected to comprise both linear and conformational determinants. In other words, it is most likely that the optimal epitopes for antibody binding after formalin fixation and antigen retrieval probably include some, but not all, linear (contiguous) protein epitopes. These data shed light on and provide an explanation for recovery of immunoreactivity after antigen retrieval. By recognizing proteins at their most basic structural level, mAbs are immunoreactive despite the fact that their tissue antigens are treated with denaturing agents such as formalin or boiled as part of the antigen retrieval process. These findings suggest that formalin fixation creates protein cross-links, leading to steric interference with antibody binding to the linear epitope. Antigen retrieval likely breaks those formalin-induced cross-links and eliminates the steric interference, thereby exposing otherwise obscured linear protein sequences for antibody binding.

---

*From <sup>1</sup>Medical Discovery Partners and the <sup>2</sup>Department of Pathology and Laboratory Medicine, Boston University School of Medicine, Boston, MA.*

*Address reprint requests to Dr Sompuram: Medical Discovery Partners LLC, 715 Albany St, L803, Boston, MA 02118.*

*Supported by grants R44CA094557 and R44CA106847 from the National Institutes of Health, Bethesda, MD (Dr Bogen).*

## References

1. Suurmeijer A, Boon M. Notes on the application of microwaves for antigen retrieval in paraffin and plastic tissue sections. *Eur J Morphol.* 1993;31:144-150.

2. Morgan J, Navabi H, Schmid K, et al. Possible role of tissue-bound calcium ions in citrate-mediated high-temperature antigen retrieval. *J Pathol.* 1994;174:301-307.
3. Gown A, Wever ND, Battifora H. Microwave-based antigenic unmasking: a revolutionary new technique for routine immunohistochemistry. *Appl Immunohistochem.* 1993;1:256-266.
4. Shi S-R, Gu J, Turens J, et al, eds. *Development of the Antigen Retrieval Technique: Philosophical and Theoretical Bases.* Natick, MA: Eaton Publishing; 2000.
5. Rait V, O'Leary T, Mason J. Modeling formalin fixation and antigen retrieval with bovine pancreatic ribonuclease A, I: structural and functional alterations. *Lab Invest.* 2004;84:292-299.
6. Rait V, Xu L, O'Leary T, et al. Modeling formalin fixation and antigen retrieval with bovine pancreatic RNase A, II: interrelationship of cross-linking, immunoreactivity, and heat treatment. *Lab Invest.* 2004;84:300-306.
7. Sompuram S, Vani K, Messana E, et al. A molecular mechanism of formalin fixation and antigen retrieval. *Am J Clin Pathol.* 2004;121:190-199.
8. Kubbutat M, Key G, Schluter C, et al. Epitope analysis of antibodies recognizing the cell proliferation associated nuclear antigen previously defined by the antibody Ki-67 (Ki-67 protein). *J Clin Pathol.* 1993;47:524-528.
9. Schmidt M, Broll R, Bruch H-P, et al. The proliferation marker Ki-67 becomes masked to MIB-1 staining after expression of its tandem repeats. *Histochem Cell Biol.* 2002;118:415-422.
10. Sompuram S, Kodala V, Zhang K, et al. A novel quality control slide for quantitative immunohistochemistry testing. *J Histochem Cytochem.* 2002;50:1425-1434.
11. Sompuram S, Vani K, Wei L, et al. A water-stable isocyanate-coated glass array substrate. *Anal Biochem.* 2004;326:55-68.
12. McLafferty M, Kent R, Ladner R, et al. M13 bacteriophage displaying disulfide-constrained micropolymers. *Gene.* 1993;128:29-36.
13. Virnekas B, Pluckthun A, Schneider K, et al. Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. *Nucleic Acid Res.* 1994;22:5600-5607.
14. Parmley S, Smith G. Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene.* 1988;73:305-318.
15. Smith G, Petrenko V. Phage display. *Chem Rev.* 1997;97:391-410.
16. Sparks A, Adey N, Cwirla S, et al. Screening phage-displayed random peptide libraries. In: Kay B, Winter J, McCafferty J, eds. *Phage Display of Peptides and Proteins.* New York, NY: Academic Press; 1996:227-253.
17. Devlin J, Panganiban L, Devlin P. Random peptide libraries: a source of specific protein binding molecules. *Science.* 1990;249:404-406.
18. Luzzago A, Felici F, Tramontano A, et al. Mimicking of discontinuous epitopes by phage displayed peptides, I: epitope mapping of human H ferritin using a phage library of constrained peptides. *Gene.* 1993;128:51-57.
19. Ewing B, Green P. Base-calling of automated sequencer traces using Phred, II: error probabilities. *Genome Res.* 1998;8:186-194.
20. Richterich P. Estimation of errors in "raw" DNA sequences: a validation study. *Genome Res.* 1998;8:251-259.
21. Kyte J, Doolittle R. A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* 1982;157:105-132.
22. Shi S, Key M, Kalra K. Antigen retrieval in formalin-fixed, paraffin-embedded tissues: an enhancement method for immunohistochemical staining based on microwave oven heating of tissue sections. *J Histochem Cytochem.* 1991;39:741-748.
23. Kuby J. *Immunology.* 2nd ed. New York, NY: WH Freeman; 1994.
24. Huebner J. Antibody-antigen interactions and measurements of immunologic reactions. In: Pier GB, Lyczak JB, Wetzler LM, eds. *Immunology, Infection, and Immunity.* Herndon, VA: ASM Press; 2004:chap 9.
25. Sompuram SR, Vani K, Bogen SA. A molecular model of antigen retrieval using a peptide array. *Am J Clin Pathol.* 2006;125:91-98.
26. CB11 [package insert]. Norwell, MA: Vision Biosystems; 2002.
27. Bevitt D, Milton I, Piggot N, et al. New monoclonal antibodies to oestrogen and progesterone receptors effective for paraffin section immunohistochemistry. *J Pathol.* 1997;183:228-232.
28. Al Saati T, Clamens S, Cohen-Knafo E, et al. Production of monoclonal antibodies to human estrogen receptor protein (ER) using recombinant ER (RER). *Int J Cancer.* 1993;55:651.
29. 6F11 [package insert]. Norwell, MA: Vision Biosystems; 2002.
30. Key G, Becker M, Baron B, et al. New Ki-67-equivalent murine monoclonal antibodies (MIB 1-3) generated against bacterially expressed parts of the Ki-67 cDNA containing three 62 base pair repetitive elements encoding for the Ki-67 epitope. *Lab Invest.* 1993;68:629-636.
31. Vojtesek B, Bartek J, Midgley C, et al. An immunochemical analysis of the human nuclear phosphoprotein p53. *J Immunol Methods.* 1992;151:237-244.
32. Press M, Spaulding B, Groshen S, et al. Comparison of different antibodies for detection of progesterone receptor in breast cancer. *Steroids.* 2002;67:799-813.