Identification and validation of cell surface antigens for antibody targeting in oncology

Paul Carter, Leia Smith and Maureen Ryan

Department of Antibody Technologies, Seattle Genetics, Inc., 21823 30th Drive SE, Bothell, Washington 98021, USA

(Requests for offprints should be addressed to P Carter; Email: pcarter@seagen.com)

Abstract

Recent clinical successes with antibodies have reinvigorated interest in the identification and validation of new antigens for antibody therapy, including cell surface proteins for targeting in oncology, the focus of this review. Target identification commonly involves the search for differences between tumor and non-tumor cell lines and/or tissue at the DNA, mRNA, protein or antibody reactivity levels. The next stage, target validation, utilizes antibodies to profile the expression of antigen in normal and tumor tissue and to verify that the antigen is selectively expressed on the surface of tumor cells. Supportive evidence for protein expression is often sought by mRNA profiling and, sometimes, analysis for genomic defects. Unfortunately, concordance between mRNA and protein levels has been found in only about 20% of cases and therefore must be evaluated for individual targets of interest. Antigens judged suitable for antibody targeting are then advanced to the next stage, namely, in vitro and then in vivo screening of antibodies for anti-tumor activities. Subsequent optimization of an antibody clinical lead for therapy is a desirable, if not obligatory, step to developing an antibody as an anti-cancer therapeutic. No single approach or even combination of methods has emerged as the preferred way to identify surface antigens suitable for targeting in oncology. Major options at each step in the process are reviewed here, including their strengths and limitations.

Endocrine-Related Cancer (2004) 11 659–687

Introduction

Antibodies are the most rapidly expanding class of human therapeutics, including their use in cancer therapy (Carter 2001, Glennie & van de Winkel 2003). Indeed, eight antibodies are now approved for oncological indications in the USA and at least twenty-four more are being pursued actively in phase I/II clinical trials or beyond (Table 1). These clinical successes have reinvigorated interest in the identification of new targets for antibody therapy.

Here we review major alternative strategies for the identification of tumor-associated cell surface antigens and the validation of their suitability for antibody targeting in oncology, including antibody generation and screening for in vitro and in vivo anti-tumor activities (Fig. 1). Several important classes of antigens potentially amenable to antibody targeting in oncology are not covered here for space considerations: soluble factors, extracellular matrix proteins, and surface proteins on tumor neovasculature.

The first of four major steps is the acquisition of biological material for target identification such as tumor and closely related non-tumor cells. Tumor biopsies and resections provide an alternative or additional material source and ideally include a normal tissue counterpart for comparison. The second step, target identification, attempts to identify differences between the tumor and non-tumor samples at one of four levels: DNA, mRNA, protein or antibody reactivity. The third major activity, target validation, requires the generation of antibodies as reagents to profile the expression of the antigen in normal and tumor tissue. Corroborating evidence is often sought by mRNA profiling and, where appropriate, analysis for genomic defects. The use of mRNA profiling data necessitates evaluation of the concordance between mRNA and protein levels for the target of interest, as such concordance is not common: ~20% of genes and
<table>
<thead>
<tr>
<th>Antigen target (common alternatives)</th>
<th>Antigen class</th>
<th>Antibody names: trade (generic)</th>
<th>Antibody type (conjugate)</th>
<th>Tumor targets</th>
<th>Status</th>
<th>Corporate sponsors</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFP</td>
<td>Oncofetal glycoprotein</td>
<td>hAFP-Y-90</td>
<td>hz IgG&lt;sub&gt;(90Y)&lt;/sub&gt;</td>
<td>Liver, germ cell</td>
<td>Phase I/II</td>
<td>Immunomedics</td>
</tr>
<tr>
<td>α&lt;sub&gt;5&lt;/sub&gt;β&lt;sub&gt;3&lt;/sub&gt; (vitronectin receptor)</td>
<td>Integrin</td>
<td>Vitek</td>
<td>hz IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Melanoma</td>
<td>Phase II</td>
<td>Medimmune</td>
</tr>
<tr>
<td>CA125 (MUC16)</td>
<td>Mucin</td>
<td>OvaRex</td>
<td>mu IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Ovarian</td>
<td>Phase IIb</td>
<td>AltaRex</td>
</tr>
<tr>
<td>CD4</td>
<td>Coreceptor for MHC class II</td>
<td>HuMax-CD4</td>
<td>hu IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>CTCL</td>
<td>Phase II</td>
<td>Genmab</td>
</tr>
<tr>
<td>CD20</td>
<td>Ca&lt;sup&gt;2+&lt;/sup&gt; channel?</td>
<td>Rituxan (rituximab)</td>
<td>ch IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>NHL</td>
<td>FDA approval 2002</td>
<td>Biogen Idec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Zevalin&lt;sub&gt;(90Y)&lt;/sub&gt; ibritumomab tuxetan</td>
<td>mu IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>NHL</td>
<td>FDA approval 2002</td>
<td>Biogen Idec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bexxar&lt;sub&gt;(131I)&lt;/sub&gt; tositumomab</td>
<td>mu IgG&lt;sub&gt;2a&lt;/sub&gt;</td>
<td>NHL</td>
<td>FDA approval 2002</td>
<td>Corixa, GlaxoSmithKline</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HuMax-CD20</td>
<td>hu IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>NHL</td>
<td>FDA approval 2002</td>
<td>Genmab</td>
</tr>
<tr>
<td>CD22 (Siglec-2)</td>
<td>Sialoadhesin, sialic acid-binding immunoglobin-like lectin</td>
<td>hCD22 (epratuzumab)</td>
<td>hz IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>NHL</td>
<td>Phase II</td>
<td>Immunomedics</td>
</tr>
<tr>
<td>CD30 (TNFRSF1)</td>
<td>TNF receptor super family</td>
<td>SGN-30</td>
<td>ch IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>HD, ALC</td>
<td>Phase II</td>
<td>Seattle Genetics</td>
</tr>
<tr>
<td>CD33 (Siglec-3)</td>
<td>Sialoadhesin, sialic acid-binding immunoglobin-like lectin</td>
<td>Mylotarg</td>
<td>hz IgG&lt;sub&gt;3&lt;/sub&gt;</td>
<td>AML</td>
<td>FDA approval 2002</td>
<td>Wyeth</td>
</tr>
<tr>
<td>CD52 (CAMPATH-1)</td>
<td>GPI-linked protein of unknown function</td>
<td>Campath-1H (alemtuzumab)</td>
<td>hz IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>B-cell CLL</td>
<td>FDA approval 2002</td>
<td>Millennium, Ilex Pharmaceuticals, Berlex</td>
</tr>
<tr>
<td>CD56 (NCAM)</td>
<td>Cellular adhesion</td>
<td>huN901-DM1</td>
<td>hz IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>SCLC</td>
<td>Phase II</td>
<td>Immunogen, Vemals</td>
</tr>
<tr>
<td>CD66e (CEA)</td>
<td>Oncofetal glycoprotein, cellular adhesion</td>
<td>hCEA-I-131</td>
<td>hz IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Metastatic CRC</td>
<td>Phase II</td>
<td>Immunomedics</td>
</tr>
<tr>
<td>CD80 (B7-1)</td>
<td>Co-regulator of T cell activation with CD86</td>
<td>IDEC-114</td>
<td>pt IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>NHL</td>
<td>Phase II</td>
<td>Biogen Idec</td>
</tr>
<tr>
<td>CD140b (PDGFRβ)</td>
<td>Growth factor receptor</td>
<td>CDP860</td>
<td>hz Fab (PEG)</td>
<td>Cancer</td>
<td>Phase II</td>
<td>Celltech</td>
</tr>
<tr>
<td>CD152 (CTLA4)</td>
<td>Negative regulator of T cell activation</td>
<td>MDX-010</td>
<td>hu IgG&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Metastatic melanoma, prostate, breast</td>
<td>Phase II</td>
<td>Medarex</td>
</tr>
<tr>
<td>Antigen target (common alternatives)</td>
<td>Antigen class</td>
<td>Antibody names: trade (generic)</td>
<td>Antibody type (conjugate)</td>
<td>Tumor targets</td>
<td>Status</td>
<td>Corporate sponsors</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------</td>
<td>---------------------------------</td>
<td>---------------------------</td>
<td>---------------</td>
<td>--------</td>
<td>-------------------</td>
</tr>
<tr>
<td>CD227 (PEM, MUC1, mucin-1)</td>
<td>Mucin</td>
<td>R1549 (pemtumomab)</td>
<td>mu IgG₁ (³²⁵Y)</td>
<td>Ovarian</td>
<td>Phase III</td>
<td>Antisoma</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brevarex</td>
<td>mu IgG</td>
<td>Multiple myeloma</td>
<td>Phase VII</td>
<td>Altarex</td>
</tr>
<tr>
<td>EGFR (HER1, ErbB1)</td>
<td>Growth factor receptor</td>
<td>Eribux (cetuximab)</td>
<td>ch IgG₁</td>
<td>Metastatic CRC</td>
<td>FDA approval 2/04</td>
<td>ImClone Systems, Bristol-Myers Squibb, Merck KGaA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ABX-EGF</td>
<td>hu IgG₂</td>
<td>Colorectal, renal cell prostate</td>
<td>Phase III</td>
<td>Abgenix, Amgen</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HuMax-EGFr</td>
<td>hu IgG₁</td>
<td>Head and neck</td>
<td>Phase II</td>
<td>Genmab</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TheraCIM hR3</td>
<td>hu IgG₁</td>
<td>Head and neck</td>
<td>Phase II</td>
<td>YM Biosciences, Oncoscience AG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IGN-101</td>
<td>mu IgG</td>
<td>NSCLC, breast cancer, metastatic CRC</td>
<td>Phase III (all)</td>
<td>Igeneon</td>
</tr>
<tr>
<td>GD3 ganglioside (anti-idiotypic vaccine)</td>
<td>Glycosphingolipid</td>
<td>BEC2 (mitumomab)</td>
<td>mu IgG₂</td>
<td>SCLC</td>
<td>Phase III</td>
<td>ImClone Systems, Merck KGaA</td>
</tr>
<tr>
<td>HER2 (HER2/neu, ErbB2)</td>
<td>Growth factor co-receptor</td>
<td>Herceptin (trastuzumab)</td>
<td>hz IgG₁</td>
<td>Metastatic breast overexpressing HER2/neu</td>
<td>FDA approval 9/98</td>
<td>Genentech</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Omnitarg (pertuzumab)</td>
<td>hz IgG₁</td>
<td>Prostate, ovarian, NSCLC and breast</td>
<td>All Phase II</td>
<td>Genentech</td>
</tr>
<tr>
<td>PSMA</td>
<td>Glycoprotein with folate hydrolase and NAALADase activities</td>
<td>MLN2704 (bivatuzumab mertansine)</td>
<td>di IgG₁ (maytansine DM1)</td>
<td>Prostate</td>
<td>Phase VII</td>
<td>Millennium</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MLN591RL</td>
<td>di IgG₁ (radionuclide)</td>
<td>Prostate</td>
<td>Phase VII</td>
<td>Millennium</td>
</tr>
<tr>
<td>Sialyl Lewis Y</td>
<td>Carbohydrate</td>
<td>SGN-15</td>
<td>ch IgG₁ (doxorubicin)</td>
<td>NSCLC</td>
<td>Phase II</td>
<td>Seattle Genetics</td>
</tr>
<tr>
<td>VEGF</td>
<td>Angiogenic growth factor</td>
<td>Avastin (bevacizumab)</td>
<td>hz IgG₁</td>
<td>Metastatic CRC</td>
<td>FDA approval 2/04</td>
<td>Genentech</td>
</tr>
</tbody>
</table>

1Antibodies in phase I/II or later oncology trials that are enrolled or completed but still under active clinical development, excluding trials with marketed products. Every effort has been made to obtain reliable data from multiple sources (company and other web sites, (Glennie & van de Winkel 2003)), but accuracy cannot be guaranteed.

ALCL, anaplastic large cell lymphoma; AFP, alphafetoprotein; AML, acute myelogenous leukemia; CEA, carcinoembryonic antigen; ch, chimeric; CLL, chronic lymphocytic leukemia; CRC, colorectal cancer; CTCL, cutaneous T-cell lymphoma; CTLA-4, cytotoxic T lymphocyte-associated protein-4; di, deimmunized; EGFR, epidermal growth factor receptor; EpCam, epithelial cell adhesion molecule; FDA, Federal Drug Administration; HER2, human epidermal growth factor receptor 2; HD, Hodgkin’s disease; hu, human; hz, humanized; MAb, monoclonal antibody; MHC, major histocompatibility complex; mu, murine; NCAM, neural cell adhesion molecule; NHL, Non-Hodgkin’s lymphoma; NSCLC, non-small cell lung cancer; PEG, polyethylene glycol; PEM, polymorphic epithelial mucin; PDGFRβ, plate-derived growth factor receptor beta; PSMA, prostate-specific membrane antigen; pt, primatized; SCLC, small cell lung cancer; TNF, tumor necrosis factor; and VEGF, vascular endothelial growth factor.
proteins in the largest study to date (Chen et al. 2002a). Beyond any bioinformatic prediction, it is imperative to establish experimentally that a protein is localized to the plasma membrane and has extracellular determinants amenable to antibody targeting. The fourth major step is the in vitro and subsequent in vivo screening of panels of antibodies for the desired anti-tumor activities. Intellectual property considerations, critical to the commercial development of antibody drugs, are outside the scope of this article.

What makes a cell surface antigen suitable for antibody targeting in oncology?

Defining desirable antigen characteristics is a critical starting point for the identification of new cell surface targets for antibodies. A favorable antigen expression profile is a multi-faceted aspect of antigen characterization that is anticipated to be a necessary but not sufficient condition for successful tumor targeting with an antibody. An ideal expression profile is abundant and homogeneous antigen expression on the external surface of all tumor cells for multiple tumor types with the majority of patients for each tumor type, and absent from normal tissue. It is not necessary to fully satisfy all of these criteria as evidenced by clinical approval of antibodies targeting HER2 (Herceptin), CD20 (Rituxan, Zevalin and Bexxar), CD33 (Mylotarg), CD52 (Campath) and epidermal growth factor receptor (EGFR) (Erbitux) (Table 1). The extent to which an antigen can deviate from an ideal expression profile and still be amenable to antibody targeting is, as yet, unknown. The complexity of this question encourages the empirical approach of establishing criteria from existing antibody drugs and refining them based upon future experience in preclinical and clinical cancer therapy. Below is a discussion of important aspects of antigen expression on tumors and normal tissue anticipated to be critical to successful antibody targeting of tumors.

A key facet of antigen expression on tumor cells that will potentially impact antibody targeting is any change in expression that occurs during the course of disease. Minimally the target antigen should be associated with advanced cancer to reflect the patient populations most commonly available to experimental drugs. Preferably, the antigen will be present in tumors at all stages of disease.
including metastatic foci, to provide a potentially broader opportunity for treating patients. Most desirable the antigen will be causally involved in disease pathogenesis, as exemplified by HER2 (Di Fiore et al. 1987, Hudziak et al. 1987), to reduce the likelihood that tumors will escape from antibody therapy by antigen down-modulation. Antigen expression restricted to early stage cancer appears impractical for antibody drug development for the following reasons. First, there is a risk that diagnosis of cancer will not occur until after antigen expression has been lost. Secondly, patients with early stage disease will often have established treatment options that preclude the use of an unproven experimental drug.

Homogeneous expression of antigen on tumor cells is desirable for antibody targeting but is not necessarily a critical factor for success. However, heterogeneous expression of antigen on tumor tissue may necessitate endowing the antibody with a means to kill bystander tumor cells, e.g. by conjugation to radionuclides or drug-containing liposomes, or fusion to cytokines or an enzyme for prodrug activation (Carter 2001, Milenic 2002, Presta et al. 2003). Expression of antigen on tumors that will support effective in vivo targeting and anti-tumor activity of an antibody. This lower limit of antigen expression for supporting antibody efficacy is unknown and arguably best defined empirically, given the multitude of parameters related to antibody, antigen and tumor that may potentially impact it. Clinical examples offer at least crude upper estimates for the minimal antigen levels required to support successful antibody targeting. Tumor cells in non-solid tumors such as leukemias and lymphomas are commonly more accessible to antibodies than those in solid tumors and this may potentially impact the minimal antigen level required for supporting successful antibody targeting.

In the case of non-solid tumors, as few as 100,000 copies of antigen per cell can be sufficient for effective anti-tumor activity of a naked antibody as judged by CD20 expression levels in non-Hodgkin’s lymphoma (Glennie & van de Winkel 2003) and the efficacy of Rituxan (Table 1). For the antibody drug conjugate, Mylotarg, anti-tumor activity against the non-solid tumor, acute myelogenous leukemia, has been observed with as few as ~5000-10,000 copies (Tanimoto et al. 1989) of the CD33 antigen per cell (Table 1).

As for successful antibody targeting of solid tumors, Herceptin is an example where patient responsiveness apparently requires very high levels of expression of the cognate antigen, HER2. Indeed, Herceptin used as a single agent in first-line treatment of metastatic breast cancer gave a response rate of 35% and 0% in 111 evaluable patients with immunohistochemistry (IHC) staining of 3+ and 2+ respectively (Vogel et al. 2002). Bench-marking these IHC data with cell lines suggests that 3+ and 2+ staining corresponds to about $2.3 \times 10^6$ and $5 \times 10^5$ copies of HER2 per cell, respectively (Ross et al. 2003). Thus clinical activity of Herceptin correlates broadly, but not invariably, with HER2 expression level.

Mathematical modeling suggests that this may be possible as judged by Omnitarg, a humanized version of MAb 2C4. Omnitarg (Table 1) has in vivo anti-tumor activity against MCF7 xenografts expressing low levels of HER2, where Herceptin has no activity (Agus et al. 2002). Partial responses were observed in 3 of 21 patients treated with Omnitarg in a phase I clinical trial (DeGrendele 2003). Omnitarg is currently in multiple phase II clinical trials (Table 1).

Beyond tumors, expression of antigen on normal tissue is a key factor in assessing the suitability of an antigen for antibody targeting in oncology. Limited or preferably no antigen expression on vital normal tissue is desirable to minimize the risk of antigen-dependent toxicities resulting from antibody therapy. This is a particularly important issue for antibodies that are ‘armed’ with drugs, toxins or radionuclides to increase the anti-tumor potency. Expression on non-vital tissue appears less problematic, provided that the antigen sink is sufficiently small to avoid major perturbation of the pharmacokinetics or pharmacodynamics of a therapeutic antibody. Antibody targeting may also potentially be possible for an antigen whose expression is restricted to a non-vital organ or cell population, even if the antigen is not differentially expressed between normal and tumor tissue. For example, CD20, CD33 and CD52 are markers for B cells, myeloid progenitors plus monocytes, and leucocytes, respectively — cell lineages that can be eliminated on a temporary basis.

Many known tumor antigens are expressed at low levels on some normal tissues and substantially higher levels on some tumor cells (Schlom 1991). Can such differential antigen expression be exploited to enhance the selectivity of antibody targeting to tumors in vivo? Mathematical modeling suggests that this may be possible by judicious tuning of the antigen-binding of an antibody such that it is high enough to enable localization to tumor cells expressing high levels of antigen but too low for significant accrual within normal tissue that expresses low level of antigen (Crombet et al. 2004). Clinical experience with anti-EGFR antibodies (Table 1) hR3 (Crombet et al. 2004), Erbitux (Mendelsohn 2000) and ABX-EGF (Foon
antigen level is very high (anti-tumor activity in one or more tumor types where the antigen is expressed). Each of these three anti-EGFR antibodies has shown anti-tumor activity in one or more tumor types where the antigen level is very high (~10^6 copies per tumor cell). In contrast, Erbitux and ABX-EGF commonly give rise to rashes on skin (~10^4 copies of EGFR per cell) that has not been observed with hR3. Antigen binding affinity is one of many possible explanations for the differences in toxicity profile between hR3 and other anti-EGFR antibodies (Crombet et al. 2004).

Shed antigen can potentially interfere with antibody therapy of tumors by several inter-related mechanisms including acceleration of antibody clearance, impairment of antibody localization to tumor and by untoward effects of any immune complexes formed. Ideally the target antigen is not shed (e.g. CD20, Einfeld 1988) or at least shed at only low (<500 ng/ml) and antibody-titratable levels in the majority of patients (e.g. HER2, Baselga et al. 1996, Cobleigh et al. 1999).

**Source material**

Selection and acquisition of biological source material such as cell lines and/or clinical samples is a critical step prior to target identification. Cell lines are the most commonly used starting material as they are readily available and propagated to provide sufficient material for any common target identification strategy. Moreover, the handling of cell lines can be closely controlled and optimized to preserve sample integrity.

A major caveat of using cell lines for target identification is that they will differ from primary tumor samples in ways and to an extent that is unknown. Indeed, both overlap and distinct differences have been observed between the gene expression profiles of primary tumor samples and tumor cell lines (Zhang et al. 1997). Thus targets will be missed if they are upregulated in primary human tumor tissue but not in tumor cell lines — false negatives. Conversely, targets may be found that are upregulated in tumor cell lines but not in primary tumor tissue — false positives.

Several strategies can be used to reduce the impact of false positive or negative findings from target identification using cell lines. First, the use of multiple cell lines in parallel may highlight findings specific to individual cell lines. Secondly, it may be possible to reduce the differences between cell lines and primary human tissue by *in vivo* passing of cell lines. Alternatively, mRNA expression profiling of cell lines alongside clinical samples may allow identification of a cell line whose mRNA expression profile most closely matches clinical specimens (K Wilson, unpublished observations). Target validation by protein and mRNA expression profiling of patient samples provides a means to eliminate false positive findings derived from the use of tumor cell lines for target discovery.

Clinical samples are the most relevant source of biological material for the identification of new tumor targets. Unfortunately, the use of patient tissue for target identification has been significantly stymied by the time, effort and cost of addressing legal and ethical issues related to the procurement and use of tissues. Ideally, one has sufficient control over tissue acquisition, handling and storage to preserve sample integrity. Minimally, one needs rigorous quality control and stringent sample selection so that poor quality materials are excluded to avoid compromising further analysis.

Another limitation of using human tissue for target identification is that biopsy material may contain a mixture of tumor and non-tumor tissue that confounds detection of molecular differences between tumor and non-tumor. The impact of sample heterogeneity can be minimized by setting stringent criteria for tissue usage, e.g. tumor represents > 70% of the sample (Beer et al. 2002). Alternatively, clinical samples can be enriched for tumor cells by laser capture micro-dissection (LCM) (Emmert-Buck et al. 1996) or by fluorescence-activated cell sorting (FACS) (Herzenberg et al. 2002) (Table 2). Enrichment for tumor cells may present challenges of obtaining human material in sufficient quantity and quality to support the target identification method of interest. Primary human tumor tissue can potentially be propagated in immune-deficient mice or limited culture in soft agar to provide more material, but at the risk of changes in antigen expression impacting subsequent analysis.

Further complexity in comparing tumor and non-tumor tissue arises from ‘field cancerization’, a concept first introduced on observation of histologically abnormal cells surrounding oral squamous cell carcinoma and proposed to explain the development of multiple primary tumors and local cancer recurrences (Slaughter et al. 1953). The presence of field cancerization with genetically altered preneoplastic cells has subsequently been reported in several different epithelial cancer types by independent investigators as reviewed by Braakhuis et al. (2003). A major clinical consequence of cancerization fields is that they may lead to local cancer recurrence if not surgically removed with the primary tumor (Tabor et al. 2004). Unfortunately, cancerization fields are not always evident at the macroscopic level as evidenced by genomic aberrations (loss-of-heterozygosity) in apparently ‘normal’ mucosa adjacent to head and neck squamous cell
<table>
<thead>
<tr>
<th>Technique</th>
<th>Description</th>
<th>Applications/Strengths</th>
<th>Limitations</th>
<th>Early references</th>
<th>Recent advances</th>
<th>Recommended reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample enrichment</strong></td>
<td>Laser capture microdissection (LCM)</td>
<td>Isolation of cells from tissue section on slide using laser excision</td>
<td>Enrichment for homogeneous population of cells for analysis; avoids infiltrating cells</td>
<td>Time-consuming; limited source of starting material, small output may limit analysis</td>
<td>Emmert-Buck et al. (1996)</td>
<td>Laser microdissection and pressure catapulting (LMPC) (Burgemeister et al. 2003)</td>
</tr>
<tr>
<td>DNA-based</td>
<td>Fluorescence in situ hybridization (FISH)</td>
<td>Detection of specific gene in cell nuclei using fluorescence-labeled cDNA as probe</td>
<td>Determination of exact copy number of target gene; objective scoring system; reliable and accurate</td>
<td>More expensive and time-consuming than IHC; requires fluorescent microscope; inability to preserve slides; difficult to identify infiltrating cells</td>
<td>Nederlof et al. (1990)</td>
<td>Multifluor-FISH (Speicher et al. 1996) and spectral karyotyping (SKY) (Liyanage et al. 1996)</td>
</tr>
<tr>
<td>Chromogenic in situ hybridization (CISH)</td>
<td>Detection of gene amplification using chromogenic-based method</td>
<td>Combines advantages of IHC and FISH; compatible with clinical diagnostic laboratories</td>
<td>May have lower sensitivity than FISH; exact enumeration of gene copies not always possible</td>
<td>Tanner et al. (2000), Zhao et al. (2002)</td>
<td>CISH for mRNA expression (Diaz et al. 2004) and GOLDFISH (Tubbs et al. 2002)</td>
<td></td>
</tr>
<tr>
<td>Comparative genomic hybridization (CGH)</td>
<td>FISH-based technique for profiling genomic DNA imbalance</td>
<td>Detects gain or loss of genome in tumor samples without prior knowledge of chromosomal abnormalities</td>
<td>Resolution limited to 10 Mbp</td>
<td>Kallioniemi et al. (1992)</td>
<td>Array CGH improves resolution to 0.5 Mbp or better (Solinas-Toldo et al. 1997, Pinkel et al. 1998)</td>
<td>Baak et al. (2003)</td>
</tr>
<tr>
<td><strong>RNA-based</strong></td>
<td>Microarrays</td>
<td>Gene-specific single-stranded DNA fragments or oligonucleotides immobilized on solid glass support chip and hybridized with 2 different fluorophore-labeled cDNA pools</td>
<td>Simultaneous expression profiling of up to tens of thousands of genes; polymorphism analysis</td>
<td>Difficult comparison of data sets across different platforms; requires prior knowledge of gene sequence and validation of genes identified by independent method; cross-hybridization problematic for genes with low expression or very similar sequence</td>
<td>Schena et al. (1995), Lockhart et al. (1996), Perou et al. (2000)</td>
<td>Exon-junction microarrays (Johnson et al. 2003)</td>
</tr>
<tr>
<td>Technique</td>
<td>Description</td>
<td>Applications/Strengths</td>
<td>Limitations</td>
<td>Early references</td>
<td>Recent advances</td>
<td>Recommended reviews</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------------------------</td>
<td>-------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Representational difference analysis (RDA)(^1)</td>
<td>Subtractive and kinetic enrichment used to purify DNA fragments present in one population and not in another</td>
<td>Sensitive technique with application to very low amounts of starting material; requires no prior knowledge of genome sequence; no special equipment required</td>
<td>Contaminating false positive transcripts that lack true differential expression</td>
<td>Lisitsyn &amp; Wigler (1993), Reiter et al. (1998)</td>
<td>LongSAGE obviates most ambiguity in transcript identification (Saha et al. 2002); MPSS (Brenner et al. 2000)(^3)</td>
<td>Bowler (2004), Ding &amp; Cantor (2004)</td>
</tr>
<tr>
<td>Serial analysis of gene expression (SAGE)(^1,2)</td>
<td>PCR-amplified short DNA tags (10-14 bp) are concatenated and sequenced</td>
<td>Detection and quantification of large number of transcripts; ready comparison between experiments</td>
<td>Limited sensitivity related to number of sequenced tags; gene must have recognition sequence for chosen enzyme; ambiguity in transcript identification</td>
<td>Velculescu et al. (1995), Rangel et al. (2003)</td>
<td></td>
<td>Hermeking (2003), Ding &amp; Cantor (2004)</td>
</tr>
<tr>
<td>Suppression subtractive hybridization (SSH)(^1)</td>
<td>PCR-based subtraction method to simultaneously amplify target and suppress non-target amplification</td>
<td>Enrichment of highly expressed genes; efficiently identifies differentially expressed genes</td>
<td>Few µg of poly (A+) RNA required; pre-amplification may result in loss of sequences</td>
<td>Diatchenko et al. (1996), Bangur et al. (2002)</td>
<td></td>
<td>Byers et al. (2000)</td>
</tr>
<tr>
<td>Cancer profiling array (CPA)(^2)</td>
<td>RNA from different tumor types and matching normal spotted onto membrane</td>
<td>Identifies different tumor types that express specific mRNA</td>
<td>Limited number of cases per tumor type</td>
<td>Zhumabayeva et al. (2001)</td>
<td></td>
<td>Cox (2001)</td>
</tr>
<tr>
<td>Quantitative real time PCR (QPCR)(^2)</td>
<td>Real-time method for detecting and quantifying nucleic acids</td>
<td>Higher accuracy, sensitivity and reproducibility than SAGE, microarrays and RDA</td>
<td>Quantification based on choice of known gene for normalization; cheaper detection alternatives to Taqman-based probes (e.g., SYBR green) lowers specificity and accuracy</td>
<td>Kruse et al. (1997), van der Velden et al. (2003)</td>
<td>Multiplex QPCR (Wittwer et al. 2001)</td>
<td>Ding &amp; Cantor (2004)</td>
</tr>
<tr>
<td>RNA-ISH(^2)</td>
<td>Isotopic labeled DNA or RNA as probe for detection of mRNA sequence in tissue section</td>
<td>Allows identification of specific cells expressing the mRNA</td>
<td>Use of fluorescent or radioactively labeled probes harder to adapt to clinical settings</td>
<td>Mogensen et al. (1991)</td>
<td>Double-color fluorescence ISH (Xi et al. 2003)</td>
<td>Xi et al. (2003)</td>
</tr>
<tr>
<td>Technique</td>
<td>Description</td>
<td>Applications/ Strengths</td>
<td>Limitations</td>
<td>Early references</td>
<td>Recent advances</td>
<td>Recommended reviews</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td></td>
<td>Isotope-coded affinity tag (ICAT)</td>
<td>Proteins are profiled using a stable isotope tagging ($^1$H, $^2$H)</td>
<td>Quantitative, high throughput, many labeling options</td>
<td>Gygi et al. (1999b)</td>
<td>Solid-phase tag transfer method (Zhou et al. 2002b)</td>
<td>Tao &amp; Aebersold (2003)</td>
</tr>
<tr>
<td>Protein-based detection</td>
<td>Immunohistochemistry (IHC)</td>
<td>In situ detection of proteins in tissues using antibody coupled to a chromogenic enzyme</td>
<td>Estimates protein expression; indicates subcellular localization; low cost, easy preservation of slides</td>
<td>Hsu et al. (1981)</td>
<td>Tissue microarrays (see line below)</td>
<td>Shi et al. (1997)</td>
</tr>
</tbody>
</table>
Table 2 continued

<table>
<thead>
<tr>
<th>Technique</th>
<th>Description</th>
<th>Applications/Strengths</th>
<th>Limitations</th>
<th>Early references</th>
<th>Recent advances</th>
<th>Recommended reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serological identification of antigens by expression cloning (SEREX)¹</td>
<td>Antigen identification method that uses serum from patients to screen a cDNA expression library made from tumor tissue</td>
<td>Tumor antigens can be directly identified through sequencing of corresponding cDNAs; identifies antigens that elicit a humoral immune response</td>
<td>Tumor antigens that fail to elicit an immune response will not be detected in the serological screen</td>
<td>Sahin et al. (1995)</td>
<td></td>
<td>Old &amp; Chen (1998)</td>
</tr>
<tr>
<td></td>
<td>Phage display¹²</td>
<td>Large (&gt; 10⁹) phage libraries of recombinant human antibodies for high throughput selection with purified antigens or cells</td>
<td>Most comprehensive method to probe epitope space; reagent human antibodies have potential to accelerate therapeutic antibody development</td>
<td>Griffiths et al. (1994), Vaughan et al. (1996)</td>
<td>Isolation of antibodies that selectively bind and enter prostate cancer cells (Liu et al. 2004)</td>
<td>Hoogenboom (2002)</td>
</tr>
<tr>
<td></td>
<td>Selected lymphocyte antibody method¹²</td>
<td>Selects a single, antibody producing cell with desired activity and captures antibody specificity through PCR-based cloning of immunoglobulin genes</td>
<td>Detects antibody activity from a single cell; does not require fusion or immortalization of antibody producing cell; is amenable to multiple species</td>
<td>Assay for single cell selection must be robust; PCR amplification of rodent and human sequences requires complex primer sets; antibody production requires recombinant expression</td>
<td>Lagerkvist et al. (1995), Babcock et al. (1996)</td>
<td></td>
</tr>
</tbody>
</table>

Applications in target ¹identification and ²validation. ³MPSS, massively parallel signature sequencing, is an alternative transcript sampling technology to SAGE that has much greater throughput and increased capacity to identify rare transcripts.
carcinoma (Tabor et al. 2001). Such field cancerization has the potential to confound comparison of tumor and non-tumor tissue in target identification and validation. Therefore, acquisition of distant but relevant normal tissue from the same individual is recommended when comparing tumor tissues with normal counterparts.

Identification of tumor-associated cell surface targets

Target identification and validation utilize a complex web of different methods to search for differences between tumor and non-tumor cell lines and/or tissue at the DNA, mRNA, protein or antibody reactivity levels (Fig. 1). Major techniques used for the discovery of tumor antigens are defined in Table 2 including their key applications, strengths, limitations, recent advances, and representative references.

Sequencing of the human genome (Lander et al. 2001, Venter et al. 2001) has greatly facilitated target discovery, e.g. by expanding the database of known and predicted proteins and providing critical mapping data for the identification of novel cancer genes. Many technologies for target discovery and validation are new and are still evolving rapidly. In contrast, the development of antibody, and other, drugs has commonly taken one or two decades from target identification through regulatory approval. Therefore, it remains to be seen which emergent technologies will have the greatest impact upon drug development.

Genomics

Alterations in the genome associated with cancer include localized changes such as substitutions, deletions and insertions within genes, medium scale aberrations as exemplified by gene amplification and deletions and larger scale defects including chromosomal translocations (Strausberg et al. 2003). In principle, all classes of cancer-associated genomic defects, with the exception of epigenetic events, can be identified by whole genome sequencing of tumor versus matched normal tissue. Currently, whole genome sequencing of tumor versus matched normal tissue is too laborious and expensive to be attractive for target discovery. Fortunately, several alternative techniques are available that allow particular types of genomic defects to be identified (Baak et al. 2003, Strausberg et al. 2003). Somatic mutations and allelic loss in tumor cells can be detected using heteroduplex analysis and loss of heterozygosity respectively. Comparative genomic hybridization (CGH) probes gene copy number and is well suited to the identification of gene amplifications and deletions, whereas spectral karyotyping (SKY) is particularly useful for the study of chromosomal translocations (Table 2).

The identification of genomic defects causally involved in cancer pathogenesis has already led to important new anti-cancer drugs as evidenced by the small molecule, Gleevec (imatinib mesylate), approved for the treatment of chronic myelogenous leukemia (Druker 2002). Gleevec inhibits the enzymatic activity of the hybrid kinase created by chromosomal translocation and fusion of the BCR and ABL genes. Genomic analysis can sometimes provide insights into the molecular basis of responsiveness to drugs as illustrated by Iressa (gefitinib), a small molecule inhibitor of the EGFR tyrosine kinase, approved for the treatment of non-small cell lung cancer (NSCLC). About 10% of NSCLC patients treated with Iressa exhibit a rapid and often dramatic anti-tumor response. DNA sequence analysis has revealed that the majority of Iressa responders have mutations in the ATP-binding pocket of the EGFR tyrosine kinase domain. In vitro, these EGFR mutations lead to enhanced tyrosine kinase activity in response to EGF and increased sensitivity to Iressa (Lynch et al. 2004, Paez et al. 2004).

The discovery of a genomic defect that ultimately led to an antibody therapeutic for cancer is illustrated by the HER2 gene and Herceptin (Table 1). The HER2 gene is amplified in about 25–30% of breast and ovarian cancer patients and is predictive of overall survival and time to relapse (Slamon et al. 1987, 1989). These seminal observations eventually led to the development of the humanized anti-HER2 antibody, Herceptin, now approved for the treatment of metastatic breast cancer (Carter et al. 2000).

Transcriptomics

Global profiling of mRNA levels — transcriptomic analysis — across large numbers of tumor samples has become commonplace using cDNA or oligonucleotide microarrays or serial analysis of gene expression (SAGE) (Table 2) (Ruan et al. 2004). Expression profiling has recapitulated prior histopathological classifications and in some cases has identified new tumor subtypes that correlate with disease progression and/or treatment outcome. Expression profiling has been undertaken for many cancers including acute leukemias (Golub et al. 1999), large B cell lymphoma (Alizadeh et al. 2000) and breast cancer (Perou et al. 2000, Sorlie et al. 2001, van’t Veer et al. 2002). By comparing transcription profiles using three different methodologies (oligonucleotide microarray, SAGE, and principal component analysis (PCA)), gene expression profiling in pancreatic ductal adenocarcinomas identified a novel set of highly expressed genes as well as validated differential expression of previously reported genes (Iacobuzio-Donahue et al. 2003). The inclusion of
PCA is another tool in the study of identified patterns from the long list of differentially expressed genes produced from transcript profiling techniques and data is stratified to highlight promising candidates for translation into clinical targets. Thus, transcriptomic analysis is a robust and broadly applicable tool for the identification of diagnostic and prognostic gene expression signatures for cancer, as well as for potential therapeutic targets.

The use of mRNA profiling to identify proteins amenable to antibody targeting assumes that mRNA levels correlate with protein levels. Unfortunately, limited concordance between mRNA and protein expression has generally (Chen et al. 2002a,b, Smith et al. 2002, Souchelnytskyi 2002), but not always (Kern et al. 2003), been observed. Indeed, in the largest study to date, only 17% of protein spots (n = 165) or 21% of genes (n = 98) had a statistically significant correlation between protein and mRNA expression (Chen et al. 2002a). The potential for discordance between mRNA and protein levels is a major drawback to the use of RNA-based methods for the identification of cell surface targets.

These limitations notwithstanding, mRNA-based methods have identified a few cell surface targets associated with cancer such as prostate stem cell antigen (PSCA) (Reiter et al. 1998), or further implicated known cell-surface proteins with cancer, including E selectin (Bhaskar et al. 2003) and EphB2 (Mao et al. 2004). PSCA was identified by representational difference analysis (RDA, Table 2) as being upregulated in an androgen-independent prostate cancer xenograft (Reiter et al. 1998). PSCA protein expression was subsequently observed in 94% (n = 112) of primary prostate tumors and all (n = 9) bone metastases (Gu et al. 2000). More recently, an anti-PSCA antibody drug (maytansinoid) conjugate was found to have robust anti-tumor activity in a tumor xenograft study (Ross et al. 2002).

**Splice variants**

A promising branch of transcriptomics is the analysis of mRNA splice variants associated with human disease, including cancer. Alternative splicing is a common mechanism that increases the number and variability of transcripts made from a single gene. A survey of > 10 000 human genes across 52 tissues and cell lines revealed that 74% of the multi-exon genes were alternatively spliced (Johnson et al. 2003). Tumor-associated splice variants have been identified for numerous genes including E-cadherin (Becker et al. 1994), EGFRvIII (Sugawa et al. 1990), CD44 (Heider et al. 1993, Wielenga et al. 1993), and HER-4 (Junttila et al. 2000, Gilbertson et al. 2001, Gilmour et al. 2001). Splice variants are detected by many methods including cDNA cloning (Gunthert et al. 1991), RT-PCR (Sturla et al. 2003), expressed sequence tags (Correa et al. 2000) and exon junction microarrays (Johnson et al. 2003).

Tumor-associated splice variants are highly attractive targets in that antibodies can be directed towards a protein domain uniquely expressed on the tumor cell surface but absent from normal tissues, as illustrated by EGFRvIII (Wikstrand et al. 1995) and E-cadherin (Becker et al. 2002). Antibodies to EGFRvIII have proved efficacious in tumor xenograft models (Luwor et al. 2001, Mishima et al. 2001, Johns et al. 2003). The EGFRvIII splice variant occurs with high frequency in gliomas and apparently in a significant subset of other tumor types (Moscatello et al. 1995, Wikstrand et al. 1995). In contrast, the exon 9 deletion variant of E-cadherin appears to be uncommon as judged by its detection in about 8% of diffuse gastric cancers (Becker et al. 1999). Thus, the prevalence of splice variants can vary widely and must be evaluated carefully in validating this class of antigens for antibody targeting.

Tumor-associated splice variants have also been detected for extracellular matrix proteins as exemplified by fibronectin (Zardi et al. 1987, Carnemolla et al. 1989). Such molecules can also provide potential targets for antibodies in oncology. For example, an antibody fragment targeting the ED-B domain of fibronectin, an extracellular matrix protein variant, has been genetically fused to tissue factor (Nilsson et al. 2001), interleukin (IL)-2 (Carnemolla et al. 2002) and IL-12 (Halin et al. 2002), allowing targeted delivery of these factors to tumors with encouraging anti-cancer efficacy.

**Proteomics**

Genomic and transcriptomic analysis are indirect methods for the discovery of cell surface targets that necessitate subsequent validation at the protein level. In contrast, proteomic methods can be tailored to search directly for targetable cell surface features that distinguish cancer cells from normal cells. A further advantage of proteomic over genomic and transcriptomic methods is that they can explore differences in post-translational modification, a potentially valuable source of targets. For example, the mucin product of the MUC1 gene is overexpressed and under-glycosylated in many human tumors (Barratt-Boyes 1996), exposing new epitopes for antibody targeting that are transparent to DNA- and RNA-based methodologies.

Mass spectrometry (MS) methods have emerged as the tools of choice for identifying disease alterations at the protein level (Hanash 2003). Profiling cancer cells and tissue to identify quantitative and qualitative differences in proteins was pioneered using 2-dimensional electrophoresis (2DE) and MS (Celis & Gromov 2003). A strong advantage of a gel-based technique such as 2DE-MS is that post-translation modifications, proteolytic
cleavage and polypeptide variation can be detected within the context of an intact, endogenous protein. 2DE has identified many proteins that are differentially expressed and/or modified in cancers including bladder carcinoma (Celis et al. 1996, 1999, Ostergaard et al. 1997), melanoma (Bernard et al. 2003), lung adenocarcinoma (Oh et al. 2001, Chen et al. 2002), breast carcinoma in situ (Wulfkuhle et al. 2002) and leukemia (Hanash et al. 1986, 1988, Melhem et al. 1997). Early application of 2DE-MS for global protein analysis has highlighted the use of protein profiles as diagnostic and/or prognostic indicators of cancer. However, downstream studies on target molecules identified through 2DE-MS studies have led to a greater understanding of their biological role in cancerous pathways. For example, oncoprotein 18 (Op18) is a phosphorylated protein that was initially identified by 2DE as being over-expressed in leukemias (Hanash et al. 1988). Subsequent studies showed that Op18 regulates cell proliferation and microtubule assembly, and warrants exploration as a potential therapeutic target for leukemia and other cancers (Hanash et al. 2002).

Advances notwithstanding, the utility of 2DE-MS for profiling the cell surface proteome is limited by the relative insolubility of membrane-associated proteins under conditions suitable for 2-DE and the inability to detect proteins that are low in abundance (Flory et al. 2002). Consequently, alternative strategies have been pursued to enrich extracts for plasma membrane targets and improve sensitivity. For example, membrane fractionation in conjunction with one-dimensional (1DE)-MS identified new, as well as known, breast cancer targets (Adam et al. 2003). Fractionation to enrich for membrane-associated proteins is also commonly used in combination with multi-dimensional liquid chromatography (LC) and MS (Han et al. 2001, Wu et al. 2003, Wu & Yates 2003). Multi-dimensional LC-MS provides the high sensitivity and throughput required for clinical sample analysis (Flory et al. 2002, Hanash 2003). Protein quantification and reduction in sample complexity can be achieved with gel-independent methods such as isotope-coded affinity tags (ICAT, Table 2) (Gygi et al. 1999a). ICAT combines affinity tags with either a ‘heavy’ or ‘light’ stable isotope label so that proteins from two different samples (e.g. tumor and non-tumor) can be compared and analyzed by MS in a quantitative fashion. ICAT methodology has been used in combination with multidimensional liquid chromatography to identify and quantify cell surface proteins in the microsomal fraction of differentiated HL-60 cells (Han et al. 2001) and to analyze the secreted and cell surface proteins from a prostate tumor line in response to androgen treatment (Martin et al. 2004).

Proteomics, like gene expression profiling, has defined new subclasses of disease and holds significant promise for the identification of new ‘drugable’ targets. However, given the inherent diversity of protein properties, proteomic methods seem unlikely to become as facile, high throughput or universal as DNA or mRNA profiling. Furthermore, even as the throughput of proteomics increases, improvements will be tempered by bioinformatic limitations and by the labor required to validate individual targets by western blot and IHC. Alternative strategies that implement either protein or antibody chips for global analysis of tumor proteins or patient serum are currently being developed and will likely further increase the impact of proteomics in medicine (Jenkins & Pennington 2001, Mitchell 2002, Mintz et al. 2003).

Antibody technologies

Antibody methods (Table 2) provide the most direct and comprehensive route to the identification of antibody-targetable differences between tumor cells and non-tumor cells. A further advantage is that antibodies generated for target discovery are multi-purpose tools that can be applied to target validation, and potentially clinical lead generation (Fig. 1). Indeed, validation of target expression is integral to the antibody screening and selection process. In contrast, other discovery methods require antibodies for target expression profiling at the validation stage. Antibody-based discovery has the potential to illuminate the full epitope space of the tumor cell surface including protein products of splice variants, post-translational modifications of proteins, and glycolipids that are invisible, or at least only opaque, to other discovery methods.

These major attractions of antibody-based discovery are significantly offset by two substantial and compounding limitations. First, antibody-based discovery has much lower throughput than many genomic, transcriptomic and some proteomic methods. Secondly, much time and effort may be extended to establish that an antibody has a favorable binding profile on tumor and non-tumor cells only to discover that the target lacks novelty from either a scientific or intellectual property perspective. This reflects the fact that one typically screens large numbers of antibodies by expression profiling prior to the identification of the cognate antigen by MS (Nasdala et al. 2002) or, less commonly nowadays, by N-terminal sequencing of the protein or expression cloning of the corresponding gene (Ridgway et al. 1999). This antigen identification bottleneck for antibody-based discovery is overcome by SEREX — serial identification of antigens by expression cloning (Sahin et al. 1995). SEREX obviates the need for target identification at the protein level since it directly identifies tumor antigens by serological screening of
cDNA expression libraries from human tumors (Table 2). The power of the SEREX technology resides in the ability to identify antigens eliciting a humoral immune response. A limitation of SEREX is that it does not directly yield a replenishable antibody supply. Specific methods of antibody generation, together with the strengths and challenges posed by each method, are highlighted in subsequent sections.

The five major technologies used for monoclonal antibody generation are murine hybridomas, human hybridomas from mice transgenic for human immunoglobulin (Ig) genes, human hybridomas derived from patients, antibody phage display libraries, and recombinant antibody production from single, antigen-selected lymphocytes (Lagerkvist et al. 1995, Babcook et al. 1996). Murine hybridoma technology is the most widely used method for MAb generation for research applications. Unfortunately, the clinical development of murine MAbs has commonly been confounded by a strong cross-species immune response. The immunogenicity of murine MAbs has often been reduced to inconsequential levels by chimerization or preferably humanization, as extensively reviewed elsewhere (Carter 2001, Brekke & Sandlie 2003, Glennie & van de Winkel 2003). Murine MAb humanization can be obviated by using human MAb from phage display libraries (Hoogenboom 2002), transgenic mice that express human antibodies (Davis et al. 2004), patients (Vollners & Brändlein 2002), or selected lymphocyte antibody methods (Babcock et al. 1996) (Table 2). Whilst human antibodies have the potential to streamline preclinical development, each method for human MAb generation poses unique challenges for research applications (see below).

**Monoclonal antibodies**

Hybridomas producing MAbs are generated from the stable fusion of immortalized myeloma cells with B cells from immunized mice (Köhler & Milstein 1975). Hybridoma technology emerged rapidly as a powerful method to identify novel cell surface markers (Williams et al. 1977). Immunophenotyping with MAbs raised to cell surface markers defined the differentiation status of hematopoietic cells and became the basis for the ‘clustering of differentiation’ or CD classification (Bernard & Boumsell 1984). Immunophenotyping of hematopoietic cells and hematologic malignancies is commonly performed by FACS analysis, a technique that can detect and quantify antigen expression on the cell surface (Table 2). FACS analysis revealed that some CD antigens commonly expressed on hematopoietic cells are also expressed on hematologic tumors. CD20, for instance, is a B cell lineage marker expressed on >90% B-cell non-Hodgkin’s lymphomas and is the target for Rituxan, Zevalin and Bexxar (Table 1).

Hybridoma development for antigen discovery typically includes a differential screen to identify antigens that are highly expressed on tumor cells and sparingly expressed, ideally absent, on non-tumor cells. Screening for cell surface antigens can be carried out using FACS analysis, antibody-cell capture, or fluorometric microvolume assay technology (FMAT). FMAT devices automate microtiter plate reading and handling, thereby facilitating high throughput screening of cell-binding or killing activities in hybridoma wells (Miraglia et al. 1999).

Dominant epitopes can skew the repertoire of hybridomas and limit the diversity of antigens detected. This problem is potentially exacerbated by the weak immunogenicity of some tumor antigens. The impact of dominant epitopes can be minimized by using a subtractive immunization approach (Hooper et al. 2003, Zijlstra et al. 2003) or by incorporating a highly selective screen.

Limitations notwithstanding, hybridoma technology has historically been the most widely used method for tumor antigen discovery. Indeed, many antigens currently being targeted by antibodies in oncology clinical trials (Table 1) were first identified using MAbs, including sialyl Lewis3 (Hellström et al. 1990) and prostate-specific membrane antigen (PSMA) (Israeli et al. 1993).

**Transgenic mice**

Human MAbs can be generated following immunization of transgenic mice that contain human immunoglobulin sequences (Fishwild et al. 1996, Mendez et al. 1997, Tomizuka et al. 2000) (Table 2). Transgenic mice offer a powerful way to generate human antibodies to validated targets for clinical development that have low risk for immunogenicity in man (Davis et al. 2004), thereby obviating the need for humanization of murine MAbs. This benefit of using transgenic over non-transgenic mice for target discovery should be balanced against several downsides (weaker immune response, more effort to generate large panels of antibodies, and high technology access cost) plus the very small proportion of reagent MAbs generated that might ultimately be deemed worthy of preclinical development.

**Human hybridomas**

Immortalizing B-cells from cancer patients has been undertaken to develop human hybridomas against cancer antigens. Preparing human hybridomas from cancer patients is advantageous because it provides a direct report on the immune response to the tumor without introducing a biased or cross-species immunization strategy. Moreover such MAbs have a low risk for
immunogenicity in man, as they are human, and often have few or no somatic mutations (Brändlein et al. 2003b). An intriguing complication is that most, if not all, tumor reactive antibodies isolated from cancer patients are IgMs (Vollmers & Brändlein 2002), necessitating either isotype switching or molecular reconfiguration to an IgG (Pancook et al. 2001) to facilitate preclinical and clinical development. Several tumor-selective IgMs from cancer patients have been used to identify their cognate antigens including SC-1 (Hensel et al. 1999) and PAM-1 (Brändlein et al. 2003a). SC-1 binds to a tumor-associated glycoform of CD55 (decay-accelerating factor) that is preferentially expressed on diffuse and intestinal gastric cancer cells and induces apoptosis (Hensel et al. 1999). PAM-1, which was isolated from a patient with stomach carcinoma, binds to a variant of cysteine-rich fibroblast growth factor receptor 1 that is expressed on the surface of many epithelial tumors.

Antibody phage libraries
The use of phage display libraries offers potentially the most direct route to tumor antigens amenable to antibody targeting, by allowing direct selection of antibodies that bind to tumor cells and counter selection against non-tumor cell binding (Cai & Garen 1995, Ridgway et al. 1999). Moreover, the use of large (>10^6 member) naïve antibody phage libraries provides a highly diverse repertoire that is anticipated to allow more comprehensive sampling of the epitope space on the surface of tumor cells than previously possible with other technologies (Liu et al. 2004).

Unfortunately, direct cell selection with phage libraries to identify new tumor antigens has been significantly stymied by non-specific binding of phage to cell surfaces and the selection of dominant antibody clones (Ridgway et al. 1999, Hoogenboom 2002). A significant advance in antibody phage library selections on cells is the use of polyvalent phage and the selection for internalizing phage (Poul et al. 2000). Non-specifically bound phage is stripped from the cell surface prior to recovering the specifically internalized phage. Recently this approach has been used to isolate >90 antibodies from a large naïve phage library that bind to prostate cancer cells but sparingly to non-tumor cells (Liu et al. 2004). Such antibodies seem particularly well suited for conjugation to cytotoxic drugs (Payne 2003) or in the construction of immunoliposomes (Nielsen et al. 2002).

Recombinant antibody production from single antigen-selected lymphocytes
Two methods for producing monoclonal antibodies from a single antibody-producing cell have been described (Lagerkvist et al. 1995, Babcock et al. 1996). Single antibody-producing cells are identified and selected using either antigen-coated magnetic beads (Lagerkvist et al. 1995) or micromanipulation of antigen-binding cells detected with the hemolytic plaque assay (Babcock et al. 1996). Once the desirable antibody-producing cell is isolated, PCR-based cloning is used to amplify the V_L and V_H sequences defining the specificity of the antibody. A major advantage of this methodology is that it can identify a rare antibody-forming cell with desirable antigen-binding characteristics. Furthermore, since this methodology retrieves antibody activity through cloning of immunoglobulin sequences, it is amenable to generating antibodies from multiple species, including humans. However, a potential challenge is that PCR-based amplification becomes increasingly complex with rodent and human immunoglobulin sequences. In addition, recombinant antibody production is more time-consuming and onerous than MAb production from hybridoma cell lines.

Target validation
The number of tumor-associated targets that are identified and advanced to the validation stage is highly dependent upon the scale and methodology of the identification process including the selection criteria. Commonly, one might obtain tens to a few hundred targets that satisfy minimal criteria for tumor association. Pragmatically, the stringency of selection criteria can be increased, if need be, to identify a smaller number of lead targets that match available capacity to validate them.

Plasma membrane localization
A critical early validation step is to confirm that a target of interest is localized to the plasma membrane and has extracellular determinants amenable to antibody targeting. Based on the gene sequence of novel targets identified through genomic or transcriptomic techniques, subcellular localization patterns for the putative protein products can be predicted using software such as PSORT II (Horton & Nakai 1997). This program searches for motifs associated with protein sorting and reports on the probability for the protein to be localized to candidate sites including the plasma membrane. Another program, TMHMM 2.0, predicts the number of transmembrane helices in a protein (Sonnhammer et al. 1998), correctly in 97–98% of cases (Krogh et al. 2001). Membrane and secreted proteins can also be identified using Gene Ontology Consortium (GO) classifications associated with RefSeq genes (Egland et al. 2003). Bioinformatic predictions of plasma membrane association are valuable but require experimental confirmation.

www.endocrinology-journals.org
Carter et al.: Target identification and validation of cell surface antigens

FACS (Table 2) using a fluorescently labeled anti-target antibody is the most commonly used method for assessing the presence of an epitope on the exterior surface of cells required for antibody targeting. Genetically fusing the gene encoding the target with an epitope tag obviates the need for an antigen-specific antibody, albeit with a minor risk of altering localization. Following transient mammalian expression of the target fusion construct, FACS is then performed with a fluorescently labeled anti-tag antibody. Alternatively, the target gene can be fused to a fluorescent tag such as green (Egland et al. 2003) or auto- (Adam et al. 2003) fluorescent protein. Following transient expression, the fluorescent product is localized by confocal fluorescence microscopy. A major drawback of fluorescent protein tagging is that it may indicate plasma membrane association but it does not show definitively that the target is accessible to an antibody on the outside of the cell.

Beyond confirming the presence of a target on the surface of cells, FACS is frequently used to screen tumor cell lines for antigen expression prior to in vitro assays. Several assays have been developed to quantitatively measure protein levels of antigens including estrogen and progesterone receptors (Gritzapis et al. 2003), Fc receptors (Christensen & Leslie 1990) and CD antigens (Barten et al. 2003). Quantitative FACS analysis can provide data on antigen density required for validation of antibody anti-tumor activity.

Target expression profiling

After establishing that a target has determinants on the outside of tumor cells, the next major step in developing anti-cancer antibodies is expression profiling of the target antigen in tumor and non-tumor, minimally at the protein level and commonly also at the mRNA level. Expression profiling with human tissue is an obligatory step in target validation that is sometimes superseded by profiling using tumor and non-tumor cell lines. A major caveat of mRNA expression profiling is that a statistically significant correlation between protein and mRNA expression in tumor samples is not common — ~20% of genes and proteins in the largest study to date (Chen et al. 2002a). Thus, for a target of interest, it is critical to evaluate the concordance between mRNA and protein levels. For antibody-based methods of target identification the validation of protein expression is accomplished, in part, during the screening and selection process.

mRNA expression profiling

Many techniques are available for mRNA expression profiling in clinical samples including microarrays, SAGE, quantitative real-time PCR (QPCR), cDNA hybridizations using multiple tissue expression arrays/cancer profiling arrays, and RNA in situ hybridizations using frozen tissue samples (Table 2). In situ hybridization is a low throughput technique suitable only for target validation, whereas most other methods have sufficient throughput to have gained significant usage both as discovery and also validation tools.

SAGE (Velculescu et al. 1995) utilizes 14-mer tags for each cDNA that are concatenated and sequenced to quantify the abundance of transcripts. Lengthening the tag length to 21-mer, so-called LongSAGE, has greatly reduced the ambiguity in identifying transcripts due to short tag size (Saha et al. 2002). SAGE experiments involve sequencing of up to 105 tags and are best suited to facilities with high throughput sequencing capability. The results from SAGE experiments, even from different laboratories, are readily compared by data normalization using the number of specific tags sequenced per million total tags sequenced. This portability of SAGE data has been capitalized upon by the National Cancer Institute funded Cancer Genome Anatomy Project, with the largest public data database of archived SAGE tag counts and an excellent online query tool (Lash et al. 2000, Boon et al. 2002). SAGE tags can efficiently identify novel transcripts or genes in the genome as well as provide data on relative expression of the target gene in normal and tumor cells. In contrast, exon splice variants commonly with identical 3’ ends are transparent to SAGE technology. SAGE expression profiling on normal and tumor tissue may help focus IHC analysis on particular organs and/or tumor types. SAGE has identified novel ovarian cancer-specific transcripts that were subsequently characterized by northern blot analysis and real time RT-PCR (Rangel et al. 2003).

QPCR is a commonly and successfully used method to corroborate expression profiles of individual genes identified by other methods. For example, genes over-expressed in small cell lung carcinoma identified by microarray expression profiling were confirmed by QPCR analysis (Bangur et al. 2002). The signal amplification intrinsic to QPCR provides exquisite sensitivity and a wide dynamic range that minimizes the need for serial dilution of samples. QPCR is a powerful tool for validation of mRNA profiles in that it provides accurate and reproducible quantification of nucleic acid and is readily amenable to automation to increase throughput. The availability of software has made the design of necessary QPCR probes simpler and more robust.

Antigen expression profiling

Evaluation of antigen expression on the surface of tumor and normal cells in human tissue samples is critical to the assessment of antigen suitability for antibody targeting. Moreover, antigen expression profiling may provide a
critical diagnostic step for patient selection for targeted therapy, as is the case for HER2 (Ross et al. 2003b).

Immunohistochemistry analysis is the most common and effective way of expression profiling antigens that is economical with limited clinical samples (Table 2). IHC involves three major and inter-related choices: recombinant protein or synthetic peptides as an immunogen, rabbit polyclonal antibodies (PAbs) or murine MAbs for antibody generation, and fresh frozen or formalin-fixed paraffin-embedded (FFPE) tissue (Goding 1996).

Antibodies, if none are available, are generated using recombinant antigen or more commonly synthetic peptides as immunogens. The synthetic peptides are designed to match portions of the extracellular domains of the target antigen that are most distinct from other known proteins. Antibodies to short peptides are likely to detect formalin-denatured antigen, but may have higher risk for cross-reactivity than antibodies to recombinant antigen, and therefore should be validated for specificity to the antigen by western blot analysis.

Rabbit PAb are more rapidly generated than murine MAbs, making them the more common choice for early target validation experiments. Another advantage of PAb over MAbs is that they recognize multiple antigenic determinants, allowing for a greater chance of detecting epitopes preserved through chemical fixation and paraffin embedding (Hoffstrom & Wayner 1994). However, if a MAb is available that is suitable for profiling of protein by IHC using FFPE tissue this will likely be preferred over a PAb because of higher specificity and affinity. IHC staining can be performed using pooled MAbs when multiple MAbs are available to distinct epitopes of the target antigen (Ziai et al. 1987). Such combinations of MAbs have been shown to increase the sensitivity of IHC without changing the percentage of stained cells.

The main advantages of FFPE tissues over fresh frozen samples for IHC are that they are commonly available in much larger numbers and retain better tissue morphology. A limitation of FFPE tissue for IHC is the potential need to search for an antigen-retrieval step to allow the primary antibody to bind antigen, with no guarantee of success (Shi et al. 2001). In contrast, fresh frozen tissue typically retains high immunoreactivity and is used for IHC when available antibodies recognize native but not formalin-denatured antigen. Analysis of large numbers of human tissues for antigen expression is a critical step for target validation and was a major bottleneck until the advent of tissue microarrays (TMAs, Table 2).

TMAs are a major advance in target validation in that they allow in situ profiling of protein, mRNA or DNA on hundreds of samples in parallel on a single slide. TMAs are constructed by acquiring cylindrical cores from either FFPE or frozen tissue specimens and arraying them at high density into a recipient block (Kallioniemi et al. 2001). Numerous TMA formats are commercially available including normal tissue arrays, cancer survey panels, and tumor specific collections with up to 1000 samples per slide.

‘Large-section’ IHC analysis and single tissue punch TMA results typically agree closely and the level of concordance may be increased by adding a few additional tissue punches to TMA (Camp et al. 2000, Rubin et al. 2002). In most instances, the high degree of standardization achieved using TMA more than compensates for minor limitations associated with the minute tissue sample size (Sauter et al. 2003).

IHC for target validation is conveniently undertaken stepwise using TMA. First, antigen expression is analyzed in a normal tissue panel to identify candidates with sparing or no staining of vital organs, and which lack major antigen sinks in other normal tissue. Secondly, prevalence and level of antigen expression in different tumor types are surveyed rapidly using common cancer panel arrays that include normal tissue counterparts from the same patient. Thirdly, the number of samples for positive tumor types is expanded to power the statistics of staining incidence. Minimally, advanced stage disease samples are analyzed to reflect the potential patient population accessible to an early stage experimental drug. More preferably ‘progression TMAs’ are used to assess changes in antigen expression at different stages of disease. Loss of expression with disease advancement and metastasis occurs for some antigens (al-Tubuly et al. 1996) and bodes poorly for antibody targeting. Additional analysis of normal tissue from different tissues and more individuals is advisable for antigens judged promising after initial TMA steps.

TMA of tumor specimens linked to clinical follow-up data, so called ‘prognosis TMA’, permits correlation of antigen expression and patient outcome. Prognosis TMA has been used to probe the association of HER2 and EGFR overexpression with patient survival in breast cancer (Sauter et al. 2003). Heterogeneity between primary tumors and their metastases can be analyzed by using a ‘metastasis TMA’ as exemplified by a study of HER2 status of primary and metastatic breast cancers (Simon et al. 2001). In situ hybridization of mRNA using TMA is a valuable way to corroborate IHC results.

Key advantages of TMA technology over conventional IHC include much greater throughput, cost efficiency, and the standardization inherent in the simultaneous analysis of up to a thousand samples per slide. Critical to TMA success has been the quantification and automated acquisition of data plus the development of sophisticated software for data analysis and archiving. Public domain TMA software is anticipated to facilitate
the analysis and exchange of large TMA data sets (Berman et al. 2003). Several commercial automated TMA analysis systems use common IHC or fluorescent dyes that provide a greater dynamic range than do peroxidase-based systems (Camp et al. 2002, Haedicke et al. 2003).

Functional target validation

Defining the function of a target antigen is desirable, but is not a prerequisite for successful antibody targeting. Readers are directed to excellent recent reviews describing the plethora of alternative functional validation techniques that include RNA interference (RNAi) (Boese et al. 2003, Lavery & King 2003), antisense (Kurreck 2003, Crooke 2004), ribozymes (Castanotto et al. 2002, Kashani-Sabet 2002), gene knock outs, transgenic animals (Tornell & Snaith 2002), viral transduction, and chemical genomics (Jung et al. 2003, Sehgal 2003). RNA interference (Fig. 1) is one of the most popular of these techniques for studying gene function, and has the potential to streamline some aspects of target validation (Boese et al. 2003, Lavery & King 2003). RNAi is the mechanism of post-transcriptional gene silencing using small RNAs naturally created from double stranded precursors by RNase III-like activity of the enzyme, Dicer (Bernstein et al. 2001, Huttvagner et al. 2001, Knight & Bass 2001). RNAi technology typically lowers the target transcript level substantially without eliminating it completely. An altered phenotype often provides significant clues as to gene function, whereas an unaltered phenotype is hard to interpret because of the possibility of residual mRNA maintaining function.

Therapeutic antibody generation

Validated antigens judged potentially suitable for antibody targeting are then advanced to the next stage, namely, antibody production and in vitro and then in vivo screening for anti-tumor activities (Fig. 1). Unfortunately, there is a dearth of convincingly validated surface antigens that is compounded by the scarcity of so-called ‘naked’ antibodies with robust anti-tumor activity. Fortunately, the potency of anti-cancer antibodies can be improved in numerous ways including arming with potent drugs (Payne 2003) or radionuclides (Goldenberg 2003). These arming technologies have been validated by the regulatory approval of the antibody drugs, Mylotarg, Zevalin, and Bexxar, conjugated with the drug calicheamicin, and the radionuclides $^{90}$Y and $^{131}$I, respectively (Table 1). Antibody enhancement strategies have been reviewed extensively elsewhere (Carter 2001, Milenic 2002, Presta 2002, Glennie & van de Winkel 2003, Ross et al. 2003a) and are covered here only to the extent that they impact the identification of a clinical lead antibody with potent in vivo anti-tumor activity. Subsequent antibody optimization, e.g. affinity maturation, to generate an antibody clinical candidate is beyond the scope of this target-centric review.

In vitro screening of antibodies for anti-tumor activity

Naked antibodies in clinical trials in oncology exhibit several different kinds of anti-tumor activity in vitro and in vivo (Carter 2001, Milenic 2002, Glennie & van de Winkel 2003, Ludwig et al. 2003) including inhibition of growth, invasion, metastasis or angiogenesis, sensitization to chemotherapy or radiotherapy, induction of apoptosis, growth factor antagonism, and secondary immune functions such as antibody-dependent cellular cytotoxicity or complement-dependent cytotoxicity (see Table 3 for representative examples). Whilst the biological complexity of human tumors cannot be recapitulated in vitro, it is both feasible and desirable to screen antibody panels for many different types of anti-tumor activity.

Numerous assays are available to assess in vitro anti-tumor activity including $[^{3}$H$]-$thymidine incorporation to assess new DNA synthesis and metabolic dye staining to evaluate the number of viable cells following a cell proliferation assay. Screening of ‘naked’ antibodies provides a good initial strategy for identifying clinical lead candidates that offers potentially the simplest preclinical and clinical development path. Multiple antigen-positive tumor cell lines should be screened to identify antibodies, such as the MAb 4D5 parent of Herceptin, that have anti-tumor activity on $< 50\%$ of antigen-positive cell lines (Lewis et al. 1993). Tumor cell lines that are antigen negative can also be stably transfected and clones expressing the antigen at different levels can be selected for use in both in vitro and in vivo assays. For example, stable transfectants of CHO, PC3, MCF7,Her2 and HCT116 cell lines expressing PSCA were generated and compared with parental lines in response to an anti-PSCA immunoconjugate for in vitro cytotoxicity assay and antitumor efficacy (Ross et al. 2002). This approach provides a model system to address the antigen density requirement and a negative control cell line with similar genetic background.

Biological knowledge of a target antigen can sometimes usefully guide the design of anti-tumor activity assays as evidenced by antibodies to EGFR. Erbitux, previously known as C225, blocks binding of EGF and transforming growth factor $\alpha$ (TGF-$\alpha$) ligands to EGFR, inhibits receptor tyrosine kinase activation by these ligands, and inhibits the growth of EGFR-positive human tumor xenografts in nude mice (Mendelsohn 2000). Erbitux is approved for the treatment of
EGFR-positive metastatic colorectal cancer (Table 1). Subsequent to many Erbitux studies, a human anti-EGFR MAb, E7.6.3, was reported that blocks binding of both EGF and TGF-α to various EGFR-expressing human tumor cell lines and abolishes ligand-dependent cell proliferation (Yang et al. 1999). MAb E7.6.3, now known as ABX-EGF, is currently in phase III clinical trials for colorectal cancer (Table 1).

For antigens with a paucity of biological understanding, it may still be possible to identify antibodies with anti-tumor activity and develop them into therapeutics, as illustrated by Herceptin. MAb 4D5 was chosen from a small panel of anti-HER2 antibodies as having the most potent spectrum of in vitro growth-inhibitory activity across multiple tumor cell lines (Hudziak et al. 1989, Fendly et al. 1990, Lewis et al. 1993). MAb 4D5 was subsequently humanized (Carter et al. 1992) and ultimately developed as Herceptin (Carter et al. 2000). Years later this panel of anti-HER2 MAb (Fendly et al. 1990) was rescreened using new-found knowledge of two additional EGFR receptor family members, HER3 and HER4, plus numerous cognate ligands. MAb 2C4, unlike Herceptin, blocks heterodimerization of HER2 with HER3 induced by the ligand, heregulin (Agus et al. 2002). MAb 2C4 inhibits the growth of tumor cells induced by heregulin (Agus et al. 2002). A humanized version of MAb 2C4, Omnitarg, is in multiple phase II clinical trials (Table 1).

Recruiting and enhancing anti-tumor activities

If MAbs are identified with only modest anti-tumor activity it is both feasible and desirable to screen them using methods to recruit or enhance activity. Screening MAbs in combination with chemotherapy is strongly recommended as many antibodies, including Herceptin (Slamon & Pegram 2001), Rituxan (Demidem et al. 1997) and Erbitux (Mendelsohn 2000) show additive or synergistic anti-tumor activity in combination with chemotherapy. Moreover, combining drugs in oncology is a widely adopted paradigm. Indeed, combining antibodies with chemotherapy can sometimes improve the anti-tumor response rate or response duration as evidenced by Herceptin (Slamon et al. 2001) and Rituxan (Herold et al. 2003). The choice of chemotherapeutic agents for antibody combination in preclinical screens is usefully guided by a consideration of common chemotherapy regimens utilized in potential clinical settings for the antibody drug.

Antibody drug conjugates commonly, but not invariably, have much greater in vitro and in vivo anti-tumor potency than their corresponding naked parent antibodies (Liu et al. 1996, Ross et al. 2002, Bhaskar et al. 2003, Doronina et al. 2003, Mao et al. 2004), thereby potentially extending the range of surface antigens amenable to antibody targeting in oncology. However, the benefits of drug conjugation must be weighed against the challenges: need for an internalizing antibody, more

<table>
<thead>
<tr>
<th>Tumor activities in vitro and in vivo</th>
<th>Representative clinical examples $^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth inhibition</td>
<td>Herceptin, Erbitux</td>
</tr>
<tr>
<td>Cell cycle arrest</td>
<td>Herceptin, Rituxan, Erbitux</td>
</tr>
<tr>
<td>Sensitization to other agents</td>
<td>Herceptin, Erbitux</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td></td>
</tr>
<tr>
<td>Radiotherapy</td>
<td></td>
</tr>
<tr>
<td>Apoptosis</td>
<td>Rituxan</td>
</tr>
<tr>
<td>Activation of proapoptotic pathway</td>
<td>anti-Trail-R1, anti-Trail-R2 agonist antibodies</td>
</tr>
<tr>
<td>Death receptor activation</td>
<td>Erbitux, Omnitarg, ABX-EGF, HuMax-EGFr</td>
</tr>
<tr>
<td>Inhibition of anti-apoptotic pathway</td>
<td>Herceptin, Rituxan, Erbitux, Campath-1H, HuMax-CD20, Rituxan, Campath-1H, HuMax-CD20,</td>
</tr>
<tr>
<td>Growth factor antagonism</td>
<td></td>
</tr>
<tr>
<td>Secondary immune functions</td>
<td></td>
</tr>
<tr>
<td>Antibody dependent cellular cytotoxicity</td>
<td>Herceptin, Rituxan, Erbitux, Campath-1H, HuMax-CD20, Rituxan, Campath-1H, HuMax-CD20,</td>
</tr>
<tr>
<td>Complement dependent cytotoxicity</td>
<td></td>
</tr>
<tr>
<td>Anti-angiogenic</td>
<td>Avastin</td>
</tr>
<tr>
<td>Angiogenic ligand or receptor antagonist</td>
<td>Herceptin, Erbitux</td>
</tr>
<tr>
<td>Indirect</td>
<td>Erbitux</td>
</tr>
<tr>
<td>Inhibition of invasion/ metastasis</td>
<td></td>
</tr>
</tbody>
</table>

Carter et al.: Target identification and validation of cell surface antigens

complex preclinical and clinical development, a possible need for a closer to ideal antigen expression profile and increased risk of toxicity in patients.

In screening MAbs for their suitability as drug conjugates it is often impractical to conjugate each individual antibody. However, it is feasible to screen in the presence of a drug conjugate of a cross-linking antibody such as an anti-light chain MAb. Importantly, the behavior of a secondary antibody drug conjugate is often predictive of the behavior of a primary antibody drug conjugate (Klussman et al. 2004).

**In vivo screening of antibodies for anti-tumor activity**

Ideally, a large panel of MAbs is screened for anti-tumor activity in a battery of different *in vitro* assays. If possible, one selects the ‘best-in-class’ for each kind of anti-tumor activity, compares their anti-tumor activities *in vivo*, and selects a clinical lead candidate. Minimally, one needs a single antibody-based agent with robust *in vitro* anti-tumor activity to warrant *in vivo* evaluation. Commonly, human tumor cell lines (or explants) are xenografted into immune-compromised mice, such as nude or SCID, and allowed to establish. Tumors are treated with test or control agents and their size followed over a few weeks to months. If necessary, the stringency of the model can be increased by allowing the tumors to grow larger (commonly 100–500 mm³), or by reducing the dose or the number of doses.

Complete responses (i.e. no detectable tumor remaining) or cures (i.e. no tumor regrowth following a complete response) under stringent conditions and in multiple models would likely provide sufficient impetus to select a lead candidate for further optimization into a clinical candidate. Whilst such a high hurdle for efficacy can sometimes be met, e.g. using antibody drug conjugates (Liu et al. 1996, Ross et al. 2002, Bhaskar et al. 2003, Doronina et al. 2003, Mao et al. 2004), it does not necessarily predicate significant benefit to patients within acceptable toxicities. Another limitation of *in vivo* tumor models is that they are unavoidably imperfect mimics of human disease. Despite these limitations, *in vivo* tumor models are critical to the anti-cancer drug development process, providing the means to filter numerous test agents with potent *in vitro* anti-tumor activity down to a small, and hopefully more promising, subset worthy of further preclinical and clinical development.

**Conclusions and perspectives**

The search for surface antigens on tumor cells amenable to antibody targeting commonly begins with a hunt for differences between tumor and non-tumor cells at the DNA, mRNA, protein or antibody reactivity levels. Thus far, no single approach or even combination of methods has emerged as the preferred way to identify surface antigens suitable for targeting in oncology. Genomics, specifically the availability of the human genome sequence (Lander et al. 2001, Venter et al. 2001), has empowered more direct means of target identification, e.g. by expanding protein databases and enabling the mapping of novel cancer-associated genes. Transcriptomic methods have proved particularly powerful in defining gene expression signatures for cancer diagnosis or prognosis, but less so to date for the identification of potential antibody targets. The most promising application of transcriptomics, from an antibody target perspective, is the discovery of tumor-specific splice variants. Low prevalence of antigens, including splice variants, is a potentially limiting factor in their exploitation for commercial antibody drug development. Proteomic methods are appealing in that they allow a direct search for cell surface features that distinguish tumor cells from their normal counterparts. Rapid progress in expanding the throughput and capabilities of mass spectrometry bode well for proteomic identification of new tumor-associated antigens.

Historically, hybridoma technology has been the most successful tool for tumor antigen identification, but the point of diminishing returns appears to have been reached for this particular application. Large libraries of antibody phage offer the most direct and potentially comprehensive method for identifying antibodies that bind to tumor cells and sparingly to non-tumor cells. Many technical obstacles have been overcome to allow successful application of phage display technology to tumor cell panning (Liu et al. 2004). However, subsequent identification of cognate antigens remains a significant bottleneck for this, and other, antibody-based approaches.

Ideally, target validation includes profiling the expression of antigens across numerous normal tissues and tumor types. For the most promising candidates it is desirable to expand the survey of positive tumors to include large numbers of specimens at each stage of disease to provide robust estimates of the extent and prevalence of antigen expression. Such extensive expression profiling has only recently become feasible with the advent of so called tissue microarrays (TMA, Table 2) that allow *in situ* profiling of protein, mRNA or DNA on hundreds of samples in parallel.

Currently, there is a dearth of convincingly validated surface antigens that is compounded by the scarcity of so-called ‘naked’ antibodies with robust anti-tumor activity. Fortunately, the potency of anti-cancer antibodies can be improved in numerous ways (Carter 2001, Milenic 2002, Presta 2002, Glennie & van de Winkel 2003,
Ross et al., 2003). Thus, for validated targets there is often a compelling rationale to enhance the activity of corresponding antibodies for human therapy and a growing box of tools with which to do so. Our hope and expectation is that the marriage of target identification with antibody enhancement technologies will ultimately translate into new and improved therapies for cancer patients.

Acknowledgements

The broad scope and limited space for this review has obliged us to pick representative examples throughout. We apologize to numerous investigators whose excellent work we have not been able to cite. The authors are employed by Seattle Genetics, a company that discovers and develops monoclonal antibody-based therapeutics, including antibody drug conjugates, to treat cancer and immunologic diseases.

References


Bowler LD 2004 Representational difference analysis of cDNA. Methods in Molecular Medicine 94 49–66.


antibody sensitizes a B cell lymphoma cell line to cell killing by cytotoxic drugs. *Cancer Biotherapy and Radiopharmaceuticals* **12** 177–186.


Carter et al.: Target identification and validation of cell surface antigens


Julius MH, Masuda T & Herzenberg LA 1972 Demonstration that antigen-binding cells are precursors of antibody-producing cells after purification with a fluorescence-activated cell sorter. PNAS 69 1934–1938.


Luwor RB, Johns TG, Murone C, Huang HJ, Cavenee WK, Ritter G, Old LJ, Burgess AW & Scott AM 2001 Monoclonal antibody 806 inhibits the growth of tumor xenografts expressing either the de-2-7 or amplified epidermal growth factor receptor (EGFR) but not wild-type EGFR. Cancer Research 61 5355–5361.


Carter et al.: Target identification and validation of cell surface antigens


Endocrine-Related Cancer (2004) 11 659–687


Zard L, Carmellona B, Siri A, Petersen TE, Paolella G, Sebastio G & Baralle FE 1987 Transformed human cells produce a


