7 Antibody Engineering and Expression

JOHN R. ADAIR
Cambridge, UK

1 Introduction 221
2 Antibody-Binding Sites 221
  2.1 Antigen-Binding Sites – Structures and Generation 221
  2.2 Sources of Antibody-Binding Sites 226
    2.2.1 Murine and Humanized Binding Sites 226
    2.2.1.1 Antibody Humanization 227
  2.2.2 Primatized Antibodies 228
  2.2.3 Human mAbs 228
    2.2.3.1 Transgenic Animals Producing Human Antibody Repertoires 228
    2.2.3.2 In vitro Production of Human mAbs 229
  2.2.4 Low Molecular Weight Antibody-Based Binding Sites 229
  2.3 Binding Site Formats 229
  2.4 Number of Binding Sites per Product 230
3 Effector Mechanism(s) 230
  3.1 Second Binding Site (Bispecific Antibodies) 231
  3.2 Antibody Fc Functions (Immune System Recruitment) 231
  3.3 Antibody–Effector Fusions and Conjugates 232
    3.3.1 Cytokines 232
    3.3.2 Enzymes 232
      3.3.2.1 Immunotoxins 232
      3.3.2.2 ADEPT/ADAPT 233
      3.3.2.3 Thrombolytic and Coagulation Enzymes 233
    3.3.3 Drugs 233
    3.3.4 Radioisotopes 234
    3.3.5 DNA (Antibody Targeted Gene Therapy) 234
4 Linking the Binding Site to the Effector Mechanism 234
5 Manipulating the Half-Life of Antibody Products 234
6 Production of Recombinant Antibodies 235
6.1 Mammalian Cell Culture 236
  6.1.1 Antibody Glycosylation and Mammalian Cell Expression 236
6.2 *E. coli* Fermentation 236
6.3 Alternative Production Systems 236
  6.3.1 Transgenic Animals 237
  6.3.2 Transgenic Plants 237
  6.3.3 Insect Cells 237
  6.3.4 GRAS (Generally Regarded as Safe) Organisms 237
7 Conclusions 237
8 References 238
List of Abbreviations

ADAPT antibody-dependent abzyme prodrug therapy
ADCC antibody-dependent cellular cytotoxicity
ADCMC antibody-dependent complement-mediated cytolysis
ADEPT antibody directed enzyme prodrug therapy
CDR complementarity determining region
CEA carcinoembryonic antigen
CpG2 carboxypeptidase G2
DT diphtheria toxin
Fab fragment antigen binding
Fc fragment crystallizable
GRAS generally regarded as safe
HAMA human anti-murine antibody
Mbp megabase pairs
PCR polymerase chain reaction
PE Pseudomonas exotoxin
RES reticuloendothelial system
RSV respiratory syncytical virus
scFv single chain Fv
TNF tumor necrosis factor

1 Introduction

Antibodies have many properties which make them attractive as the starting point for novel biopharmaceutical products. They occur naturally in serum and mucosal surfaces at high concentrations and have long intrinsic half-lives. They combine high affinity and specificity for ligand (antigen) with innate effector elements. These functions are located in separate protein domains within the antibody which can be isolated by genetic engineering techniques and can be rearranged into novel combinations with other non-antibody elements. These novel molecules can be produced in a range of expression systems.

There are now several licensed monoclonal antibody (mAb) therapeutics and diagnostic products (Tab. 1) and a much larger number of murine, human and recombinant antibody candidates in research, preclinical and clinical development phases. It is anticipated that the rate of product licensing will continue to increase.

This chapter summarizes recent progress in the development of antibody-based products. The product concept and targets for antibody-based products are outlined and basic antibody structure, and the underlying genetic organization which allows easy antibody gene manipulation, and the isolation of novel antibody binding sites are then described. Features considered in the design and construction of antibody-based products are then summarized and finally, recent developments in options for production are noted.

2 Antibody-Binding Sites

Recombinant DNA technologies allow a modular approach to the design of antibody-based products. The product can be viewed as comprising one or more binding or targeting domains, one or more effector domains or elements, and suitable linking element(s), as shown schematically in Fig. 1. The composition and derivation of these elements are discussed in more detail later. Basic antibody protein structure is summarized in Fig. 2.

This section summarizes how the antibody-binding site structure, which provides the affinity and specificity suitable for its role as the binding and targeting domain, is generated.

2.1 Antigen-Binding Sites – Structures and Generation

The binding site is formed by non-covalent association of the N-terminal “variable” domains of the heavy and light chains (\(V_H\) and \(V_L\)) respectively (see Fig. 2). Structural studies of antibodies complexes with antigen (Wilson and Stanfield, 1994; MacCallum et al., 1996) show that the antigen-binding site is composed from the solvent exposed side chains of amino acids in 6 loops, 3 from \(V_H\) and 3 from \(V_L\), which emerge from the underlying \(\beta\)-sandwich frameworks of the
<table>
<thead>
<tr>
<th>Disease Area</th>
<th>Licensed mAb Product or Development Stage</th>
<th>Product Type</th>
<th>Date of First Product Launch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transplantation</td>
<td>Orthoclone-OKT®3 (muromonab-CD3)</td>
<td>murine mAb</td>
<td>1986</td>
</tr>
<tr>
<td></td>
<td>Zenapax™ (daclizumab)</td>
<td>humanized mAb</td>
<td>1997</td>
</tr>
<tr>
<td></td>
<td>Simulect® (basiliximab)</td>
<td>chimeric mAb</td>
<td>1998</td>
</tr>
<tr>
<td>Cancer</td>
<td>Panorex® (edrecolomab)</td>
<td>murine mAb</td>
<td>1995</td>
</tr>
<tr>
<td></td>
<td>Rituxan™ (MabThera® (rituximab)</td>
<td>chimeric mAb</td>
<td>1997</td>
</tr>
<tr>
<td></td>
<td>Herceptin® (trastuzumab)</td>
<td>humanized mAb</td>
<td>1998</td>
</tr>
<tr>
<td></td>
<td>Bexxar™ (131I Therapeutic) (tositumomab)</td>
<td>murine mAb</td>
<td>application late 1998</td>
</tr>
<tr>
<td></td>
<td>OncoScint® CR/OV (111In Diagnostic)</td>
<td>murine mAb</td>
<td>1992</td>
</tr>
<tr>
<td></td>
<td>(Satumomab Pendetide)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ProstaScint® (111In Diagnostic) (Capromab Pendetide)</td>
<td>murine mAb</td>
<td>1996</td>
</tr>
<tr>
<td></td>
<td>CEA-Scan® (99mTc Diagnostic) (arcitumomab)</td>
<td>murine Fab</td>
<td>1996</td>
</tr>
<tr>
<td></td>
<td>Verluma™ (99mTc Diagnostic)</td>
<td>murine Fab</td>
<td>1996</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>MyoScint® (111In Diagnostic) (Imicromab Pentetate)</td>
<td>murine mAb</td>
<td>1989</td>
</tr>
<tr>
<td></td>
<td>ReoPro® (abciximab)</td>
<td></td>
<td>1994</td>
</tr>
<tr>
<td>Infectious disease</td>
<td>Synagis™ (palivizumab)</td>
<td>humanized mAb</td>
<td>1998</td>
</tr>
<tr>
<td>Auto-immune disease</td>
<td>Remicade™ (infliximab)</td>
<td>chimeric mAb</td>
<td>1998</td>
</tr>
<tr>
<td>Allergy/asthma</td>
<td>phase III (RhuMab E25)</td>
<td>humanized mAb</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 1a. Idealized antibody-based product, comprised of 3 components, a binding site domain, a linker and an effector domain. These components may be derived from a variety of sources as noted in the figure and described in the text. b IgG antibody product, comprising 2 identical binding sites, an Ig hinge linker and one Fc region.

variable domains (Fig. 3). Of these loops, H3, which is most variable in sequence and length, and L3, occupy the center of the binding site and tend to make most of the key contacts with antigen. These 6 loop regions are substantially, but not exactly, similar in extent to the previously defined complementarity determining regions (CDRs) (WU and KABAT, 1970). The CDR were initially identified as regions within the variable domains which exhibited even greater between antibody sequence variability (hypervariability).

For 5 out of the 6 loops, but not yet H3, it has been possible to identify a small number of canonical conformations that each loop can adopt depending on sequence, length and the identity of key framework residues which contact the loops (CHOTHIA et al., 1989; CHOTHIA and LESK, 1987; TRAMONTANO et al., 1990). Using this information predictions can be made about the conformation for a newly derived antibody sequence. Some estimate of the likely structure of the binding site, from sequence analysis or calculation or structure determination, is important for many antibody engineering projects (e.g., humanization or affinity modification).

The sequences of H3 and L3 are generated during the process of genomic recombination and somatic mutation which occurs during B cell development. The coding sequence of H3 is formed when one member from each of 3 genetic segments, the VH minigenes, D_H (diversity) segments and J_H (joining) segments, is brought together during the maturation of the Ig heavy chain. The coding sequence for L3 is formed by recombination of one each from a set of V_L minigenes and J_L segments during the formation of the light chain. H1 and H2, and L1 and L2 are derived from the coding sequence in the VH and V_L minigenes, respectively, but can also be altered later during the somatic mutation process (WABL and STEINBERG, 1996).

The number of these sequences in humans, primates and rodents, their organization into sequence subfamilies and their relative utilization in the immune response is of interest to antibody engineering projects. For example, the human immunoglobulin loci have
Fig. 2a. Antibody structure. The basic immunoglobulin unit, as revealed by electron microscopy and X-ray crystallography experiments, adopts a Y or T shaped form. The binding site “arms” (termed Fab, for fragment antigen binding) can be isolated by digestion of IgG with enzymes. Papain digestion releases individual Fab fragments by cleavage on the N-terminal side of the interheavy chain disulfide bridges. Fab comprises the light chain and the Fd fragment, the 2 N-terminal domains of the heavy chain. The non-covalent interactions between Fd and light chain lead to a strong association ($k_D \approx 10^{11} \text{ M}^{-1}$), although the affinity of the individual interacting domains is much less ($k_D \approx 10^6–7 \text{ M}^{-1}$) (e.g., HORNE et al., 1982 and references therein). Pepsin digestion releases the F(ab)\(_2\) fragment by digestion on the C-terminal side of the interheavy chain disulfide bridges. F(ab)\(_2\), therefore, contains both light chains and 2 Fd hinge heavy chain fragments linked by disulfide bridges. The Fd hinge heavy chain fragment is known as the Fd'.

The N-terminal variable domains together are termed Fv (fragment variable). In a small number of cases the Fv can be isolated from IgG with pepsin (GIVOL, 1991). The associating surface between the 2 domains of Fv derives from CDR and framework sequences and, therefore, the stability of FVs varies from antibody to antibody. The other major fragment of IgG formed after pepsin digestion is termed the Fe (fragment crystallizable). The Fe contains the domains which interact with immunoglobulin receptors on cells and with the initial elements of the complement cascade. Pepsin sometimes also cleaves before the third constant domain (CH3) of the heavy chain to give a large fragment F(abc) and a small fragment pFc'. These terms are also used for analogous regions of the other immunoglobulins. b Antibody fragments which derive from IgG and which can be produced by means of genetic manipulation are shown. The scFv comprises the 2 domains in Fv genetically linked with a peptide which links the C-terminus of one domain to the N-terminus of the other linker.

been substantially cloned and mapped. There are believed to be approximately 95 $V_H$ segments, although only about half of these would be able to contribute to forming functional antibody; 30 (approx.) $D_H$ and 6 $J_H$ sequences over a 1.1 megabase pairs (Mbp) region of chromosome 14q (COOK and TOMLINSON, 1995; MATSUDA and HONJO, 1996). There are 2 light chain loci $k$ and $l$. At the $k$ locus there are 76 $V_k$, of which 32 are potentially functional, and 5 $J_k$ followed by a single $C_k$ gene in 2 Mbp of chromosome 2q (ZACH AU, 1993) (Fig. 4). The $\lambda$ locus extends over 1.15 Mbp on chromosome 22q (WILLIAMS et al., 1996). There are 52 (approx.) $V_\lambda$, an undefined number of these are pseudogenes. The $J_\lambda$ and $C_\lambda$ are linked and duplicated. There are 7 $J_\lambda-CA$ miniloci although 3 are pseudogenes. V-J recombination, therefore, determines which $C_\lambda$ will be used.

Heavy chain recombination occurs before light chain reorganization and the heavy variable domain appears on the surface of pre-B cells as membrane bound $\mu H$ in association with a surrogate light chain. Later in B cell development $\kappa$ locus rearrangement occurs,
and if no productive κ chain is made at either of the κ alleles then λ rearrangement occurs until a functional light chain is formed (Fig. 5). A functional light chain then appears on the B cell surface with a heavy chain. Exposure to antigen and selection for B cells expressing binding sites with reasonable affinity for antigen then occurs.

Within a species the V minigene sequences can be grouped into a number of families according to amino acid or nucleotide sequence homology. In some cases interspecies homology between these families appear higher than intraspecies homology. The interspecies homologies can range from \( \sim 40\% \) to \( \sim 80\% \) between mouse and man, and are very high between primates and man. The sequence differences are sufficient to cause murine antibodies to be recognized as foreign by humans.

### 2.2 Sources of Antibody-Binding Sites

A number of sources of suitable antibody-binding sites is available. These are:

1. murine mAbs and humanized binding sites
2. primatized antibodies
3. human mAbs
   - transgenic animals producing human antibody repertoires
   - \textit{in vitro} production of human mAbs
4. lower molecular weight antibody-based binding sites

#### 2.2.1 Murine and Humanized Binding Sites

The variable domain coding sequences can be obtained from among the many thousands of murine mAbs which have been generated since 1975, including many with clinical applications. Methods for production of hybridomas are reviewed by Donohue et al. (1995).

Murine mAbs may provide binding sites which are suitable for single use, \textit{in vivo} diagnostic or therapeutic purposes. However, in most but not all cases, sequence differences between murine and human antibodies lead...
to an immune response (human anti-murine antibody, HAMA) when murine mAbs are used as diagnostics or therapeutics, even in immunocompromised patients. This HAMA response is against both variable and constant regions so constructing a chimeric antibody (MORRISON et al., 1984; NEUBERGER et al., 1985; LOBUGLIO et al., 1989) by attaching the constant regions of a human antibody to the binding site domains of a murine antibody does not always remove the problem. Clinical data with chimeric antibodies are reviewed by SALEH et al. (1994). The HAMA response can, in principle, be avoided by various strategies, including coadministration of tolerizing anti-CD4 antibodies (MATHIESON et al., 1990), conjugation with polyethylene glycol ("pegluation") (NUCCI et al., 1991; INADA et al., 1995) or transferring the binding site from the murine mAb and transplanting it to a human mAb (humanization).

2.2.1.1 Antibody Humanization

Antibody humanization involves the substitution of sufficient residues from the variable domains of a non-human mAb into the variable domains from a human mAb so as to reconstitute in the human mAb the binding affinity and specificity of the non-human mAb. This notion was first raised at around that time chimeric antibodies were initially being described (MUNRO, 1984) because it was anticipated that chimeric antibodies would still be immunogenic in man due to their foreign variable domains. When humanization was suggested many of the structural features of antibody variable domains discussed in Sect. 2.1 above were not yet known, and the relative importance of the framework residues, including sequences introduced during somatic mutation, in determining binding site affinity and specificity was unknown. It was also
Fig. 5. IgGκ transcription and translation.

Humanized antibody construction—heavy chain

Murine MAb

Human MAb

Humanized MAb heavy chain gene

Humanized MAb heavy chain gene

Fig. 6. Humanized antibody construction – heavy chain.
not clear to what extent any conformational adjustments of the CDRs after interaction with antigen, or solvent effects, contributed to net binding affinity, or, more importantly, how to take these factors into account during a humanization process (Fig. 6).

Nevertheless, the first successful experiments involving simple CDR transfers from a mouse mAb to a human mAb were soon described (Jones et al., 1986; Verhoeven et al., 1988) and were rapidly followed by the first humanization of a clinically relevant mAb, CAMPATH-1 (CD52), (Riechmann et al., 1988) The latter subsequently entered clinical investigation (Hale et al., 1988).

These and many other humanization experiments confirmed that certain murine framework residues were indeed necessary for restitution of significant binding activity. This has led to the development of a number of processes of humanization and descriptions of humanized mAb products (Mountain and Adair, 1992). Each in its own way seeks to identify those key amino acids (mainly buried within each variable domain) which interact with or affect, directly or indirectly, the conformation of the binding site. The success of these procedures can be measured by the fact that there are now well over a hundred disclosed examples of humanized mAbs, with a considerable number in clinical evaluation (Adair and Bright, 1995).

2.2.2 Primatized Antibodies

As an alternative to murine systems, cynomolgus macaques have been used to develop mAbs to human antigens (Newman et al., 1992). These primates are sufficiently related to humans to generate antibodies very similar in sequence and structure to humans, but are distinct enough to recognize many human antigens as foreign. Hybridomas can be generated and the binding site coding sequences extracted. First examples of their use have been macaque variable domain–human constant domain chimeric antibodies. One of these “primatized” antibodies, a CD4, chimeric mAb, is currently in human phase II clinical studies for rheumatoid arthritis, and a number of others are in preclinical development.

2.2.3 Human mAbs

The human immune system is a further source of useful binding sites. Volunteers or vaccinees immunized with particular foreign antigens, or patients convalescing after an infection, or cancer patients can be the source of B cells from which human hybridomas expressing human mAbs to a range of antigens can be prepared (James, 1994). A number of human antiviral and antitumor IgG mAbs from human hybridomas has been tested in the clinic. Alternatively the coding sequences for the antibody binding sites can be extracted for later genetic manipulation (Larrick et al., 1992).

2.2.3.1 Transgenic Animals

Producing Human Antibody Repertoires

An alternative approach to human mAb production is to transplant sufficient of the human Ig loci into the genome of another species and then to develop hybridomas after hyperimmunization with the antigen of interest. Mice are obvious candidates because of the established hybridoma technology. Lonnberg (1994) and Bruggemann and Neuberger (1996) and Bruggemann and Taussig (1997) have recently reviewed developments in the production of mAbs from transgenic mice. High affinity human mAbs with therapeutic potential have successfully been developed by this approach (Fishwild et al., 1996; Mendez et al., 1997).

2.2.3.2 In vitro Production of Human mAbs

There have been many attempts to derive B cells expressing specific antibodies by immunizing in vitro which have met with little success. Primary or secondary in vitro immu-