REVIEW

ANTIBODY ENGINEERING: ADVANCES FROM THE USE OF *ESCHERICHIA COLI* EXPRESSION SYSTEMS

Andreas Plückthun

Genzentrum der Universität München, c/o Max-Planck-Institut für Biochemie, Am Klopferspitz, D-8033 Martinsried, FRG.

Synthesis in *Escherichia coli* of correctly folded antibody fragments that bind antigen with the same affinity as the whole antibody is now possible. Here I review the techniques for achieving this and the physical properties of the various fragments produced. This technology not only facilitates antibody engineering but is also the basis of screening libraries for binding activity. Although the immunization of animals has not been made unnecessary in the production of monoclonal antibodies, steps toward this goal are now feasible.

Antibodies hold a firm place in biological research, and will have an increasingly important role in medical and industrial applications. They are highly selective binding agents and can be generated against almost any substance by standard approaches. They usually have rather high binding affinities for the target substance, and they have numerous biological effector functions mediating the immune response subsequent to binding antigen. Furthermore, as antibodies with widely different specificities are structurally similar molecules, standard ways of detection, derivatization and coupling with marker enzymes or other proteins such as toxins can be devised[1].

The invention of monoclonal antibodies[2] has provided general access to homogeneous antibodies of predefined specificity. The next step, the redesign of the antibody molecule itself, requires the use of the methods of gene technology. Initially the crucial problem was to find convenient ways of expressing the protein. The first expression systems reported with which antibodies could be obtained in their native state were for mammalian cells [for reviews see refs. 3 and 4]. This review will focus on facilitating native antibody expression using a bacterial system, *Escherichia coli*[5], and on the new possibilities arising from this technology.

OPPORTUNITIES OFFERED BY ANTIBODY EXPRESSION IN *E. COLI*

The expression of antibody fragments in *E. coli* brings the arsenal of techniques of bacterial gene technology to antibodies. There are several ways to express antibody fragments in *E. coli* (see below), and there are some advantages common to all the various approaches. These include a well established and convenient gene technology, which allows constructions to be made easily and directly in the expression vectors. The fast growth of *E. coli* and its comparatively simple fermentation make the large-scale production of antibody fragments relatively convenient. Fragments of the antibody can be produced directly, obviating proteolytic preparation. The facile production of small fragments on a large scale and the rapid mutagenesis now possible benefit functional studies[6] and structural work such as NMR and X-ray crystallography[7–9], and therefore ultimately the development of more rational engineering strategies.

Antibody engineering may also include the design of bifunctional molecules. Examples include the gene fusion of an antigen binding fragment with a toxin[10], a reporter enzyme[11,12] or an affinity handle[13]. While expression in *E. coli* is not essential for producing such bifunctional molecules, it facilitates the work considerably.

Some other opportunities arising from antibody production in *E. coli* require, however, that the antibody protein is produced in the native state, i.e. that no *in vitro* refolding is necessary[14–16]. Examples are the screening of binding activity directly from bacterial colonies or in bacteriophage plaques, which is the prerequisite for all work with gene libraries, be it a collection of random mutants or an assortment of clones from the immunological repertoire. If catalytic antibodies[17] are to be screened directly by their "enzymatic" activities, they must also be produced in the native state. Much more research needs to be carried out on these screening strategies in order to be able to handle larger numbers, weaker binding constants and, for catalytic antibodies, the detection of typically low catalytic activities. Yet the work with libraries will probably be among the most promising approaches that *E. coli* has to offer. The production of single species of antibody fragments also benefits from a native expression, since the one-step purification by antigen affinity chromatography depends on the antibody fragment being in the native state.

There is no single expression strategy and no particular antibody fragment (Fig. 1) that is the optimal solution for all conceivable applications. The choices (see below) depend very critically on the intended application, be it the

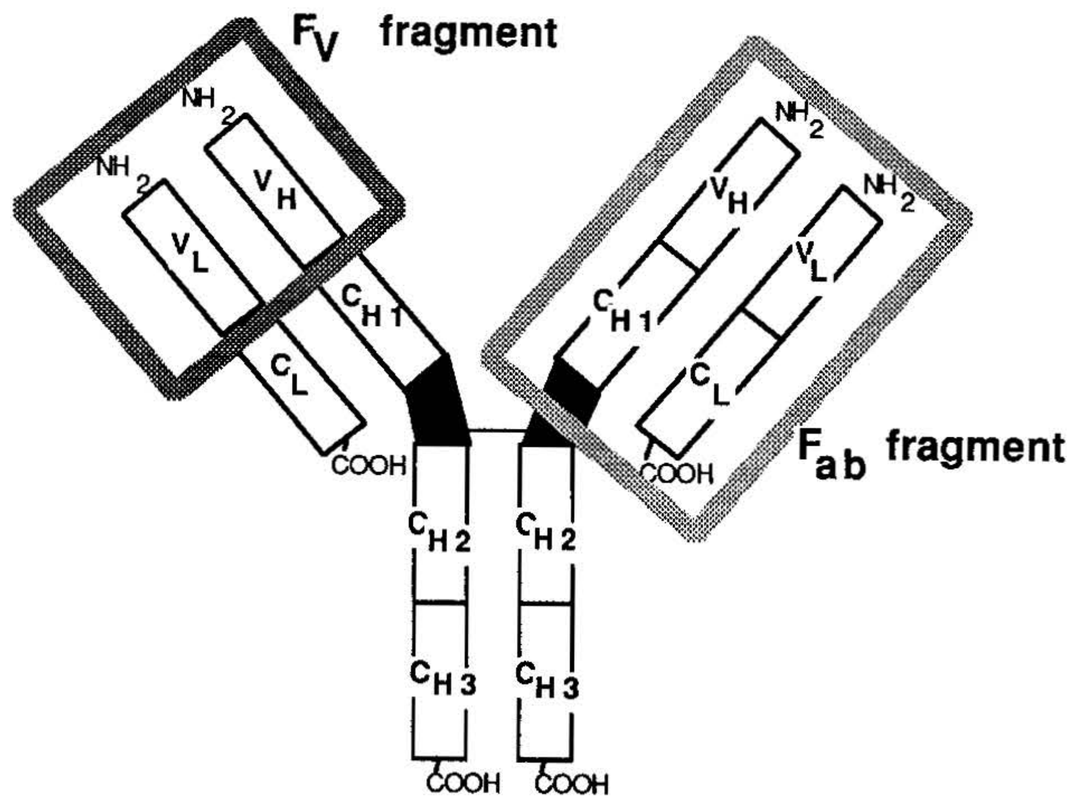

FIGURE 1 Antigen binding fragments of an antibody. The F_v fragment is a heterodimer of only the variable domains of the heavy chain and the light chain, whereas the F_{ab} fragment also contains the constant domain of the light chain and the first constant domain of the heavy chain.

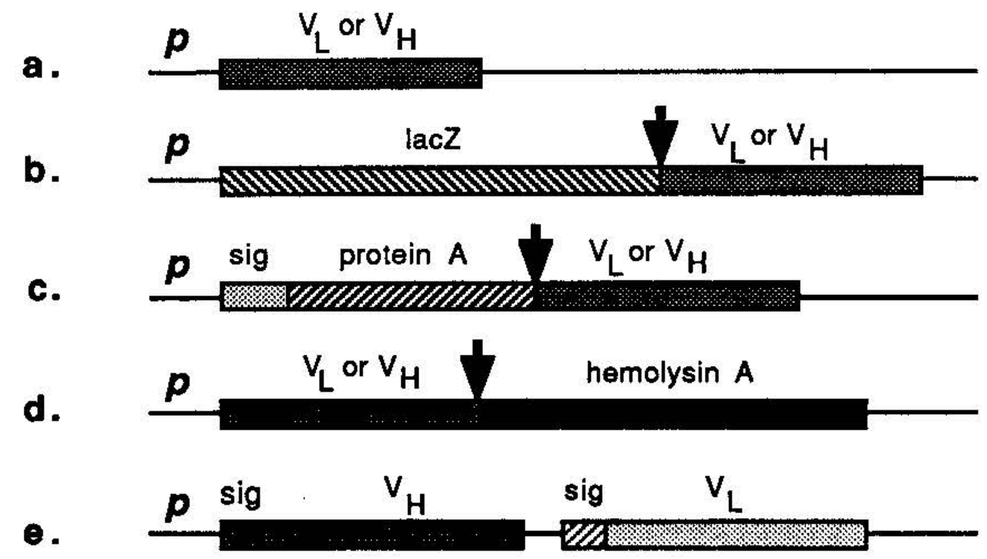

FIGURE 2 The arrangement of genes for the various strategies of antibody expression in *E. coli*. In all cases, *p* denotes a promoter. All boxes are coding regions, and all coding regions are preceded by appropriate translation initiation signals (not shown). Arrows denote sites for specific proteases, to allow *in vitro* cleavage of the fusion proteins. V_L and V_H are the variable domains of the antibody, but other antibody fragments may be used analogously. *sig* denotes a signal sequence. (a): direct cytoplasmic expression of antibody domains, (b): expression of cytoplasmic fusion proteins, to be cleaved *in vitro*, (c): secreted fusion proteins, to be cleaved *in vitro*, (d): secreted fusion protein with C-terminal (uncleaved) fusion signal that must be removed *in vitro*, (e): simultaneous secretion of both chains in the same cell, by employing an artificial operon and two bacterial signal sequences, requiring no *in vitro* refolding.

mass production of a single antibody species, the rapid engineering of an antibody, its structure determination, the testing of many variants or the screening of libraries. The requirements will also differ for antibodies intended for human or animal use as opposed to those intended for *in vitro* research or industrial purposes only. In the evaluation of the suitability of different expression systems for mass production, it must also be taken into account that the *E. coli* expression system for antibody fragments has been developed only very recently, has not yet been optimized, and probably not approached its limits.

GLYCOSYLATION

Antibodies are glycosylated in the organisms that produce them naturally. It appears that most antibody classes carry sugar residues in the C_H2[18] domain, with the known exception of mouse IgA[19]. Some classes (IgM and IgA) are also glycosylated in C_H1[19–21]. IgM is glycosylated in all constant domains, and some IgA subclasses are glycosylated in the hinge region. The glycosylation patterns also seem to vary between antibodies from different species. In rare cases, the variable regions can fortuitously contain glycosylation sites, either in V_H[22,23] or in V_L[24]. These rare, unusual glycosylation events may improve[23] or abolish[25] antigen binding.

Ordinarily, however, glycosylation has no influence on the antigen-binding properties of the antibody[26–30]. This question has been examined both by the inhibition of glycosylation with tunicamycin[26,27,31,32] or by using the methods of gene technology[30,33]. Specifically, the expression of the F_{ab} fragment of the mouse IgA McPC603 in *E. coli* gave an identical hapten binding constant as the proteolytic material from mouse ascites, which is glycosylated in C_H1[30].

Glycosylation may play a critical role in the secretion of proteins from the eukaryotic cell, and may affect the solubility and degradation of the protein. The secretion of murine IgA and IgE, and to some degree also IgM, is severely affected by the glycosylation inhibitor tunicamycin, whereas IgG and IgD appear to be only slightly affected[20,26–29,31–36]. Since IgG can thus be produced in non-glycosylated form, most studies on the effects of glycosylation on functions other than antigen binding have been carried out with IgG. Briefly, binding of IgG by protein A is not hampered by the lack of glycosylation, whereas three other effector functions are severely reduced: the ability to activate complement, the ability to bind to the F_c receptor and the ability to induce cellular cytotoxicity[26,27,31,32].

As most antibody classes have no glycosylation sites within the F_{ab} fragments, *identical* F_{ab} molecules can be produced in *E. coli* and in mammalian cells. Even for classes that are glycosylated within the F_{ab} fragment, antigen binding does not seem to be affected. The production and folding of whole antibodies in *E. coli* has not yet been fully investigated. Such unglycosylated molecules would be expected to bind the antigen, but they may have folding or stability deficiencies in the F_c part and would not be able to fulfill the functions listed above that depend on glycosylation. The production of such whole, unglycosylated, antibodies from *E. coli* would therefore mainly be useful for biochemical investigations. It may be noted, however, that the two arms of the antibody can often bind to the same surface and thus generate a sufficiently high avidity when the thermodynamic affinity is actually rather weak. This may make whole antibodies or $(F_{ab})_2$ fragments useful in those cases where very low affinity is a problem and the antigen is polymeric or surface bound.

In summary, antibody glycosylation does not influence antigen binding, but the lack of glycosylation does affect some of the effector functions. For industrial and most *in vitro* uses of antibodies, glycosylation is thus immaterial. If applications are envisaged, where glycosylation would be vital, the stepwise engineering of the binding site may still be carried out more conveniently using *E. coli*, and the final version of the antibody may then be expressed in mammalian cells.

EXPRESSION STRATEGIES FOR *E. COLI*

All procedures that have been used for the expression of other heterologous proteins in *E. coli* have also been used with antibodies or antibody fragments, and shall be briefly summarized here. However, only the simultaneous secretion of both chains (or the use of a secreted single chain fragment) gives the advantages of folding to the native, functional state in *E. coli*, as summarized below. This method will therefore be discussed in more detail. An overview of the arrangements of genes for the various strategies is given in Figure 2.

Direct cytoplasmic expression. Conceptually, the most straightforward approach is the direct expression of the antibody chains without a signal sequence[10,37–47]. The yields reported for different antibody chains, fragments and constructions are rather different and critically depend on the design of the translation initiation region, the proteolytic stability of the fragment and the host strain. Usually, inclusion bodies have been obtained with this approach, provided the expression was high enough to overcome the proteolytic degradation by *E. coli* proteases. Therefore, the resulting polypeptides are denatured, mostly reduced, and may carry a methionine or even a formyl-methionine at the N-terminus[48]. They must therefore be renatured *in vitro*, a reaction that needs to be empirically optimized with respect to protein concentration, redox-buffer, salt concentration, and temperature. While such an expression strategy has been reported for a variety of different fragments and domains expressed in the bacterial cell[10,37–47], only very recently has a systematic study of refolding appeared[49]. It remains to be investigated, however, whether the same conditions will be optimal for other antibody fragments.

Expression as a cytoplasmic fusion protein. Cleavable fusion proteins with N-terminal fragments of highly expressed cytoplasmic proteins can be constructed to over-

come several problems. An otherwise inefficient translational initiation can be made more efficient, the N-terminal portion may help the protein precipitate and thus protect it from proteolytic degradation, and a precise *in vitro* cleavage of the fusion protein may also lead to the desired N-terminus. The V_L domain[44] has been expressed as a gene fusion to the *c*II repressor of phage λ. It was cleaved with the blood clotting protease factor Xa and reconstituted with antibody-derived V_H. Both V_H and V_L have also been fused to a truncated β-galactosidase, cleaved with factor Xa, refolded and purified to homogeneity by hapten affinity chromatography (R. Glockshuber and A. Plückthun, unpublished). This strategy can be useful for the construction of variants that are very unstable *in vivo*, for the expression of a single variable or constant domain, or for the expression of single domains of related molecules from the immune system. The success of this strategy depends on the availability of an efficient cleavage, and, just as with direct expression, on an efficient refolding procedure and on the purification of the correctly folded and assembled molecular species by antigen affinity chromatography.

Secreted fusion proteins. Another strategy that can be used to express antibody domains is to fuse them to another secreted protein. Both N-terminal and C-terminal fusions to antibody domains have been constructed[13,15,50]. The fusion partner used at the N-terminus was *Staphylococcus aureus* protein A and the fusion partner at the C-terminus was either V_H or V_L (ref. 15 and I. Pfitzinger and A. Plückthun, unpublished) or a single-chain F_v fragment[13]. The resulting hybrid protein is secreted to the periplasm and, after partial lysis of the outer membrane, to the medium. The fusion proteins can be purified by IgG-affinity chromatography.

The use of a C-terminal secretion signal, namely that of hemolysin A[50] has also been investigated. Here the antibody domain was expressed without an N-terminal signal sequence and genetically fused to the C-terminal fragment of hemolysin via a cleavable linker sequence. In this case, the linker contained the recognition sequence for factor Xa. Secretion depends on the specialized transport system for hemolysin, which must be supplied on a separate plasmid[50]. The fusion proteins of V_H and V_L with hemolysin A were secreted to the medium, but the purification and handling of the rather hydrophobic fusion proteins is not trivial, and this method will require further development before being competitive as a general production method for antibody domains.

Functional expression by secretion. The simultaneous secretion of both chains of the antibody fragment gives rise to native and functional antibody fragments, and is therefore the prerequisite for many of the attractive features of a bacterial expression system, notably the screening of binding activity without prior *in vitro* folding. It was first developed for the F_v fragment[14] and the F_{ab} fragment[15] in our laboratory and independently by Better et al. for the F_{ab} fragment[16].

The essence of the strategy is to reproduce in *E. coli* the normal folding and assembly pathway of antibodies within the eukaryotic cell. In antibody producing cells, the two chains are expressed separately as precursors with N-terminal signal sequences and separately transported to the lumen of the endoplasmic reticulum (ER). There, the signal sequences are cleaved by a membrane-bound signal peptidase. In the lumen of the ER, folding of the protein, disulfide bond formation and assembly of the light and heavy chain to the complete antibody take place[29].

The crucial hypothesis in the design of the secretory expression system for antibody fragments was that protein transport to the periplasm of *E. coli* is functionally

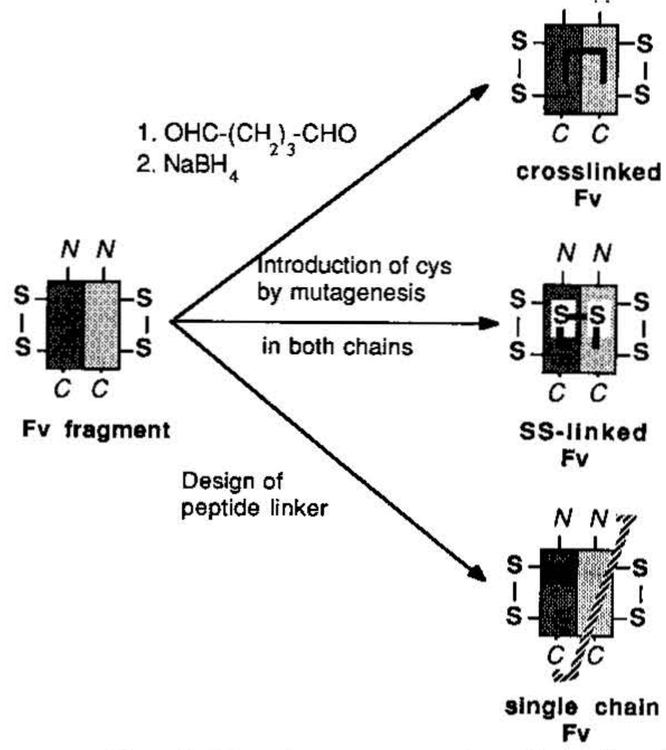

FIGURE 3 Scheme of stabilization strategies for the F_v fragment. The top arrow describes the chemical crosslinking of both chains. The middle arrow describes the design of an artificial additional disulfide bond connecting the two domains. The bottom arrow describes the design of a peptide linker connecting the C-terminus of one chain to the N-terminus of the other chain.

equivalent to the transport of a protein to the lumen of the ER. A system was thus designed that directs both chains of the antigen binding fragment of the antibody to the periplasm of the same *E. coli* cell. It has been known for some time that many heterologously expressed secretory proteins can fold correctly when transported to the periplasm of *E. coli* (for a review see ref. 51). The expression of antibody fragments was the first example, however, in which the assembly of a periplasmic heterodimer was required.

There are several critical steps in the expression and secretion pathway that must occur correctly for the F_v or F_{ab} fragment to assemble, and these shall be discussed: (1) Approximately stoichiometric amounts of both chains must be synthesized. This was accomplished by constructing an artificial operon, in which both genes are under the control of the same promoter and both coding regions are preceded by Shine-Dalgarno sequences. (2) Both chains must be transported to the periplasm. This was achieved by precise fusions to two different *E. coli* signal sequences, one from the outer membrane protein A (*ompA*) and the other from alkaline phosphatase (*pho*A). It has been found, however, that a variety of signal sequences will function on both chains (A. Skerra and A. Plückthun, submitted). (3) Both signal sequences must be cleaved at the correct position to yield the identical N-terminus as in the original antibody molecule. The precise fusion of the signal sequences is necessary but not always sufficient for this to occur[51], and the location of cleavage had to be verified experimentally for both chains. This was done for the F_v fragment of McPC603 by the determination of the N-terminal protein sequence on both chains, and the correct cleavage was confirmed for both chains[14]. (4) The folding to globular domains must then occur, (5) the intramolecular disulfide bonds must form, and (6) the two chains must assemble to form the required heterodimer. While the *in vivo* folding has not yet been investigated in detail, it probably is crucial that the folding of either chain occurs in the presence of the other, to give a heterodimer.

There are several advantages to this secretory expression system: First, it directly leads to an assembled functional product with correctly formed disulfide bonds without the need to refold the protein *in vitro*. Second, the problem of protease degradation is greatly diminished, as there are fewer proteases in the periplasm than in the cytoplasm. Protection is also achieved by the folding to globular domains, accompanied by the oxidation of the S-S-bonds in the periplasm. Since the exposure to the oxidizing milieu of the periplasm during the folding

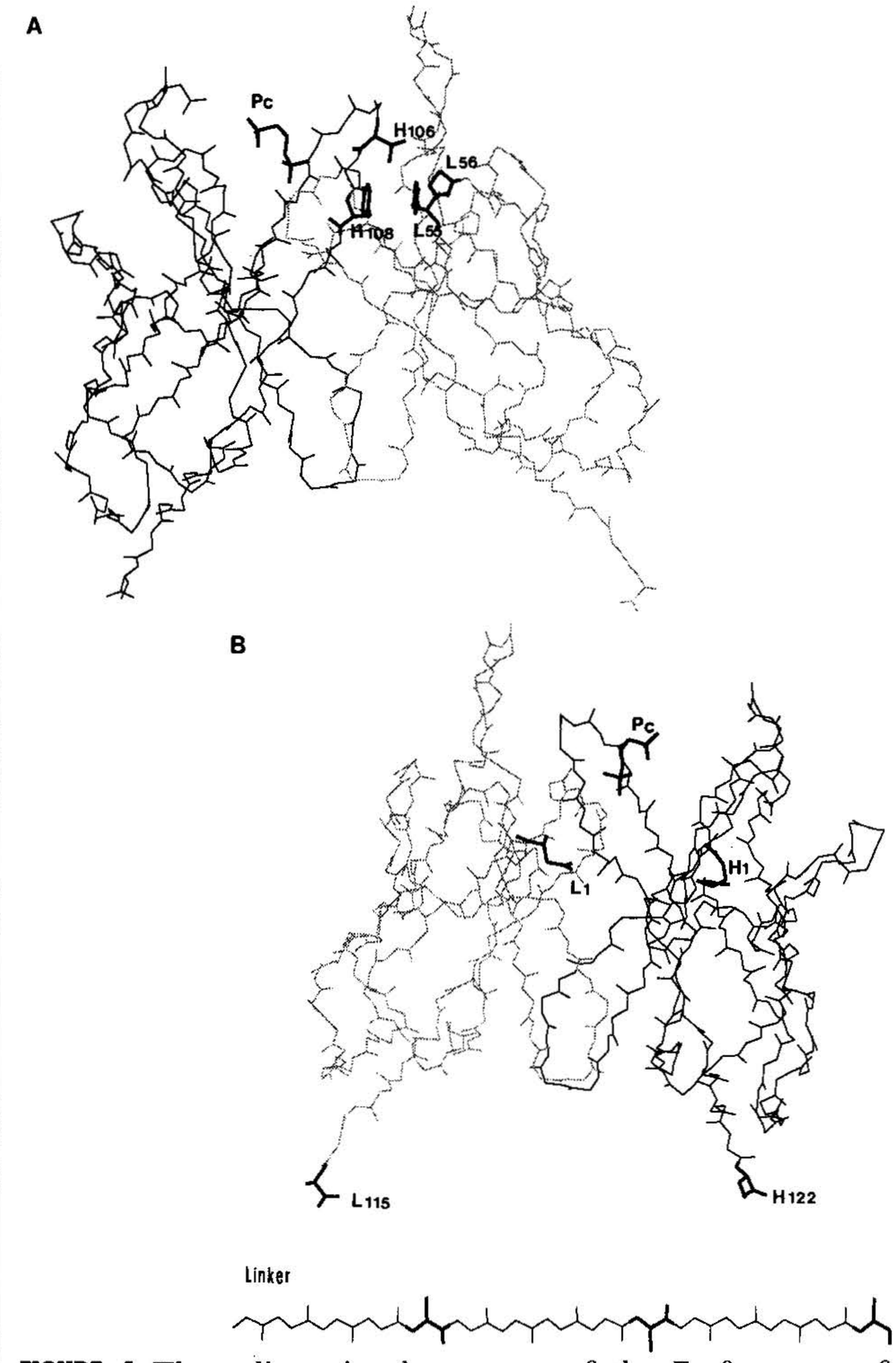

FIGURE 4 Three-dimensional structure of the F_v fragment of McPC603 illustrating the geometric relations in the covalent linking strategies. The heavy chain is shown in black, and the light chain is shown in grey. The hapten phosphorylcholine is labeled and emphasized by thick lines on the top. The numbering is that of McPC603. (A): residues that have been replaced to obtain successful *in vivo* formation of an intermolecular disulfide bond, (B): geometric relationships in a single-chain F_v fragment. The N-termini of both chains are labeled and denoted by thick lines. The linker with the sequence $(GGGGS)_3$ is shown on the same scale in completely extended form on the bottom. The serine residues are emphasized by thick lines. In McPC603, it was used to connect H122 to L1.

process is necessary for the disulfide bonds to form, the secretion of both chains is an essential requirement.

The expression rate must be commensurate with the rates for the transport, folding and assembly steps or the protein will accumulate as insoluble material in either the periplasm or the cytoplasm. It is not automatically useful, therefore, to employ extremely strong promoters, or problems with cell lysis may be encountered (see below). In our laboratory the *lac* promotor/operator was found to be useful. The design of our initial vectors has been reported[14,15], and improved vectors have since been constructed[52] according to the same principles, but with several new convenient features and improved properties. They are described in detail elsewhere[52].

The correct assembly of both the F_v and the F_{ab} fragment was proven experimentally. The most critical test is the determination of the binding constant and stoichiometry. In the case of the F_v and the F_{ab} fragment of McPC603 prepared from *E. coli*, an affinity constant identical to that of the whole antibody prepared from ascites or its proteolytic F_{ab} fragment could be demonstrated, as well as a stoichiometry of binding of one mole hapten per mole F_v fragment[14,30,53].

It is generally found advantageous, especially in working with large volumes for preparative purposes, to achieve a secretion to the periplasm, and not to the medium. The volumes to be processed are smaller, and fewer problems with the separation of cell debris are usually observed. Leakage of periplasmic proteins to the culture medium is a phenomenon that is sometimes observed when heterologous proteins are secreted in *E. coli*. To avoid premature leakage, the operon expressing the two chains of the antibody fragment was put under the control of an inducible promotor (*lac*). At higher temperatures, the outer membrane of the cells begins to become permeable shortly after induction[15]. This results in the liberation of periplasmic proteins, e.g. β-lactamase, to the growth medium, and in the simultaneous "leakage" of the F_v or the F_{ab} fragment. There is no indication that this phenomenon is related to the choice of a particular signal sequence (A. Skerra and A. Plückthun, submitted), and it would be incorrect to speak of a "secretion to the medium". Notably at lower growth temperatures, this phenomenon can be minimized[53] (A. Skerra and A. Plückthun, submitted). The biochemical basis of this phenomenon is only poorly understood. It is not a consequence of the expression of F_{ab} versus F_v fragments either, but it does depend on the strain and growth conditions. In essence, it leaves the investigator the choice of obtaining the fragments from the medium or the periplasm, depending on host strain, plasmid, induction time, temperature, and probably several other yet unidentified parameters.

ANTIBODY FRAGMENTS FUNCTIONALLY EXPRESSED IN *E. COLI* AND THEIR PROPERTIES

The fragment of the antibody needed (Fig. 1) will depend on the application envisaged. If whole antibodies are desired, mammalian cells may currently still be the preferred production host[3,4]. If antibodies are intended for human use, it will have to be decided whether the natural effector functions of the whole antibody are desired or whether instead artificial effectors should be engineered into much smaller fragments of the antibody, which are particularly easily obtained from *E. coli* and may have particular advantages in tumor diagnostics and therapy (see below).

In the following sections, some of the properties of the various fragments will be discussed. The mouse IgA McPC603 is a particularly well characterized antibody and its crystal structure with and without bound antigen is known[54–56]. From studies of this antibody, of which many different fragments have been obtained and compared side by side, some general features can be noted.

F_{ab} fragments. F_{ab} fragments of antibodies can easily be prepared by proteolytic cleavage and have been well characterized. They have frequently been documented as having the same antigen binding activities as the whole antibody[1,18]. The recombinant F_{ab} fragment of McPC603 was designed (according to the defined electron density in the crystal structure) to end before the hinge region, and it was compared to the proteolytic material[30]. The proteolytic IgA fragment obtained by pepsin cleavage is an F_{ab}' fragment, where the prime denotes that it contains the hinge region. Furthermore, it is glycosylated in C_H1. Yet, neither of these differences have any influence on the antigen binding constant[30]. Binding constants of other recombinant F_{ab} fragments have been investigated by solid phase methods, and similar binding as in the proteolytic material was also found[16].

There is an interesting difference between the F_{ab} fragment and the F_v fragment, however. It appears that the yield of folding and/or assembly of the F_{ab} fragment is lower in *E. coli* than that of the F_v fragment, while no dramatic differences in expression or secretion can be

detected (A. Skerra and A. Plückthun, submitted).

F_v fragments and single-chain fragments. F_v fragments are currently the smallest conceivable antibody fragments still containing the complete antigen binding site. F_v fragments are difficult and often impossible to prepare by proteolysis, and therefore, only few reports were available about their properties (reviewed in ref. 57) before they became available through gene technology[14]. There was one report about a V_H/V_LC_L hybrid apparently not showing full functionality[58]. However, a detailed analysis of the F_v fragment of McPC603 has shown[14,53] that it does indeed have the same binding properties as the F_{ab} fragment or the whole antibody. Complications in binding measurements may arise, however, since F_v fragments have tendency to dissociate into V_H and V_L upon dilution[53]. The exact equilibrium constants for domain dissociation will vary from antibody to antibody, since the hypervariable loops contribute to this interaction. For McPC603, the dissociation constant was found to be about 10^{-6} M (ref. 53). This dissociation does not interfere with all applications, and the relatively small size of the molecule makes it an attractive target for structural studies by X-ray crystallography[9] and NMR. An important application for F_v fragments or covalently linked F_v fragments (see below) might be in tumor diagnostics and therapy. The small size may allow the penetration of dense tumor tissue and provide for low antigenicity, but tumor localization and *in vivo* pharmacokinetic properties of these fragments still need to be carefully analyzed.

Several strategies have now been developed (Fig. 3) to covalently link the two chains of the F_v fragment to overcome the dissociation problem. The first approach was to chemically crosslink the two chains with glutaraldehyde. In the case of McPC603, this is accomplished without any loss of binding affinity. These crosslinked species have been used in mutagenesis studies to determine the influence of different binding interactions[6]. The second strategy was to design an intermolecular disulfide bond (Fig. 4). It could be shown that the additional disulfide bridge indeed forms spontaneously in the periplasm, and the molecule could thus be obtained in fully functional form from *E. coli*[53]. Antigen binding affinity was only marginally affected in the case of McPC603.

The third strategy was to connect the two domains by a peptide linker (Figs. 3,4). The resulting fragment has been called "single-chain F_v fragment". Such peptide-linked single-chain F_v fragments had first been obtained as insoluble inclusion bodies that had to be refolded *in vitro*[10,45–47], but the single-chain F_v fragment can also be secreted in functional form[53]. In the cases investigated, the affinities were almost identical to the complete antibody. It remains to be seen, however, whether this will turn out to be a general result. It is conceivable that the linker may obstruct or disturb certain antigen binding sites.

The linking of the domains in the secreted single-chain fragment does not increase the yield of functional fragment when compared to the secreted F_v fragment under exactly the same conditions, and it is thus unlikely that the association of V_L and V_H is a kinetic problem in *E. coli*. Rather, the yield even appears to be slightly lower than in the normal F_v fragment under identical conditions, possibly because of some proteolysis within the linker.

Both orientations of the linker (the C-terminus of V_H connected to the N-terminus of V_L or the C-terminus of V_L connected to the N-terminus of V_H) have been constructed[10,45–47,53] and both arrangements gave rise to fragments that can be refolded to give binding activity. Two different types of linker designs have been reported. In one, a flexible, hydrophilic linker was chosen[46,53],

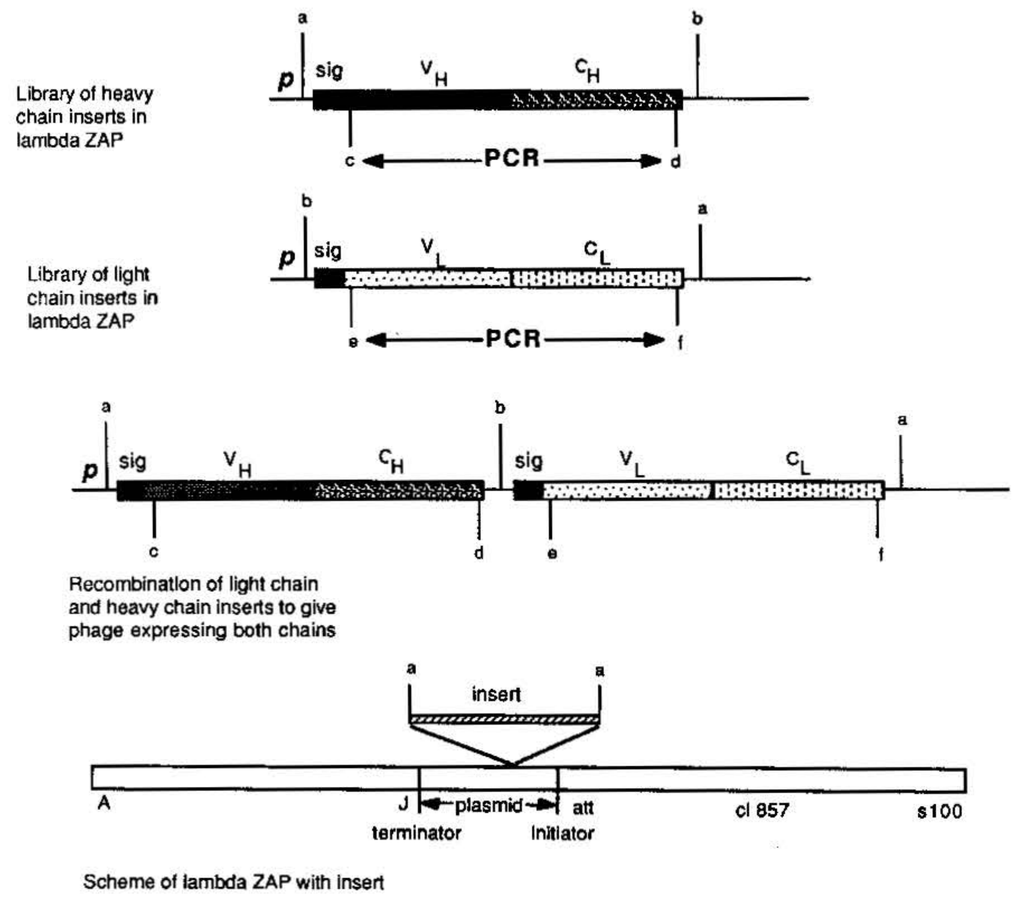

FIGURE 5 The combinatorial approach to expression of F_{ab} fragments in λ-ZAP derivatives. *p* denotes a promoter, *sig* a signal sequence, V_HC_H and V_LC_L are the genes for the two chains of the F_{ab} fragment. PCR between the two arrows denotes the region which has been obtained by PCR amplification. The small letters *a–f* denote restriction sites. At the bottom, the λ-ZAP phage is shown schematically, with some phage markers labeled underneath. (Adapted from ref. 65).

which was shown to be compatible with secretion and *in vivo* folding. In the other case[45], a fragment was taken from the protein structure database, which was known to connect residues at a similar distance and orientation. The *in vivo* folding in *E. coli* of the latter single-chain fragment has not been reported.

Fusions of the C-terminus of the single-chain F_v fragment to a toxin[10] and a fragment of the *Staphylococcus aureus* protein A have been constructed (I. Pfitzinger and A. Plückthun, unpublished). In these hybrid proteins, both the antibody domains and the additional domains seemed to be functional. Functionality of both parts was also found when the protein A portion was fused to the N-terminus of the single-chain F_v fragment[13]. This underlines again that antibody domains often fold independently and that the normally present C-terminal constant domains may be removed or replaced by a variety of different domains without endangering the binding activity. The domain to be attached to the antibody fragment must, however, be able to fold without disturbing the folding of the variable domains and still be functional when it carries an antigen binding fragment as an N-terminal extension.

Isolated domains. The physical characteristics of isolated domains will vary from antibody to antibody. In some cases, antigen binding will be observed, in others not. We have investigated the properties of isolated V_H and V_L domains from McPC603. The V_H domain is poorly soluble at temperatures above 4°C, and is unable to bind antigen (R. Glockshuber and A. Plückthun, unpublished). The use of V_H domains as a general substitute for antibody combining sites[59] therefore faces some technical challenges. It is possible that the solubility problem can be alleviated by appropriately engineering the framework, but it remains to be shown that satisfactory binding and, most importantly, the narrow specificity, which is the hallmark of antibodies, can be maintained in such modified fragments.

Isolated V_L domains have a tendency to dimerize. The precise dissociation constant varies depending on the specific antibody but usually lies between 10^{-3} M and 10^{-6} M. V_L of McPC603 does not bind the antigen, either. On the other hand, this recombinant domain was recently shown to give rise to very well ordered crystals of the

dimer[7], with a structure essentially identical to the V_L domain in the F_{ab} fragment. This offers the prospect of rapidly obtaining a structural database of characteristic complementarity determining regions by protein engineering, as many variants crystallize isomorphously (B. Steipe, R. Huber and A. Plückthun, unpublished).

Purification. In the periplasmic secretion system, considerable enrichment of the antibody protein is possible by cell fractionation, which can be carried out on a preparative scale[52]. The periplasm of *E. coli* contains only a small part of the total cell protein, and therefore subsequent purification is simplified. Alternatively, the protein can be purified from a cell homogenate. But if the protein is purified from the medium, it must be enriched from large volumes and be separated from cell debris [see e.g. ref. 50].

Antibodies and their fragments can of course be purified by classical techniques such as ion exchange chromatography, but particularly advantageous is the use of antigen affinity chromatography[60]. This technique is the same as used for the purification of whole antibodies from ascites or cell culture. In the case of anti-hapten antibodies, the fragment can be eluted with soluble hapten[14], while for protein antigens other mild displacement techniques are usually required[1,60]. One of the consequences of expression in *E. coli* of native fragments is that they are directly amenable to affinity-based purifications.

Another general strategy is the design of an affinity handle. In antibody fragments, the C-terminus is the preferred location[52]. A single-chain fragment can be purified to homogeneity in a single step using immobilized metal affinity chromatography, and only a minimal structural perturbation of 3 additional C-terminal residues is necessary. This affinity tail does not disturb the binding of the antibody fragment nor does it interfere with secretion in *E. coli*[52]. Both a C-terminal and an N-terminal fusion of protein A (ref. 13, and I. Pfitzinger and A. Plückthun, unpublished) to a single-chain F_v fragment have also been exploited for affinity-chromatography with IgG as the affinity ligand, but this constitutes a much larger structural perturbation than the metal-affinity tail.

Libraries. Two essential components were necessary to express and screen antibody libraries in *E. coli*. First, native expression and assembly *in vivo* had to be established for single genes[14–16], after which it could be extended to libraries. Second, a system for rapidly obtaining the required sequences was needed. For this, the development of the polymerase chain reaction (PCR) proved crucial. Antibodies have, even within their variable domains, framework regions that are fairly well conserved[61]. In addition, sequences in the constant regions can be used for priming. Several authors have reported such a PCR amplification of antibody variable domains and suitable primers for it[59,62–67]. Experiments with expression libraries reported so far started from immunized animals. In this case, the mRNA producing the newly elicited antibodies is very highly enriched. The library therefore does not have to be unduly large to find the desired clones.

Two strategies have been reported for screening the binding activity. One was the secretion of only V_H domains[59], whose properties have been discussed above. Such screening must be carried out with great care to insure that affinity is not due to hydrophobic binding to the domain interface in the absence of V_L. Furthermore, the approach implicitly assumes that enough binding interactions are actually located on V_H for the particular antibody desired. The *selectivity* (which is one of the essential features of the antibody) is likely to be lower for such fragments than for the whole binding pocket made up of V_H and V_L, but this has not yet been investigated in detail.

Another approach was therefore developed to avoid this problem, which is based on F_{ab} fragments. A λ-phage was designed based on "λ-ZAP"[68], built around a plasmid vector exactly analogous to the secretion vectors described initially [14–16] (Figs. 2,5). A library of such λ-phages was created from the recombination of separate libraries of V_H and V_L genes obtained by PCR. The resulting plaques contained enough F_{ab} fragment to detect binding activities by binding radioactively labeled antigen to plaque lifts on filters. For producing the fragment for further characterization, the plasmid can be excised from the phage and the production methods described above used. The PCR based approach must now be subjected to a critical analysis of the actual diversity contained in such libraries. It should be stressed that the work with libraries is not limited to the combinatorial approach. Improved plasmid vectors[52] allow random mutagenesis approaches and other strategies for generating libraries.

The final goal will be to create libraries approaching the diversity of the immune system. This will, however, probably not be possible by a simple extension of existing approaches. At the very least, all existing approaches will have to be combined. In a "naive" library, i.e. the antibody library of an animal not previously immunized with the antigen under scrutiny, the mRNAs encoding suitable molecules may be extremely rare. Therefore, libraries with very large diversity must be generated. The screening of such libraries will only be possible if it can be done in the same multistep fashion as in the immune system. Therefore, future approaches will have to aim at rebuilding analogous multistep selection processes containing such features as selection and somatic mutation, but on the basis of bacterial cells or bacteriophages.

One of the obvious applications for the bacterial expression and screening technology would be to produce human monoclonal antibodies. One must keep in mind, however, the various problems to be overcome. Not only must the desired antibody bind with high affinity, but equally important is its selectivity, i.e. its *narrow* specificity. Antibodies with broad specificities would be unlikely candidates for the clinic. Finding antibodies with a desired binding affinity in a library is therefore only the beginning of the research necessary in this area.

References

1. Weir, D. M. (Ed.) 1986. Immunochemistry, Handbook of Experimental Immunology, 4th edition, Vol. 1. Blackwell Scientific Publications, Oxford, UK.
2. Köhler, G. and Milstein, C. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature **256**:495–497.
3. Morrison, S. L. and Oi, V. T. 1989. Genetically engineered antibody molecules. Adv. Immunol **44**:65–92.
4. Brüggemann, M. and Neuberger, M. S. 1988. Novel antibodies by DNA manipulation. Prog. Allergy **45**:91–105.
5. Plückthun, A. 1990. Antibodies from *Escherichia coli*. Nature **347**:497–498.
6. Glockshuber, R., Stadlmüller, J. and Plückthun, A. 1990. Mapping and modification of an antibody hapten binding site: a site-directed mutagenesis study of McPC603. Biochemistry, *In press*.
7. Glockshuber, R., Steipe, B., Huber, R. and Plückthun, A. 1990. Crystallisation and preliminary X-ray studies of the V_L domain of the antibody McPC603 produced in *Escherichia coli*. J. Mol. Biol. **213**:613–615.
8. Boulot, G., Eiselé, J.-L., Bentley, G. A., Bhat, T. N., Ward, E. S., Winter, G. and Poljak, R. J. 1990. Crystallization and preliminary X-ray diffraction study of the bacterially expressed F_v from the monoclonal anti-lysozyme antibody D1.3 and of its complex with the antigen lysozyme. J. Mol. Biol. **213**:617–619.
9. Bhat, T. N., Bentley, G. A., Fischmann, T. O., Boulot, G. and Poljak, R. J. 1990. Small rearrangements in structures of F_v and F_{ab} fragments of antibody D1.3 on antigen binding. Nature **347**:483–485.
10. Chaudhary, V. K., Queen, C., Junghans, R. P., Waldmann, T. A., FitzGerald, D. J. and Pastan, I. 1989. A recombinant immunotoxin consisting of two antibody variable domains fused to *Pseudomonas* exotoxin. Nature **339**:394–397.
11. Casadei, J., Powell, M. J. and Kenten, J. H. 1990. Expression and secretion of aequorin as a chimeric antibody by means of a mammalian

expression vector. Proc. Natl. Acad. Sci. USA **87**:2047–2051.
12. Neuberger, M. S., Williams, G. T. and Fox, R. O. 1984. Recombinant antibodies possessing novel effector functions. Nature **312**:604–608.
13. Tai, M., Mudgett-Hunter, M., Levinson, D., Wu, G., Haber, E., Oppermann, H. and Huston, J. S. 1990. A bifunctional fusion protein containing F_c-binding fragment B of staphylococcal protein A amino terminal to antidigoxin single-chain F_v. Biochemistry **29**:8024–8030.
14. Skerra, A. and Plückthun, A. 1988. Assembly of a functional immunoglobulin F_v fragment in *Escherichia coli*. Science **240**:1038–1041.
15. Plückthun, A. and Skerra, A. 1989. Expression of functional antibody F_v and F_{ab} fragments in *Escherichia coli*. Meth. Enzymol. **178**:497–515.
16. Better, M., Chang, C. P., Robinson, R. R. and Horwitz, A. H. 1988. *Escherichia coli* secretion of an active chimeric antibody fragment. Science **240**:1041–1043.
17. Shokat, K. M. and Schultz, P. G. 1990. Catalytic antibodies. Ann. Rev. Immunol. **8**:335–363.
18. Nisonoff, A., Hopper, J. E. and Spring, S. B. 1975. The Antibody Molecule. Academic Press, New York.
19. Young, N. M., Jackson, G. E. D. and Brisson, J. R. 1990. The glycopeptides of the mouse immunoglobulin A T15. Mol. Immunol. *In press*.
20. Davis, A. C. and Shulman, M. J. 1989. IgM—Molecular requirements for its assembly and function. Immunol. Today **10**:118–128.
21. Robinson, E. A. and Appella, E. 1980. Complete amino acid sequence of a mouse immunoglobulin α chain (MOPC511). Proc. Natl. Acad. Sci. USA **77**:4909–4913.
22. Spiegelberg, H. L., Abel, C. A., Fishkin, B. G. and Grey, H. M. 1970. Localization of the carbohydrate within the variable region of light and heavy chains of human γG myeloma proteins. Biochemistry **9**:4217–4223.
23. Wallic, S. C., Kabat, E. A. and Morrison, S. L. 1988. Glycosylation of a V_H residue of a monoclonal antibody against α(1→6) dextran increases its affinity for antigen. J. Exp. Med. **168**:1099–1109.
24. Sox, H. C. Jr. and Hood, L. 1970. Attachment of carbohydrate to the variable region of myeloma immunoglobulin light chains. Proc. Natl. Acad. Sci. USA **66**:975–982.
25. Margni, R. A. and Binaghi, R. A. 1988. Nonprecipitating asymmetric antibodies. Ann. Rev. Immunol. **6**:535–554.
26. Nose, M. and Wigzell, H. 1983. Biological significance of carbohydrate chains on monoclonal antibodies. Proc. Natl. Acad. Sci. USA **80**:6632–6636.
27. Leatherbarrow, R. J., Rademacher, T. W., Dwek, R. A., Woof, J. M., Clark, A., Burton, D. R., Richardson, N. and Feinstein, A. 1985. Effector functions of a monoclonal aglycosylated mouse IgG2a: binding and activation of complement component C1 and interaction with human monocyte F_c receptor. Mol. Immunol. **22**:407–415.
28. Granato, D. A. and Neeser, J.-R. 1987. Effect of trimming inhibitors on the secretion and biological activity of a murine IgE monoclonal antibody. Mol. Immunol. **24**:849–855.
29. Wall, R. and Kuehl, M. 1983. Biosynthesis and regulation of immunoglobulins. Ann. Rev. Immunol. **1**:393–422.
30. Skerra, A., Glockshuber, R. and Plückthun, A. 1990. Structural features of the McPC603 F_{ab} fragment not defined in the crystal structure. FEBS Lett. **271**:203–206.
31. Leatherbarrow, R. J. and Dwek, R. A. 1983. The effect of aglycosylation on the binding of mouse IgG to staphylococcal protein A. FEBS Lett. **164**:227–230.
32. Walker, M. R., Lund, J., Thompson, K. M. and Jefferis, R. 1989. Aglycosylation of human IgG1 and IgG3 monoclonal antibodies can eliminate recognition by human cells expressing FcγRI and/or FcγRII receptors. Biochem. J. **259**:347–353.
33. Taylor, A. K. and Wall, R. 1988. Selective removal of α heavy-chain glycosylation sites causes immunoglobulin A degradation and reduced secretion. Mol. Cell. Biol. **8**:4197–4203.
34. Hickman, S., Kulczycki, A., Jr., Lynch, R. G. and Kornfield, S. 1977. Studies of the mechanism of tunicamycin inhibition of IgA and IgE secretion by plasma cells. J. Biol. Chem. **252**:4402–4408.
35. Vasilov, R. G. and Ploegh, H. L. 1982. Biosynthesis of murine immunoglobulin D: heterogeneity of glycosylation. Eur. J. Immunol. **12**:804–813.
36. Tartakoff, A. and Vassalli, P. 1979. Plasma cell immunoglobulin M molecules. Their biosynthesis, assembly, and intracellular transport. J. Cell. Biol. **83**:284–299.
37. Boss, M. A., Kenten, J. H., Wood, C. R. and Emtage, J. S. 1984. Assembly of functional antibodies from immunoglobulin heavy and light chains synthesised in *E. coli*. Nucl. Acids Res. **12**:3791–3806.
38. Wood, C. R., Boss, M. A., Patel, T. P. and Emtage, S. J. 1984. The influence of messenger RNA secondary structure on expression of an immunoglobulin heavy chain in *Escherichia coli*. Nucl. Acids Res. **12**:3937–3950.
39. Kenten, J., Helm, B., Ishizaka, T., Cattini, P. and Gould, H. 1984. Properties of a human immunoglobulin ε-chain fragment synthesized in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **81**:2955–2959.
40. Cabilly, S., Riggs, A. D., Pande, H., Shively, J. E., Holmes, W. E., Rey, M., Perry, L.J., Wetzel, R. and Heyneker, H. L. 1984. Generation of antibody activity from immunoglobulin polypeptide chains produced in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **81**:3273–3277.
41. Liu, F., Albrandt, K. A., Bry, C. G. and Ishizaka, T. 1984. Expression of a biologically active fragment of human IgE ε chain in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **81**:5369–5373.
42. Field, H., Rees, A. R. and Yarranton, G. T. 1988. A functional recombinant immunoglobulin variable domain from polypeptides produced in *Escherichia coli*. Vaccines **88**:29–34.
43. Field, H., Yarranton, G. T. and Rees, A. R. 1989. Expression of mouse immunoglobulin light and heavy chain variable regions in *Escherichia coli* and reconstitution of antigen-binding activity. Protein Eng. **3**:641–647.
44. Baldwin, E. and Schultz, P. G. 1989. Generation of a catalytic antibody by site-directed mutagenesis. Science **245**:1104–1107.
45. Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S., Lee, T., Pope, S. H., Riordan, G. S. and Whitlow, M. 1988. Single-chain antigen-binding proteins. Science **242**:423–426.
46. Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R. and Oppermann, H. 1988. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain F_v analogue produced in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **85**:5879–5883.
47. Condra, J. H., Sardana, V. V., Tomassini, J. E., Schlabach, A. J., Davies, M. E., Lineberger, D. W., Graham, D. J., Gotlib, L. and Colonno, R. J. 1990. Bacterial expression of antibody fragments that block human rhinovirus infection of cultured cells. J. Biol. Chem. **265**:2292–2295.
48. Sherman, F., Stewart, J. W. and Tsunasawa, S. 1985. Methionine or not methionine at the beginning of a protein. Bioessays **3**:27–31.
49. Buchner, J. and Rudolph, R. Renaturation, purification and characterization of recombinant F_{ab} fragments produced in *E. coli*. 1991. Bio/Technology **91**:157–164.
50. Holland, I. B., Kenny, B., Steipe, B. and Plückthun, A. 1990. Secretion of heterologous proteins in *Escherichia coli*. Meth. Enzymol. **182**:132–143.
51. Briggs, M. S. and Gierasch, L. M. 1986. Molecular mechanisms of protein secretion: the role of the signal sequence. Adv. Protein Chem. **38**:109–180.
52. Skerra, A., Pfitzinger, I. and Plückthun, A. 1991. The functional expression of antibody F_v fragments in *Escherichia coli*: improved vectors and a generally applicable purification strategy. Bio/Technology **9**:273–278.
53. Glockshuber, R., Malia, M., Pfitzinger, I. and Plückthun, A. 1990. A comparison of strategies to stabilize immunoglobulin F_v fragments. Biochemistry **29**:1362–1367.
54. Segal, D. M., Padlan, E. A., Cohen, G. H., Rudikoff, S., Potter, M. and Davies, D. R. 1974. The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin F_{ab} and the nature of the antigen binding site. Proc. Natl. Acad. Sci. USA **71**:4298–4302.
55. Satow, Y., Cohen, G. H., Padlan, E. A. and Davies, D. R. 1986. Phosphocholine binding immunoglobulin F_{ab} McPC603. An X-ray diffraction study at 2.7 Å. J. Mol. Biol. **190**:593–604.
56. Perlmutter, R. M., Crews, S. T., Douglas, R., Sorensen, G., Johnson, N., Nivera, N., Gearhart, P. J. and Hood, L. 1984. The generation of diversity in phosphorylcholine-binding antibodies. Adv. Immunol. **35**:1–37.
57. Givol, D. 1979. The antibody combining site, p. 71–125. *In:* Int. Rev. Biochem. Vol. 23. Defense and Recognition IIB. Structural Aspects. Lennox E. S. (Ed.). University Park Press, Baltimore, MD.
58. Sen, J. and Beychok, S. 1986. Proteolytic dissection of a hapten binding site. Proteins **1**:256–262.
59. Ward, E. S., Güssow, D., Griffiths, A. D., Jones, P. T. and Winter, G. 1989. Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. Nature **341**:544–546.
60. Dean, P. D. G., Johnson, W. S. and Middle, F. A. (Eds.). 1985. Affinity Chromatography, a Practical Approach. IRL Press Ltd., Eynsham, UK.
61. Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M. and Gottesmann, K. S. 1987. Sequences of Proteins of Immunological Interest, 4th ed. US Dept. of Health and Human Services, Public Health Service, Natl. Inst. of Health, Bethesda, MD.
62. Orlandi, R., Güssow, D. H., Jones, P. T. and Winter, G. 1989. Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. Proc. Natl. Acad. Sci. USA **86**:3833–3837.
63. Sastry, L., Alting-Mees, M., Huse, W. D., Short, J. M., Sorge, J. A., Hay, B. N., Janda, K. D., Benkovic, S. J. and Lerner, R. A. 1989. Cloning of the immunological repertoire in *Escherichia coli* for generation of monoclonal catalytic antibodies: construction of a heavy chain variable region-specific cDNA library. Proc. Natl. Acad. Sci. USA **86**:5728–5732.
64. Chaudhary, V.K., Batra, J. K., Gallo, M. G., Willingham, M. C., Fitz-Gerald, D. J. and Pastan, I. 1990. A rapid method of cloning functional variable-region antibody genes in *Escherichia coli* as single-chain immunotoxins. Proc. Natl. Acad. Sci. USA **87**:1066–1070.
65. Huse, W. D., Sastry, L., Iverson, S. A., Kang, A. S., Alting-Mees, M., Burton, D. R., Benkovic, S. J. and Lerner, R. A. 1989. Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. Science **246**:1275–1281.
66. Caton, A. J. and Koprowski, H. 1990. Influenza virus hemagglutinin-specific antibodies isolated from a combinatorial expression library are closely related to the immune response of the donor. Proc. Natl. Acad. Sci. USA **87**:6450–6454.
67. Larrick, J. W., Danielsson, L., Brenner, C. A., Wallace, E. F., Abrahamson, M., Fry, K. E. and Borrebaeck, C. A. 1989. Polymerase chain reaction using mixed primers: cloning of human monoclonal antibody variable region genes from single hybridoma cells. Bio/Technology **7**:934–938.
68. Short, J. M., Fernandez, J. M., Sorge, J. A. and Huse, W. D. 1988. λ-ZAP: a bacteriophage λ expression vector with *in vivo* excision properties. Nucl. Acid Res. **16**:7583–7600.