# Parallel Processing of Complex Biomolecular Information: Combining Experimental and Computational Approaches

Jestin Jean-Luc, Lafaye Pierre



# 1

# Parallel Processing of Complex Biomolecular Information: Combining Experimental and Computational Approaches

Jestin Jean-Luc and Lafaye Pierre
*Institut Pasteur*
*France*

## 1. Introduction

While protein functions such as binding or catalysis remain very difficult to predict computationally from primary sequences, approaches which involve the parallel processing of diverse proteins are remarkably powerful for the isolation of rare proteins with functions of interest.
Stated using a Darwinian vocabulary, a repertoire of proteins can be submitted to selection according to a function of interest for isolation of the rare fittest proteins. Parallel processing strategies rely mainly on the design of *in vitro* selections of proteins. To ensure that complex molecular information can be extracted after selection from protein populations, several types of links between the genotype and the phenotype have been designed for the parallel processing of proteins: they include the display of nascent proteins on the surface of the ribosome bound to mRNA, the display of proteins as fusions with bacteriophage coat proteins and the fusion of proteins to membrane proteins expressed on the surface of yeast cells. In the first two display strategies, covalent and non covalent bonds define chemical links between the genotype and the protein, while in the last case compartmentation by a membrane provides the link between the protein and the corresponding gene.
While parallel processing strategies allow the analysis of up to $10^{14}$ proteins, serial processing is convenient for the analysis of tens to thousands of proteins, with the exceptions of millions of proteins in the specific case where fluorescent sorting can be adapted experimentally.
In this review, the power of parallel processing strategies for the identification of proteins of interest will be underlined. It is useful to combine them with serial processing approaches such as activity screening and the computational alignment of multiple sequences. These molecular information processing (MIP) strategies yield sequence-activity relationships for proteins, whether they are binders or catalysts (Figure 1).

## 2. Parallel processing strategies

Display technologies *in vitro* are based on the same « idea »: the creation of large diverse libraries of proteins followed by their interrogation using display technologies *in vitro*. An

antibody fragment (single-chain Fv (scFv), camelids single domain antibodies (VHH) or Fab fragments) (Figure 2) or an enzyme can be presented on phage or yeast surfaces as well as on ribosomes, while the encoding nucleotide sequence is incorporated within or is physically attached.

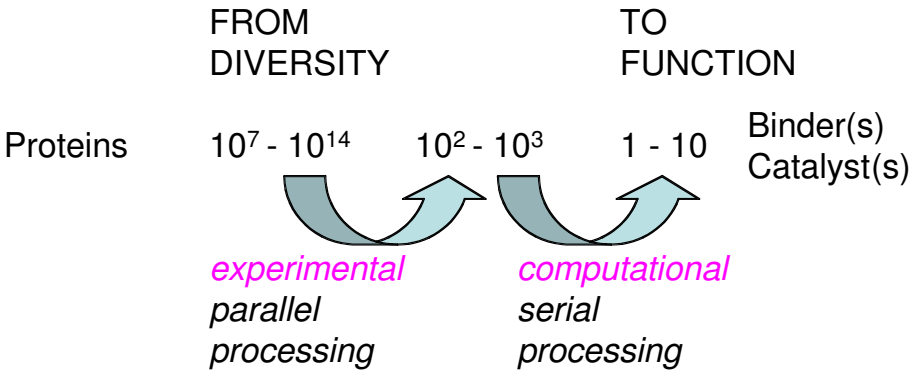



Fig. 1. Parallel and experimental processing combined with serial and computational processing prior to thermodynamic and kinetic characterization allow protein engineering towards new functions.

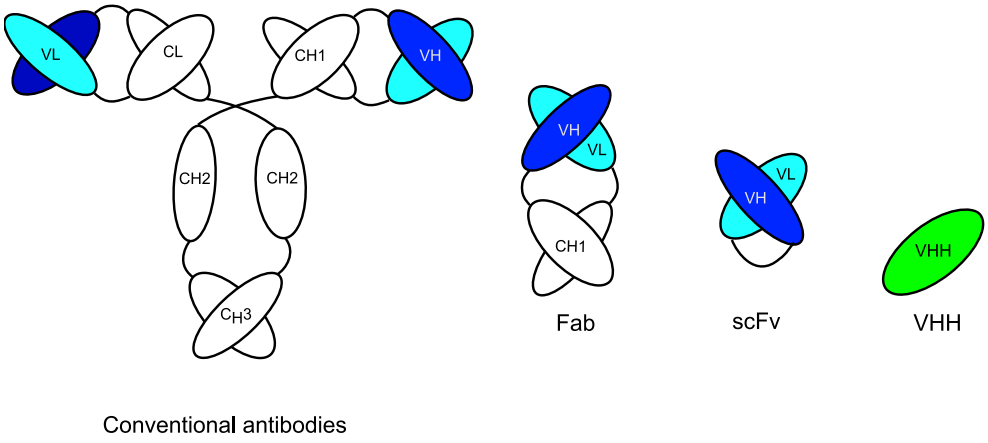



Fig. 2. Representation of mammalian antibodies and synthetic fragments: Fab, scFv and VHH.

This link of phenotype to genotype enables selection and enrichment of molecules with high specific affinities or exquisite catalytic properties together with the co-selected gene (Figure 3). Consequently, the need for serial screening is reduced to a minimum.

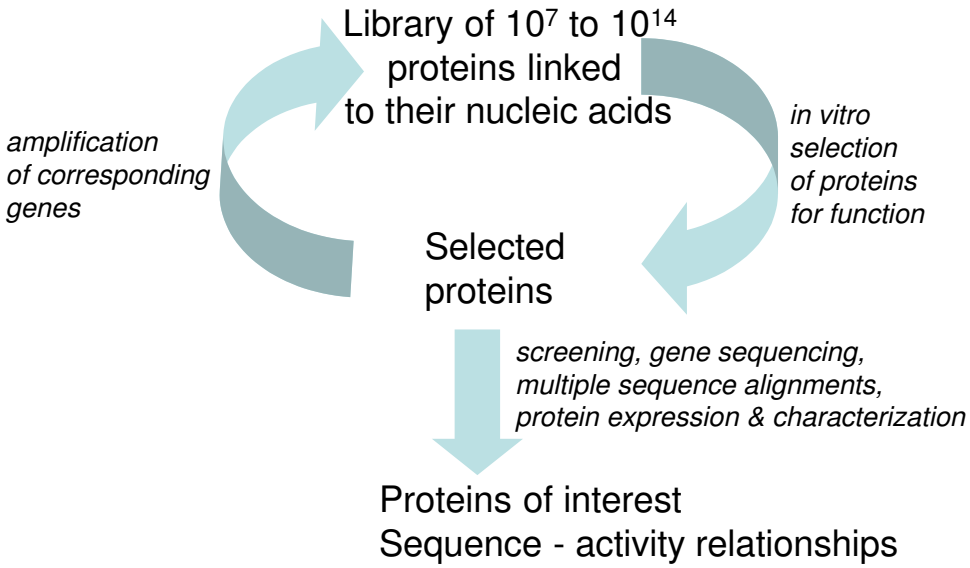



Fig. 3. Directed protein evolution cycles yield sequence-activity relationships for proteins. A cycle consists of the selection of proteins according to their function and of the amplification of their corresponding nucleic acids which are linked to the proteins. Iteration of the cycles diminishes the background of the selection and yields a selected population enriched in proteins with functions of interest. Characterization of these selected proteins and their genes establishes sequence-activity relationships.

### 2.1 Phage display

In 1985, M13 phage displaying a specific peptide antigen on its surface was isolated from a population of wild type phage, based on the affinity of a specific antibody for the peptide (Smith, 1985). Antibody variable domain were successfully displayed by McCafferty et al in 1990, enabling the selection of antibodies themselves (McCafferty et al., 1990) (Figure 4).

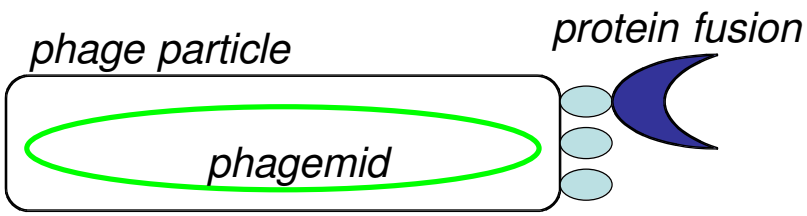



Fig. 4. Bacteriophage particle highlighting the link between a protein fused to a phage coat protein and its corresponding gene located on the phagemid. In the case of *Inovirus*, the filamentous phage particle is a cylinder with a diameter of three to five nanometers, which is about one micrometer long.

Phage display technology (Figure 4) enables the selection from repertoires of antibody fragments (scFv, Fab, VHH) displayed on the surface of filamentous bacteriophage (Smith, 1985). VHH domains are displayed by fusion to the viral coat protein, allowing phage with antigen binding activities (and encoding the antibody fragments) to be selected by panning on antigen. The selected phage can be grown after each round of panning and selected again, and rare phage ($< 1/10^6$) isolated over several rounds of panning.
The antibody fragments genes population is first isolated from lymphocytes then converted to phage-display format using PCR. The PCR products are digested and ligated into phage vector. Subsequent transformation usually yield libraries of $10^6$ to $10^{11}$ clones, each clone corresponding to a specific antibody fragments (VHH, scFv, Fab). This library is panned against the antigen then expression of selected clones is performed. Their biochemical characteristics are analyzed (purity, affinity, specificity) as well as their biological characteristics.
The major advantages of phage display compared with other display technologies are its robustness, simplicity, and the stability of the phage particles, which enables selection on cell surfaces (Ahmadvand et al., 2009), tissue sections (Tordsson et al., 1997) and even *in vivo* (Pasqualini & Ruoslahti, 1996). However, because the coupling of genotype and phenotype (i.e. protein synthesis and assembly of phage particles) takes place in bacteria, the encoding DNA needs to be imported artificially. Library size is therefore restricted by transformation efficiency. Despite great improvements in this area, the largest reported libraries still comprise no more than $10^{10}$ to $10^{11}$ different members. Moreover, the amplification of selected variants *in vivo* can lead to considerable biases. Antibody fragments that are toxic for the host, poorly expressed or folded, inefficiently incorporated into the phage particle or susceptible to proteolysis or aggregation slow down the bacterial growth and display less efficiently. This reduces the library diversity and enables a low potency but fast growing clone to dominate a whole population after just a few rounds of selection.

### 2.2 Ribosome display

Ribosome display was first developed by Dower *et al* (Mattheakis et al., 1994) where mRNA, ribosome and correctly folded functional peptide in a linked assembly could be used for screening and selection (Figure 5).

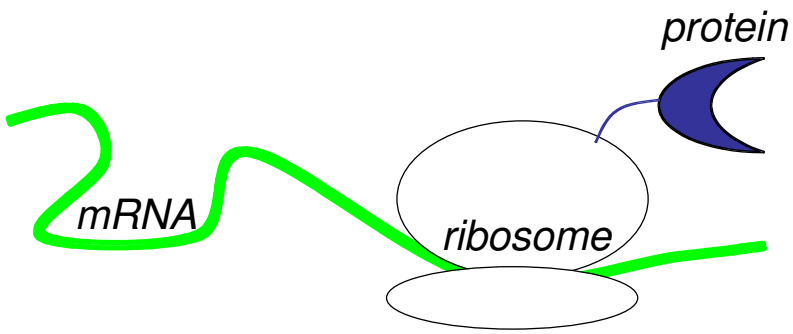



Fig. 5. Ribosome providing the link between the nascent protein and the corresponding mRNA. Such complexes are stabilized in the absence of stop codons and at low temperatures.

In ribosome display, a DNA library coding for particular proteins, for instance scFv or VHH fragments of antibodies, is transcribed *in vitro*. The mRNA is purified and used for *in vitro* translation. Because the mRNA lacks a stop codon, the ribosome stalls at the end of the mRNA, giving rise to a ternary complex of mRNA, ribosome and functional protein. A library of these ternary complexes is tested against the potential ligand (in the case of the antibodies, against the antigen). The binding of the ternary complexes (ribosome, mRNA and protein) to the ligand allows the recovery of the encoding mRNA that is linked to it and that can be transcribed into cDNA by reverse transcriptase-PCR (RT-PCR). Cycles of selection and recovery can be iterated both to enrich rare ligand-binding molecules, and to select molecules with the best affinity.
Ribosome display has been used for the selection of proteins, such as scFv antibody fragments and alternative binding scaffolds with specificity and affinity to peptides (Hanes & Pluckthun, 1997), proteins (Hanes et al., 2000; Knappik et al., 2000; Binz et al., 2004; Lee et al., 2004; Mouratou et al., 2007) and nucleic acids (Schaffitzel et al., 2001). Using transition-state analogs or enzyme inhibitors that bind reversibly to their enzyme (suicide substrates), ribosome display can also be used for the selection for enzymatic activity.
As it is entirely performed *in vitro*, there are two main advantages over other selection strategies. First, the diversity is not limited by the transformation efficiency of bacterial cells, but only by the number of ribosomes and different mRNA molecules present in the test tube. According to the fact that the functional diversity is given by the number of ribosomal complexes that display a functional protein, this number is limited by the number of functional ribosomes or different mRNA molecules, whichever is smaller. An estimate representing a lower limit, of the number of active complexes with folded protein was determined as $2.6 \times 10^{11}$ per milliter of reaction (Zahnd et al., 2007) and probably is about $10^{13}$. Second, random mutations can be introduced easily after each selection rounds, as no library must be transformed after any diversification steps. This allows facile directed evolution of binding proteins over several generations.
However, ribosome display suffers some drawbacks because RNA is extremely labile to ubiquitous Rnases, because the ternary RNA-ribosome-protein complex is very sensitive to heat denaturation and to salt concentration and because large proteins such as DNA polymerases cannot necessarily be produced by *in vitro* translation.

### 2.3 Yeast surface display

Yeast surface display (YSD) was first demonstrated as a method to immobilize enzymes and pathogen-derived proteins for vaccine development. The β-galactosidase gene from *Cyamopsis tetragonoloba* was fused to the C terminal half of α-agglutinin, a cell wall anchored mating protein in *S. cerevisiae* (Schreuder et al., 1993).
Increased stability was seen for the enzyme when linked to the cell wall, compared with direct secretion of the full β-galactosidase enzyme into the media. Early work also used the flocculin Flo1p as an anchor to attach β-galactosidase to the cell wall, with similar results (Schreuder et al., 1996). Both α-agglutinin and flocculin, along with cell wall proteins such as Cwp1p, Cwp2p, Tip1p, and others, belong to the glycosylphosphatidylinositol (GPI) family of cell wall proteins that can be used directly for display (Kondo & Ueda, 2004). These proteins are directed to the plasma membrane via GPI anchors and subsequently are linked directly to the cell wall through a β-1,6-glucan bridge for incorporation into the mannoprotein layer (Kondo & Ueda, 2004). These large intact proteins as well as their C-

terminal fragments have been demonstrated to mediate display of a range of heterologous proteins upon protein fusion.
The α-agglutinin system developped by Wittrup *et al* (Boder & Wittrup, 1997; Boder et al., 2000; Boder & Wittrup, 2000) uses Aga2p as the display fusion partner. A disulfide linkage between Aga1p, a GPI/β-1,6-glucan-anchored protein, and Aga2p anchors the protein to the cell wall. Thus, coexpression of Aga1p with an Aga2p fusion leads to cell wall-anchored protein on the surface of yeast via disulfide bonding. The majority of applications of YSD utilize now the Aga2p anchor system.
In the yeast surface display system (Figure 6), the antibody fragment (scFv for example) is fused to the adhesion subunit of the yeast agglutinin protein Aga2p, which attaches to the yeast cell wall through disulfide bonds to Aga1p. Each yeast cell typically displays $1.10^4$ to $1.10^5$ copies of the scFv, and variations in surface expression can be measured through immuno-fluorescence labeling of either the hemagglutinin or c-Myc epitope tag flanking the scFv.

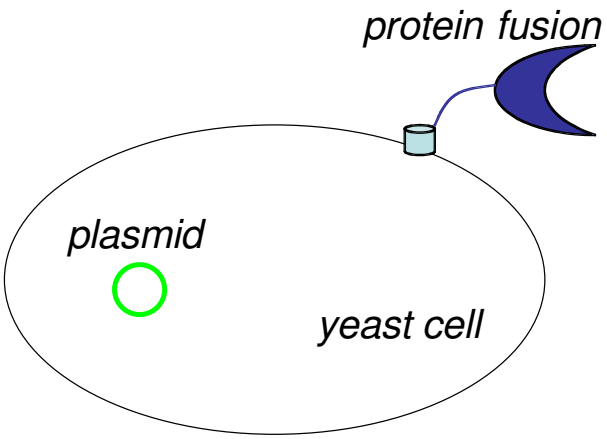



Fig. 6. Yeast cell providing the link between a protein of interest fused to a membrane protein and its corresponding gene located on a plasmid.

In the yeast surface display system (Figure 6), the antibody fragment (scFv for example) is fused to the adhesion subunit of the yeast agglutinin protein Aga2p, which attaches to the yeast cell wall through disulfide bonds to Aga1p. Each yeast cell typically displays $1.10^4$ to $1.10^5$ copies of the scFv, and variations in surface expression can be measured through immuno-fluorescence labeling of either the hemagglutinin or c-Myc epitope tag flanking the scFv. Likewise, binding to a soluble antigen of interest can be determined by labeling of yeast with biotinylated antigen and a secondary reagent such as streptavidin conjugated to a fluorophore. The display of scFv antibody fragments on the surface of yeast is a powerful and robust technique for the selection of affinity reagents (van den Beucken et al., 2003). Using yeast display for probing immune libraries offers one major advantages over alternative systems. The main advantage is that the scFv displaying yeast can be isolated by FACS and characterized by flow cytometry. The use of FACS in the selection procedure allows the visualization of antigen binding in real-time and the enrichment of each step in the selection can be easily quantified using statistical analyses. Modern flow cytometers can

easily screen millions of binding events in only a few minutes, and the precision of sorting antigen-binding yeast while eliminating nonspecific interactions facilitates large enrichments in a relatively short period of time. In addition, following selection of scFv clones, YSD allows the determination of steady-state kinetic parameters by flow cytometry ($K_D$ value determination (VanAntwerp & Wittrup, 2000)).
However, current yeast display technology is limited by the size of libraries that can be generated and, typically, only libraries of between $10^6$ and $10^7$ mutants are routinely possible using conventional *in vitro* cloning and transformation.

## 3. Binders analyzed by parallel processing

This chapter will focus on antibodies, the major class of known binding proteins.

### 3.1 Introduction on the natural diversity of immunoglobulins

One characteristic of the immune response in vertebrate is the possibility to raise immunoglobulin (Ig) against any type of antigen (Ag), known or unknown. An Ig contains two regions: the Variable domain involved in the binding with the Ag and the Constant domain with effector functions. Each Ig is unique and the variable domain, which is present in each heavy and light chain of every antibody, differ from one antibody to an other. Differences between the variable domains are located on three loops known as complementarity determining regions CDR1, CDR2 and CDR3. CDRs are supported within the variable domains by conserved framework regions. The variability of Ig is based on two phenomena: somatic recombination and somatic hypermutation (SHM).
Somatic recombination of Ig, also known as V(D)J recombination, involves the generation of a unique Ig variable region. The variable region of each immunoglobulin heavy or light chain is encoded in several gene segments. These segments are called variable (V), diversity (D) and joining (J) segments. V, D and J segments are found in Ig heavy chains, but only V and J segments are found in Ig light chains. The IgH locus contains up to 65 VH genes, 27 D genes and 6 J genes while the IgL locus contains 40 V genes and 4-5 J genes, knowing that there are two light chains kappa and lambda. In the bone marrow, each developing B cell will assemble an immunoglobulin variable region by randomly selecting and combining one V, one D and one J gene segment (or one V and one J segment in the light chain). For heavy-chains there are about 10530 potential recombinations (65x27x6) and for light chains 360 potential recombinations (200+160). Moreover some mutations (referred as N-diversity somatic mutations) occur during recombination increasing the diversity by a factor $10^3$. These two phenomena, recombination and somatic mutations, lead to about $10^6$-$10^7$ possibilities for heavy chains and 3.5 $10^5$ possibilities for light chains generating the formation of about $2.10^{12}$ different antibodies and thus different antigen specificities (Figure 7) (Jones & Gellert, 2004).
Following activation with antigen, B cells begin to proliferate rapidly. In these rapidly dividing cells, the genes encoding the variable domains of the heavy and light chains undergo a high rate of point mutation, by a process called somatic hypermutation (SHM). The SHM mutation rate is about $10^{-3}$ per base pair and per cell division, that is approximately one million times above the replicative mutation rate. As a consequence, any daughter B cells will acquire slight amino acid differences in the variable domains of their antibody chains.

This serves to increase the diversity of the antibody pool and impacts the antibody's antigen-binding affinity. Some point mutations will result in the production of antibodies that have a weaker interaction (low affinity) with their antigen than the original antibody, and some mutations will generate antibodies with a stronger interaction (high affinity). It has been estimated that the affinity of an immunoglobulin for an antigen is raised by a factor 10 to 100 (Kepler & Bartl, 1998). B cells that express high affinity antibodies on their surface will receive a strong survival signal during interactions with other cells, whereas those with low affinity antibodies will not, and will die by apoptosis. The process of generating antibodies with increased binding affinities is called affinity maturation (Neuberger, 2008).

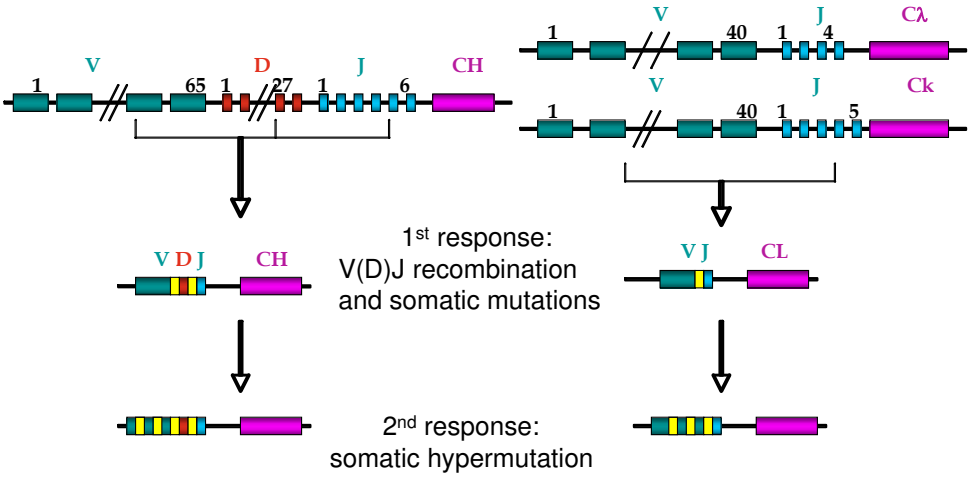



Fig. 7. Recombination and hypermutation of immunoglobulins. A yellow rectangle represents a point mutation. Recombination and somatic hypermutation are shown for heavy chains (left) and for light chains (right).

This quite complex process for generation of highly specific antibodies is a challenge for the obtention of recombinant antibodies. Many factors influence the quality of the recombinant antibodies: starting or not from an immunized animals or humans, the size and the quality of the libraries, the possibility to mutate the antibodies.

### 3.2 Antibody libraries

#### 3.2.1 Recombinant antibody libraries

Recombinant antibody libraries have been constructed by cloning antibody heavy- or light-chain variable genes directly from lymphocytes of animals or human and then expressing as a single-chain fragment variable (scFv) single-domain antibodies (VHH) or as an antigen-binding fragment (Fab) using various display technologies. The recombinant antibody technology, an alternative to traditional immunization of animals, facilitates to isolate target specific high affinity monoclonal antibodies without immunization by virtue of combination with high throughput screening techniques.

A strategy for creation of a combinatorial antibody library is very important to isolate high specificity and affinity antibodies against target antigens. To date, a variety of different antibody libraries have been generated, which range from immune to naive and even synthetic antibody libraries (Table 1). Immune libraries derived from IgG genes of immunized donors (Sanna et al., 1995) are useful if immunized patients are available but have the disadvantage that antibodies can only be made against the antigens used for immunization. In contrast, antibodies against virtually any antigen, including self-, non-immunogenic, or toxic antigens, can be isolated from naive or synthetic libraries. Naive libraries from non-immunized donors have been generated by PCR-cloning Ig repertoires from various B-cell sources (Marks et al., 1991; Vaughan et al., 1996; Sheets et al., 1998; de Haard et al., 1999)) derived from human or camel germ line genes and randomized only in the CDR3 regions (Hoogenboom & Winter, 1992; Nissim et al., 1994; de Kruif et al., 1995). Synthetic libraries have been generated from a repertoire of 49 human germline VH genes segments rearranged *in vitro* to create a synthetic CDR3 region (Hoogenboom & Winter, 1992) or derived from a single V-gene with complete randomization of all CDRs (Jirholt et al., 1998; Soderlind et al., 2000) (Table 1).

| | **Synthetic** | **Naive** | **Immune** |
|---|---|---|---|
| **V-gene source** | Unrearranged V-gene segments | Rearranged-V genes from Ig pool | Rearranged V-genes from specific IgG pool |
| **Contents** | controlled | uncontrolled | uncontrolled |
| **Repertoire construction** | Once (single pot) | | New repertoire for every antigen |
| **Affinity of antibodies** | Depending on library size : μM from standard size repertoire ($10^7$) nM from very large repertoire ($10^{10}$) | | Biased for high affinity (nM if antigen is immunogenic) |
| **specificity** | Any | Originally biased against self | Immunodominant epitopes, biased against self |

Table 1. Comparison between Synthetic, Naive and Immune libraries (according to (Hoogenboom, 1997))

### 3.2.2 Immune libraries

Efficient isolation of specific high affinity binders from relatively small sized libraries was shown using immune antibody libraries constructed from B lymphocytes of immunized mice, camels or patients with a high antibody titer for particular antigens, in our laboratory and by others: a targeted immune library contained typically about $10^6$ clones (Burton, 1991; Barbas et al., 1992; Barbas et al., 1992) (Table 1). However, the construction of an immune library is not always possible due to the difficulty in obtaining antigen-related B lymphocytes.

The quality of the immune response will likely dictate the outcome of the library selections. It is generally accepted that early in the immune response the repertoire of immunoglobulins is diverse and of low affinity to the antigen. The process of SHM through successive rounds of selection ensures that the surviving B cells develop progressively higher affinities, but probably at the expense of diversity. The balance between diversity and

affinity is something that may be exploited by researchers depending on the goal of their study.

### 3.2.3 Non-immune libraries

The most important parameter in the non-immune antibody library is the library diversity, i.e., library size, in an aspect that, in general, the larger the library, the higher the likelihood is to isolate high affinity binders to a particular antigen. Typically, a $10^9$ to $10^{11}$ library diversity has been reported to generate specific high affinity binders with dissociation constants in the 1–1000 nM range (Table 1). For example, scFvs against crotoxin, a highly toxic-β-neurotoxin isolated from the venom of the rattlesnake, *Crotalus durrissus terrificus*, have been selected from two non-immune scFv libraries which differ by their size; respectively $10^6$ (Nissim et al., 1994) and $10^{10}$ diversity (Vaughan et al., 1996). The affinity of anti-crotoxin scFvs is in the micromolar range in the first case and in the nanomolar range in the second case. Moreover, these latter scFvs possessed an *in vivo* neutralizing activity against a venom toxin.

However, creating a large antibody library is time consuming and does not always guarantee to isolate high affinity binders to any given antigen. Therefore, many attempts have been undertaken to make the library size as big as possible, and site-specific recombination systems have been created to overcome the library size limitations given by the conventional cloning strategies. Besides library generation, the panning process itself limits also the library size that can be handled conveniently.

Therefore, it is important to generate libraries with a high quality of displayed antibodies, thus emphasizing the functional library size and not only the apparent library size. For instance, one limitation of phage display is that it requires prokaryotic expression of antibody fragments. It is well known that there is an unpredictable expression bias against some eukaryotic proteins expressed from *Escherichia coli* because the organism lacks foldases and chaperones present in the endoplasmic reticulum of eukaryotic cells that are necessary for efficient folding of secreted proteins such as antibody fragments. Even minor sequence changes such as single point mutations in the complementarity determining regions (CDRs) of Fab fragments can completely eliminate antibody expression in *E. coli* (Ulrich et al., 1995), and a random sampling of a scFv library showed that half of the library had no detectable level of scFv in the culture supernatant (Vaughan et al., 1996). Because the protein folding and secretory pathways of yeast more closely approximate those of mammalian cells, it has been shown that yeast display could provide access to more antibodies than phage display (Bowley et al., 2007). In this study, the two approaches were directly compared using the same HIV-1 immune scFv cDNA library expressed in phage and yeast display vectors and using the same selecting antigen (HIV-1 gp120). After 3 to 4 rounds of selection, sequence analysis of individual clones revealed many common antibodies isolated by both techniques, but also revealed many novel antibodies derived from the yeast display selection that had not previously been described by phage display. It appears that the level of expression of correctly folded scFv on the phage surface is one of the most important criteria for selection.

VHH libraries may be an advantageous alternative because VHH are highly soluble, stable, easily expressed in *E. coli* and because they do not tend to aggregate (Muyldermans, 2001; Harmsen & de Haard, 2007). Moreover due to their small size (15 kDa compared to 25-30 kDa for a scFv and 50 kDa for a Fab), VHH could diffuse easily in tissues, bind to poorly accessible epitopes for conventional antibody fragments (Desmyter et al., 1996; Stijlemans et

al., 2004) and bind non-conventional epitopes (Behar et al., 2009). Chen *et al* (Chen et al., 2008) have prepared a phage displayed VH-only domain library by grafting naturally occurring CDR2 and CDR3 of heavy chains on a VHH-like scaffold. From this library (size 2.5 $10^{10}$) they have selected high quality binders against viral and cancer-related antigens. From a non-immune VHH library of $10^8$ diversity, VHH have been selected against various viral protein by phage display. These VHH had an affinity in the nanomolar range but more interestingly the $k_{off}$ is very low (about $10^4$ to $10^5$ $s^{-1}$) allowing them to be suitable for crystallographic studies (Lafaye –personnal communication).

### 3.3 Affinity optimization

With tools such as phage, yeast and ribosome display available to isolate rapidly specific high-potency antibodies from large variant protein populations, a major key to efficient and successful *in vitro* antibody optimization is the introduction of the appropriate sequence diversity into the starting antibody. Generally, two approaches can be taken: either amino acid residues in the antibody sequence are substituted in a targeted way or mutations are generated randomly.

#### 3.3.1 Affinity increase by targeted mutations

Antibodies are ideal candidates for targeted sequence diversification because they share a high degree of sequence similarity and their conserved immunoglobulin protein fold is well studied. Many *in vitro* affinity maturation efforts using combinatorial libraries in conjunction with display technologies have targeted the CDRs harbouring the antigen-binding site. Normally, amino acid residues are fully randomized with degenerate oligonucleotides. If applied to all positions in a given CDR, however, this approach would create far more variants than can be displayed on phage, on yeast or even ribosomes – saturation mutagenesis of a CDR of 12 residues, for example, would result in $20^{12}$ different variants. In addition, the indiscriminate mutation of these residues creates many variants that no longer bind the antigen, reducing the functional library size. Scientists have therefore restricted the number of mutations by targeting only blocks of around six consecutive residues per library (Thom et al., 2006) or by mutating four variants in all the CDRs (Laffly et al., 2008) or by mutating only the CDRs 1 and 2 (Hoet et al., 2005). Mutagenesis has also been focussed on natural hotspots of SHM (Ho et al., 2005). In other works, the residues to be targeted were chosen based on mutational or structural analyses as well as on molecular models (Yelton et al., 1995; Osbourn et al., 1996; Chen & Stollar, 1999). Further affinity improvements have been achieved by recombining mutations within the same or different CDRs of improved variants (Jackson et al., 1995; Yelton et al., 1995; Chen & Stollar, 1999; Rajpal et al., 2005). Despite some substantial gains, such an approach is unpredictable. As an alternative, CDRs were sequentially mutated by iterative constructions and pannings of libraries, starting with CDR3, in a strategy named « CDR walking » (Yang et al., 1995; Schier et al., 1996). Although this results in greater improvements, it is time consuming and permits only one set of amino acid changes to recombine with new mutations.

#### 3.3.2 Affinity increase by random mutations

In addition to the targeted strategies, several random mutagenesis methods can be used to improve antibody potency. One is the shuffling of gene segments, where VH and VL populations, for example, can be randomly recombined with each other (Figini et al., 1994;

Schier et al., 1996) or be performed with CDRs (Jirholt et al., 1998; Knappik et al., 2000). An alternative approach is the possibility that independent repertoires of heavy chain (HC) and light chain (LC) can be constructed in haploid yeast strains of opposite mating type. These separate repertoires can then be combined by highly efficient yeast mating. Using this approach, Blaise *et al* (Blaise et al., 2004) have rapidly generated a human Fab yeast display library of over $10^9$ clones, allowing the selection of high affinity Fab by YSD using a repeating process of mating- driven chain shuffling and flow cytometric sorting.
Another approach is the indiscriminate mutation of nucleotides using the low-fidelity Taq DNA polymerase (Hanes et al., 2000), error-prone PCR (Hawkins et al., 1992; Daugherty et al., 2000; Jermutus et al., 2001; van den Beucken et al., 2003), the error-prone Qbeta RNA replicase (Irving et al., 2001) or *E. coli* mutator strains (Irving et al., 1996; Low et al., 1996; Coia et al., 2001) before and in-between rounds of selection. Shuffling and random point mutagenesis are particularly useful when used in conjunction with targeted approaches because they enable the simultaneous evolution of non-targeted regions (Thom et al., 2006); in addition, they are powerful when performed together because individual point mutations can recombine and cooperate, again leading to synergistic potency improvements. This has created some of the highest affinity antibodies produced so far, with dissociation constants in the low picomolar range (Zahnd et al., 2004) and in a study using yeast display, even in the femtomolar range (Boder et al., 2000). When performed separately, random mutagenesis can help identify mutation hotspots, defined as amino acid residues mutated frequently in a population. To this end, a variant library generated by error-prone PCR, for example, might be subjected to affinity selections followed by the sequencing of improved scFvs. In a manner similar to somatic hypermutation, this method leads to the accumulation of mutations responsible for potency gains mainly in CDRs, despite having been introduced randomly throughout the whole scFv coding sequence (Thom et al., 2006).

### 3.3.3 Affinity increase by selection optimization

Mutant libraries are often screened under conditions where the binding interaction has reached equilibrium with a limiting concentration of soluble antigen to select mutants having higher affinity. When labelled with biotin, for example, the antigen and the bound scFv–phage, scFv-yeast or scFv–ribosome–mRNA complexes can be pulled down with streptavidin-coated magnetic beads. The antigen concentration chosen should be below the $K_D$ of the antibody at the first round of selection and then reduced incrementally during subsequent cycles to enrich for variants with lower $K_D$ (Hawkins et al., 1992; Schier et al., 1996). Selections have been performed in the presence of an excess of competitor antigen or antibody, resulting specifically in variants with lower off-rates (Hawkins et al., 1992; Jermutus et al., 2001; Zahnd et al., 2004; Laffly et al., 2008).
Protein affinity maturation has been one of the most successful applications of YSD. Initial studies led by Wittrup *et al.* used an anti-fluorescein scFv to show the effectiveness of YSD in protein affinity maturation (Boder et al., 2000; Feldhaus & Siegel, 2004). Since each yeast cell is capable of displaying $10^4$ to $10^5$ copies of a single scFv (Boder & Wittrup, 1997), fluorescence from each cell can be readily detected and accurately quantified by flow cytometry. This feature of YSD allows not only precise and highly reproducible affinity measurement, but also rapid enrichment of high-affinity populations within mutant libraries (Boder et al., 2000). Moreover, on-rate selections have been realized only with yeast display, which profits from using flow cytometric cell sorting to finely discriminate variants with specified binding kinetics (Razai et al., 2005).

The selected antibodies can be tested for increased affinity but should preferentially be screened for improved potency in a relevant cell-based assay because the sequence diversification and selection process might also have enriched variants with increased folding efficiency and thermodynamic stability, both contributing to potency and, ultimately, efficacy.

### 3.4 Conclusions on the parallel processing of binders

Phage, yeast and ribosome display were proven to be powerful methods for screening libraries of antibodies. By means of selection from large antibody repertoires, a wide variety of antibodies have been generated in the form of scFv, VHH or Fab fragments. After a few rounds of panning or selection on soluble antigens and subsequent amplification in *E. coli*, large numbers of selected clones have to be analyzed with respect to antigen specificity, and binding affinity. Analysis of these selected binders is usually performed by ELISA. Hopefully, the introduction of automated screening methods to the display process provides the opportunity to evaluate hundreds of antibodies in downstream assays. Secondary assays should minimally provide a relative affinity ranking and, if possible, reliable estimates of kinetic or equilibrium affinity constants for each of the hits identified in the primary screen.

Surface plasmon resonance (SPR) methods has been used to measure the thermodynamic and kinetic parameters of antigen-antibody interactions. An SPR secondary screening assay must be capable of rapidly analyzing all the unique antibodies discovered in the primary screen. The first generations of widely used commercial systems from Biacore process only one sample at a time and this limits the throughput for antibody fragments screening to approximately 100 samples per day. Recently however, several biosensors were introduced to increase the number of samples processed with different approaches for sample delivery (Wassaf et al., 2006) (Safsten et al., 2006; Nahshol et al., 2008).

To reduce the number of antibodies tested and so far the amount of antigen used, it is crucial to analyze the diversity of the antibody fragments after the first screening performed by ELISA. Usually after few rounds of selection, a limited number of clones, found in several copies, are obtained. In that case, it is un-necessary to analyze such redundant clones. It is the reason why we have decided in our laboratory to sequence the clones after the first screening, then to analyze only the unique clones by SPR in a secondary screening.

Despite the growing knowledge around antibody structures and protein–protein interactions, and the rapid development of *in silico* evolution, molecular modelling and protein–protein docking tools, it is still nearly impossible to predict the multitude of mutations resulting in improved antibody potency. Moreover, specific structural information – on the antibody to be optimized (paratope), its antigen (epitope) and their interaction – can lack the high resolution required to determine accurately important details such as side-chain conformations, hydrogen-bonding patterns and the position of water molecules. Therefore, the most effective way to improve antibody potencies remains the use of display technologies to interrogate large variant populations, using either targeted or random mutagenesis strategies.

## 4. Catalysts analyzed by parallel processing

### 4.1 Enzyme libraries

To isolate rare catalysts of interest for specific chemical reactions, the parallel processing of millions of mutant enzymes turned out to be a successful strategy (Figures 3&8). Various

types of protein libraries can be constructed. Almost random protein sequences have been designed and submitted to selection for the isolation of nucleic acid ligases (Seelig & Szostak, 2007). Given that most enzymes have more than 50 amino acids, and that each amino acid can be one out of twenty in the standard genetic code, $20^{50}$ distinct sequences can be considered. The parallel or serial processing of so many proteins cannot be conceived experimentally. A useful strategy then relies on the directed evolution of known enzymes, which catalyze chemical reactions that are similar to the reactions of interest (Figure 8).
Enzyme libraries have been constructed by random mutagenesis of the corresponding genes. This can be achieved by introduction of manganese ions within PCR mixtures during amplification of the gene encoding the enzyme. Manganese ions alter the fidelity of the DNA-dependent DNA polymerase used for amplification and provided their concentration is precisely adjusted, the average number of base substitutions per gene can be accurately evaluated (Cadwell & Joyce, 1994). Concentrations of deoxynucleotides triphosphates can be further adapted so as to define the relative rates of different base substitutions (Fromant et al., 1995).

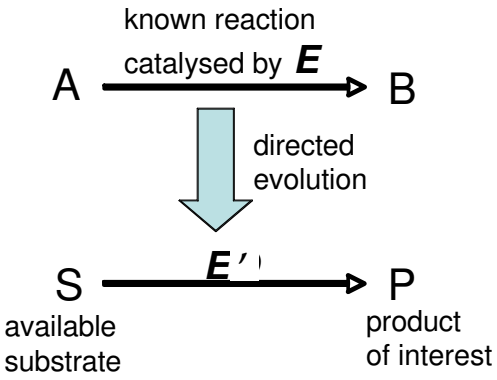



Fig. 8. Strategy for the isolation of a catalyst of interest E′ for the reaction from S to P by directed evolution of an enzyme E catalyzing a similar chemical reaction converting A into B.

Other enzyme libraries have been constructed by directed mutagenesis at specific sites within proteins, for example in our laboratory. Known x-ray crystal structures of enzymes in complex with their substrates can be used as a basis to identify the specific amino acids known to bind the substrates at the active site (Figure 9).
Oligonucleotides can be further synthesized with random mutations introduced specifically at the very few codons coding amino acids known to interact with the substrates. PCR assembly of such oligonucleotides can then be used to reconstitute full-length open reading frames coding for mutant proteins. Experience from our laboratory indicates that protein libraries designed by introduction of quasi-random mutations over an entire protein domain yield a higher number of catalysts of interest than protein libraries carefully designed by introduction of mutations at specific sites within the active site. This strategy requires nevertheless an efficient parallel processing strategy for analysis of millions of protein mutants.

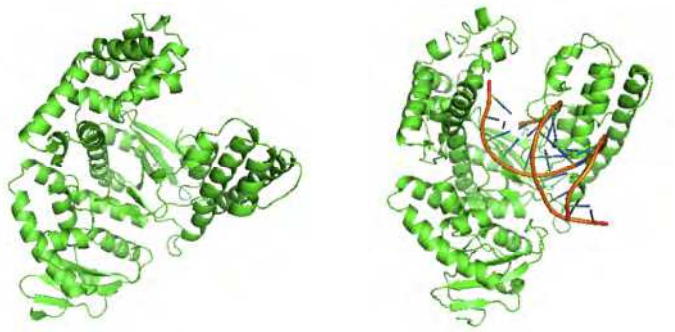


Fig. 9. Comparison of *Thermus aquaticus* DNA polymerase I's Stoffel fragment structures with (2ktq) and without a DNA duplex (1jxe) at the active site.

### 4.2 Selections from enzyme libraries

Design of selections for the isolation of catalysts from large protein repertoires has been far from obvious. The various parallel processing strategies to identify active enzymes rely generally on selections for binding. Selections for binding to suicide inhibitors were first tested (Soumillion et al., 1994). Selection of protein mutants for binding to transition state analogues yield in principle catalysts. This approach remains delicate, possibly because of the rough similarity between transition states and transition state analogues whose stability is required for the selections, and because of the time required to synthesize transition state analogues by organic synthesis. Successful parallel processing strategies for the isolation of catalysts relied on the selection of multiple products bound to the enzyme complex that catalyzed the chemical reaction. These *in vitro* selections are furthermore selections for the highest catalytic turnovers (Figure 10). Populations of enzymes with the highest catalytic efficiencies are thereby isolated.

Sequencing of the genes encoding hundred variants of the selected population then allows multiple sequence alignments to be carried out for the identification of recurrent mutations which characterize the catalytic activity change or improvement. Further characterization of isolated catalysts consists of the measurement of the kinetic parameters for the chemical reactions studied. Improvements of the catalytic efficiencies by several orders of magnitude have been described in the literature for several enzymes. These results have important applications in the field of biocatalysis.

Alternatively, for substrate-cleaving reaction, the concept of catalytic elution was reported (Pedersen et al., 1998): complexes between enzymes displayed on the surface of bacteriophages and their substrates bound to a solid phase are formed. Activation of the enzyme results in release of the phage-enzyme from the solid phase if the enzyme is active, while inactive enzymes remain bound to the solid phase (Soumillion & Fastrez, 2001).

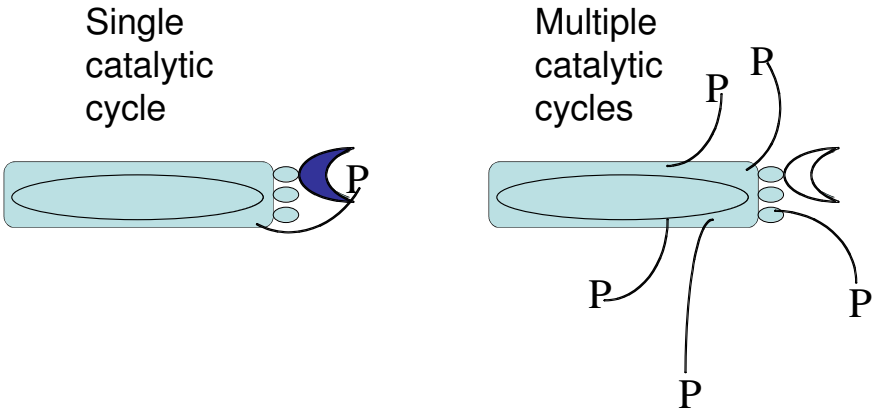



Fig. 10. Comparison of a highly active enzyme (white) efficiently captured by affinity chromatography for the product with a protein catalyzing a single substrate to product conversion (blue) unlikely to be isolated by affinity chromatography for the product.

### 4.3 Conclusion for enzymes

The parallel processing of molecular information on the catalytic activity of proteins (« Is the protein a catalyst or not ? ») is remarkably achieved by *in vitro* selection from large libraries of millions or billions of mutant proteins. Reduction of the large diversity into a small diversity of hundred(s) of variant proteins with the catalytic activity of interest allows characterization by serial processing to be accomplished. The sequencing of the corresponding genes for hundred(s) of variants allows computation of alignments for multiple sequences. The yield of protein production and the catalytic efficiencies for tens of selected variants allow the most promising variant protein to be identified. These results define sequence-activity relationships for enzymes. If enzyme-substrate complex structures are available, the sequence-structure-activity relationships that can be derived provide the central information for use in further biocatalytic applications.

## 5. Conclusion

Molecular biology, bioinformatics and protein engineering reached in the last decades a state allowing the isolation of proteins for desired functions of interest. Proteins can be isolated with a binding specificity for a given target, while enzymes can be isolated for given chemical reactions. Binding proteins and antibodies in particular found remarkable applications in the field of therapeutics. Enzymes turn out to be extremely useful in the field of biocatalysis for the production of chemicals at industrial scales within a sustainable environment.

Over the last twenty years, the use of antibodies has increased greatly, both as tools for basic research and diagnostics, and as therapeutic agents. This has largely been driven by ongoing advances in recombinant antibody technology. Today, more than 20 recombinant antibodies are widely used in clinic such as the human anti-TNFα antibody marketed as Humira® and many more antibodies are currently in clinical trials.

Satisfying industrial needs in the field of biocatalysis requires efficient enzymes to be isolated. While natural enzymes rarely fulfill industrial needs, and as long as computational approaches alone do not allow the sequences of protein catalysts to be designed, experimental methods such as the parallel processing strategies relying on *in vitro* selection combined with computational approaches for the characterization of catalysts may well be the most powerful strategies for the isolation of enzymes for given chemical reactions. Most notably, these new biocatalysts act in aqueous solutions without organic solvents at large scale and are ideally suited for green industrial processes.

A highly efficient design of binders and catalysts according to function can make use of a unique strategy: selection from large repertoires of proteins according to a function yield secondary protein repertoires of high interest, which can then be processed in series for their characterization due to their reduced diversity. Characterization involves sequencing of the corresponding genes for alignment of numerous protein sequences so as to define consensus sequences. This is the major advantage of molecular information parallel processing (MIPP) strategies: *defining conserved amino acids within protein scaffolds tightly linked to function.*

In conclusion, the parallel processing of biomolecular information (« Does the protein bind the target ? » or « Is the protein a catalyst for the chemical reaction ? ») is so far best achieved experimentally by using repertoires of millions or billions of proteins. Analysis of hundred(s) of protein variants is then best done computationally: use of multiple sequence alignment algorithms yields then sequence-activity relationships required for protein applications. Further biochemical and biophysical characterization of proteins (« Does the protein tend to form dimers or to aggregate ? », « Can the protein be produced at high level ? » , « What is the protein's pI ? ») is essential for their final use which may require high level soluble expression or cell penetration properties. In this respect, the development of algorithms analyzing protein properties remains a major challenge.

## 6. References

Ahmadvand, D., Rasaee, M. J., Rahbarizadeh, F., Kontermann, R. E. & Sheikholislami, F. (2009). Cell selection and characterization of a novel human endothelial cell specific nanobody. *Molecular Immunology,* Vol. 46, No. 8-9, pp. 1814-1823.

Barbas, C. F. d., Bjorling, E., Chiodi, F., Dunlop, N., Cababa, D., Jones, T. M., Zebedee, S. L., Persson, M. A., Nara, P. L., Norrby, E. & Burton, D. R. (1992). Recombinant human Fab fragments neutralize human type 1 immunodeficiency virus in vitro. *Proceedings of the National Academy of Sciences USA,* Vol. 89, No. 19, pp. 9339-9343.

Barbas, C. F. d., Crowe, J. E., Jr., Cababa, D., Jones, T. M., Zebedee, S. L., Murphy, B. R., Chanock, R. M. & Burton, D. R. (1992). Human monoclonal Fab fragments derived from a combinatorial library bind to respiratory syncytial virus F glycoprotein and neutralize infectivity. *Proceedings of the National Academy of Sciences USA,* Vol. 89, No. 21, pp. 10164-10148.

Behar, G., Chames, P., Teulon, I., Cornillon, A., Alshoukr, F., Roquet, F., Pugniere, M., Teillaud, J. L., Gruaz-Guyon, A., Pelegrin, A. & Baty, D. (2009). Llama single-domain antibodies directed against nonconventional epitopes of tumor-associated carcinoembryonic antigen absent from nonspecific cross-reacting antigen. *Febs Journal,* Vol. 276, No. 14, pp. 3881-3893.

Binz, H. K., Amstutz, P., Kohl, A., Stumpp, M. T., Briand, C., Forrer, P., Grutter, M. G. & Pluckthun, A. (2004). High-affinity binders selected from designed ankyrin repeat protein libraries. *Nature Biotechnology,* Vol. 22, No. 5, pp. 575-582.

Blaise, L., Wehnert, A., Steukers, M. P., van den Beucken, T., Hoogenboom, H. R. & Hufton, S. E. (2004). Construction and diversification of yeast cell surface displayed libraries by yeast mating: application to the affinity maturation of Fab antibody fragments. *Gene,* Vol. 342, No. 2, pp. 211-218.

Boder, E. T., Midelfort, K. S. & Wittrup, K. D. (2000). Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proceedings of the National Academy of Sciences USA,* Vol. 97, No. 20, pp. 10701-10705.

Boder, E. T. & Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nature Biotechnology,* Vol. 15, No. 6, pp. 553-557.

Boder, E. T. & Wittrup, K. D. (2000). Yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods Enzymology,* Vol. 328, No. pp. 430-444.

Bowley, D. R., Labrijn, A. F., Zwick, M. B. & Burton, D. R. (2007). Antigen selection from an HIV-1 immune antibody library displayed on yeast yields many novel antibodies compared to selection from the same library displayed on phage. *Protein Engineering Design Selection,* Vol. 20, No. 2, pp. 81-90.

Burton, D. R. (1991). Human and mouse monoclonal antibodies by repertoire cloning. *Tibtech,* Vol. 9, No. pp. 169-175.

Cadwell, R. C. & Joyce, G. F. (1994). Mutagenic PCR. *PCR Methods & Applications,* Vol. 3, No. 6, pp. 136-140.

Chen, W., Zhu, Z., Feng, Y., Xiao, X. & Dimitrov, D. S. (2008). Construction of a large phage-displayed human antibody domain library with a scaffold based on a newly identified highly soluble, stable heavy chain variable domain. *Journal of Molecular Biology,* Vol. 382, No. 3, pp. 779-789.

Chen, Y. & Stollar, B. D. (1999). DNA binding by the VH domain of anti-Z-DNA antibody and its modulation by association of the VL domain. *Journal of Immunology,* Vol. 162, No. 8, pp. 4663-4670.

Coia, G., Hudson, P. J. & Irving, R. A. (2001). Protein affinity maturation in vivo using E. coli mutator cells. *Journal of Immunological Methods,* Vol. 251, No. 1-2, pp. 187-193.

Daugherty, P. S., Chen, G., Iverson, B. L. & Georgiou, G. (2000). Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies. *Proceedings of the National Academy of Sciences USA,* Vol. 97, No. 5, pp. 2029-2034.

de Haard, H. J., van Neer, N., Reurs, A., Hufton, S. E., Roovers, R. C., Henderikx, P., de Bruine, A. P., Arends, J. W. & Hoogenboom, H. R. (1999). A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *Journal of Biological Chemistry,* Vol. 274, No. 26, pp. 18218-18230.

de Kruif, J., Boel, E. & Logtenberg, T. (1995). Selection and application of human single chain Fv antibody fragments from a semi-synthetic phage antibody display library with designed CDR3 regions. *Journal of Molecular Biology,* Vol. 248, No. 1, pp. 97-105.

Desmyter, A., Transue, T. R., Ghahroudi, M. A., Thi, M. H., Poortmans, F., Hamers, R., Muyldermans, S. & Wyns, L. (1996). Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nature Structural Biology,* Vol. 3, No. 9, pp. 803-811.

Feldhaus, M. & Siegel, R. (2004). Flow cytometric screening of yeast surface display libraries. *Methods Molecular Biology,* Vol. 263, No. pp. 311-332.

Figini, M., Marks, J. D., Winter, G. & Griffiths, A. D. (1994). In vitro assembly of repertoires of antibody chains on the surface of phage by renaturation. *Journal of Molecular Biology,* Vol. 239, No. 1, pp. 68-78.

Fromant, M., Blanquet, S. & Plateau, P. (1995). Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Analytical Biochemistry,* Vol. 224, No. 1, pp. 347-353.

Hanes, J., Jermutus, L. & Pluckthun, A. (2000). Selecting and evolving functional proteins in vitro by ribosome display. *Methods Enzymology,* Vol. 328, pp. 404-430.

Hanes, J. & Pluckthun, A. (1997). In vitro selection and evolution of functional proteins by using ribosome display. *Proceedings of the National Academy of Sciences USA,* Vol. 94, No. 10, pp. 4937-4942.

Harmsen, M. M. & de Haard, H. J. W. (2007). Properties, production, and applications of camelid single-domain antibody fragments. *Applied Microbiology and Biotechnology,* Vol. 77, No. 1, pp. 13-22.

Hawkins, R. E., Russell, S. J. & Winter, G. (1992). Selection of phage antibodies by binding affinity. Mimicking affinity maturation. *Journal of Molecular Biology,* Vol. 226, No. 3, pp. 889-896.

Ho, M., Kreitman, R. J., Onda, M. & Pastan, I. (2005). In vitro antibody evolution targeting germline hot spots to increase activity of an anti-CD22 immunotoxin. *Journal of Biological Chemistry,* Vol. 280, No. 1, pp. 607-617.

Hoet, R. M., Cohen, E. H., Kent, R. B., Rookey, K., Schoonbroodt, S., Hogan, S., Rem, L., Frans, N., Daukandt, M., Pieters, H., van Hegelsom, R., Neer, N. C., Nastri, H. G., Rondon, I. J., Leeds, J. A., Hufton, S. E., Huang, L., Kashin, I., Devlin, M., Kuang, G., Steukers, M., Viswanathan, M., Nixon, A. E., Sexton, D. J., Hoogenboom, H. R. & Ladner, R. C. (2005). Generation of high-affinity human antibodies by combining donor-derived and synthetic complementarity-determining-region diversity. *Nature Biotechnology,* Vol. 23, No. 3, pp. 344-348.

Hoogenboom, H. R. (1997). Designing and optimizing library selection strategies for generating high-affinity antibodies. *Trends Biotechnology,* Vol. 15, No. 2, pp. 62-70.

Hoogenboom, H. R. & Winter, G. (1992). By-passing immunisation. Human antibodies from synthetic repertoires of germline VH gene segments rearranged in vitro. *Journal of Molecular Biology,* Vol. 227, No. 2, pp. 381-388.

Irving, R. A., Coia, G., Roberts, A., Nuttall, S. D. & Hudson, P. J. (2001). Ribosome display and affinity maturation: From antibodies to single V-domains and steps towards cancer therapeutics. *Journal of Immunological Methods,* Vol. 248, No. 1-2, pp. 31-45.

Irving, R. A., Kortt, A. A. & Hudson, P. J. (1996). Affinity maturation of recombinant antibodies using E. coli mutator cells. *Immunotechnology,* Vol. 2, No. 2, pp. 127-143.

Jackson, J. R., Sathe, G., Rosenberg, M. & Sweet, R. (1995). In vitro antibody maturation. Improvement of a high affinity, neutralizing antibody against IL-1 beta. *Journal of Immunology,* Vol. 154, No. 7, pp. 3310-3319.

Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J. & Pluckthun, A. (2001). Tailoring in vitro evolution for protein affinity or stability. *Proceedings of the National Academy of Sciences USA,* Vol. 98, No. 1, pp. 75-80.

Jirholt, P., Ohlin, M., Borrebaeck, C. A. K. & Soderlind, E. (1998). Exploiting sequence space: shuffling in vivo formed complementarity determining regions into a master framework. *Gene,* Vol. 215, No. 2, pp. 471-476.

Jones, J. M. & Gellert, M. (2004). the taming of a transposon: V(D)J recombination and the immune system. *Immunological Review,* Vol. 200, No. 1, pp. 233-248.

Kepler, T. B. & Bartl, S. (1998). Plasticity under somatic mutations in antigen receptors. *Current Topics in Microbiology & Immunology,* Vol. 229, pp. 149-162.

Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Pluckthun, A. & Virnekas, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *Journal of Molecular Biology,* Vol. 296, No. 1, pp. 57-86.

Kondo, A. & Ueda, M. (2004). Yeast cell-surface display--applications of molecular display. *Applied Microbiology Biotechnology,* Vol. 64, No. 1, pp. 28-40.

Laffly, E., Pelat, T., Cedrone, F., Blesa, S., Bedouelle, H. & Thullier, P. (2008). Improvement of an antibody neutralizing the anthrax toxin by simultaneous mutagenesis of its six hypervariable loops. *Journal of Molecular Biology,* Vol. 378, No. 5, pp. 1094-1103.

Lee, M. S., Kwon, M. H., Kim, K. H., Shin, H. J., Park, S. & Kim, H. I. (2004). Selection of scFvs specific for HBV DNA polymerase using ribosome display. *Journal of Immunology Methods,* Vol. 284, No. 1-2, pp. 147-157.

Low, N., Holliger, P. & Winter, G. (1996). Mimicking somatic hypermutation: affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain. *Journal of Molecular Biology,* Vol. 260, No. 3, pp. 359-368.

Marks, J. D., Hoogenboom, H. R., Bonnert, T. P., McCafferty, J., Griffiths, A. D. & Winter, G. (1991). By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *Journal of Molecular Biology,* Vol. 222, No. 3, pp. 581-597.

Mattheakis, L. C., Bhatt, R. R. & Dower, W. J. (1994). An in vitro polysome display system for identifying ligands from very large peptide libraries. *Proceedings of the National Academy of Sciences USA,* Vol. 91, No. 19, pp. 9022-9026.

McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature,* Vol. 348, No. 6301, pp. 552-554.

Mouratou, B., Schaeffer, F., Guilvout, I., Tello-Manigne, D., Pugsley, A. P., Alzari, P. M. & Pecorari, F. (2007). Remodeling a DNA-binding protein as a specific in vivo inhibitor of bacterial secretin PulD. *Proceedings of the National Academy of Sciences USA,* Vol. 104, No. 46, pp. 17983-17988.

Muyldermans, S. (2001). Single domain camel antibodies: current status. *Reviews in Molecular Biotechnology,* Vol. 74, pp. 277-302.

Nahshol, O., Bronner, V., Notcovich, A., Rubrecht, L., Laune, D. & Bravman, T. (2008). Parallel kinetic analysis and affinity determination of hundreds of monoclonal antibodies using the ProteOn XPR36. *Analytical Biochemistry,* Vol. 383, No. 1, pp. 52-60.

Neuberger, M. S. (2008). Antibody diversity by somatic mutation: from Burnet onwards. *Immunology Cell Biol,* Vol. 86, No. 2, pp. 124-132.

Nissim, A., Hoogenboom, H. R., Tomlinson, I. M., Flynn, G., Midgley, C., Lane, D. & Winter, G. (1994). Antibody fragments from a 'single pot' phage display library as immunochemical reagents. *Embo Journal,* Vol. 13, No. 3, pp. 692-698.

Osbourn, J. K., Field, A., Wilton, J., Derbyshire, E., Earnshaw, J. C., Jones, P. T., Allen, D. & McCafferty, J. (1996). Generation of a panel of related human scFv antibodies with high affinities for human CEA. *Immunotechnology,* Vol. 2, No. 3, pp. 181-196.

Pasqualini, R. & Ruoslahti, E. (1996). Organ targeting in vivo using phage display peptide libraries. *Nature,* Vol. 380, No. 6572, pp. 364-346.

Pedersen, H., Hölder, S., Sutherlin, D. P., Schwitter, U., King, D. S. & Schultz, P. G. (1998). A method for directed evolution and functional cloning of enzymes. *Proceedings of the National Academy of Sciences USA,* Vol. 95, No. 18, pp. 10523-10528.

Rajpal, A., Beyaz, N., Haber, L., Cappuccilli, G., Yee, H., Bhatt, R. R., Takeuchi, T., Lerner, R. A. & Crea, R. (2005). A general method for greatly improving the affinity of antibodies by using combinatorial libraries. *Proceedings of the National Academy of Sciences USA,* Vol. 102, No. 24, pp. 8466-8471.

Razai, A., Garcia-Rodriguez, C., Lou, J., Geren, I. N., Forsyth, C. M., Robles, Y., Tsai, R., Smith, T. J., Smith, L. A., Siegel, R. W., Feldhaus, M. & Marks, J. D. (2005). Molecular evolution of antibody affinity for sensitive detection of botulinum neurotoxin type A. *Journal of Molecular Biology,* Vol. 351, No. 1, pp. 158-169.

Safsten, P., Klakamp, S. L., Drake, A. W., Karlsson, R. & Myszka, D. G. (2006). Screening antibody-antigen interactions in parallel using Biacore A100. *Analytical Biochemistry,* Vol. 353, No. 2, pp. 181-190.

Sanna, P. P., Williamson, R. A., De Logu, A., Bloom, F. E. & Burton, D. R. (1995). Directed selection of recombinant human monoclonal antibodies to herpes simplex virus glycoproteins from phage display libraries. *Proceedings of the National Academy of Sciences USA,* Vol. 92, No. 14, pp. 6439-6443.

Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H. J. & Pluckthun, A. (2001). In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proceedings of the National Academy of Sciences USA,* Vol. 98, No. 15, pp. 8572-8577.

Schier, R., Bye, J., Apell, G., McCall, A., Adams, G. P., Malmqvist, M., Weiner, L. M. & Marks, J. D. (1996). Isolation of high-affinity monomeric human anti-c-erbB-2 single chain Fv using affinity-driven selection. *Journal of Molecular Biology,* Vol. 255, No. 1, pp. 28-43.

Schreuder, M. P., Brekelmans, S., van den Ende, H. & Klis, F. M. (1993). Targeting of a heterologous protein to the cell wall of Saccharomyces cerevisiae. *Yeast,* Vol. 9, No. 4, pp. 399-409.

Schreuder, M. P., Mooren, A. T., Toschka, H. Y., Verrips, C. T. & Klis, F. M. (1996). Immobilizing proteins on the surface of yeast cells. *Trends Biotechnology,* Vol. 14, No. 4, pp. 115-120.

Seelig, B. & Szostak, J. W. (2007). Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature,* Vol. 448, No. 7155, pp. 828-831.

Sheets, M. D., Amersdorfer, P., Finnern, R., Sargent, P., Lindquist, E., Schier, R., Hemingsen, G., Wong, C., Gerhart, J. C. & Marks, J. D. (1998). Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens. *Proceedings of the National Academy of Sciences USA,* Vol. 95, No. 11, pp. 6157-6162.

Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science,* Vol. 228, No. 4705, pp. 1315-1317.

Soderlind, E., Strandberg, L., Jirholt, P., Kobayashi, N., Alexeiva, V., Aberg, A. M., Nilsson, A., Jansson, B., Ohlin, M., Wingren, C., Danielsson, L., Carlsson, R. & Borrebaeck, C. A. (2000). Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nature Biotechnology,* Vol. 18, No. 8, pp. 852-856.

Soumillion, P. & Fastrez, J. (2001). Novel concepts for selection of catalytic activity. *Current Opinion in Biotechnology,* Vol. 12, No. 4, pp. 387-394.

Soumillion, P., Jespers, L., Bouchet, M., Marchand-Brynaert, J., Winter, G. & Fastrez, J. (1994). Selection of beta-lactamase on filamentous bacteriophage by catalytic activity. *Journal of Molecular Biology,* Vol. 237, No. 4, pp. 415-422.

Stijlemans, B., Conrath, K., Cortez-Retamozo, V., Van Xong, H., Wyns, L., Senter, P., Revets, H., De Baetselier, P., Muyldermans, S. & Magez, S. (2004). Efficient targeting of conserved cryptic epitopes of infectious agents by single domain antibodies. African trypanosomes as paradigm. *Journal of Biological Chemistry,* Vol. 279, No. 2, pp. 1256-1261.

Thom, G., Cockroft, A. C., Buchanan, A. G., Candotti, C. J., Cohen, E. S., Lowne, D., Monk, P., Shorrock-Hart, C. P., Jermutus, L. & Minter, R. R. (2006). Probing a protein-protein interaction by in vitro evolution. *Proceedings of the National Academy of Sciences USA,* Vol. 103, No. 20, pp. 7619-7624.

Tordsson, J., Abrahmsen, L., Kalland, T., Ljung, C., Ingvar, C. & Brodin, T. (1997). Efficient selection of scFv antibody phage by adsorption to in situ expressed antigens in tissue sections. *Journal of Immunological Methods,* Vol. 210, No. 1, pp. 11-23.

Ulrich, H. D., Patten, P. A., Yang, P. L., Romesberg, F. E. & Schultz, P. G. (1995). Expression studies of catalytic antibodies. *Proceedings of the National Academy of Sciences USA,* Vol. 92, No. 25, pp. 11907-11911.

van den Beucken, T., Pieters, H., Steukers, M., van der Vaart, M., Ladner, R. C., Hoogenboom, H. R. & Hufton, S. E. (2003). Affinity maturation of Fab antibody fragments by fluorescent-activated cell sorting of yeast-displayed libraries. *FEBS Letters,* Vol. 546, No. 2-3, pp. 288-294.

VanAntwerp, J. J. & Wittrup, K. D. (2000). Fine affinity discrimination by yeast surface display and flow cytometry. *Biotechnology Prog,* Vol. 16, No. 1, pp. 31-37.

Vaughan, T. J., Williams, A. J., Pritchard, K., Osbourn, J. K., Pope, A. R., Earnshaw, J. C., McCafferty, J., Hodits, R. A., Wilton, J. & Johnson, K. S. (1996). Human Antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nature Biotechnology,* Vol. 14, No. 3, pp. 309-314.

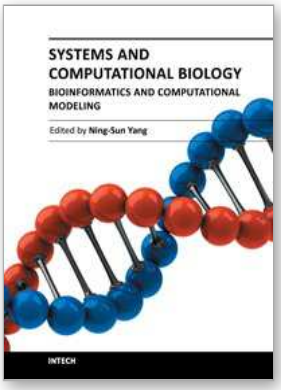
Wassaf, D., Kuang, G., Kopacz, K., Wu, Q. L., Nguyen, Q., Toews, M., Cosic, J., Jacques, J., Wiltshire, S., Lambert, J., Pazmany, C. C., Hogan, S., Ladner, R. C., Nixon, A. E. & Sexton, D. J. (2006). High-throughput affinity ranking of antibodies using surface plasmon resonance microarrays. *Analytical Biochemistry,* Vol. 351, No. 2, pp. 241-253.

Yang, W. P., Green, K., Pinz-Sweeney, S., Briones, A. T., Burton, D. R. & Barbas, C. F. r. (1995). CDR walking mutagenesis for the affinity maturation of a potent human anti-HIV-1 antibody into the picomolar range. *Journal of Molecular Biology,* Vol. 254, No. 3, pp. 392-403.

Yelton, D. E., Rosok, M. J., Cruz, G., Cosand, W. L., Bajorath, J., Hellstrom, I., Hellstrom, K. E., Huse, W. D. & Glaser, S. M. (1995). Affinity maturation of the BR96 anti-carcinoma antibody by codon-based mutagenesis. *Journal of Immunology,* Vol. 155, No. 4, pp. 1994-2004.

Zahnd, C., Amstutz, P. & Pluckthun, A. (2007). Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target. *Nature Methods,* Vol. 4, No. 3, pp. 269-279.

Zahnd, C., Spinelli, S., Luginbuhl, B., Amstutz, P., Cambillau, C. & Pluckthun, A. (2004). Directed in vitro evolution and crystallographic analysis of a peptide-binding single chain antibody fragment (scFv) with low picomolar affinity. *Journal of Biological Chemistry,* Vol. 279, No. 18, pp. 18870-18877.




**Systems and Computational Biology - Bioinformatics and Computational Modeling**

Edited by Prof. Ning-Sun Yang

ISBN 978-953-307-875-5
Hard cover, 334 pages
**Publisher** InTech
**Published online** 12, September, 2011
**Published in print edition** September, 2011

Whereas some â€œmicroarrayâ€ or â€œbioinformaticsâ€ scientists among us may have been criticized as doing â€œcataloging researchâ€, the majority of us believe that we are sincerely exploring new scientific and technological systems to benefit human health, human food and animal feed production, and environmental protections. Indeed, we are humbled by the complexity, extent and beauty of cross-talks in various biological systems; on the other hand, we are becoming more educated and are able to start addressing honestly and skillfully the various important issues concerning translational medicine, global agriculture, and the environment. The two volumes of this book present a series of high-quality research or review articles in a timely fashion to this emerging research field of our scientific community.

**How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jestin Jean-Luc and Lafaye Pierre (2011). Parallel Processing of Complex Biomolecular Information: Combining Experimental and Computational Approaches, Systems and Computational Biology - Bioinformatics and Computational Modeling, Prof. Ning-Sun Yang (Ed.), ISBN: 978-953-307-875-5, InTech, Available from: http://www.intechopen.com/books/systems-and-computational-biology-bioinformatics-and-computational-modeling/parallel-processing-of-complex-biomolecular-information-combining-experimental-and-computational-app

INTECH
open science | open minds

**InTech Europe**
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

**InTech China**
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821