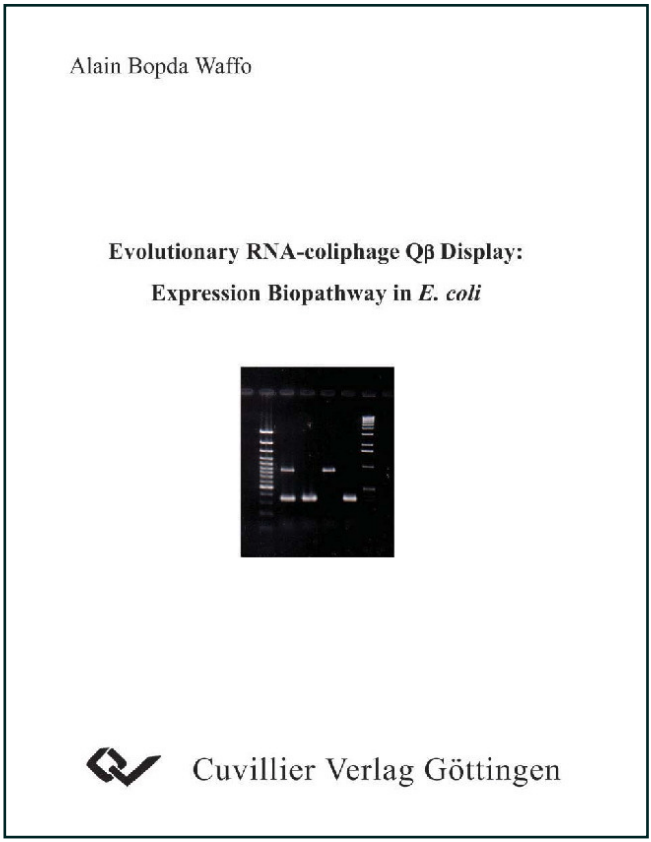
Alain Bopda Waffo (Autor)

Evolutionary RNA-coliphage Qß Display: Expression Biopathway in E.coli



https://cuvillier.de/de/shop/publications/3054

CHAPTER 1

Introduction

Coined by George Smith in 1985, the phage display technique or phage exposition (previously called phage fusion) consists of an insertion of a foreign DNA fragment (coding for either an enzyme, a receptor, an antibody, an antigen, an agonist or an antagonist) into a structural phage gene to create a fusion protein (with the foreign sequence in the middle) which is incorporated into a virion that retains its infectivity and exposes the foreign peptides in an accessible form.

The phage display technique has mostly been applied using the phage M13 (Scott and Smith, 1990; Smith and Scott, 1993). The gene sequence of a functional foreign protein is fused with one of the minor phage coat protein (pIII) gene by the means of recombinant DNA techniques. The cDNA resulting from large expression libraries can also be fused to the structural phage gene. After the transformation of the fusion constructed gene (*in vivo* translation and assambly), the peptide or peptide libraries are displayed on the surface of a virus. The hybrid phages, product of the *in vitro* replication coupled to *in vivo* translation are made up of a building block displaying the peptide (phenotype) and the recombinant DNA code (genotype). There is a physical connection between the peptide display and the coding recombinant DNA. This connection allows, through the use of one of a wide range of empirical selection or screening procedures, the isolation and enrichment of variants that exhibit required properties to a particular target molecule.

An M13 phage vector in phage display is a DNA-filamentous bacteriophage with a larger genome size (6407 bp) thereby allowing different sized DNA molecules or

libraries to be inserted into its genome. However, the DNA-replication system has very low mutation rates, which strongly affect (prevent) a rapid evolutive adaptation.

In contrast to the DNA-replication system, the RNA-replication system of RNA phages possesses very noticeable features including high mutation rates, high population size and short replication times that can be exploited for a rapid evolutive study (Domingo and Holland, 1997). Their high adaptability would make RNA phages ideal vectors for an evolutive phage display. Another important feature of RNA phages is the likely absence of recombination that is not a strict requirement for adaptability. One limiting step of RNA manipulation is the production and handling *in vitro* of the recombinant RNA. To overcome this difficulty, RNA can be retrotranscribed into DNA that allows for performing recombinant DNA techniques followed by retranscription of the recombinant DNA to RNA (Weissmann *et al.,* 1979; Domingo *et al.,* 2001).

The full cDNA of the RNA-coliphage Qβ has been successfully cloned into several plasmids among which is a pBR322-derivative plasmid (Taniguchi and Weissmann, 1978; Barrera *et al.,* 1993). The Qβ phages are liberated from cells transformed with the cloned cDNAs of the phage genome. Qβ cDNA can be used for recombinant DNA techniques in vitro followed by *in vivo* translation and phage production.

The work described here focuses on evolutionary protein engineering using the Qβ RNA-coliphage display system. The overall goal is to learn how to make molecules superior to natural proteins. Our aim was to expose functional proteins: receptor and enzyme on the surface of phages and optimize their function by an evolutionary approach.

CHAPTER 2

Theoretical background

2.1. The concept of evolution: mutation, selection and amplification

Evolution is defined as any change through time, usually viewed in a biological sense. This definition can be oriented as gradual and permanent change in form and function of successive generations of adult living organisms over a long period of geological time. Evolutionary processes are clearly abundant in biology and these processes provide some insight to the origin of the life. Charles Darwin developed a theory of "Evolution by Natural Selection" that can be summarized in the three following postulates: i) individuals within species are variable ii) some of these variations are passed on to offspring iii) individuals having the most favorable variations are the most successful, reproductively (preferentially represented in later generations) and are naturally selected.

With the development of biology, scientific studies have been used to try to support Darwin's principles. These include Ronald Fisher's theory of population genetics and Gregor Mendel's laws of heredity. The success and efficiency of Darwinian postulates in evolution is based on the dichotomy of genotype and phenotype (Schuster and Stadler, 1994), which are respectively the object under variation, and the target of selection. Evolutionary optimization is applied to all living organisms including unicellular, procaryotic and eucaryotic organisms, but the time span required for direct observation of evolutionary phenomena for higher organism are at least hundreds to thousands of years implying a long time for experiments (Eigen and

Schuster, 1977; Eigen *et al.,* 1989; Schuster and Stadler, 1994). The cellular metabolic key component of the phenotypic expression of an organism normally involves several complex pathways, which are also prone to an evolutionary process. Genotype and phenotype are clearly distinct features of an organism and the distinction between them is subtle.

Based on the replication system of population of RNA molecules, Sol Spiegelman and his coworkers pioneered experiments on evolution of RNA in the test-tube (Spiegelman, 1971). RNA molecules or RNA viruses' replication seems to be the model system for evolution study. RNA virus replication possesses some distinctive features, namely: high mutation rates (enormous plasticity), high yields (great population size) and short replication times (Domingo and Holland, 1997). In most RNA virus replication system, the main factor that contributes to such high mutation rates is the absence of proofreading repair activities in RNA replicases and transcriptases (Steinhauer *et al*., 1992). One consequence of this lack of proofreading activities in enzyme is the replication of viruses as complex, dynamic distribution of non-identical but related replicons called quasispecies. The term of quasispecies was coined by Eigen (1971) who first developed the theoretical treatment of replication with limited copying fidelity as a model system for early life forms on earth (Domingo and Holland, 1997).

Theoretically, quasispecies mean that population molecules are in steady-state (equilibrium), involving organized distribution of polynucleotide sequences (the error copies or closely related sequences) of a master sequence. In sequence space, the fittest genotype or master sequence is present at the highest concentration and may or may not coincide with the average or consensus sequence. In this light, previous population biology studies was seen as the consideration of the wild type not as an individual genome (with a defined nucleotide sequence) but as an ensemble of closely related genomes (Eigen and Biebricher, 1988). Mutation could also be the consequence of environmental parameters (Biebricher, 1996). The quasispecies interact and are subject to continuous mutation and competition for resources (Biebricher, 1999) during replication and the fit variant would survive ("Survival of fittest" of Darwin). Mutation rate of error-prone replication is the central point

(causal) of quasispecies and has been determined for some biological systems.

Species name	Genome size (bases)	Generation time [s] size	Typical population size	Mutation rate (substitution/nucleotide
Pisum sativum	1×10^{10}	3.15×10^{7}	5×10^{2}	10^{-5} - 10^{-12}
Drosophila melanogaster	2.6×10^{8}	1.2×10^{6}	5×10^{3}	10^{-5} - 10^{-10}
Escherichia coli	3.4×10^{6}	1.0×10^{3}	10^{11}	10^{-4} - 10^{-8}
phage Qβ	4.2×10^{3}	240	10^{14}	3×10^{-4}
MNV11RNA	86	40	10^{16}	3×10^{-4}

Table1: Data of some biological system used in evolutionary research (from Domingo *et al*., 2001)

The average mutation rate appears to be inversely proportional to genome size (Drake, 1991). The everlasting change in RNA virus genomes makes them more adaptive than other living beings in occurrence DNA-based organisms. Due to this high adaptation potential, RNA viruses can be exploited for evolutionary biotechnology precisely for phage display. This is possible by working at cDNA level, since the manipulation of RNA molecules is technically difficult. The full cDNA of the RNA coliphage Qβ was cloned into plasmid, which liberated Qβ phage after transformation of bacteria cells. This constructed vector can also be used to clone foreign functional protein genes in order to display a pool of hybrid phages presenting the foreign proteins on the phage surface. The most adapted hybrid phage can then be selected by artificial selection against a known target, e.g. biopanning (Jongsma *et al*., 1995).

2.2. Principle of display technology and applications

The last decade has seen the emergence of new research field of increasing impact known as evolutionary biotechnology. This research field is based on Darwinian evolutionary principles, which are applied to error-prone replicating molecules like RNA or DNA, which together with proteins are the key components of display technology.

In general display technology refers to a collection of methods for creating libraries of