DEPARTAMENTO DE CIÊNCIAS DA VIDA

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Evolving the Thermostability of PpAzoR, an Azoreductase from *Pseudomonas putida* MET94, by Directed Evolution

Dissertação apresentada à Universidade de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Bioquímica, realizada sob a orientação científica da Professora Doutora Lígia Martins (Universidade Nova de Lisboa) e da Professora Doutora Paula Morais (Universidade de Coimbra)

Nádia Gonçalves

2012

Acknowledgments

First I would like to express my deepest gratitude to the Instituto de Tecnologia Química e Biológica (ITQB), especially the Microbial and Enzymatic Technology (MET) group, for providing up to date facilities, conditions and a great work environment which made this work possible.

I am sincerely grateful to my supervisor, Dr. Lígia Martins, whose expertise, rigor, knowledge and guidance added considerably in the completion of this project. I am also grateful for her encouragement, bright suggestions, support, patience and understanding in difficult moments, when no solution seemed possible, teaching me how important it is to take one step at a time and be confident about your work.

I am sincerely grateful to Dr. Vânia Brissos for her tremendous support, encouragement, stimulating suggestions and advices, and her vast skills and knowledge providing the direction and all the support in directed evolution experiments. And for making the daily work much more pleasant with her patience and kindness.

I am grateful for the support throughout the year of Dr. Paula Morais and for accepting being my co-supervisor.

I deeply appreciate the support, friendship and companionship from all the members of the MET group: Ana Santos, Maura Ferreira, Pedro Bernardo, Zhenjia Chen, whom made me feel like being part of the family, especially to Sónia Mendes for the great support, suggestions and help in the beginning and throughout the year.

I would like to thanks my family, whom are too many to name all of them, for their tremendous support throughout this past two years, but I would like to address a special thanks to my grandfather Pedro, for his support, encouragement and advices; to my brother Igor, for being supportive and always being there for me whenever I needed; and my father for is support.

To all my friends, whom I will never thank enough, for their constant presence and support, and for all the moments we shared, making me feel the luckiest person for having them in my life.

Last, but by no means least, to my mother Joana, for her unconditional love, support and encouragement, and although she is not physically here, I can still fell it every single day.

Abstract

Azo dyes are the major group of synthetic colorants used in industries and are serious environmental pollutants. PpAzoR is a FMN-dependent NADPH azoreductase from *Pseudomonas putida* MET94, able to decolourize several azo dyes to their corresponding aromatic amines, under anaerobic conditions. This property makes PpAzoR very attractive for the biological treatment of wastewaters, but its thermal stability is rather low (with a half life of 30 min at 40°C) representing a serious drawback for its biotechnological applications. In this study it is described the improvement of PpAzoR thermostability by implementing directed evolution approaches. Directed evolution has being reported as a useful tool for improving enzymatic properties. PpAzoR mutant libraries were generated using error-prone PCR and variants were grown in 96 well-plates. To identify hits with improved thermostability and/ or activity a high-throughput screening assay was developed and optimised based on enzymatic measurements before and after a heat treatment. Four rounds of error-prone PCR and high-throughput screenings (≈ 10,000 clones) yielded 20 improved hits. With DNA sequencing of the variants, 18 mutations were identified as being involved in the thermostability, mostly in protein loops, one single mutation, located close to the FMN binding site (Y179H) is most probably responsible for the increase in enzymatic activity. A variant with the highest improvements (13G10) showing a 12-fold higher thermostability and a 2-fold higher activity relative to the wild type, had 6 mutations. These variant obtained, are a suitable starting point for further evolution to improve activity or to engineer a PpAzoR NADH-dependent.

Keywords: PpAzoR, thermostability, directed evolution, error-prone PCR, high-throughput screening.

Resumo

Os corantes azo são o maior grupo de corantes sintéticos utilizados na indústria, e a sua libertação no ambiente representa uma das maiores fontes de poluição a nível mundial. A PpAzoR é uma azoreductase NADPH-dependente de FMN da *Pseudomonas putida* MET94, capaz de reduzir vários corantes azo, em anaerobiose, às suas respetivas aminas aromáticas. Esta propriedade faz com que esta enzima seja um excelente candidato para o aplicações biotecnológicas, mas o facto de possuir uma estabilidade térmica muita baixa (tempo de meia vida a 40°C de 30 min) diminui claramente o seu potencial de utilização. Neste estudo é descrito o melhoramento da estabilidade térmica da PpAzoR por técnicas de evolução dirigida. A evolução dirigida tem mostrado ser uma ferramenta importante na otimização de algumas propriedades enzimáticas. Com base na introdução de mutações aleatórias na sequência de DNA do gene que codifica para a PpAzoR, quatro bibliotecas de mutantes foram criadas utilizando "error-prone PCR". A identificação dos mutantes com aumento na estabilidade e/ ou na atividade foi feita através "high-throughput screening" em placas de 96 poços, através da medição da atividade enzimática antes e depois de uma incubação a temperaturas elevadas. Como resultado da aplicação destas metodologias aproximadamente 10.000 clones foram avaliados o que resultou na seleção de 20 variantes apresentando actividade enzimática e termoestabilidade superiores à enzima nativa. Através da sequenciação do DNA foram identificadas 18 mutações envolvidas na termoestabilidade da enzima, localizadas principalmente em loops, e uma (Y179H), perto da posição da zona de ligação do FMN, responsável pelo aumento verificado na atividade. O variante 13G10 contém 6 mutações, apresentando os valores mais elevados da termoestabilidade e de atividade relativamente à enzima nativa: 12 vezes mais estável e 2 vezes mais ativo. Os variantes obtidos constituem um ponto de partida adequado para o melhoramento da atividade enzimática ou para a evolução de uma enzima dependente de NADH.

Palavras-chave: PpAzoR, termoestabilidade, evolução dirigida, error-prone PCR, high-throughput screening.

Table of contents

Abbreviations

AQS – anthraquinone-2-sulfonic acid

BZ – 1,4-benzoquinone

Ep-PCR – Error prone PCR

FMN – flavin mononucleotide

HTS - High-throughput Screening

IPTG – isopropyl β-D-1-thiogalactopyranoside

LB – Luria-Bertani

NADH – nicotinamide adenine dinucleotide

NADPH – nicotinamide adenine dinucleotide phosphate

NSA – 1,2-naphthoquinone-4-sulfonate

OD_{600nm} – optical density at 600 nm

RB5 – reactive black 5

SDS-PAGE – sodium dodecyl sulfate polyacrylamide gel electrophoresis

SOB – Super Optimal Broth

TB – Terrific Broth

Introduction

Azo dyes: general aspects

Azo dyes are the most versatile group of synthetic dyes, being extensively used in the paper and pulp, cosmetics and textile industries (Chen, 2006). They are characterized by one or more azo bonds (-N=N-), and it is estimated that more than 7 x 10^5 tons of these dyes are produced annually (dos Santos et al., 2007; Rai et al., 2005).

In textile industries the large-scale production and application of dyes, leads to increased amount of complex mixtures of dyestuff being released into the environment in the wastewaters which is one of the main sources of pollution worldwide. It is estimated that 1-10 % of dyestuff are lost and discharged directly with aqueous effluents, mainly during the dye production process (2 %) and during colourisation process (10-50 %), due to the incomplete exhaustion of the dyes on to the fibres (Forgacs et al., 2004; Kandelbauer and Guebitz, 2005; Pearce et al., 2003; Rai et al., 2005).

The release of azo dyes into the environment is highly undesirable, not only because of the aesthetic merit, transparency and gas solubility of water bodies that colour imposes, but also because some of the azo dyes and/or their breakdown products, mainly aromatic amines, are toxic or potentially carcinogenic (van der Zee et al., 2002). These dyes are conceived to be resistant to light, temperature, microbial degradation, which makes them highly recalcitrant (Rodriguez-Couto, 2009). Due to this stability, sometimes they are found in a chemical unchanged form, even after wastewater treatment (Suzuki et al., 2001).

Technologies used for colour removal from wastewaters

Chemical/Physical treatment

Currently, the treatment systems for coloured wastewaters are based mainly on chemical methods and physical methods using filtration, adsorption, reverse osmosis, flocculation and coagulation (dos Santos et al., 2007; Rai et al., 2005). Although these methodologies exhibit some advantages, the drawbacks are also considerable: they are very expensive due to intensive energy requirements, there are formation of hazardous byproducts and production of high amounts of concentrated sludge creating a disposal problem, and sometimes may even represent a secondary pollution due to addition of excessive chemicals into the environment (Chen, 2006; Forgacs et al., 2004; Pearce et al., 2003; Stolz, 2001).

Biological treatment

The biological treatment of wastewaters is seen as an attractive solution, because of its low-cost and eco-friendly reputation. In the past years several microorganisms have shown their ability to degrade azo dyes under certain environmental conditions, mainly bacteria (*Enterococcus* sp., *Streptococcus faecallis*, *Proteus* sp., *Bacillus* sp., *Pseudomonas* sp., etc.) and fungi (*Phanerochaete chrysosporium*, *Trametes versicolor*, *Aspergillus* sp.) (Kandelbauer and Guebitz, 2005; Rai et al., 2005). Under aerobic conditions azo dyes are not readily metabolized; however under anaerobic conditions many bacteria reduce the dye molecules to produce colourless aromatic amines. These amines are resistant to further anaerobic mineralization and are potentially toxic. Fortunately these are good substrates for aerobic degradation (Pearce et al., 2003; Stolz, 2001). In such cases a complete removal of azo dyes from wastewaters can be achieved using a sequential or integrated anaerobic-aerobic treatment, (Rai et al., 2005; van der Zee et al., 2002). Therefore the combination of anaerobic and aerobic processes using whole cells is a very promising technology since they enable degradation of dyes and aromatic amines product, representing the most appropriate bio-systems for biodegradative processes (Kandelbauer and Guebitz, 2005; Mendes et al., 2011a).

Enzymatic treatment

With the strong demand for new recycling technologies, and to reduce the enormous water consumption in the textile finishing, enzymatic processes are particularly sought because of their specificity: they only attack the dyes molecules, while valuable dyeing additives or fibres are kept intact and can potentially be reused. There is a limited range of dyes that can be decolourised by a single enzyme, but this issue can be overcome by using a cocktail of enzymes (Kandelbauer and Guebitz, 2005). Importantly, although dye molecules display vast differences in chemical structures, they are mainly degraded by a few redox-active enzymes, all showing relatively wide substrate specificity: laccases, peroxidases, azoreductases, and some mono- or di-oxygenases (Kandelbauer and Guebitz, 2005). Laccases and peroxidases are oxidative enzymes typically produced by fungi, and azoreductases are reductases mainly found in bacteria (Rai et al., 2005; Stolz, 2001).

Laccases are oxidoreductases belonging to the multicopper oxidase family, and are able to reduce a range of aromatic compounds using the available oxygen as electron acceptor. Although most of these enzymes were first identified in white-rot fungi, a laccase from *Bacillus subtilis*, CotA-laccase, was extensively studied and shown to be able to decolourise a variety of structurally different synthetic dyes at an alkaline pH (Pereira et al., 2009a; Pereira et al., 2009b).

Azoreductase is a generic name given enzymes involved in the reduction of azo bonds. Utilizing nicotinamide adenine dinucleotide (NADH) or nicotinamide adenine dinucleotide phosphate (NADPH) as electron donor azoreductases can decolourises azoic dyes into the corresponding by reductive cleavage of the azo bond under aerobic or anaerobic conditions (McMullan et al., 2001; Ryan et al., 2010). Currently there are three types of azoreductases characterized in bacteria: flavin-dependent NADH preferred azoreductase, flavin-dependent NADPH preferred azoreductase and flavin free azoreductases. The currently accepted mechanism of azo dye reduction is known as a ping-pong bi-bi mechanism, which requires two cycles of NADPH-dependent reduction of FMN to $FMNH_2$, which reduces the azo substrate to a hydrazine in the first cycle, and reduces the hydrazine to two amines in the second cycle (Ryan et al., 2010). Azoreductases activity are mainly harboured by FMN-dependent enzymes, and they have being identified in several microorganisms:, *Escherichia coli*, *Bacillus* sp. SF, *Pseudomonas putida* MET 94, *Rhodobacter sphaeroides*, *Bacillus* sp. OY1-2, *Pseudomonas aeruginosa* (Liu et al., 2009; Maier et al., 2004; Mendes et al., 2011b; Nakanishi et al., 2001; Suzuki et al., 2001; Wang et al., 2007). Microorganisms are exposed to a variety of toxic and antimicrobial compounds in the environment, which induce general and specific stress response in the growing cells. Flavin-dependent azoreductases share strong similarities with regard to sequence, structure, and reaction mechanism with the larger family of flavin-dependent quinone reductases that includes Lot6p from *Saccharomyces cerevisiae* and the mammalian NQO1 (Deller et al., 2008). These enzymes are involved in the reduction of quinones, quinoneimines, azo dyes, and nitro groups to protect the cells against the toxic effects of free radicals and reactive oxygen species arising from electron reductions. They are assumed to take part in the organism's enzymatic detoxification systems; e.g., the azoreductases from *E. coli* and *B. subtilis* were recently implicated in the cellular response to thiol-specific stress (Leelakriangsak et al., 2008; Liu et al., 2009; Towe et al., 2007) and Lot6p, the azoreductase homologue in *S. cerevisiae* has been implicated in the response to oxidative stress (Sollner et al., 2009; Sollner et al., 2007). Furthermore, as additional members of this family of enzymes are discovered, the list of transformed substrates continues to grow. Evolutionarily, these enzymes may provide a selective advantage to bacteria under various conditions of environmental stress (Khersonsky and Tawfik, 2010). Functional promiscuity, multi-specificity and structural plasticity are inherent to proteins, since when new protein functions are needed, promiscuous functions, as well as alternative conformations mediating them, provide the most accessible solution (Nobeli et al., 2009; Tawfik, 2010).

PpAzoR, an azoreductase from *Pseudomonas putida* MET94

Aiming at investigating the enzymatic biodegradation and bioconversion of synthetic dyes, a library of 48 microorganisms was screened for their ability to decolourize azo dyes and *Pseudomonas putida* MET94 was identified as the most efficient decolourising strain (Mendes et al., 2011b). By *in silico* screening, the enzyme, PpAzoR, involved in the decolourisation process was cloned and biochemically characterized. Crystals of PpAzoR that diffracted up to 1.6Å resolution were obtained, using the vapor diffusion method, and the crystal structure has been solved by the Molecular Replacement Method (Correia et al., 2011). Moreover this enzyme was shown to be a flavin mononucleotide (FMN)-dependent homodimer with a molecular weight around 40 kDa, exhibiting a higher specificity for NADPH, rather than for NADH as electron donor. PpAzoR exhibited a broad substrate specificity for azo dyes reduction, with all the azo dyes tested being decolourised at a considerable extent under anaerobic conditions by a ping-pong bi-bi mechanism (Mendes et al., 2011b).

Using the synergistic action of PpAzoR and CotA-laccase, Mendes et al. (2011b) evaluated the enzymatic degradation of 18 azo dyes and 3 model wastewaters and the final toxicity of the reaction mixtures (Mendes et al., 2011a). A sequential anaerobic-aerobic treatment was performed and the toxicity was evaluated over the growth of *Saccharomyces cerevisiae* and the reproduction of *Caenorhabditis elagans*. This treatment resulted in 100 % decolourisation of all dyes tested and a reduction level of toxicity was 50-95 %, as compared with a single treatment with PpAzoR. For the 3 model wastewaters evaluated, the sequential treatment resulted in complete decolouration (values very close to 100 %) and relatively low toxicity levels. These results are very promising, since they show that a combined system with PpAzoR and CotA can allow the simultaneous degradation and detoxification of multi-dye mixtures.

PpAzoR broad substrate specificity towards the degradation of a range of structurally diverse azo dyes makes it attractive for biological treatment of wastewaters. However, its thermal stability is rather low (half-life of 30 minutes at 40°C) (unpublished results). Therefore, the evolution of PpAzoR towards a higher stability was proposed using directed evoltion. Several enzymatic properties have being successfully improved using directed evolution approaches, but thermostability is one of the most desired, with some examples of successfully evolved enzymes: a lipase, a β-glucoronidase, ligninolytic oxidoreductases, a xylanse, a *p*-nitrobenzil esterase, a cytochrome P450 peroxygenase, a phytase, etc (Acharya et al., 2004; Flores and Ellington, 2002; Garcia-Ruiz et al., 2010; Miyazaki et al., 2006; Salazar et al., 2003; Shivange et al., 2012). Thermostable enzymes are much more attractive because thermal stability is normally correlated with a longer life-time, a higher tolerance to harsh conditions frequently found in industrial processes, and also to higher reaction yields and solubilisation of

polymeric or hydrophobic compounds. And ultimately, enzymes with higher thermostability are typically tolerant to many other harsh conditions often sought, such as the presence of organic co-solvents, extreme pHs, high salt concentration, high pressure, etc (Garcia-Ruiz et al., 2010). Besides, using directed evolution approaches can offer some insights into the molecular determinants for protein thermostability.

Directed Evolution as a tool for improving enzymatic properties

Enzymes are nature's catalysts, capable of carrying out a tremendous range of biochemical functions, but with an ever increasing demand for new and improved biocatalysts, sometimes their efficiency and precision do not meet the requirements for industrial applications. One approach to engineering enzymes is to make specific modifications, but this demands a detailed and frequently unattainable understanding of the relationship between sequence and function (Bloom et al., 2005). Meanwhile, evolutionary engineering, also termed "directed evolution", bypasses this problem and has emerged as key technology for biomolecular engineering and generated impressive results in the functional adaptation of single genes and proteins (Reetz et al., 2008).

Rational Design vs Directed evolution

Although relationships between the sequence, structure and function of proteins are continuously being elucidated, the extensive knowledge that is necessary for the application of rational engineering approaches is available for only a tiny fraction of known enzymes (Kuchner and Arnold, 1997). Therefore attempts to engineer enzymes rationally have being limited (Dalby, 2003). Conversely, the general methods used in directed evolution mimic natural evolution processes such as random mutagenesis or recombination, and so similarly permit to engineer enzymes without understand the structure and function in great detail (Dalby, 2003).

Directed evolution has proven to be a useful and powerful method to optimize the catalytic properties of enzymes by entailing the accumulation of beneficial mutations (Tracewell and Arnold, 2009). Essentially like natural evolution, it comprises the iterative implementation of (1) the generation of a "library" of mutated genes, (2) its functional expression, and (3) a sensitive assay to identify individuals showing the desired properties, either by selection or screening (Brakmann, 2001).

Directed evolution strategies

Directed evolution normally starts with a parent enzyme and an engineering goal (Bloom and Arnold, 2009). The gene for the parent enzyme is mutagenized to create a library of mutant genes. Gene products that show improvement with respect to the

desired property or set of properties are identified by screening or selection, and the genes encoding those enzymes are subjected to further cycles of mutation and screening in order to accumulate beneficial mutations. The evolution therefore can involve few or many generations, depending on how far one wishes to progress and the effects of mutations observed in each generation (Kuchner and Arnold, 1997).

There is no single "best" mutagenesis method. A highly effective and efficient directed evolution strategy is to gradually accumulate single beneficial mutations, either sequentially or by recombination while applying (often increasing) selection pressure (Tracewell and Arnold, 2009).

Error-prone PCR (ep-PCR) protocols are modifications of the standard PCR methods, conceived to alter and enhance the natural error rate of *Taq* polymerase. These protocols typically contain higher concentrations of $MgCl_2$, in order to stabilize non-complementary pairs, $MnCl_2$ can also be added to increase the error-rate, as well as, varying ratios of nucleotides in the reaction, etc (Arnold and Georgiou, 2003b). Due to its simplicity and versatility, this technique has emerged as the most common technique which can result in mutation frequencies as high as 2 % per nucleotide position (Brakmann, 2001). However, the number of amino acids substitutions accessible when using ep-PCR is limited because this reaction biases the distribution of mutation type in favor of transitions (A↔G, T↔C), and also multiple substitutions within a single codon are extremely rare (Brakmann, 2001).

Saturation mutagenesis enables the creation of a library of mutants containing all possible mutations at one or more pre-determined target positions in a gene sequence. This is achieved by introducing all possible base triplets at a given codon position, thereby resulting in the formation of all 20 amino acids at this position of the protein. Since random point mutations do not access a large fraction of the enzyme sequence, this technique is normally used to increase the number of amino acid substitutions accessible by random mutagenesis (Arnold and Georgiou, 2003b). This method restricts the random mutations to predicted sites in the enzyme creating therefore what is called focused libraries, but it requires structural information in order for the correct mutagenesis sites to be chosen (Reetz et al., 2008). It does not require construction and screening of very large libraries, and with this can come the possibility of using screens that are higher in quality and more likely to accurately interrogate the desired properties (Tracewell and Arnold, 2009).

Recombination represents an alternative or additional approach for generating genetic diversity based on the mixing of genetic material from a number of parent sequences. Comparatively to random mutagenesis, recombination may be advantageous in concentrating beneficial mutations, which have arisen independently and may be additive, and likewise, in concentrating deleterious mutations which

subsequently might be more efficiently purged from the population by selection (Brakmann, 2001).

DNA Shuffling was the first technique introduced for random in vitro recombination of gene variants created by random mutagenesis. It employs the PCR reassembly of whole genes from a pool of short overlapping DNA sequences which are generated by random enzymatic fragmentation of different parental genes (Brakmann, 2001).

Currently there is no such thing as the "best" mutagenesis method but a very common practice is to use sequentially ep-PCR and saturation mutagenesis or DNA shuffling, depending on the engineering goal. Since there are many ways to create sequence diversity the goal of choosing a mutagenesis strategy is to minimize the screening requirements and increase the chances of finding beneficial mutations (Tracewell and Arnold, 2009).

Selection vs high-throughput screening is used to identify improved mutants, and it depends on the engineering goal (Bloom and Arnold, 2009). The most challenging step in directed evolution is to develop a screening or selection scheme that is sensitive to the properties of interest (Brakmann, 2001). In some cases, the protein property of interest can be coupled to the survival of the host-cell, thereby allowing for direct genetic selection of cells carrying improved mutants. More frequently, it is not possible to design an effective selection method, and mutants must be assayed directly in a high-throughput screening (HTS) (Bloom and Arnold, 2009). Screening measures enzymatic activity or another property via calorimetric or other easily detectable reactions (Bommarius et al., 2006). When compared to selection, the screening approach enables a better control of the applied constraints, and it is also more versatile. The number of individual mutants that can be screened in a certain period of time (throughput) will depend strongly on the enzymatic reaction and the sensitivity of the applied detection principle, therefore it is extremely important to choose the selective constraints that reflect precisely the desired property (Brakmann, 2001). These assays can be performed in agar plates, however most screens are performed in a multi-well microplate and the enzymatic activity can be assayed using a microplate reader spectrophotometer (Arnold and Georgiou, 2003a).

Directed evolution of PpAzoR

PpAzoR broad substrate specificity towards the degradation of a range of structurally diverse azo dyes makes it attractive for biological treatment of wastewaters. However, its thermal stability is rather low (half-life of 30 minutes at 40°C) (unpublished results) representing a serious drawback to its biotechnological application.

Thermal stability has being considered by many as a key feature when it comes to protein robustness, evolvability and catalytic function (Bloom et al., 2006). Several enzymatic properties have being successfully improved using directed evolution approaches, but thermostability is one of the most desired, with some examples of successfully evolved enzymes: a lipase, a β-glucoronidase, ligninolytic oxidoreductases, a xylanse, a *p*-nitrobenzil esterase, a cytochrome P450 peroxygenase, a phytase, etc (Acharya et al., 2004; Flores and Ellington, 2002; Garcia-Ruiz et al., 2010; Miyazaki et al., 2006; Salazar et al., 2003; Shivange et al., 2012). Thermostable enzymes are much more attractive because thermal stability is normally correlated with a longer life-time, longer life-time during use and higher temperature optimum for activity and also a higher tolerance to harsh conditions frequently found in industrial processes, and also to higher reaction yields and solubilisation of polymeric or hydrophobic compounds (Salazar et al., 2003). And ultimately, enzymes with higher thermostability are typically tolerant to many other harsh conditions often sought, such as the presence of organic co-solvents, extreme pHs, high salt concentration, high pressure, etc (Garcia-Ruiz et al., 2010). Besides, using directed evolution approaches can offer some insights into the molecular determinants for protein thermostability.

Therefore, the evolution of PpAzoR towards a higher stability was proposed using directed evolution. After setting up a reliable screening condition, error-prone PCR was used to introduce random mutations in the *ppAzoR* sequence and the resulting mutants were stressed under high temperatures and explored for activity and thermostability using high-throughput screening (HTS).

Material and Methods

Bacterial strains, plasmids and media

The *Escherichia coli* strain DH5α (Novagen) was used for routine propagation and amplification of plasmids constructs. The *E. coli* Tuner (DE3, Novagen) and KRX (Promega) strains were used to express the *ppAzoR* and variant genes cloned in pET-21a (+) plasmid (Novagen); in the Tuner strain the target genes are under the control of T7 promotor, induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) and in the KRX strain, the genes are under the control of the *rha*P_{BAD} promoter, induced by rhamnose. Luria-Bertani medium (LB) was used for maintenance and growth of *E. coli* strains, with appropriate antibiotic selection when needed. LB medium contains the following components (per liter): 10 g of tryptone, 5 g of yeast extract and 10 g of $NaCl_2$. In some experiments cells were cultured in Terrific Broth medium (TB) or Super Optimal Growth medium (SOB). TB contains the following components (per liter): 12 g of tryptone, 24 g of yeast extract, 3.4 g of glycerol, 2.3 g of KH_2PO_4 and 12.5 g of K_2HPO_4. SOB contains (per liter): 20 g of tryptone, 5 g of yeast extract, 0.584 g of $NaCl_2$ and 0.186 g of KCl_2. All culture media were sterilized by autoclaving and stored at room temperature until use.

Random mutagenesis by error-prone PCR (ep-PCR)

Variation in the *ppAzoR* gene was generated by using error-prone PCR (ep-PCR). Primers 5' GGAGAGTCATATGAAACTGTTGC 3' (PpaF) and 5' CAACCAAAGGATCCCTT GATCAGG 3' (PpaR) flank the gene beyond the *Nde*I and *Bam*HI sites were used for amplification. Ep-PCR was carried out in 50 µL reaction volume containing 3 ng of DNA template, 0.5 µM of primers, 200 µM of dNTPs, 7 mM $MgCl_2$, *Taq* polymerase buffer, and 5 U of *Taq* polymerase (Fermentas). The influence of $MnCl_2$ was studied at different concentrations (200-250 µM). After an initial denaturation period of 10 min at 94°C, the following steps were repeated for 30 cycles in a thermal cycler (MyCycler™ thermocycler, Biorad): 1 min at 94°C, 1 min at 55°C and 45 s at 72°C and at the end 10 min at 72°C. The amplified product (5 µL) was visualized by agarose (1%) electrophoresis, purified using the GFX PCR DNA and the Gel Band Purification kits (GE Healthcare) and eluted with milli-Q H_2O. Ten microliters of purified PCR product and 10 µL of plasmid pET-21a (+) (Novagen) were digested with 10 U of *Nde*I (Fermentas) at 37°C for 2 h. The mixture was placed at 65°C for 20 min and then dialyzed against milli-Q H_2O for 30 min using MF™ filters (Millipore). A second digestion was performed with 10 U of *Bam*HI (Fermentas) for 2 h at 37°C, and pET-21a (+) was simultaneously dephosphorylated with 1 U Alkaline Phosphatase (FastAP, Fermentas) to prevent plasmid self-ligation. The products were purified using the GFX PCR DNA and Gel Band Purification kits (GE Healthcare). The ligations were performed with 0.5 U of T4 DNA ligase (Fermentas) using a ratio of 1:8 of vector to insert. This solution was incubated overnight, at room temperature and then incubated at 65°C for 10 min. The

preparation was then dialyzed against milli-Q H_2O for 60 min using MF^{TM} filters; 5 µL of this solution was used to transform electrocompetent *E. coli* KRX cells.

Preparation of competent cells

A LB agar plate was streaked out with a frozen stock of *E. coli* cells and incubated overnight at 37°C. A single colony was picked, used to inoculate 50 mL of SOB medium and then incubated overnight at 37°C at 120 rpm. One liter of SOB medium was inoculated in order to start with an optical density at 600 nm (OD_{600nm}) of 0.05 and incubated at 37°C at 120 rpm. When the $OD_{600nm} \approx 0.8$ the cells were transferred to pre-cooled centrifuge bottles and spin down at 5,000 rpm (15 min at 4°C). The supernatant was discarded, and the cell pellet was treated accordingly to the heat shock or electroporation protocols. In the **heat shock protocol**, the cell pellet was first washed with an ice-cold 100 mM $MgCl_2$ solution. After centrifugation (5,000 rpm, 15 min at 4°C) the supernatant was discarded and the pellet washed using an iced-cold 100 mM $CaCl_2$ solution and kept on ice for 2 h. After this period a second centrifugation was performed, the supernatant discarded and the pellet suspended in 1 mL of 100 mM $CaCl_2$ containing 15 % of glycerol. Aliquots of 150 µL were frozen on nitrogen and stored at -80°C. In the **electroporation protocol**, the cell pellet was washed with 500 mL of a sterile ice-cold 10 % glycerol solution. The cells were centrifuged, re-washed in the same solution, centrifuged for a second time and resuspended in wash buffer to a final volume of 2 mL. Aliquots of 100 µL were frozen on nitrogen and stored at -80°C.

Transformation of *Escherichia coli* cells

For the transformation of *E. coli* cells two different protocols were used: heat shock and electroporation. In the **heat shock protocol**, 2 µL of plasmidic DNA was added to an aliquot of 150 µL of *E. coli* competent cells and placed on ice for 30 min. This mixture was incubated at 42°C for 90 s and placed on ice for 5 min. One mililiter of LB medium was added to the cells and incubated at 37°C for 1 h at 200 rpm. After this period cells were centrifuged at 5,000 rpm for 5 min; 1 mL of supernatant was discarded and the cells were ressuspended in the remaining medium. The cells were spread on a LB agar plate, supplemented with 100 µg/mL of ampicillin, and incubated overnight at 37°C. In the **electroporation protocol**, 2 µL of DNA was added to an aliquot of 100 µL electrocompetent cells, previously thawed on ice, mixed and placed on ice for 5 min. This mixture was transferred to a sterile and pre-cooled electroporation cuvette, which was placed in the Xcell ShockPod chamber (Gene Pulser Xcell^{TM}, Biorad) and pulsed using the set C = 25 µF, PC = 200 Ω, V = 2.5 kV. One milliliter of LB medium was immediately added and the suspension was transferred to a tube, and incubated at 37°C for 1 h at 200 rpm. The cells were then centrifuged at 5,000 rpm

for 5 min; 1 mL of supernatant was discarded and the cells were resuspended in the remaining medium. The cells were spread on a LB agar plate, supplemented with 100 µg/mL of ampicillin and incubated overnight at 37°C.

Selection of the appropriate enzyme substrate for activity screenings

Growth of recombinant strains and *ppAzor* overexpression

The plasmid pLP-1 (Mendes *et al.*, 2011), containing the *ppAzor* gene, was transformed into *E. coli* Tuner (DE3, Novagen) or *E. coli* KRX (Promega), producing LOM528 and LOM531 strains, respectively. Single colonies were used to inoculate 20 mL of LB medium supplemented with 100 µg/mL ampicillin, grown overnight at 37°C, 160 rpm. Fresh cultures were transferred to 100 mL of LB medium supplemented with 100 µg/mL ampicillin, in order to start the growth with an $OD_{600nm} = 0.05$. Cultures were incubated at 30°C, 160 rpm and when $OD_{600nm} \approx 0.6$, 100 µM IPTG was added to LOM528, and 0.1 % Rhamnose to LOM531. After 24 h of cultivation, cells were collected by centrifugation (10,000 rpm, 15 min at 4°C). The cell pellets were suspended in 1.5 mL of 20 mM Tris-HCl buffer (pH 7.6), containing 5 mM $MgCl_2$, 1 U/mL of DNAse I, and 2 µL/mL of a mixture of protease inhibitors: antipain and leupeptin. Cells were disrupted by French Press (Thermo EFC) and then centrifuged at 18,000 rpm for 2 h at 4°C. The supernatants (cell crude extracts) were collected and used to perform enzymatic assays. The protein concentration was determined using the Bradford assay with bovine serum albumin (BSA) as standard. SDS-PAGE electrophoresis was performed to visualize protein overproduction in crude extracts.

Enzymatic assays

Enzymatic activities were measured by using one azo-dye, reactive black 5 (RB5) and five quinone substrates, anthraquinone-2-sulfonic acid (AQS), 1,4-benzoquinone (BZ), 1,2-dihydroxybenzene (catechol), 2-hydroxy-1,4-naphthoquinone (Lawsone), 1,2-naphthoquinone-4-sulfonate (NSA) (Mendes, S., unpublished results). The substrate concentration used in the assays was 100 µM, except for BZ where a 50 µM concentration was used, and NADPH was used at a 250 µM concentration. Enzymatic assays were performed at 30°C, in 100 mM sodium phosphate buffer (pH 7). The reactions were initiated by the addition of crude extracts and followed by monitoring the decrease in absorbance of NADPH at 340 nm (ε = 6,220 M^{-1} cm^{-1}) on a Nicolet Evolution 300 spectrophotometer (Thermo Industries). For reactions with RB5, the cuvettes containing the buffer and substrates were sealed with a suba-seal rubber, and made anaerobic by Argon bubbling. Reactions were initiated by directly injecting anoxic cell crude extracts with a syringe.

Overexpression of *ppAzoR* in 96-well plates

Single LOM531 colonies were picked with a sterile toothpick and transferred to a 96 well-plate (Greiner Bio-One) containing 200 µL of LB medium supplemented with 100 µg/mL ampicillin. Two approaches were used: a first in which all the 96 wells were used, and a second using only the 60 interior wells, therefore filling the peripheral wells with water. In both cases, cultures were grown for 24 h at 30°C, 750 rpm in a Titramax 1000 shaker (Heidolph). Twenty microliters aliquots were used to inoculate 180 µL of TB medium supplemented with 100 µg/mL ampicillin. The 96 well-plates were sealed with parafilm and incubated at 30°C, 750 rpm. After 3 h of incubation, 0.1 % Rhamnose was added to induce gene expression, and the growth proceeded for a 24 h period. The cells were harvested by centrifugation at 4,000 rpm, for 20 min at 4°C.

Cell disruption

Three different approaches were tested for cell disruption, a chemical, an enzymatic and a physical method. For the **chemical** cell disruption, 100 µL of a 40 % Bacterial Protein Extraction Reagent (B-PER®) lysis solution (Thermo Scientific), was used to resuspend the cell pellets. In the **enzymatic** method the cell pellets were incubated at -80°C for 1 h and then resuspended in 100 µL of a lysis buffer consisting of 20 mM Tris-HCl (pH 7.6), lyzozyme (2 mg/mL) and DNAse I (2 U/mL) and incubated for 1 h at 32°C. Freezing and thawing was the **physical** cell disruption method applied; the 96 well-plates were submerged in liquid nitrogen or incubated at -80°C for 15 min, and then thawed at room temperature for 5 min. After 3 cycles of freeze and thaw, cell pellets were resuspended in 100 µL of 20 mM Tris-HCl (pH 7.6) and DNAse I (2 U/mL). After cell disruption, plates were centrifuged at 4,000 rpm, for 30 min, at 4°C and the supernatant (cell crude extracts) used for enzymatic activity measurements.

Enzymatic assays

Aliquots of 20 µL of cell crude extracts were transferred to a new 96-well plate and the reaction was initiated by adding 180 µL of 100 mM sodium phosphate buffer (pH 7) containing AQS and NADPH in a final concentration of 100 µM and 250 µM, respectively. The oxidation of NADPH at 340 nm was followed on a Synergy 2 (Biotek) micro plate reader.

High-throughput screening for activity and thermostability

From a fresh agar plate, individual colonies were randomly picked and transferred to a 96 well-plate containing 200 µL of LB medium supplemented with 100 µg/mL of ampicillin. In every plate, three wells were used to inoculate the parent of each generation. Only the interior wells were used for cultivation with the perimeter wells

being filled with water. Cultures were incubated at 30°C for 24 h, at 750 rpm in Titramax 1000 shaker (Heidolph). Twenty microliters of these cultures were used to inoculate a new 96 well-plate containing 180 µL TB medium supplemented with 100 µg/mL of ampicillin and incubated at 30°C for 3h at 750 rpm. After this period 0.1% rhamnose was added to induce gene expression, and cells were harvested by centrifugation (4,000 rpm for 20 min at 4°C) after 24 h of growth. Cells were disrupted using one cycle of freezing and thawing. Cell crude extracts (20 µL) were transferred into two identical 96 well-plates. One plate was assayed directly for initial activity (A_i). The second plate was incubated at high temperatures for a defined time period, cooled on ice for 5 min and then allowed to come to room temperature for 5 min and assayed for residual activity (A_r). The thermostability values were calculated using the ratio of the residual activity to the initial activity of the mutant (m), normalized to the parent type (p) – $(A_r/A_i)_m/(A_r/A_i)_p$. The activity relative to the parent was calculated using the ratio of the initial activity of the mutant to the parent type (Ai_m/Ai_p). The mutants exhibiting the highest activity or stability were selected and re-screened, to rule out false positives. The selected hit mutants from each generation were grown overnight in 10 mL LB medium supplemented with 100 µg/mL of ampicillin, at 37°C, 120 rpm. The plasmidic DNA was extracted by using QIAprep Spin Miniprep Kit (QIAGEN) and mutations identified by DNA sequencing analysis. To identify the codons exchanges and amino acids substitutions, the Basic Local Alignment Search Tool (BLAST) was used. In each generation, the mutant with the highest stability yet unchanged activity was chosen to be the parent for the next generation.

Results and Discussion

Validation of high-throughput screenings

Implementing a screening protocol, in which a reduced risk of selecting false positives is achieved, requires an optimization of several steps throughout the process. This normally demands a validation of the screening protocol in order to identify possible variability sources which can compromise the correct data interpretation (Arnold and Georgiou, 2003a). The variability among assays must be minimized during each step of the screening process, cell growth, gene expression, lysis and activity assays. In the HTS validation experiments, multiple colonies of a single wild type clone are grown and processed in 96-well plates, just as the mutants that will be cultivated and assayed (Arnold and Georgiou, 2003a).

Selection of appropriate recombinant strains and substrates for enzymatic assays

PpAzoR is able to reduce a broad range of structurally diverse azo dyes under anaerobic conditions, but it exhibits a higher activity towards quinones when compared to azo dyes (Mendes et al., 2011b). High-throughput screenings involving the reduction of azo dyes would involve the need to work in a glove box since the presence of dioxygen leads to the inhibition of the decolourisation activity, and this experimental limitation could impose significant restrictions on the number of mutants that could be screened. Therefore cell crude extracts (Figure 1) of two recombinant *E. coli* strains overexpressing *ppAzoR*, LOM528 (*E. coli* Tuner) and LOM528 (*E. coli* KRX), were tested for the reduction of one azo dye, reactive black 5 (RB5) and five quinones, anthraquinone-2-sulfonic acid (AQS), 1,4-benzoquinone (BZ), 1,2-dihydroxybenzene (catechol), 2-hydroxy-1,4-naphthoquinone (Lawsone) and 1,2-naphthoquinone-4-sulfonate (NSA) (Table 1).

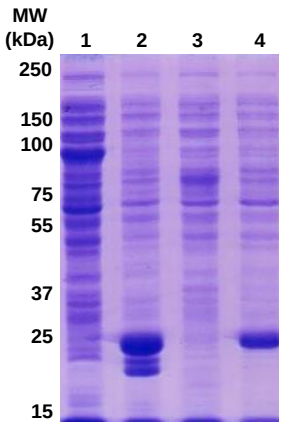

Figure 1 – SDS-PAGE analysis of cells crude extracts from (Lane1) LOM528 non-induced; (Lane 2) LOM528 induced; (Lane 3) - LOM531non-induced; (Lane 4) - LOM531 induced. The molecular mass of PpAzoR monomer is 23 kDa.

Table 1 – Specific activity of PpAzoR towards different substrates by using cell crude extracts of LOM528 and LOM531. The enzymatic assays were performed at 30°C in 100 mM sodium phosphate buffer (pH 7), in the presence of 250 µM NADPH, 100 µM of substrate, except for BZ where a 50 µM concentration was used. The reaction was followed by monitoring the decreasing in absorbance of NADPH at 340 nm.

Specific Activity (nmol/min.mg)				
	LOM528		LOM531	
Substrates	Induced	Non-Induced	Induced	Non-Induced
RB5	26 ± 1	5 ± 0.4	42 ± 8	0.3 ± 0.005
AQS	423 ± 52	253 ± 21	846 ± 106	21 ± 7
BZ	2494 ± 354	3713 ± 689	2924 ± 300	3387 ± 326
Cathecol	7 ± 0.2	14 ± 0.2	16 ± 2	6 ± 0.2
Lawsone	530 ± 42	192 ± 40	729 ± 199	173 ± 107
NSA	327 ± 9	397 ± 94	610 ± 150	23 ± 3

In crude cell extracts of induced vs. non-induced LOM528 strain only 5- and 2-fold higher activities were measured for RB5 and AQS and Lawsone reduction, respectively. No major differences in the activities are observed between the induced and non induced cell extracts for the other substrates tested. These should be related with low levels of expression of the heterologous gene in this strain and/or to high levels of intrinsic activity of *E.coli* azoreductase (Liu et al., 2009). Cell crude extracts of strain LOM531 shows 3 to 140 times higher activities for all the substrates tested with the exception of BZ in the induced when compared to the non induced cell extracts. These results lead to the decision of selecting strain LOM531 for further studies. Moreover, *E. coli* KRX strain has the advantage when compared with the Tuner strain to act as a cloning as well as an expression strain, avoiding additional steps of plasmid transference among strains in the evolution process. For the LOM531 strain, the RB5 is the substrate that shows a higher difference in the specific activity between the induced (42 ± 8 nmol/min.mg) and the not induced strain (0.3 ± 0.005 nmol/min.mg). However, the specific activity for AQS is 20 times higher when compared to RB5 (846 ± 106 nmol/min.mg) and also show a 40-fold between the induced and the not induced extracts. Moreover, while the assays for RB5 need anaerobic conditions, reduction of AQS occurs in the presence of dioxygen. Overall our results show that the strain

LOM531 is the most appropriate for the overproduction of PpAzoR and AQS is the most suitable substrate for a HTS protocol using 96-well plates.

Cell growth and cell disruption in 96-well plates

The second step was to choose the more suitable medium for PpAzoR overproduction in 96-well plates. LOM531 was grown in LB and TB medium and harvested by centrifugation. A cell disruption procedure was applied since PpAzoR is an intracellular protein. The B-PER® lysis solution, containing a mild and non ionic detergent was used, because it is an easy and fast method already established in the laboratory. Protein concentration and activities using AQS as substrate were tested in cell crude extracts. However, we could not detect any difference in the enzymatic activity in crude extracts of induced or non induced cells. We have analysed the extracts by SDS-PAGE and a band with the molecular mass of PpAzoR was observed in induced extracts. In order to rule out any problems associated with enzymatic assays induced extracts from cultures grown at a larger scale and disrupted in a French-Press were tested in 96-well plates. A significant difference in enzymatic activities was measured between the induced (68 ± 11 nmol/min.mg) and the non induced (3 ± 0.4 nmol/min.mg) cell crude extracts. Therefore it seemed that the detergent used for cell disruption in 96-well plates could have lead to PpAzoR inactivation. To test this hypothesis cells were disrupted using Lysozyme, an enzyme that hydrolyses the 1,4-β linkages connecting N-acetylmuramic acid and N-acetylglucosamine in peptidoglucan, found in bacterial cell walls. Using this method we found that induced crude extracts of cells grown in TB media, shows 10 times higher activities when compared to the non induced strains (65 ± 30 vs. 6 ± 2 nmol/min.mg). On the other hand, no difference between the induced and not induced crude extracts was observed when LB was used (8 nmol/min.mg).

Cell disruption by lysozyme is a relative costly and time consuming method that would affect the number of mutants that could be screened simultaneously. Therefore, a physical method, in which the cell pellets are subject to cycles of freezing and thawing was tested. The application of this method results in the formation of small pores in the cytoplasmic membrane which allows soluble proteins produced in high levels to escape from the intracellular milieu (Johnson and Hecht, 1994). The 96 well-plates were submerged in liquid nitrogen or incubated at -80°C for 15 minutes, and thawed at room temperature. The protein concentration in crude extracts was lower than the one observed when enzymatic lysis was used (Table 2). Accordingly to Johnson and Hecht (1994), the lower protein yield is due to limitations of the membrane pores size induced by freeze and thaw that would retain proteins with larger molecular masses. This selective cut-off eventually explains the higher specific activity achieved in induced crude extracts subjected to this treatment (Table 2). Besides, these are easy, quickly and cheaper protocols when compared to the chemical or enzymatic applied.

Consequently this method was chosen to be used in the HTS. Freezing with liquid nitrogen was preferred since it is less time consumption, relatively to incubation at -80°C.

Table 2 – Protein content and activity in cell crude extracts. Cells were disrupted using chemical (B-PER® lysis solution), enzymatic (Lysozyme) or physical (freeze/thaw using Liquid Nitrogen or incubation at -80°C) methods.

Disruption method		Protein (mg/mL)	Specific activity (nmol/min.mg)
Chemical	B-PER	0.4 ± 0.1	16 ± 4
Enzymatic	Lysozyme	1.2 ± 0.2	65 ± 30
Physical	Freeze/Thaw (LN)	0.2 ± 0.1	266 ± 55
	Freeze/Thaw (-80°C)	0.2 ± 0.1	243 ± 88

LN – Liquid Nitrogen

Activity assays in 96-well plates

After these various optimization steps we have tested the reproducibility of *ppAzoR* expression system using six 96-well plates. We have measured OD_{600nm}, total protein content and enzymatic activities for AQS using crude extracts of LOM531 cells grown in TB medium and using the freeze/thaw method to disrupt the cells. We initially found very high coefficients of variance (CV = standard deviation/mean x 100 %) for cell concentration and protein content (Table 3).

Table 3 – Growth, protein and specific activity parameters from cells grown in 96 well plates. Data obtained from the whole 96 wells (A), 60 interiors wells (B) and after incubation at 55°C during 1 h (C).

	Cell concentration		Protein		Specific Activity	
	OD_{600nm}	CV (%)	(mg/mL)	CV (%)	(nmol/min.mg)	CV (%)
A	0.6 ± 0.2	33	0.2 ± 0.1	53	266 ± 55	21
B	1.0 ± 0.1	10	0.5 ± 0.1	15	542 ± 91	17
C	-	-	-	-	174 ± 30	17

Higher OD_{600nm} and protein concentration levels in wells located in the perimeter of the plate were found when compared to the interiors wells. This variability was

previously reported (Arnold and Georgiou, 2003a; Garcia-Ruiz et al., 2010) and represents a high risk of choosing false positives. We have filled with H_2O the perimeter cells and the plates were sealed with a foil and a plastic cover, closed with parafilm. Noteworthy when only the inner wells were used for cell cultivation, the CV values were reduced more than 50% (Table 3). Besides, a better growth throughout the plate was observed, which resulted in doubling the average of the protein produced and a CV of ≈ 15% was achieved.

For the set-up of the thermostability assays several incubation conditions were tested and we have selected the conditions that lead to a residual activity is about one-third of the initial activity as proposed by Arnold and Georgiou (2003a): 55°C for 1 h. This condition was selected for screening the first generation (Table 3).The final values of thermostability came from the ratio of residual (r) activity to initial (i) activity (Ar/Ai) normalised with the corresponding parent strain and a CV of ≈ 17% was achieved.

Overall our data show the importance of performing a validation step before implementing a HTS protocol for a directed evolution approach. Following the approaches described it was possible to improve the growth conditions, the protein production, the activities and to reduce the coefficient of variation (Table 3) and will help on the correct interpretation of the collected data.

Library construction by epPCR

Libraries of mutants were constructed by error-prone PCR using Taq DNA polymerase. The mutation rate was tuned by varying the $MnCl_2$ concentration from 0.1 to 0.25 mM of $MnCl_2$. The objective was to produce libraries with 30-45% of clones showing less than 10% of activity of the Wt activity; this usually corresponds to 1-3 amino acid changes (1-5 nucleotide substitution/gene) (Arnold and Georgiou, 2003a).

Enzymatic assays were performed in the four libraries constructed (Figure 2). We found that increasing the concentration of $MnCl_2$ results in increased numbers of inactive clones in the libraries; 10, 15, 30 and 73% of clones with less than 10% of wild type activity was found in libraries with 0.1, 0.15, 0.2 and 0.25 mM of $MnCl_2$, respectively.

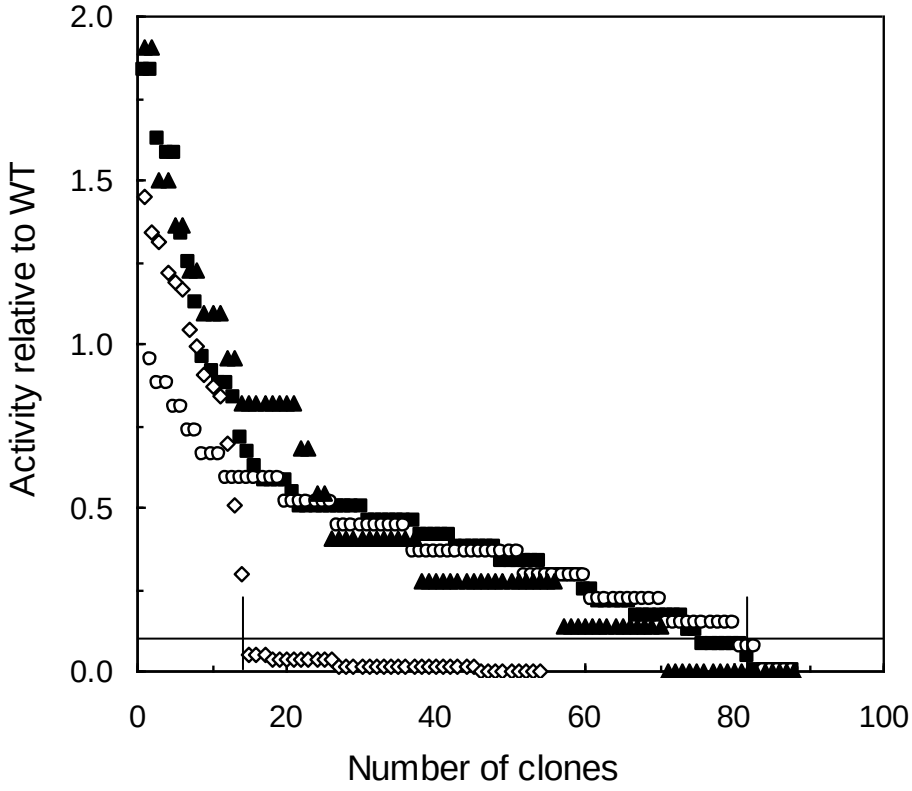

Figure 2 – Landscape for four mutant libraries using different $MnCl_2$ concentration: (■) 0.1 mM, (□) 0.15 mM, (▲) 0.2 mM, and (◊) 0.25 mM. Activity of clones relative to the wild type is plotted in descending order.

Based in these results we have decided to use 0.2 mM of $MnCl_2$ in the ep-PCR protocol to create the PpAzoR mutants libraries since this concentration allows to obtain the desired number of inactive mutants (30%) presumably having to 1-3 amino acid changes in the PpAzoR molecule.

Directed evolution of PpAzoR towards increased thermostability

Almost 10,000 clones were explored in four rounds of molecular evolution to improve the thermostability of PaAzoR.

In the first generation, 2214 clones were screened (Figure 3A) and 18 variants were identified with improvements in the thermostability or in the initial activity. In order to rule out false positives, and confirm the reliability of these variants, a re-screening was performed in the most promising clones (Figure 3B).

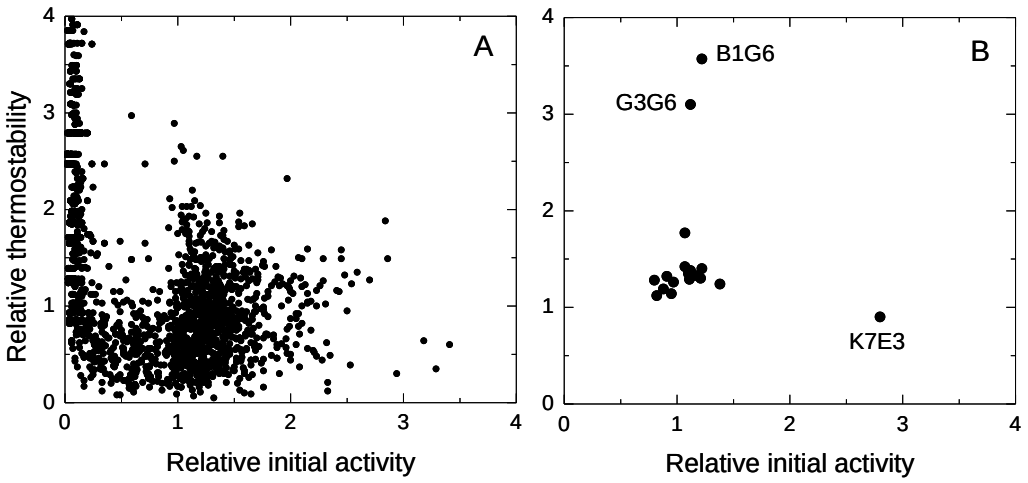

Figure 3 – Directed evolution landscape for the first generation mutant library. A – Initial activity *vs* thermostability of 2214 clones screened relative to the wild type. B – Re-screening of the best mutants identified in the first generation. Stability is measured by the ratio of residual activity following incubation at 55°C for 1 h to initial activity.

Three variants were found to be positive hits: two variants, named B1G6 and G3G6, exhibited an improvement in thermostability of around 3.5-fold when compared to the wild type, while maintaining a similar enzymatic activity (Figure 3B and Table 4). On the other hand, it was possible to select a variant (K7E3) showing a 3-fold higher activity than the wild type strain without a decreased thermal stability.

The sequence information of the variants showed that the mutation rate was one to two nucleotide substitutions per gene. The higher thermostable variants, B1G6 and G3G6, shared a common amino acid substitution, Q192R, and the most active variant, K7E3, had the Y179H mutation (Table 4). We have decided to use the B1G6 variant as parent for the second generation of directed evolution.

In the second generation, 2052 clones were screened for higher stability and the heat treatment was adjusted to 55°C for 90 min (Figure 4A). In this generation, 9 variants exhibited improvement in the thermostabililty or in the initial activity and were re-screened (Figure 4B).

Table 4 – Summary of library screening conditions, amino acids substitutions accumulated in PpAzoR variants and initial activity and thermostability relative to their parents. The parent for the next generation are in bold.

Generation	Temperature and incubation period	Variants	Mutations	Initial activity relative to parent	Thermostability relative to parent
1st	55°C, 60 min	**B1G6**	Q192R, 1 silence	1.2 ± 0.05	3.7 ± 0.5
		G3G9	Q192R	1.2 ± 0.07	3.2 ± 0.5
		K7E3	Y179H	2.8 ± 0.2	0.8 ± 0.3
2nd	55°C, 90 min	**16B7**	Q192R, A46P, V159A, 2 silence	0.7 ± 0.1	3.2 ± 0.9
		37C9	Q192R, Y120C, 1 silence	0.3 ± 0.1	2.9 ± 0.7
		12B8	Q192R, Y179H, 2 silence	5.4 ± 0.5	0.6 ± 0.08
		12F9	Q192R, Y179H, 3 silence	5.6 ± 0.5	0.6 ± 0.1
3rd	60°C, 45 min	**23C10**	Q192R, A46P, V159A, C129S, 3 silence	1.4 ± 0.05	2.5 ± 0.09
		16B6	Q192R, A46P, V159A, A48P, 2 silence	0.9 ± 0.05	2.6 ± 0.4
		19 E4	Q192R, A46P, V159A, Y179H, 3 silence	3.2 ± 0.2	1.4 ± 0.14
		31D6	Q192R, A46P, Y179H, 6 silence	4.0 ± 0.3	0.8 ± 0.2
4th	80°C, 60 min	13G10	Q192R, A46P, V159A, C129S, D7H, A178D, 3 silence	0.8 ± 0.07	3.9 ± 0.2
		6F10	Q192R, A46P, V159A, C129S, N14D, L143Q, 5 silence	1.0 ± 0.1	3.8 ± 0.2
		27E4	Q192R, A46P, V159A, C129S, L161M, L169P, 3 silence	0.9 ± 0.01	3.1 ± 0.2
		23C5	Q192R, A46P, V159A, C129S, K74E, A88G, 3 silence	0.9 ± 0.07	2.9 ± 0.4
		10F8	Q192R, A46P, V159A, C129S, N131D, 3 silence	0.8 ± 0.1	2.8 ± 0.5
		1C11	Q192R, A46P, V159A, C129S, E36D, L143Q, 4 silence	1.6 ± 0.4	2.8 ± 0.3
		14D4	Q192R, A46P, V159A, C129S, A77T, N131D, 3 silence	1.0 ± 0.1	2.5 ± 0.2
		32F5	Q192R, A46P, V159A, C129S, I6V, T79R, Y179H, 4 silence	2.5 ± 0.2	1.4 ± 0.2
		23E4	Q192R, A46P, V159A, C129S, Y179H, 3 silence	2.1 ± 0.2	0.9 ± 0.2

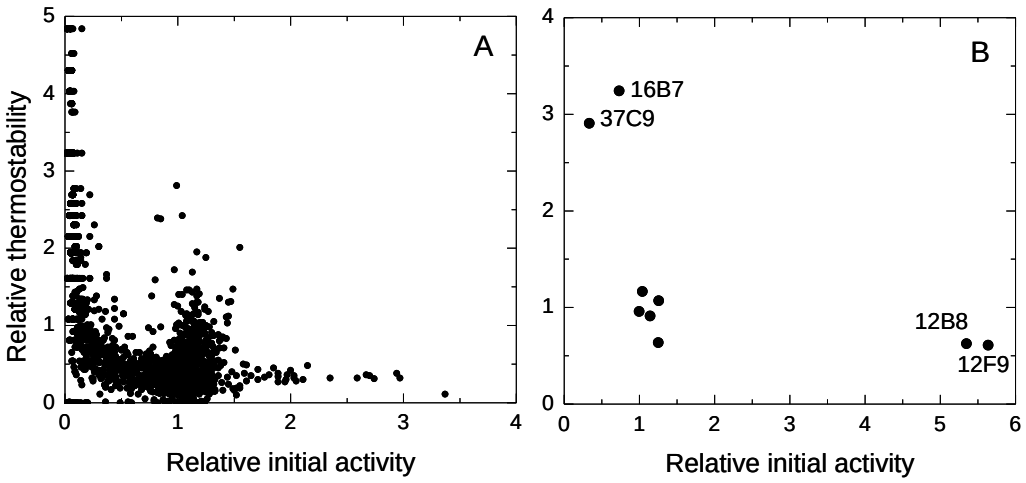

Figure 4 – Directed evolution landscape for the second generation mutant library. A – Initial activity *vs* thermostability of 2052 clones screened relative to the B1G6. B – Re-screening of the best mutants identified. Stability is measured by the ratio of residual activity following incubation at 55°C for 90 min to initial activity.

Four variants were positive hits: the 16B7 and 37C9 variants with 3.2-fold and 2.9-fold higher thermostability than the B1G6 variant, respectively and the 12B8 and 12F9 variants with 5.4-fold and 5.6-fold higher initial activity than the B1G6 variant, respectively (Figure 4B and Table 4). However, for all these variants the improvement in one property came with the cost of reducing the other (following the well known tradeoff that usually appears between activity and stability for many single point mutations (Garcia-Ruiz et al., 2010). The initial activity of 16B7 and 37C9 variants was 0.7-fold and 0.3-fold, respectively, relative to the parent for this generation, variant B1G6, and the thermostability of 12B8 and 12F9 variants was 0.6-fold of the B1G6 variant. The sequence analysis of these variants revealed the following mutations Q192R, A46P, V159A for variant 16B7, Q192R, Y120C for variant 37C9 and Q192R, Y179H for variants 12B8 and 12F9. The parent variant for the third generation was 16B7, since it was the one that retains a higher activity among the identified hits.

In the third generation 2160 clones were screened and the heat treatment was adjusted to 60°C for 45 min (Figure 5A). Nine variants showed higher thermostability or higher initial activity as compared with the parental strain. Their re-screening allowed identifying 4 positive hits (Figure 5B).

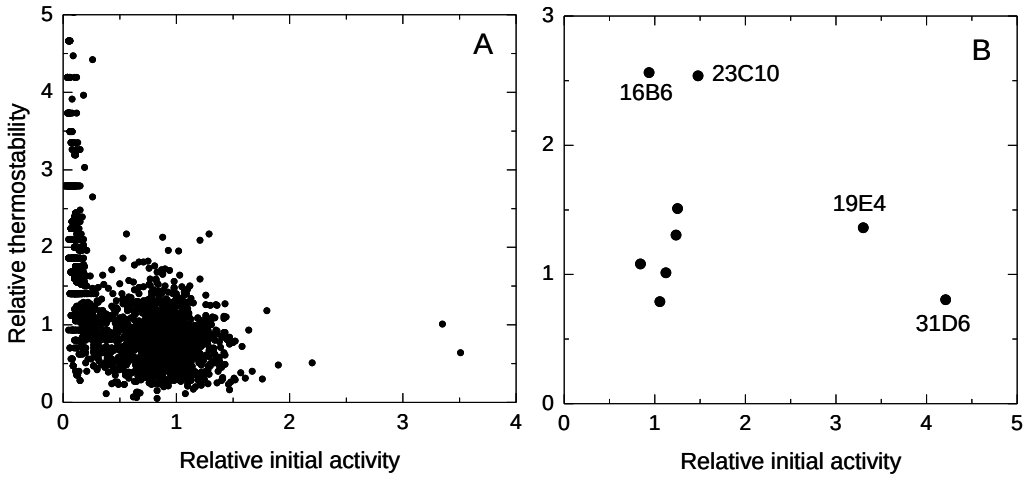

Figure 5 – Directed evolution landscape for the third generation mutant library. A – Initial activity *vs* thermostability of 2160 clones screened relative to 16B7. B – Re-screening of the best mutants identified. Stability is measured by the ratio of residual activity following incubation at 60°C for 45 min to initial activity.

Two variants, named 16B6 and 23C10, exhibited higher thermostability relative to the parental 16B7 variant, with an increase of 2.6-fold and 2.5-fold, respectively, and two other variants, named 19E4 and 31D6, had a 3.2-fold and 4-fold higher initial activity, respectively when compared to the 16B7 variant (Figure 5B and Table 4). In this generation, all new mutations incorporated did not show a negative impact either in the stability or in the activity. Interestingly, the 23C10 variant showed even a slightly improvement in the initial activity (1.4-fold *vs* parent type, 16B7) and the 19E4 active variant exhibited an improvement in the thermostability (1.4-fold *vs* parent type, 16B7). The sequence analysis of these mutants revealed the following mutations Q192R, A46P, V159A, A48P (16B6), Q192R, A46P, V159A, C129S (23C10), Q192R, A46P, V159A, Y179H (19E4) and Q192R, A46P, Y179H (31D6). The 23C10 variant showed the highest thermostability and also a small improvement in activity, and therefore it was selected to be the parent for the fourth generation.

In the fourth generation 2160 clones were screened and the heat treatment was adjusted to 80°C for 1 h (Figure 6A). Eighteen variants showed a higher thermostability or initial activity than the parental strain (23C10 variant) and a re-screening was performed. All the variants were positive hits (Figure 6B) and 9 variants (seven with higher stability and two with higher activity than 23C10 variant) were sequenced.

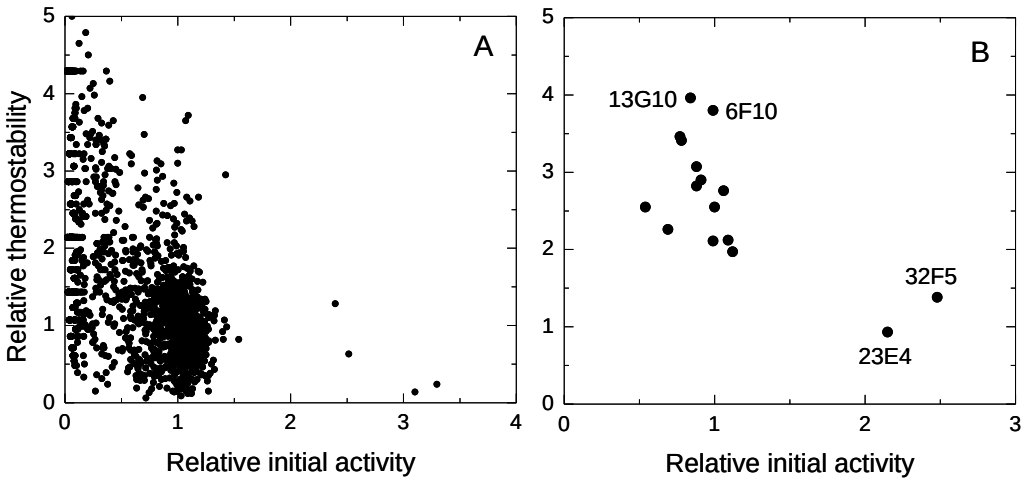

Figure 6 – Directed evolution landscape for the fourth generation mutant library. A – Initial activity *vs* thermostability of 2160 clones screened relative to the 23C10. B – Re-screening of the best mutants identified in the first generation. Stability is measured by the ratio of residual activity following incubation at 80°C for 1h to initial activity.

The 1C11, 6F10, 10F8, 13G10, 14D4, 23C5 and 27E4 variants had an improvement in thermostability while the activity was maintained (Table 4). The most thermostable variants were the 6F10 and the 13G10, with 3.8-fold and 3.9-fold when compared to the 23C10 variant, respectively (Figure 6B and Table 4). The two variants with higher activity, named 23E4 and 32F5, showed an increase of 2.1-fold and 2.5-fold relative to 23C10 variant, respectively, without losing stability. Sequencing these mutants revealed the following changes Q192R, A46P, V159A, C129S, E36D, L143Q (1C11), Q192R, A46P, V159A, C129S, N14D, L143Q (6F10), Q192R, A46P, V159A, C129S, N131D (10F8), Q192R, A46P, V159A, C129S, D7H, A178D (13G10), Q192R, A46P, V159A, C129S, A77T, N131D (14D4), Q192R, A46P, V159A, C129S, K74E, A88G (23C5), Q192R, A46P, V159A, C129S, L161M, L169P (27E4), Q192R, A46P, V159A, C129S, Y179H (23E4) and Q192R, A46P, V159A, C129S, I6V, T79R, Y179H (32F5).

Evaluation of activity and stability of the best variants of each generation

In summary, after four rounds of random mutagenesis by ep-PCR 18 hits were selected with increased initial activity and/or thermostability (Figure 7). The selection pressure was progressively increased from 55°C in 1st generation to 80°C in the 4th generation. This approach was possible because the improvements accomplished in each generation were sufficient to retain more than 30% of residual activity. Interestingly, even if the initial goal of this work was to improve the thermostability of PpAzoR, it was also possible to select variants with higher activity than the wild type strain.

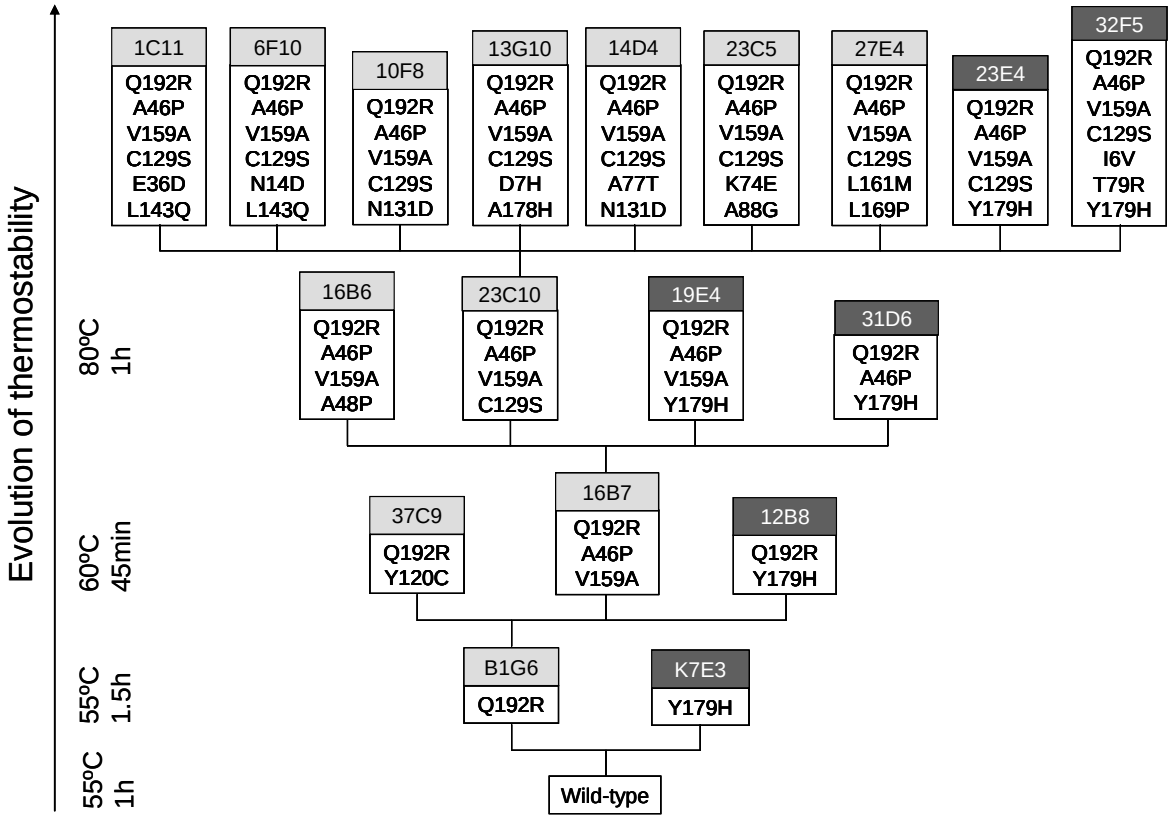

Figure 7 – Lineage of PpAzoR mutants generated in this study. Only non-synonymous mutations are shown. The mutants with higher stability are in light grey and the mutants with higher activity are in dark grey.

The variants with the highest initial activity (K7E3, 12B8, 31D6 and 32F5) and the highest thermostability (B1G6, 16B7, 23C10 and 13G10) of each generation were grown in a larger scale and enzymatic reactions were performed using the crude cell extracts (Figure 8 and 10). It was noticed that after cells disruption the color of the crude extracts of variants were slightly less yellow than the wild enzyme. For this reason, the enzymatic reactions were also performed in the presence of flavin mononucleotide (FMN).

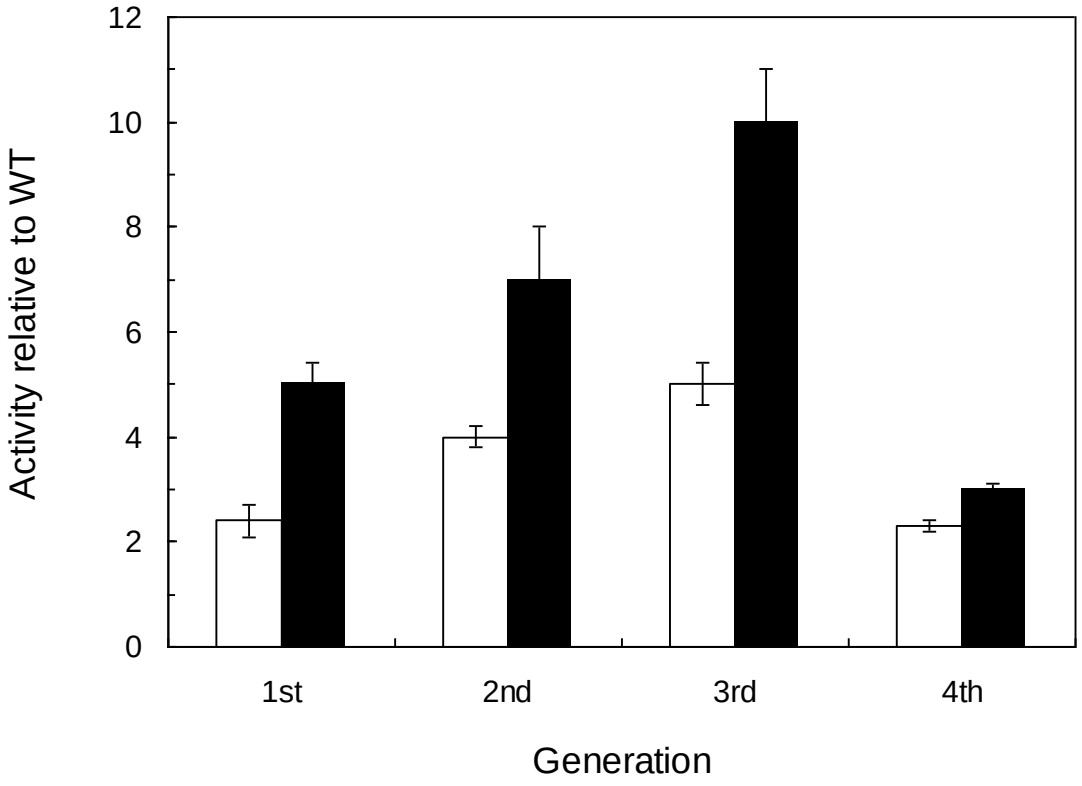

Figure 8 – Initial activity relative to wild type from the variants with the highest activity of each generation without addition of FMN (white) or with 20 µM of FMN (black): 1st (K7E3), 2nd (12B8), 3rd (31D6) and 4th (32F5). Reactions were performed in triplicates in 100 mM sodium phosphate buffer (pH 7), 100 µM AQS, 250 µM NADPH.

All variants tested (K7E3, 12B8, 31D6 and 32F5) showed a 2 to 5-fold increased initial activity when compared to the wild type strain and the variant 31D6 from the third generation is the one with the highest activity measured (Figure 7). Moreover, all variants, except 32F5 (4th generation), show a 2-fold increase in activity when FMN is added to the reaction. The effect of FMN in the activity of the wild type enzyme was also evaluated, but its activity remained unchanged (data not shown). Possibly the introduced mutations caused some altered structural conformations that result in proteins showing a lower FMN occupancy than wild type (Maier et al., 2004; Wang et al., 2007). In the case of the 32F5 variant (4th generation), the mutations I6V and/or T79R could had contributed to reduce this constrain, however they also lead to lower activity.

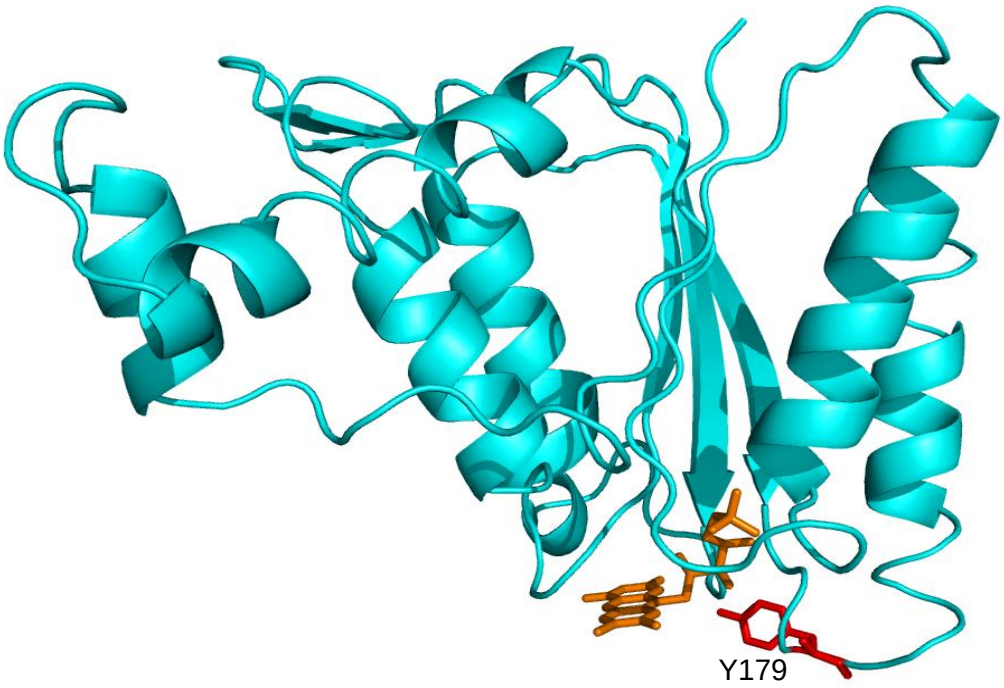

Figure 9 – Overview of the amino acid Y179 identified to be important in activity (in red) in PpAzoR. The FMN group is represented in orange.

Noteworthy, all the variants exhibiting highest activities share the Y179H mutation. The analysis of PpAzoR crystal structure (Figure 9) shows that this residue is located in close proximity (6.6 Å) to the FMN binding site. Presumably, the smaller size of the His residue in comparison to Tyr may enlarge the cavity surrounding FMN, enhancing the accessibility to FMN, and thus contributing to an increased PpAzoR activity.

The variants that exhibited the highest thermostability in each generation (B1G6, 16B7, 23C10 and 13G10) were analysed in terms of their thermostability in relation to the wild type, after incubation at 55°C for 30 min (Figure 10).

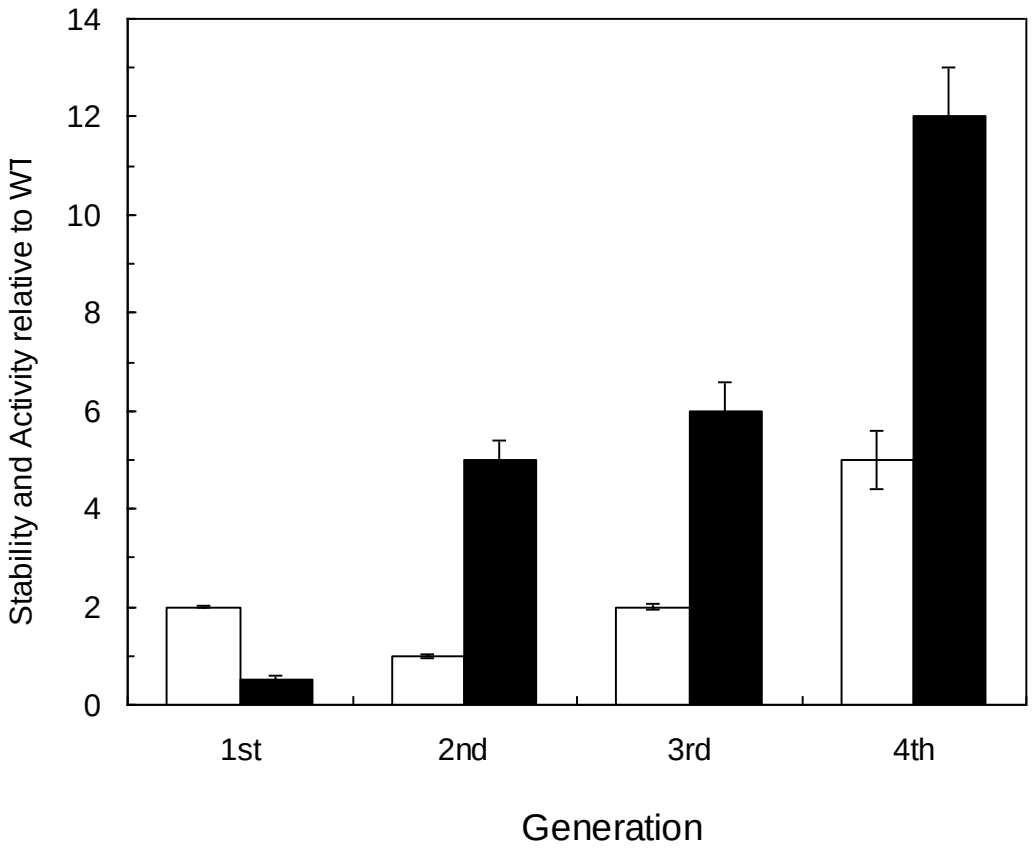

Figure 10 – Activity (white) and stability (back) of variants with the highest stability from each generation in relation to the wild type enzyme: 1st (B1G6), 2nd (16B7), 3rd (23C10) and 4th (13G10). Reactions were performed in triplicates in 100 mM sodium phosphate buffer (pH 7), 100 µM AQS, 250 µM NADPH and 20 µM FMN. The stability was determined after incubation at 55°C for 30 min.

The B1G6 variant, from the first generation, has a very low residual activity, similar to the wild type after incubation at 55°C for 30 min. In order to check the thermostabilty improvement, both wild type and variant were incubated at 50°C for 30 min. At these conditions a 1.5-fold improvement in stability was measured for the B1G6 variant. It was possible to confirm an increase in stability in each generation as observed before in the screenings performed in 96-well plates.

In nature, mutations that join activity and stability are rare because of the genetic drift and the low frequency of simultaneous selective pressure towards both features (Arnold and Georgiou, 2003a). Indeed, it has been proved quite difficult for protein engineers to find single mutations which improve both properties simultaneously. Remarkably in this work it was possible to increase the activity and stability at the same time. A trade-off between stability and activity was noticed only in the second generation. Somehow the interactions caused by the A46P and V159A mutations decreased the activity of the 16B7 variant while increasing 5-fold its stability. The best variant from the third generation, 23C10, acquired the C129S mutation improving 6-

fold its stability as compared with the wild type. The 13G10 variant, from the fourth generation, was the one with the highest stability (12-fold), as previously mentioned, and it also showed a 2-fold increase in its activity. The initial activity of these four variants also doubled in the presence of FMN.

Structural interpretation of mutations that improve stability

The 18 mutations introduced by directed evolution and responsible for an increase in thermostability were mapped using the crystal structure of PpAzoR isolated from *Pseudomonas putida* MET94 (Figure 11).

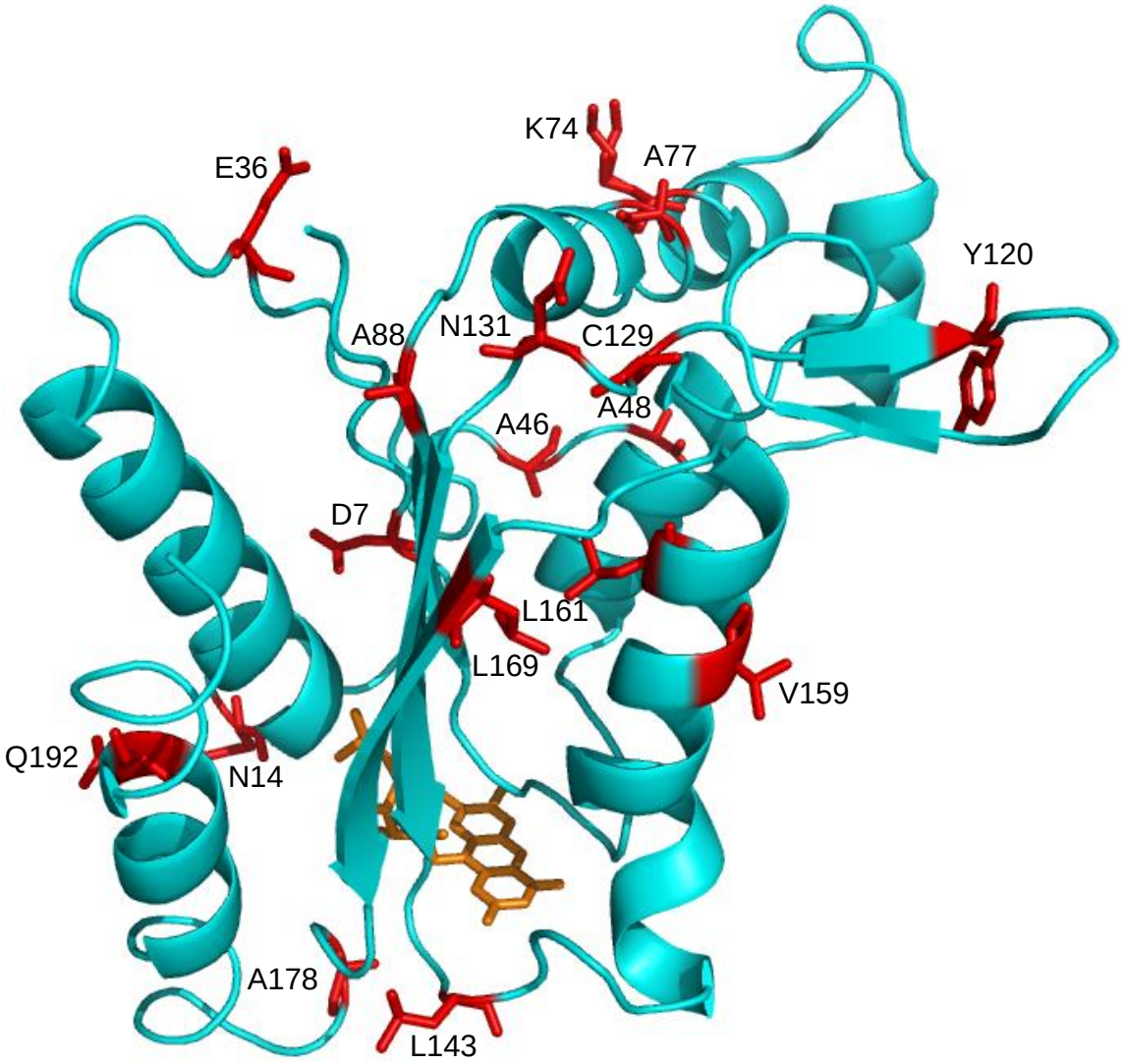

Figure 11 – Structural mapping of the amino acids identified to be important in thermostability (in red) in PpAzoR. The FMN group is represented in orange.

All the mutations identified are distributed in different regions of the protein but located on or near the surface of the enzyme (Figure 11 and Table 5). In addition the majority (12) occur in loops. Previous studies found that the thermostability of some

enzymes could be improved by the stabilization of large loop structures (Spiller et al., 1999).

Table 5 – Structural mapping of the amino acids with importance in the thermostability of PpAzoR.

Mutations	Localization	Secondary structure
Asp7His	Interface	Loop
Asn14Asp	Surface	Loop
Glu36Asp	Surface	Loop
Ala46Pro	Surface	Loop
Ala48Pro	Surface	Loop
Lys74Glu	Surface	Loop
Ala77Thr	Surface	Loop
Ala88Gly	Interface	Sheet
Tyr120Cys	Surface	Loop(NAD(P)H site)
Cys129Ser	Surface	Loop(NAD(P)H site)
Asn131Asp	Surface	Loop(NAD(P)H site)
Leu143Gln	Surface	Loop(NAD(P)H site)
Val159Ala	Surface	Helix
Leu161Met	Interface	Helix (dimerisation site)
Leu169Pro	Surface	Sheet
Ala178Asp	Surface	Loop
Tyr179His	Surface	Loop
Gln192Arg	Surface	Helix

While some of the mutations found during the evolution protocol can be easily interpreted on a structural basis, others are somehow difficult.

The first mutation (Q192R) that was picked up by the screening protocol is located in a α-helix. Arginine residues are known to enhance polar interactions such as hydrogen bonds and/or salt bridges and consequently to increase the stability. The A46P mutation picked up in the second generation (16B7 variant) occurs in a region of the protein that as the longest stretch of 20 amino acid residues without any regular secondary structure. Proline residues are known to stabilise the folded states of proteins compared to the unfolded state due to the constraints they introduce in the backbone conformation. A similar mutation (A48P) was found in the mutant 16B6 in the third generation. The stabilization by the introduction of Pro residues into loop regions is quite well-documented (Matthews, 1993; Watanabe et al., 1996). Thus, the replacement of alanine to proline in the longest loop of the molecule might contribute

to the stability of the protein. The Y120C mutation picked up in the second generation (37C9 variant) occurs in the putative NAD(P)H binding motif (Mendes et al., 2011b) and contributes to an increased protein stabilization but simultaneously cause a severe drop in its activity. Several other mutations were found in this region such as, C129S in the third generation (23C10 variant), N131D (10F8 and 14D4 variants) and L143Q (1C11 and 6F10) in the fourth generation. These findings suggest that the stability of the NAD(P)H binding site leads to an overall improvement of the protein thermostability. Interestingly, in the last generation six from the seven variants with higher thermostability (1C11, 6F10, 13G10, 14D4, and 23C5) have a mutation to a negatively charged residue on the protein surface. It seems that this type of mutation, together with ones accumulated in the first three generations, confer stability to the protein. The 27E4 variant shows two new mutations, L161M located in the region of the helix α5 putatively involved in the dimerisation of the two monomers of the enzyme and L169P located in a β-sheet and 3.8 Å from the L161. Mutation in this region of the enzyme may stabilise the dimer formation, resulting in a more stable enzyme.

The use of random mutagenesis to explore the mechanism of thermostability is an attractive approach, since this methodology introduces a small number of amino acid changes, which could be directly linked to increase stability. By accumulating a relatively small number of mutations both the thermostability and the activity of PpAzoR were increased.

Conclusions

Directed evolution by random mutagenesis has proven to be an important tool for improving the thermostability of PpAzoR, by allowing the accumulation of small number of amino acid changes in each generation. By using this approach and using highly controlled screening conditions, we were able to improve not only the thermostability but also the activity of PpAzoR. Throughout the process we were able to identify and eliminate sources of variability normally responsible for misleading data interpretation, and these allowed us to create an efficient HTS protocol. After four rounds of ep-PCR we identified 20 hits with improved thermostability or activity. By DNA sequencing a total of 18 amino acids substitutions throughout the PpAzoR sequence were identified as thermostable since they were found in variants with higher stability. On the other hand, a single amino acid substitution was identified in mutants with higher activity, Y179H, which is located in close proximity with the FMN binding site.

It has being reported that in principle, it is easier to improve the thermostability while maintaining the activity than vice versa, due to the trade-off that usually appears between the activity and stability for many single mutations (Garcia-Ruiz et al., 2010). Although this trade-off was identified in hits from the second generation, the vast majority of hits showed improvements in one of the property while maintaining the other one, and we identified a variant (13G10) in the 4^{th} generation with improved thermostability and activity, which is the best hit identified so far.

Future work

In this project we were able to improve the PpAzoR thermostability, activity or both properties simultaneously and ep-PCR enabled the accumulation of mutations that lead to these improvements. Therefore it would be interesting to use DNA shuffling of some variants of the fourth generation to find mutants with improved activity towards azo dyes or to engineer a PpAzoR NADH-dependent.

References

Acharya, P., Rajakumara, E., Sankaranarayanan, R., Rao, N.M., 2004. Structural Basis of Selection and Thermostability of Laboratory Evolved Bacillus subtilis Lipase. Journal of Molecular Biology 341, 1271-1281.

Arnold F.H., Georgiou G., 2003. Directed Enzyme Evolution: screening and selection methods. Methods in Molecular Biology Humana Press, Totowa, New Jersey, 230.

Arnold F.H., Georgiou G., 2003. Directed Evolution Library creation: methods and protocols. Methods in Molecular Biology Humana Press, Totowa, New Jersey, 231.

Bloom, J.D., Arnold, F.H., 2009. In the light of directed evolution: Pathways of adaptive protein evolution. Proceedings of the National Academy of Sciences 106, 9995-10000.

Bloom, J.D., Labthavikul, S.T., Otey, C.R., Arnold, F.H., 2006. Protein stability promotes evolvability. Proceedings of the National Academy of Sciences 103, 5869-5874.

Bloom, J.D., Meyer, M.M., Meinhold, P., Otey, C.R., MacMillan, D., Arnold, F.H., 2005. Evolving strategies for enzyme engineering. Current Opinion in Structural Biology 15, 447-452.

Bommarius, A.S., Broering, J.M., Chaparro-Riggers, J.F., Polizzi, K.M., 2006. High-throughput screening for enhanced protein stability. Current Opinion in Biotechnology 17, 606-610.

Brakmann, S., 2001. Discovery of Superior Enzymes by Directed Molecular Evolution. ChemBioChem 2, 865-871.

Chen, H., 2006. Recent advances in azo dye degrading enzyme research. Curr Protein Pept Sci 7, 101-111.

Correia, B., Chen, Z., Mendes, S., Martins, L.O., Bento, I., 2011. Crystallization and preliminary X-ray diffraction analysis of the azoreductase PpAzoR from Pseudomonas putida MET94. Acta Crystallographica Section F 67, 121-123.

Dalby, P.A., 2003. Optimising enzyme function by directed evolution. Current Opinion in Structural Biology 13, 500-505.

Deller, S., Macheroux, P., Sollner, S., 2008. Flavin-dependent quinone reductases. Cell Mol Life Sci 65, 141-160.

dos Santos, A.B., Cervantes, F.J., van Lier, J.B., 2007. Review paper on current technologies for decolourisation of textile wastewaters: Perspectives for anaerobic biotechnology. Bioresource Technology 98, 2369-2385.

Flores, H., Ellington, A.D., 2002. Increasing the thermal stability of an oligomeric protein, beta-glucuronidase. Journal of Molecular Biology 315, 325-337.

Forgacs, E., Cserháti, T., Oros, G., 2004. Removal of synthetic dyes from wastewaters: a review. Environment International 30, 953-971.

Garcia-Ruiz, E., Mate, D., Ballesteros, A., Martinez, A., Alcalde, M., 2010. Evolving thermostability in mutant libraries of ligninolytic oxidoreductases expressed in yeast. Microbial Cell Factories 9, 17.

Johnson, B.H., Hecht, M.H., 1994. Recombinant Proteins Can Be Isolated from E. coli Cells by Repeated Cycles of Freezing and Thawing. Nat Biotech 12, 1357-1360.

Kandelbauer, A., Guebitz, G.M., 2005. Bioremediation for the Decolorization of Textile Dyes - A Review. Environmental Chemistry, 269-288.

Khersonsky, O., Tawfik, D.S., 2010. Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu Rev Biochem 79, 471-505.

Kuchner, O., Arnold, F.H., 1997. Directed evolution of enzyme catalysts. Trends in Biotechnology 15, 523-530.

Leelakriangsak, M., Huyen, N.T., Towe, S., van Duy, N., Becher, D., Hecker, M., Antelmann, H., Zuber, P., 2008. Regulation of quinone detoxification by the thiol stress sensing DUF24/MarR-like repressor, YodB in Bacillus subtilis. Mol Microbiol 67, 1108-1124.

Liu, G., Zhou, J., Fu, Q., Wang, J., 2009. The Escherichia coli AzoR is involved in resistance to thiol.specific stress caused by electrophilic quinones. J Bacteriol 191, 6394-6400.

Maier, J., Kandelbauer, A., Erlacher, A., Cavaco-Paulo, A., Gubitz, G.M., 2004. A new alkali-thermostable azoreductase from Bacillus sp. strain SF. Appl Environ Microbiol 70, 837-844.

Matthews, B.W., 1993. Structural and Genetic Analysis of Protein Stability. Annual Review of Biochemistry 62, 139-160.

McMullan, G., Meehan, C., Conneely, A., Kirby, N., Robinson, T., Nigam, P., Banat, I.M., Marchant, R., Smyth, W.F., 2001. Microbial decolourisation and degradation of textile dyes. Applied Microbiology and Biotechnology 56, 81-87.

Mendes, S., Farinha, A., Ramos, C.G., Leitão, J.H., Viegas, C.A., Martins, L.O., 2011a. Synergistic action of azoreductase and laccase leads to maximal decolourization and detoxification of model dye-containing wastewaters. Bioresource Technology 102, 9852-9859.

Mendes, S., Pereira, L., Batista, C., Martins, L., 2011b. Molecular determinants of azo reduction activity in the strain Pseudomonas putida MET94. Applied Microbiology and Biotechnology 92, 393-405.

Miyazaki, K., Takenouchi, M., Kondo, H., Noro, N., Suzuki, M., Tsuda, S., 2006. Thermal Stabilization of Bacillus subtilis Family-11 Xylanase by Directed Evolution. Journal of Biological Chemistry 281, 10236-10242.

Nakanishi, M., Yatome, C., Ishida, N., Kitade, Y., 2001. Putative ACP phosphodiesterase gene (acpD) encodes an azoreductase. J Biol Chem 276, 46394-46399.

Nobeli, I., Favia, A.D., Thornton, J.M., 2009. Protein promiscuity and its implications for biotechnology. Nat Biotech 27, 157-167.

Pearce, C.I., Lloyd, J.R., Guthrie, J.T., 2003. The removal of colour from textile wastewater using whole bacterial cells: a review. Dyes and Pigments 58, 179-196.

Pereira, L., Coelho, A., Viegas, A.C., Ganachaud, C., Lacazio, G., Tron, T., Robalo, M., Martins, L.O., 2009a. On the mechanism of biotransformation of the anthraquinonic dye Acid Blue 62 by laccases. Adv Synth Catal 351, 1857-1865.

Pereira, L., Coelho, A.V., Viegas, C.A., Santos, M.M., Robalo, M.P., Martins, L.O., 2009b. Enzymatic biotransformation of the azo dye Sudan Orange G with bacterial CotA-laccase. J Biotechnol 139, 68-77.

Rai, H.S., Bhattacharyya, M.S., Singh, J., Vats, P., Banerjee, U.C., 2005. Removal of dyes from the effluent of textile and dyestuff manufacturing industry: a review of emerging techniques with reference to biological treatment. Critical Rev. Environ Sc. Tecnhol. 35, 219-238.

Reetz, M.T., Kahakeaw, D., Lohmer, R., 2008. Addressing the Numbers Problem in Directed Evolution. ChemBioChem 9, 1797-1804.

Rodriguez-Couto, S., 2009. Enzymatic biotransformation of synthetic dyes. Curr Drug Metab 10, 1048-1054.

Ryan, A., Laurieri, N., Westwood, I., Wang, C.J., Lowe, E., Sim, E., 2010. A novel mechanism for azoreduction. J Mol Biol 400, 24-37.

Salazar, O., Cirino, P.C., Arnold, F.H., 2003. Thermostabilization of a Cytochrome P450 Peroxygenase. ChemBioChem 4, 891-893.

Shivange, A., Serwe, A., Dennig, A., Roccatano, D., Haefner, S., Schwaneberg, U., 2012. Directed evolution of a highly active Yersinia mollaretii phytase. Applied Microbiology and Biotechnology 95, 405-418.

Sollner, S., Deller, S., Macheroux, P., Palfey, B.A., 2009. Mechanism of flavin reduction and oxidation in the redox-sensing quinone reductase Lot6p from Saccharomyces cerevisiae. Biochemistry 48, 8636-8643.

Sollner, S., Nebauer, R., Ehammer, H., Prem, A., Deller, S., Palfey, B.A., Daum, G., Macheroux, P., 2007. Lot6p from Saccharomyces cerevisiae is a FMN-dependent reductase with a potential role in quinone detoxification. FEBS J 274, 1328-1339.

Spiller, B., Gershenson, A., Arnold, F.H., Stevens, R.C., 1999. A structural view of evolutionary divergence. Proceedings of the National Academy of Sciences 96, 12305-12310.

Stolz, A., 2001. Basic and applied aspects in the microbial degradation of azo dyes. Appl Microbiol Biotechnol 56, 69-80.

Suzuki, Y., Yoda, T., Ruhul, A., Sugiura, W., 2001. Molecular cloning and characterization of the gene coding for azoreductase from Bacillus sp. OY1-2 isolated from soil. J Biol Chem 276, 9059-9065.

Tawfik, D.S., 2010. Messy biology and the origins of evolutionary innovations. Nature Chemical Biology 6, 692-696.

Towe, S., Leelakriangsak, M., Kobayashi, K., Van Duy, N., Hecker, M., Zuber, P., Antelmann, H., 2007. The MarR-type repressor MhqR (YkvE) regulates multiple dioxygenases/glyoxalases and an azoreductase which confer resistance to 2-methylhydroquinone and catechol in Bacillus subtilis. Mol Microbiol 66, 40-54.

Tracewell, C.A., Arnold, F.H., 2009. Directed enzyme evolution: climbing fitness peaks one amino acid at a time. Current Opinion in Chemical Biology 13, 3-9.

van der Zee, F.P., Bisschops, I.A.E., Lettinga, G., Field, J.A., 2002. Activated Carbon as an Electron Acceptor and Redox Mediator during the Anaerobic Biotransformation of Azo Dyes. Environmental Science & Technology 37, 402-408.

Wang, C.-J., Hagemeier, C., Rahman, N., Lowe, E., Noble, M., Coughtrie, M., Sim, E., Westwood, I., 2007. Molecular cloning, characterization and ligand-bound structure of an azoreductase from Pseudomonas aeruginosa. . J. Mol. Biol. 373, 1213-1228.

Watanabe, K., Kitamura, K., Suzuki, Y., 1996. Analysis of the critical sites for protein thermostabilization by proline substitution in oligo-1,6-glucosidase from Bacillus coagulans ATCC 7050 and the evolutionary consideration of proline residues. Applied and Environmental Microbiology 62, 2066-2073.