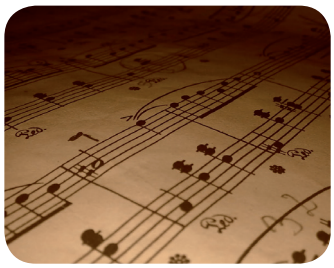
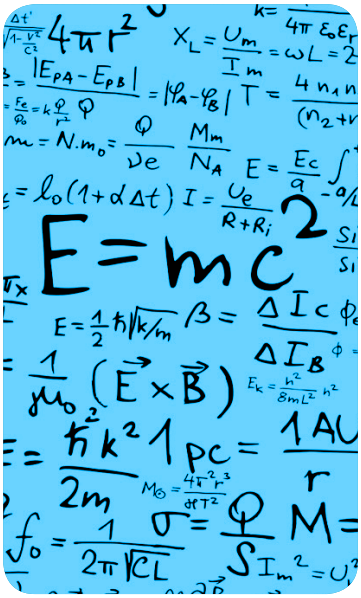




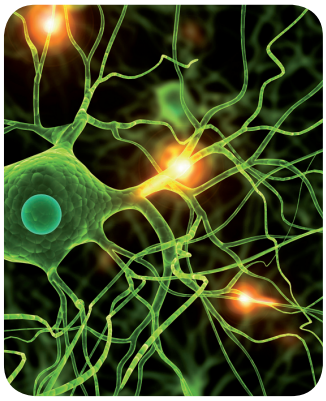
MÁSTERES de la UAM

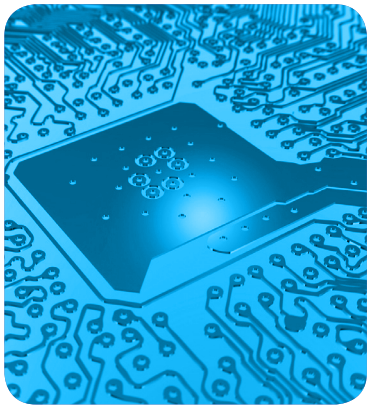
Facultad de Ciencias
/ 14-15

Microbiología

UNIVERSIDAD AUTONOMA DE MADRID

excelencia Campus Internacional UAM CSIC





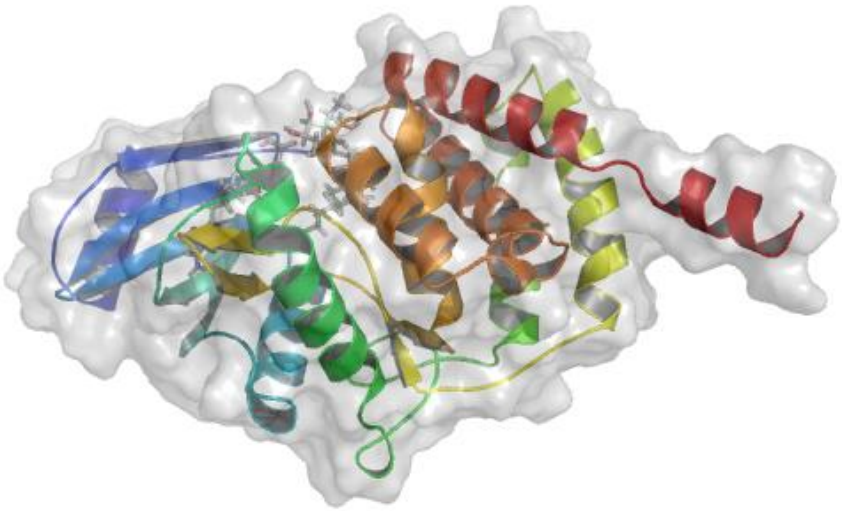
Thermostabilization of hygromycin B phosphotransferase from *Escherichia coli*

Sandra Bosch Reñé

UAM EDICIONES

UNIVERSIDAD AUTONOMA DE MADRID

Thermostabilization of hygromycin B phosphotransferase from *Escherichia coli*



Trabajo fin de máster

Director: Aurelio Hidalgo

Sandra Bosch Reñé

Madrid, septiembre 2015

Máster de Microbiología

ABSTRACT

Enzymes have an important role for biocatalysis in an industrial process, outstripping traditional chemistry. Although thermostable enzymes offers more advantages than mesophilic enzymes, because of their superior stability at the working temperatures and but also against organic solvents and detergents. However, thermostable enzymes of thermophilic organisms are not always available or do not always meet the performance requirements of the industrial process in which they are involved. Therefore, methods to thermostabilize enzymes from mesophilic organisms are needed. The thermostability of an enzyme can be enhanced by directed evolution using the principle of folding interference at high temperatures in a thermophilic host (*Thermus thermophilus*) through the expression of a fusion between the target protein (N terminus) and a thermostable (C terminus) antibiotic selection marker.

Nevertheless, the currently used thermostable selection markers showed false positives, such as the kanamycin nucleotidyltransferase (Kat), and the bleomycin binding protein (ShBle), while the hygromycin B phosphotransferase (Hph5), was not thermostable enough to provide selection.

In this work, we have developed variant Hph17, from Hph5 that allowed cells to grow at 71 °C. This enzyme has 5 more amino acids substitutions, which implies a slight difference between both structures. However, the melting temperature (T_m) only changed by 0.4 °C comparing with Hph5. It would suggest that *in vitro* the thermostability would not be affected, but other factors, such activity or differences in the codon usage, would enhance survival of Hph17-expressing transformants.

INDEX

1. INTRODUCTION

1.1. Extremophiles

An extreme environment is a biotope which presents extreme physical factors (temperature, radiation or pressure) as well as extreme geochemical factors (desiccation, salinity, pH or metal) (Rothschild and Mancinelli, 2001). The study of such extreme environments permitted to find organisms inhabiting them, which are called extremophiles and define the physicochemical limits to life. These organisms do not just "tolerate" their extreme environment but they "require" it to grow (Madigan *et al.,* 2010).

Temperature is probably the most important environmental factor affecting the growth of microorganisms in two opposing ways. On one hand, as temperature increases, chemical and enzymatic reactions proceed at more rapid rates and growth becomes faster. On the other hand, cell components may be irreversibly damaged (Madigan *et al.,* 2010).

According to their growth temperature, microorganisms can be divided into four large groups *(Figure 1)*: psychrophiles (having an optimum temperature <15 °C), mesophiles (15-45 °C), thermophiles (>45 °C) and hyperthermophiles (>80 °C).

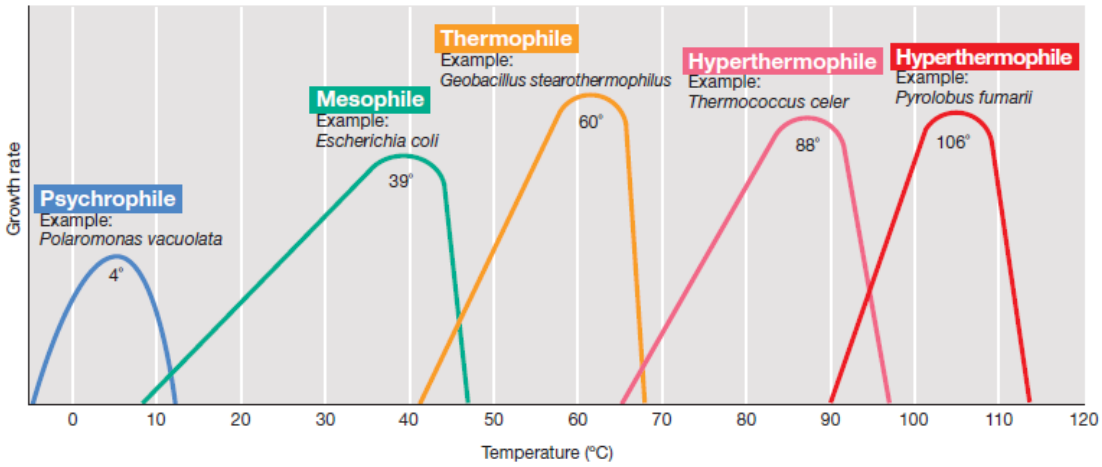


Figure 1. Temperature and growth response in different classes of microorganisms (Madigan et al., 2010).

1.1.1. Thermophiles

Thermophiles and hyperthermophiles only include members of domains Archaea and Bacteria (Trivedi *et al.,* 2006). The high temperatures for thermophiles are found in diverse biotopes, such as soils subject to full sunlight, fermenting materials, etc. Hyperthermophiles are often associated with volcanic phenomena, in particular, hot springs (Madigan *et al.,* 2010).

Thermophiles and hyperthermophiles have the following adaptation mechanisms to high temperatures (Rothschild and Mancinelli, 2001):

- **Membranes**. To maintain an optimal fluidity, the cell must adjust the composition (amount and type) of the membrane.
- **DNA**. Some salts protect the DNA from depurination and hydrolysis and enhance the stability of DNA, because these salts screen the negative charges of the phosphate groups.
- **Compatible solutes**. They are accumulated in the cytoplasm playing an important role in the protection of proteins and other cellular components against heat denaturation. Mannosylglycerate is one of the most general solutes in thermophiles (Faria *et al.,* 2008).
- **Proteins**. Proteins are under constant destabilizing effects of high temperatures. Several reasons have been attributed to higher stability including better hydrogen bonding, better Van der Waals interaction, shorter loops, increase in aromatic interaction, etc. (Sadeghi *et al.,* 2005).

1.1.2. Thermozymes for biocatalysis

The use of enzymes for biocatalysis in an industrial process has some advantages above traditional chemistry. Enzymes are part of white biotechnology (defined as the use of organisms and enzymes for industrial processing and production of materials, chemicals and energy) (Urbieta *et al.,* 2015). Two of their most important advantages are stereoselectivity and regioselectivity. The ability to catalyze reactions under mild conditions (ambient temperature, atmospheric pressure and neutral pH) is another benefit, as is its biodegradability. These features make enzymes environmentally friendly or "green" catalysts (Rozzell, 1999). To further expand the industrial use of enzymes, it would be desirable to obtain stable enzymes against high temperatures and organic solvents (Choi *et al.,* 2015).

Thermostable enzymes (and thermozymes, in particular) generally confer these advantages compared with mesophilic enzymes for industrial applications:

- Decrease in viscosity and increase in the solubility and diffusion coefficient (Niehaus *et al.,* 1999).
- Higher rigidity compared to mesophilic enzymes at room temperature, which protects them from unfolding and preserves their catalytically active structure (Bruins *et al.,* 2001).
- Decrease the risk of contamination from mesophiles (Niehaus *et al.,* 1999).

- Possibility to produce thermozymes in mesophilic hosts, and easy purification by thermal denaturation (Niehaus *et al.,* 1999).
- Higher stability against heat and organic solvents (Choi *et al.,* 2015).
- Increased half-life, reducing the replacement of inactivated enzymes in an industrial process (Elleuche *et al.,* 2015).

These thermostable enzymes can be obtained by optimizing enzymes either from mesophiles (Elleuche *et al.,* 2015), from DNA libraries (including the genomes of unculturable organisms) (Lopez-Lopez *et al.,* 2014) or directly from thermophilic microorganisms, such as the genus *Thermus*.

1.1.3. *Thermus thermophilus* as biological model

Thermus spp is one of the most abundant genus of thermophilic bacteria (Cava *et al.,* 2009), which can be found in hot environments, such in a deep gold mine, heating piles, industrial heating systems, thermal spring or geothermal abyssal (Balkwill *et al.,* 2004).

All species of *Thermus* are Gram negative bacteria, which grow at an optimum temperature range from 62 to 75 °C. Their shape is slender bacillary and they have an orange or yellow pigmentation due to the carotenoids present in the membrane. Their motility is reduced because they do not have flagellum. *Thermus* divide by binary fussion when they are at stationary phase and form a septum. The content of G+C in *Thermus* is 69%, higher than mesophiles microorganisms. The DNA of *Thermus* is divided between a chromosome of 1.9 Mbp and a megaplasmid of 0.23 Mbp (Cava *et al.,* 2009).

Thermus thermophilus, at salt stress conditions, they accumulate trehalose (used at high levels of NaCl), mannosylglycerate (used at low levels of NaCl) or both (Cava *et al., 2009).* The strain *T. thermophilus* HB27 has type IV pili, which is important for attachment, surface colonization, twitching motility and natural transformation. This strain cannot grow under low-oxygen conditions, in contrast to the strain *T. thermophilus* HB8 (Henne *et al.,* 2004).

Generally, they use energy-coupled systems for substrate uptake and they have catabolic pathways for most amino acids and genes for a complete urea cycle. It has been found genes for β-oxidation, the Embden-Meyerhof pathway, complete tricarboxylic acid cycle, gluconeogenesis and glyoxylate bypass (Henne *et al.,* 2004).

T. thermophilus HB27 exhibits high competence for natural transformation (Friedrich *et al.,* 2002). It expresses this competence constitutively and it has a high efficiency, up to 40 kb/s and cell (Schwarzenlander and Averhoff, 2006). This natural competence system and the easily grow

under laboratory condition allowed the development of *T. thermophilus* HB27 as a genetic tool, permitting their manipulation to limits so far reached for mesophilic bacteria models (Cava *et al.,* 2009).

1.1.4. Selection system of thermostable proteins

To predict which changes increase the thermostability of a protein is a difficult task even if the three-dimensional structure is known. Despite many efforts to understand the structural basis of protein stability, there is still no universal strategy to stabilise any protein by a limited number of rationally designed mutations (Eijsink *et al.*, 2005). An alternative to rational design of proteins is directed evolution, consisting on iterative cycles of mutagenesis and screening/selection. Furthermore, it does not need prior knowledge of the structure or other characteristics of the protein. Consequently, this strategy does not infer the thermostability of a protein *a priori* and each sequence generated should be explored independently, requiring high-throughput robotic formats, with the corresponding high costs and time consumption (Chautard *et al.,* 2007).

To avoid an independent exploration, there are selection methods representing a cost-effective alternative to high-throughput robotic assays. However, they require the enzymatic property of interest to be linked to the survival of the recombinant host. While it is easy to implement metabolic selection assays, there are only two methods for thermoselection reported in the literature: Proside (an *in vitro* selection method for proteins with increased stability) (Martin and Schmid, 2003) and THR. The method named THR is based on the "folding interference principle" and consists in the expression of a fusion of two proteins, the protein of interest at N terminus and a thermostable antibiotic selection marker (the reporter) at C terminus. If the protein of interest is not correctly folded, the reporter will not fold properly, thus the clone will not survive. Therefore, it effectively translates protein stability into a life-or-death signal at high temperatures (Chautard *et al.,* 2007).

Candidate thermostable reporters used to correlate protein stability with host survival are the engineered kanamycin nucleotidyltransferase from *Staphylococcus aureus* (Kat), bleomycin binding protein (ShBle) from *Streptoalloitechus hindustanus* and the engineered hygromycin B phosphotransferase *(*Hph5*)* from *Escherichia coli*. However, all three reporters showed drawbacks. Kanamycin exhibited positive but weak correlation between survival and stability of the passenger protein, due to the dimeric nature of Kat. Bleomycin did not provide selection, due to the overstabilization caused by the tetrameric nature of ShBle. Hygromycin did not yield

any selection due to the lability of the marker Hph5, which is not stable enough (unpublished data).

1.2. Hygromycin B phosphotransferase

1.2.1. Hygromycin B

Hygromycin B is an aminoglycoside antibiotic (AGA) produced originally in *Streptomyces hygroscopicus* (Borovinskaya *et al.,* 2008). Similar to other antibiotics, they are secondary metabolites that allows competition for the same ecological niche with other bacteria or fungi, killing them (Becker and Cooper, 2012).

Chemically, hygromycin B contains 2-deoxystreptamine as a central core with substitutions in the 5- and 4-positions. It also has a dual ether linkage between its second and fourth rings, forming a third ring *(Figure 2).*

Figure 2. Chemical structure of aminoglycoside hygromycin B (Borovinskaya et al., 2008).

Hygromycin B has a binding site in rRNA of the small ribosome subunit (30S), next to the aminoacyl-tRNA binding site. It only binds to a single site on the ribosome, exhibiting monophasic inhibition (Borovinskaya *et al.,* 2008). Hygromycin B allows other noncognate tRNA to bind and leads to a misreading of the mRNA and synthesis of faulty proteins (Becker and Cooper, 2012).

1.2.2. Mechanisms of resistance

The target of hygromycin B is 30S rRNA, even so the resistance mechanism is not the mutation or modification of this rRNA, but instead the enzymatic modification of the hygromycin B.

There are 3 different groups of aminoglycoside modifying enzymes, including ATP-dependent aminoglycoside phosphotransferases (APHs), acetyl-CoA-dependent aminoglycoside acetyltransferases and ATP-dependent aminoglycoside nucleotidyltransferases. These enzymes

are usually encoded on plasmids, transposons or integrons, which makes them highly mobile and facilitates the spreading of resistance (Becker and Cooper, 2012).

The enzymes that confer resistance to hygromycin B are APHs. The phosphorylation of hydroxyl groups in hygromycin B introduces a negative charge in the molecule, changing effectively the ability to bind to rRNA. In the group of APHs, there are APH(4) and APH(7'') that give resistance to hygromycin B (Becker and Cooper, 2012; Ramirez and Tolmasky, 2010). *Figure 3* shows the sites of enzymatic modifications of APH(4) and APH(7'') on hygromycin B.

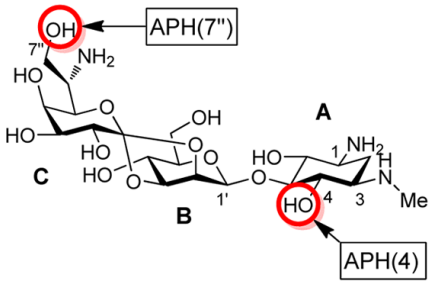


Figure 3. Sites of enzymatic modifications on hygromycin B (Becker and Cooper, 2012).

APH(4). There are 2 different enzymes in this subclass: APH(4)-Ia (also named Hph*)* and APH(4)-Ib (Hyg). These enzymes are not clinically relevant, but they have been used in the construction of cloning vehicles for prokaryotes and eukaryotes.

APH(7''). This enzyme was isolated from *S. hygroscopicus* and the gene has been cloned and engineered to been used in molecular genetic analysis.

1.2.3. Thermostable hygromycin B phosphotransferase

The first time that hygromycin B phosphotransferase *(hph)* from *E. coli* was thermostabilized was using *Sulfolobus solfataricus* as a host-vector system (Cannio *et al.,* 1998). The method used to introduce mutations in the gene was error-prone PCR (epPCR). They suggested that the amino acid substitutions S52T and W238C, retained its enzyme activity up to 65 °C. The epPCR has some differences comparing with the conventional PCR (Brackmann and Schwienhorst, 2006; Arnold and Georgiou, 2003):

- The polymerase used is Taq DNA which lacks 3'-5' exonuclease activity and exhibits a higher error rate.

- Increased Mg^{2+} concentration for stabilizing noncomplementary base pairs. An increasing Mg^{2+} concentration, decreases the specificity of the polymerase (Markoulatos *et al.,* 2009).
- Addition of Mn^{2+} for reducing the base pairing specificity, Mn^{2+} alters the conformation of the polymerase so as to affect nucleotide selection. An increasing Mn^{2+} concentration, increases the mutation rate of the polymerase (Beckman *et al.,* 1985).
- Unbalanced dNTPs stoichiometry for forcing misincorporation and helps to reduce the natural error bias of Taq DNA polymerase.
- Increased polymerase concentration for enhancing the probability of elongation of misprimed termini.

From the mutations S52T and W238C, Nakamura *et al.,* 2005 thermostabilized *hph*, using a host-vector system of *T. thermophilus.* In this case, the methodology used was an *in vivo*-directed evolutionary strategy, which it consists in an increasing selection temperature, allowing the growth of the cells that contains new mutations in *hph* that confers higher thermostability *(Figure 4)*.

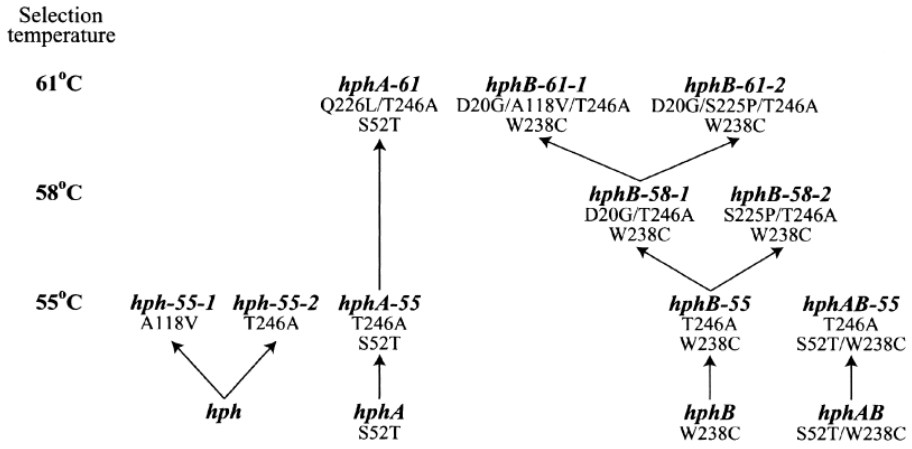


Figure 4. In vivo-directed evolution pathway of the hph mutant genes. The starting genes used were hph from E. coli, hphA, hphB and hphAB. The selection temperature and the amino acid substitutions are shown (Nakamura et al., 2005).

The amino acid substitution S52T and W238C did not confer resistance at 55 °C; however 5 spontaneously generated, independent mutants were obtained by selection of the transformants at this temperature. The mutation introduced at 55 °C was A118V or T246A, when the temperature was increased to 58 and then 61 °C, three more amino acid substitutions were acquired: D20G, S225P and Q226L. *T. thermophilus* carrying all the 5 amino acid substitutions was able to grow at up to 67 °C, but the transformation efficiency decreased to about half

comparing with 65 °C (*Table 1)*. It suggested that Hph5 protein at 67 °C was not sufficient to enable the host to form a colony from a single cell in the presence of hygromycin B.

Table 1. Effect of the amino acid substitutions of Hph E. coli the T. thermophilus temperature growth.

Protein	Amino acid substitution	Temperature (°C)
Hph wt	-	40
HphAB	S52T, W238C	< 55
Hph-55-1, Hph-55-2	T246A, A118V	55
Hph-B-58-1, Hph-B-58-2	D20G, S225P	58
HphA-61	Q226L	61
Hph5	T246A, A118V, D20G, S225P, Q226L	65

Concluding that the *hph5* gene can be used as a selection marker in the *T. thermophilus* host-vector system, the selection can be achieved up to 65 °C, when the gene is either located on a plasmid or integrated in the chromosome (Nakamura *et al.,* 2005).

The protein obtained by Nakamura *et al.,* 2005 was expressed to compare with the Hph from *E. coli.* Both proteins showed a similar specific activity and enzyme activity and presented one-step unfolding profiles. The temperature at which the protein unfolds was calculated as melting temperatures, and it was 37.2 and 58.8 °C for Hph and Hph5, respectively (Nakamura *et al.,* 2008).

This thermostable protein has been used as a reporter in the method of folding interference, having unsuccessful results due to the lability (unpublished data).

1.3. Aim of this work

This work aims to further thermostabilize the hygromycin B phosphostransferase, Hph5, from *E. coli* allowing growth of *Thermus* above 65 °C, in order that obtain a reporter that correlates protein thermostability with host survival. To that extent, we will use directed evolution and create mutations in the parental *hph5* by epPCR to screen the library in *T. thermophilus* at higher temperatures. The best variants or "hits" will be confirmed by assays on plate and show correlation in the folding interference assay between the stability of the target protein and the host survival.

2. MATERIALS AND METHODS

2.1. Cultivation

2.1.1. Strains

The bacterial strains used in this work are described in *Table 2.*

Table 2. Strains used.

Strain	Utilization
Escherichia coli DH5α	To amplify plasmid and to obtain a library.
Escherichia coli BL21 (DE3)	To express proteins.
Thermus thermophilus HB27ES	Thermophile model.

2.1.2. Growth

E. coli was grown in LB (Luria Bertani Lysogeny broth) medium *(Table 3)* and *T. thermophilus* was grown in TB (*Thermus* Broth) medium *(Table 4).*The medium compounds were mixed with distilled water for LB medium and with mineral water *Sierra de Cazorla* for TB medium and pH was adjusted to 7 for LB medium and to 7.5 for TB medium with NaOH 2 M. If solid medium was required, 1.5% of agar was added. After autoclaving and upon cooling down, media were supplemented with a final concentration of 100 µg/mL of ampicillin, 30 µg/mL of kanamycin or 100 µg/mL of hygromycin.

E. coli was usually grown at 37 °C for 12 h. *T. thermophilus* was usually grown at 60-70 °C for 24-48 h. Liquid cultures were shaken in an orbital incubator at 180-200 rpm.

Table 3. Composition of LB medium.

Compound	Concentration (g/L)
Tryptone	10
Yeast extract	5
NaCl	5

Table 4.Composition of TB medium.

Compound	Concentration (g/L)
Tryptone	8
Yeast extract	4
NaCl	3

2.1.3. Transformation

For each transformation of *E. coli* DH5α, 5 µL of ligation or 100-200 ng of plasmid were added to 50 µL of competent cells. After 30 min incubation on ice, the tubes were heat shocked at 42 °C for 90 s and subsequently incubated on ice for 5 min. Afterwards, 200 µL of SOC medium *(Table 5)* were added and incubated at 37 °C and 180 rpm for 1 h if the antibiotic used was ampicillin or hygromycin or 3 hours if it was kanamycin. Finally, the suspension was spread on selection plates with the antibiotic.

Electroporation of *E. coli* BL21 was carried out by mixing 45 µL of competent cells with 100-200 ng of plasmid and subjecting the cells to a short pulse (5 ms) at 12500 V/cm (EasyjectPlus D2000; 2500 V, 201 Ω and 25 µF, using cuvettes of 0.2 cm). Immediately after the pulse, 500 µL of LB supplemented with SOC medium were added and incubated under the same conditions as *E. coli* DH5α.

Table 5.Composition of SOC medium.

Compound	**Concentration**
Tryptone	20 g/L
Yeast extract	5 g/L
NaCl	0.5 g/L
KCl	0.2 g/L
$MgCl_2$	10 mM
$MgSO_4$	20 mM
Glucose	0.3 % (m/v)

T. thermophilus HB27ES was transformed by natural competence adding 0.1 µg of DNA to 0.5 mL of cell culture at exponential phase (OD_{600} 0.3-0.4) in TB medium. Independently of the antibiotic used, cultures were incubated 4 h after the addition of DNA and the suspension was spread on selection plates with antibiotic.

2.2 Plasmids and oligonucleotides

2.2.1. Plasmids

The thermostable gene hygromycin B phosphotransferase, *hph5,* was initially in the shuttle plasmid pMH184, which contains two replication origins, one for *E. coli* and another for *T. thermophilus* and a gene encoding Hph5 as selection marker.

The parental *hph5* gene and its mutant derivative *hph17* were cloned into plasmid pET28b, which affords their expression as His-tagged products under the control of the T7 promoter.

2.2.2. Oligonucleotides (primers)

Amplification of *hph5* gene was carried out using primers, ep_HPH_fw and ep_HPH_rv *(Table 6)*. For sequencing, hph_sec1 was used. To subclone *hph17* into pET28b, 2 primers with the restriction sites NdeI and EcoRI were designed.

Table 6.Oligonucleotides used.

Name	Sequence	Use
ep_HPH_fw	GCCCACGACCCTTAAGGA	To amplify *hph5*
ep_HPH_rv	CTGGCACAGATGGTCATAACC	To amplify *hph5*
hph_sec1	CGGAGAGGAAACGCGGCAAC	To sequence
hph_fw_Nde	AAAAAACATATGAAAAAGCCTGAACTCACCG	To subclone *hph5*
hph_rv_Eco	AAAAAACATATGAAAAAGCCTGAACTCACCG	To subclone *hph5*

2.3 Hph variant library construction

2.3.1. Error-prone PCR

A library of *hph5 variants* was obtained by error-prone PCR (epPCR) using a *BioRad C1000 Touch* thermocycler. The reaction mix and the thermocycler program for amplification are described in *Tables 7* and *8,* respectively. Three different concentrations of Mn^{2+} were assayed.

Table 7.Reaction mix for epPCR of hph5.

Component	Final concentration
Nuclease-free water	Up to 100 µL
10x Mutant buffer	1x
dNTPs unbalanced	0.2 mM A; 0.2 mM G; 1 mM C; 1 mM T
$MnCl_2$	0.1/0.2/0.3 mM
ep_HPH_fw	0.5 µM
ep_HPH_rv	0.5 µM
Template DNA	5 ng/100 µL
NZYTaq DNA polymerase (Nzytech)	0.05 U/µL

Table 8.Thermocycler program for the randomization of hph5 by epPCR.

Cycle	Step	Temperature (°C)	Time
1	Denaturation	95	3 min
2 to 31	Denaturation	95	30 s
	Annealing	65	30 s
	Elongation	72	75 s
32	Elongation	72	10 min

The PCR products were purified using the kit *Wizard® SV Gel and PCR Clean-Up System* (Promega).

2.3.1. Cloning

Firstly, the plasmid pMH184 and the PCR products were digested with the restriction enzymes BglII and NdeI using Thermo Scientific FastDigest enzymes, according to the manufacturer's instructions. The plasmid was digested and the corresponding band was purified from an agarose electrophoresis gel, while the PCR products were directly purified using the kit *Wizard® SV Gel and PCR Clean-Up System* (Promega).

Afterwards, the plasmid and the PCR products were ligated with *T4 DNA ligase* (Promega). A typical reaction setup contained 50 ng of plasmid, 1:3 molar ratio (plasmid:insert) of PCR product, 0.5 µL of ligase, and 1 µL of ligase buffer. The final volume was adjusted to 10 µL with Milli-Q water and the reaction was incubated overnight at 16 °C.

E. coli DH5α was transformed with the ligation product, as mentioned above, until 40 CFU/ng DNA were obtained. For each Mn^{2+} concentration, 10 colonies were grown in LB broth supplemented with hygromycin, the plasmid was extracted using the *GeneJet Plasmid MiniPrep kit* (Thermo Scientific) and they were sequenced, to obtain a quick overview of the mutational spectrum.

If the mutational spectrum was satisfactory, the colonies were pooled with LB and the plasmids were extracted.

2.4 Selection of thermostable mutants

2.4.1. Selection conditions

The selection conditions in *T. thermophilus* were fixed using temperature and antibiotic concentration as selective pressure.

The temperatures tested were 65, 67.5 and 70 °C. To register the temperature inside the heater, a thermo button (Plug&Track) was used.

Additionally four hygromycin concentrations were checked: 50, 100, 150 and 200 mg/L.

2.4.2. Relative thermostability assay on plate

The thermostability of some mutants was checked in *T. thermophilus* spreading 10 µL drops of different serial dilutions of these mutants.

The cellular concentration was equalized through optical density (OD), assuming 1 OD unit equates to 10^9 CFU/ml. The highest OD used was 1 and the dilutions were made up to $OD=10^{-5}$.

As negative control, *T. thermophilus* with the plasmid containing the parental *hph5* was utilised.

The mutants that exhibited better growth at lower dilutions, were sequenced.

2.5 Biochemical characterization of Hph5 and variant Hph17

2.5.1. Expression of Hph5 and Hph17

Constructions on pMH184 vectors harbouring genes *hph5* and *hph17* were amplified by PCR using primers "hph_fw_Nde" and "hph_rv_Eco" to introduce NdeI and EcoRI restriction sites. The reaction mix and the thermocycler program for amplification are described in *Tables 9* and *10,* respectively.

Table 9. Reaction mix for PCR.

Component	Final concentration
Nuclease-free water	Up to 50 µL
5x Phusion® HF buffer	1x
dNTPs	0.2 mM
hph_fw_Nde	0.5 µM
hph_rv_Eco	0.5 µM
Template DNA	5 ng/100 µL
Phusion® High Fidelity DNA polymerase (New England Biolabs)	0.02 U/µL

Table 10.Thermocycler program for amplification.

Cycle	Step	Temperature (°C)	Time
1	Denaturation	98	2 min
2 to 31	Denaturation	98	20 s
	Annealing	65	30 s
	Elongation	72	30 s
32	Elongation	72	10 min

After purifying the PCR product, *hph5, hph17* and plasmid pET28b were digested with NdeI and EcoRI. The digested genes and plasmid were ligated, *E. coli* DH5α was transformed with the ligation product and a transformant with the insert was grown in LB liquid with kanamycin.

The generated pET28b-hph5 and pET28-hph17 constructs were transformed in *E. coli* BL21. One colony of each insert was grown in LB liquid with kanamycin until the OD was approximately 0.4-0.6. Then, 1.5 mL of culture medium was inoculated in 100 mL of lactose autoinduction medium *(Table 11)* and incubated at 37 °C overnight for protein expression.

Table 11. Composition of lactose autoinduction medium.

Compound	Concentration (g/L)
Tripticase	9.28
Yeast extract	4.64
Glycerol	0.5
Glucose	0.5
Lactose	2
$(NH_4)_2SO_4$	3.3
KH_2PO_4	6.8
Na_2HPO_4	7.1
$MgSO_4$	0.132

Cells were harvested by centrifugation (4000 xg for 10 minutes) and stored at -20°C until further use.

2.5.2. Enzyme purification

Cells were resuspended in 25 mL of binding Ni-NTA (Ni(II)- nitrilotriacetic acid) buffer *(table 12)* and lysed with a pressure homogenizer *Niro SOavi* (GEA). Soluble proteins were separated from the cell debris by centrifugation at 15,000 xg for 30 min.

The cell extract was mixed with 1 mL of Ni-NTA resin (Ni-NTA Superflow, Qiagen) slurry, incubated at 4 °C in a shaker for 60 min and packed in a plastic column. The column was then washed with binding Ni-NTA buffer. Finally, the elution of the protein was carried out with Ni- elution NTA buffer *(table 12)*. The proteins were concentrated using Amicon Ultra-15 concentrators with a 30 kDa cutoff (Merck Millipore).

Table 12. Buffers used to purify the enzyme with a Ni-NTA resin.

Ni-NTA Binding buffer	Ni-NTA Elution buffer
20 mM imidazole	500 mM imidazole
50 mM sodium phosphates	50 mM sodium phosphates
500 mM NaCl	500 mM NaCl
pH=7.4	pH=7.4

After protein concentration, purity was checked by SDS-PAGE (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis) *(table 13)* and protein concentrations were determined using *Bio-Rad Protein Assay* (Bio-Rad), according to the manufacturer's protocol using bovine serum albumin (BSA) as a standard.

Table 13.Compositions of separation and stacking gels used for 2 SDS-PAGE gels.

12% separation gel	Quantity
H_2O	1.13 mL
40% acrylamide/1% bisacrylamide	3 mL

2 M Tris/HCl pH 8.8	5 mL
10% SDS	100 µL
TEMED	20 µL
1% APS	250 µL
6% Stacking gel	**Quantity**
H_2O	1.3 mL
30% acrylamide/1% bisacrylamide	1 mL
2 M Tris/HCl pH 6.8	2.5 mL
10% SDS	50 µL
TEMED	10 µL
1% APS	125 µL

2.5.3. Determination of melting temperature

To determine whether the improvement in the survival of *hph17*-expressing transformants was due to increased thermostability or increased activity, thermal denaturation assays were performed to determine the melting temperature (T_m).

The thermostability of each protein was determined in a *Rotor Gene™ 6000* (Corbett Life Sciences) real time thermocycler by differential scanning fluorimetry (DSF). For this assay (*Table 14*), 20 micromolar protein in buffer (100 mM sodium phosphates pH=7.0, 150 mM NaCl) was supplemented with 5x *SYPRO Orange* (Sigma-Aldrich) and subjected to a temperature ramp from 35 to 95 °C at 1 °C/min. The assay was carried out in triplicate.

Table 14. Composition of the reaction mix to determine Tm.

Compound	Concentration
Buffer	Up to 20 µL
Protein	20 µM
SYPRO Orange	5x

2.6. Bioinformatics analysis

The alignment and translation of genomic sequences were performed using the program Vector NTI 10 (Invitrogen).

The structural model of the mutant of Hph5 was created using the program Yasara using the default automated macro. The templates were selected automatically, homology modelling was carried out with the templates PDB code 3w0o, 3w0p, 3w0q, 3w0r and 3w0s. Loops were modelled using 50 loops having the same sequence. The best parts of the models were combined to obtain a hybrid model. Images were generated with PyMol v0.99 (DeLano Scientific).

3. RESULTS AND DISCUSSION

3.1. Obtaining functional mutants of the protein Hph5 in *T. thermophilus*

3.1.1. Hph variant library construction in *E. coli* by epPCR

Variants of the thermostable hygromycin B phosphotransferase (*hph5*) were obtained by error-prone PCR (epPCR).

The epPCR was carried out with 3 different concentrations of Mn^{2+} (0.1, 0.2 and 0.3 mM), this cation alters the conformation of the polymerase so as to affect nucleotide selection. (Beckman *et al.,* 1985). The corresponding product was digested with BglII and NdeI, ligated into pMH184 precut with the same enzymes and transformed in *E. coli* DH5α. Increasing Mn^{2+} concentrations afforded increasing mutation rates (*Figure 5*), and the percentage of the transitions and the transversions (*Figure 6*) was calculated to estimate the variability and mutational spectrum balance of the library.

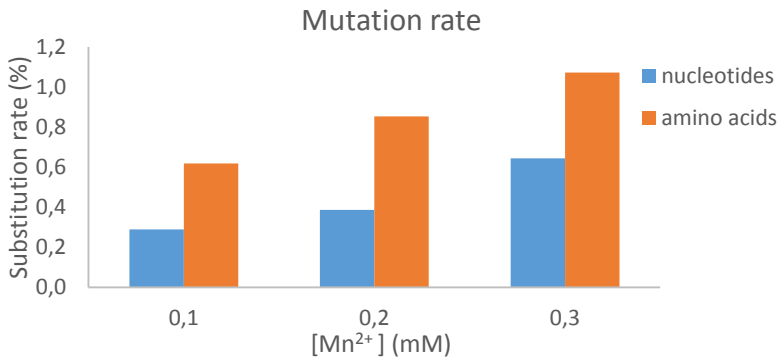


Figure 5. Mutation rate of the gene hph5 and the protein at different Mn^{2+} concentration.

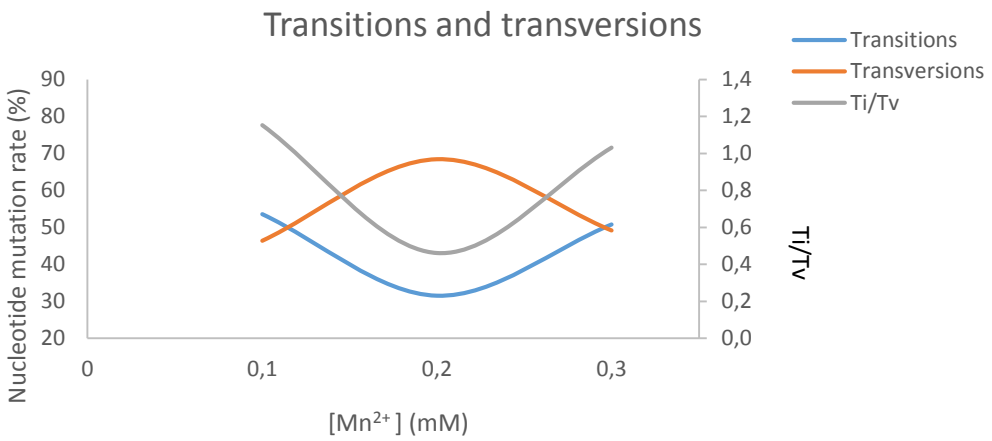


Figure 6. Percentage variation of the transitions (Ti) and transversions (Tv) at different Mn^{2+} concentration. The ratio Ti/Tv is indicated as a proportion.

The library characterization shows a direct correlation between Mn^{2+} concentration and the mutation rate calculated per nucleotide or amino acid. The average number of mutations also depends on the template length. In most cases, the mutagenized DNA will encode a protein, the fraction of mutated amino acids will be higher than the fraction of mutated nucleotides by a factor of 2.2, this is because a mutation in any of the three position of a codon may result in an amino acid substitution. A mutation at the first position of a codon will cause an amino acid change 96% of the time; mutation at the second and third position will cause amino acid changes 100% and 23% of the time, respectively (Wilson and Keefe, 2001)

As it was already observed that bias towards transitions exists at Mn_{2+} concentration of 0.5 mM (Shivange *et al.,* 2009; Wong *et al.,* 2007), we tried to determine the sequence bias of this library. The transition/transversion rate had the lowest value when the Mn_{2+} concentration is 0.2 mM, this ratio increased when the Mn^{2+} concentration varied from 0.2 mM.

In most epPCR conditions employing Taq DNA polymerase, a bias towards transitions is observed. Due to only 10 clones were analysed for each Mn_{2+} concentration, a bias towards transition, could not be demonstrated. It is important to mention that a biased occurrence in gene sequences would limit significantly the diversity of generated protein sequence libraries, because the genetic code is optimized that transitions lead, more frequently than transversions, often to chemically similar or identical amino acids (fraction of conserved aminoacids: 34.5% transition bias and 14.2% transversion bias) (Shivange *et al.,* 2009).

A Mn^{2+} concentration of 0.1 mM created an average of 2 amino acid substitutions in the whole protein (342 amino acids). Due to this low substitution rate, only concentrations of 0.2 and 0.3

mM of Mn^{2+} were used to generate libraries of $4.5 \cdot 10^4$ and $2.3 \cdot 10^4$ individuals, respectively. The cells were pooled, diluted and the plasmid library was extracted using conventional methods.

3.1.2. Hph variant library selection in *T. thermophilus*

The selection conditions were fixed at the minimum temperature and antibiotic concentration (hygromycin) at which transformants expressing parental Hph5 could not grow. The temperatures tested were 65, 67.5 and 70 °C and the antibiotic concentrations assayed were 50, 100, 150 and 200 mg/L (*Figure 7)*. Temperatures were selected according to Nakamura *et al.,* 2005, where they could grow *T. thermophilus* up to 67 °C.

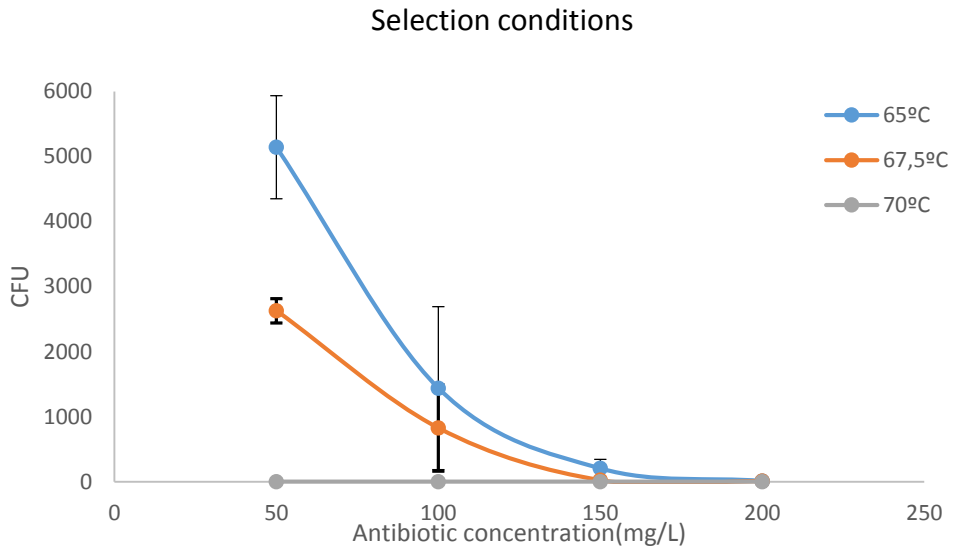


Figure 7. Growth of T. thermophilus transformants expressing parental Hph5 at different temperature and hygromycin concentration. CFU represented is the average of triplicates and bars indicate the standard deviation.

As shown in *Figure 7*, increasing the selection temperature from 65 to 67.5 °C, reduced the survival rate by half, while raising the temperature to 70 °C, completely abolished growth. Survival of the transformants inversely correlated with increasing hygromycin concentration, achieving no growth at 200 mg/L hygromycin, regardless of the incubation temperature. The purpose was to obtain a thermostable mutant and not a more active mutant, thus a hygromycin concentration that permitted the growth was chosen, it was 100 mg/L because it is an intermediate concentration. The selection conditions were set to 70 °C and 100 mg/L of hygromycin concentration.

After determining the selection conditions, *T. thermophilus* was transformed with *hph5* variant library and grown at 70 °C *(Table 15).* Because of the high number of transformants obtained,

the transformation was repeated, increasing the temperature to 71 °C and thus, increasing the selective pressure.

Table 15. T. thermophilus growth (CFU) at different temperature. Results are expressed in total CFU obtained at each transformation and in CFU obtained per µg of plasmidic DNA.

Transformation	**Temperature**					
	70 °C	70 °C	71 °C	71 °C	60 °C	60 °C
	Total CFU	CFU/µg DNA	Total CFU	CFU/µg DNA	Total CFU	CFU/µg DNA
pMH184, parental *hph5*	0	0	0	0		
pMH184, mutated *hph5* with 0.2 mM Mn^{2+}	58	32.2	5	2.5	9961	$5 \cdot 10^{4}$
pMH184, mutated *hph5* with 0.3 mM Mn^{2+}	1150	639	187	93.5	8645	$4.3 \cdot 10^{4}$

After increasing the temperature 1 °C, the number of CFU decreased 12 and 6 times for 0.2 mM and 0.3 mM, respectively. At both temperatures, the CFU obtained is higher with 0.3 mM than 0.2 mM. This trend indicates that the lower Mn^{2+} concentration yields fewer amino acid mutations and it leads to a reduced enhancement in the thermostability of Hph5, compared to the 0.3 Mn library. However, in this library, the mutation rate seems to be high enough to afford a larger proportion of survivors.

A measure of the stringency of the selection is given by the selection factor, which is defined as the ratio of thermostable variants over the total number of variants obtained (ratio between the number of transformants at 70-71 and 60 °C). At a temperature of 70 °C, the ratio is $6 \cdot 10^{-4}$ for 0.2 mM of Mn^{2+} and $1.5 \cdot 10^{-2}$ for 0.3 Mn^{2+}. This factor decreased to $5 \cdot 10^{-5}$ for 0.2 mM and $2 \cdot 10^{-4}$ for 0.3 mM when the temperature was increased to 71 °C. The difference between the ratios obtained at 70 and 71 °C would suggest that when the temperature increases, so does the relative fitness of the individuals in the 0.3 Mn library.

3.1.3. Verification and sequence analysis of hits

Some of the selected clones (or "hits") were verified using a serial dilution assay on plate. Transformant cultures were serially diluted and 10 µL drops of each were spread. Twenty different hits of each Mn^{2+} concentration were verified initially at 71 °C, keeping in mind that hits used were obtained from the library at 71 °C *(Figure 8)*.

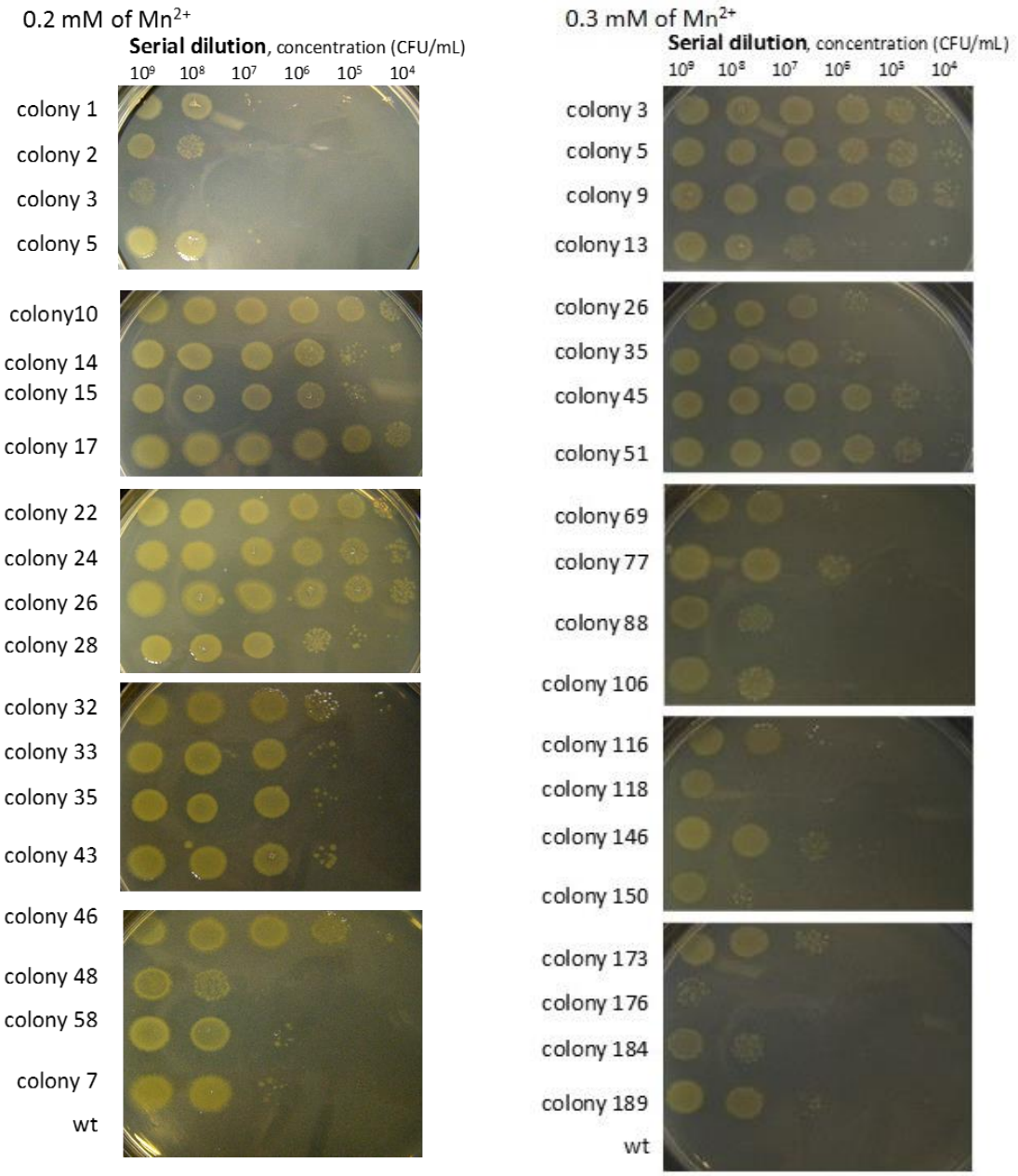


Figure 8. Relative thermostability assay at 71 °C. The cells containing the plasmid pMH184 with hph5 do not grow at this conditions, cells containing hph5 mutated show different resistance degree.

Due to the low differences showed by transformants containing the mutated *hph5* hits, the assay was repeated increasing the selective pressure to 74 °C, the highest temperature possible due to instrumental limitations. Only few colonies could grow, all of them corresponded to 0.2 mM of Mn^{2+}, such as colony 7, 14, 17 (the highest growth) and 26, and all of them only grew when the concentration was 10^9 CFU(mL (*Figure 9*).

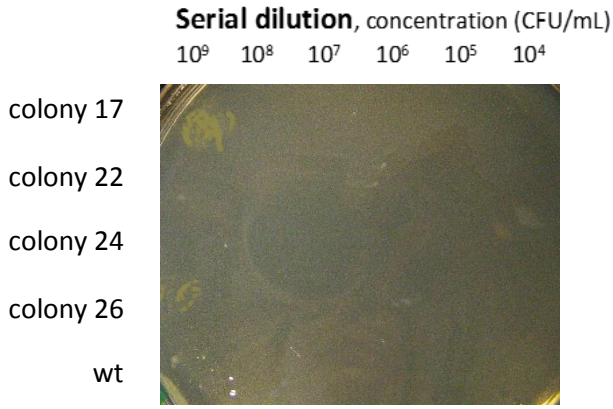


Figure 9. Relative thermostability assay at 74 °C. The cells containing the plasmid pMH184 with hph5 do not grow at these conditions and colonies 17 and 26 could grow at the concentration of 10^9 CFU/mL.

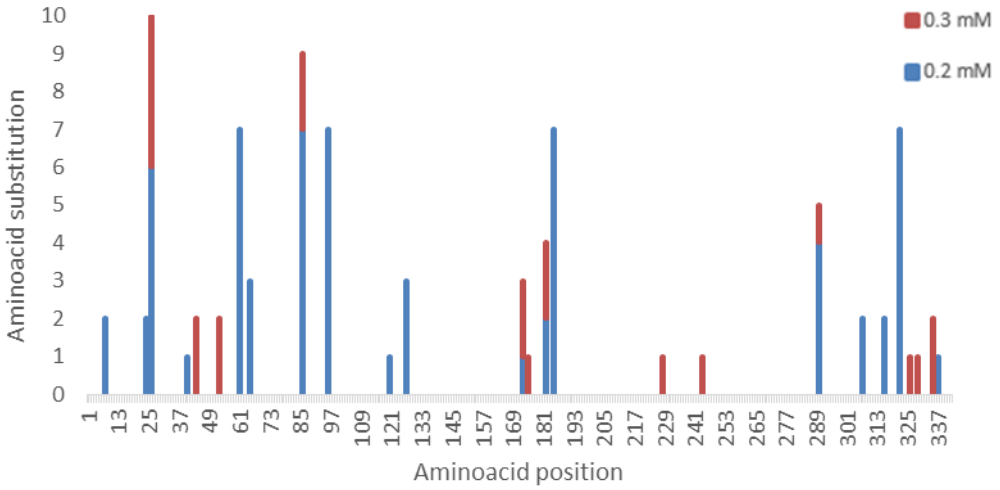


Figure 10. Amino acid substitution position along the hph5 sequence. The 0.2 mM of Mn^{2+} is shown In blue, whereas the 0.3 mM library is shown in ted. Total sequences analysed were 30, 20 of 0.2 mM of Mn^{2+} and 10 of 0.3 mM.

The 30 confirmed variants that performed best in the dilution assay were sequenced: 20 of them corresponded to the 0.2 mM of Mn^{2+} library and 10 to 0.3 mM Mn^{2+} library. The position and abundance of the substitutions along the sequence is shown in *Figure 10* and *Table 16.*

Table 16. Amino acid substitution at different Mn^{2+} concentration. It is also indicated the position of the original amino acid.

Amino acid substitution	**Number of times observed**		**Amino acid substitution**	**Number of times observed**	
	0.2 mM	0.3 mM		0.2 mM	0.3 mM
T8A	2	-	N24D	2	-
K26M	5	4	K26T	1	-

N40D	1	-	C44R	-	2
S53C	-	2	H61R	7	-
Q65R	3	-	G86S	7	2
P96Q	7	-	E120D	1	-
I127L	3	-	F173S	1	-
T173S	1	2	T175A	-	1
L182M	2	2	V185A	7	-
D228E	-	1	L246A	-	1
L290Q	4	1	K307T	2	-
A316V	2	-	E322V	7	-
P326A	-	1	E329G	-	1
S335T	-	2	L337P	1	-

As shown in *Table 16*, 26 different amino acids from the original protein, Hph5, were substituted. More importantly, 4 amino acid substitutions, are present in both libraries which were generated independently for each Mn^{2+} concentration. These "hotspots", would suggest that these positions are important for the thermostability, and would need to be further explored by saturation mutagenesis.

In this case, it becomes obvious the difference in the substitution position where it has high probability to find substitutions, such as at N terminus, at the middle (173-185) and at C terminus. The probability decreases in some other points, but it is also possible. The amino acids that have a binding site for hygromycin B are Gln101, Asp198, Ser201, Asp216, Asn231, Trp235, Trp238, Leu239, Gln273 and Asp285 (data from PDB). None of the mutations obtained were localized in a binding site for the substrate, suggesting *a priori* that the activity may not have been influenced.

The protein that showed highest relative thermostability on plate was used to model its structure and determine its melting temperature (T_m) was the gene obtained from the colony 17 of 0.2 mM of Mn^{2+} and it had the substitutions R61H, S86G, Q96P, A185V and V322E regarding Hph5. As mentioned above, it grew at 74 °C and it has the same sequence as other 6 clones.

3.2. Analysis of Hph17

3.2.1. Homology modelling and structural analysis of Hph17

A structural model was generated by YASARA based on the crystal structure of Hph5 (PDB: 3w0o, 3w0p, 3w0q, 3w0r and 3w0s) However, neither the structure file, nor the homology model show the substitution E322V because the last amino acid in the structure is Ser313, probably due to typical flexibility of the last amino acids of a protein. The location of rest of the substitutions was analysed and R61H is located in a α-helix, S86G and Q96P are in loops and A185V is in another α-helix *(Figure 11)*.

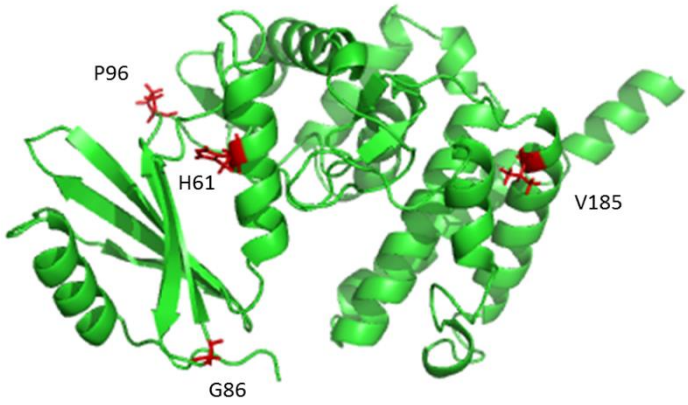


Figure 11. Hph17, amino acid substitutions are shown in red.

A difference found is in the substitution S86G in Hph17. This mutation is part of a loop between 2 β-sheets, increasing the distance between these β-sheets (5.11 Å), comparing with the parental Hph5 (4.91 Å) *(Figure 12).*

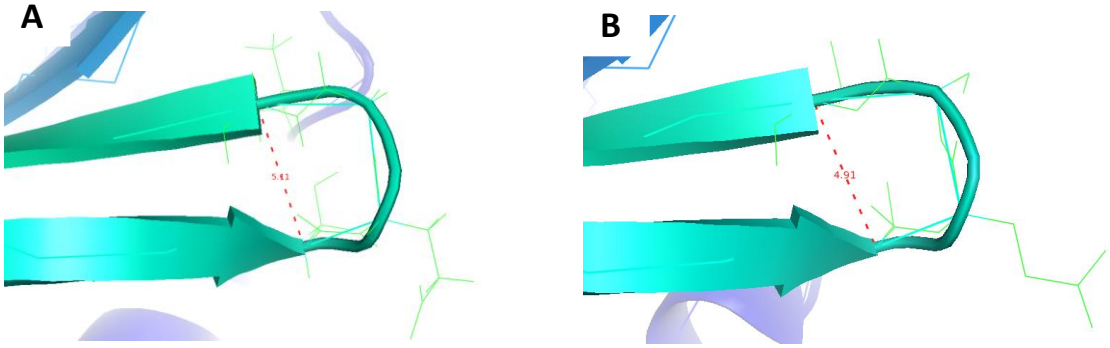


Figure 12. Loop between 2 β-sheets containing the amino acids SEGL in Hph17 (A) and SESL in Hph5 (B). In red, it is shown the distance between the β-sheets.

As shown in *Figure 13*, the substitution at position 96 (Q96P) changes the interaction between the amino acid and its close amino acids. Pro96 binds with Ala95, Gly97, Asp207 and Thr212, while Glu96 binds with Gly97, Asp207 and Thr212. The substitution Q96P reduces the distance between Pro96 and the other amino acids compared to Glu96. This reduction of the distance may contribute to the stabilization of the protein.

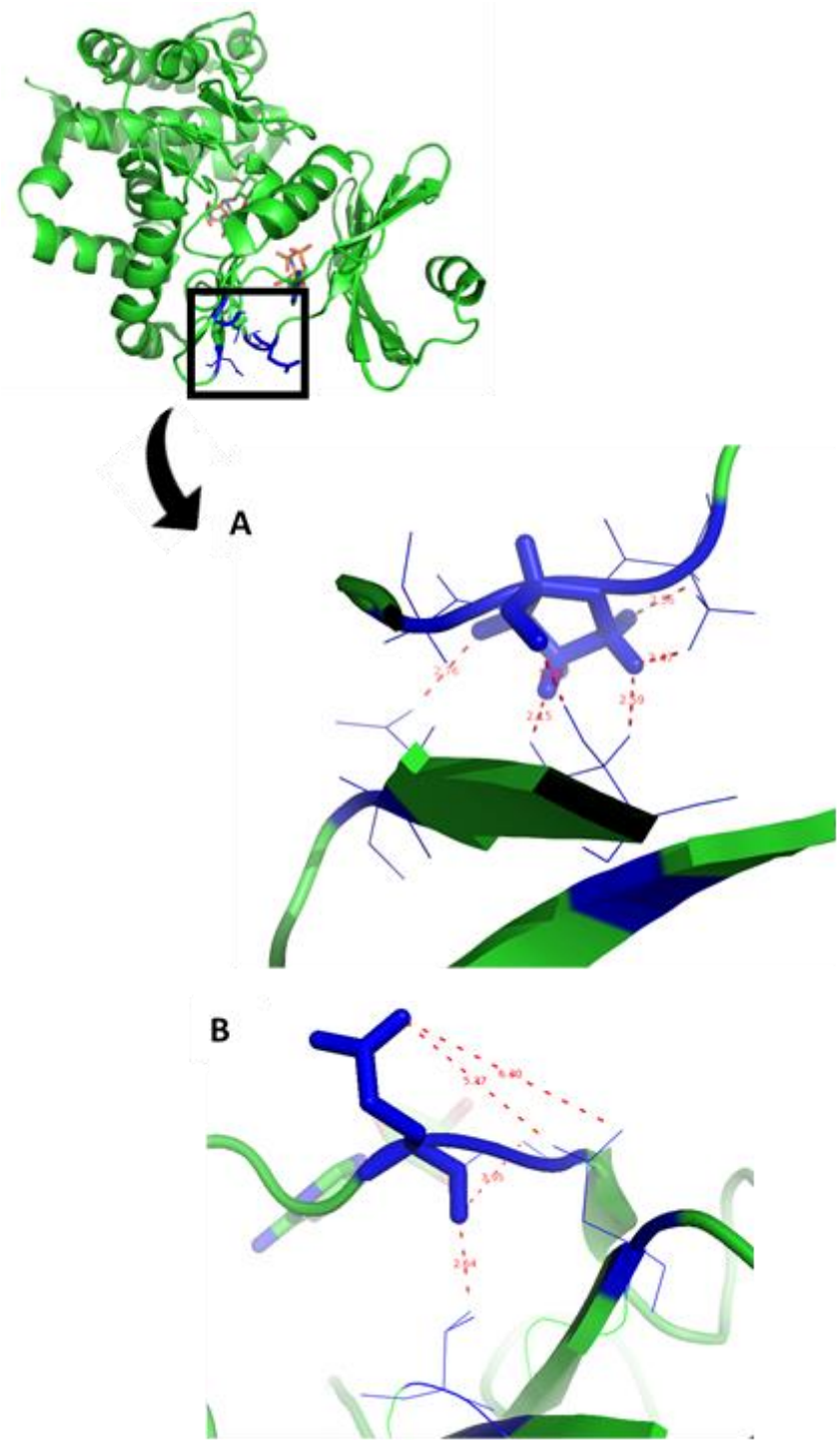
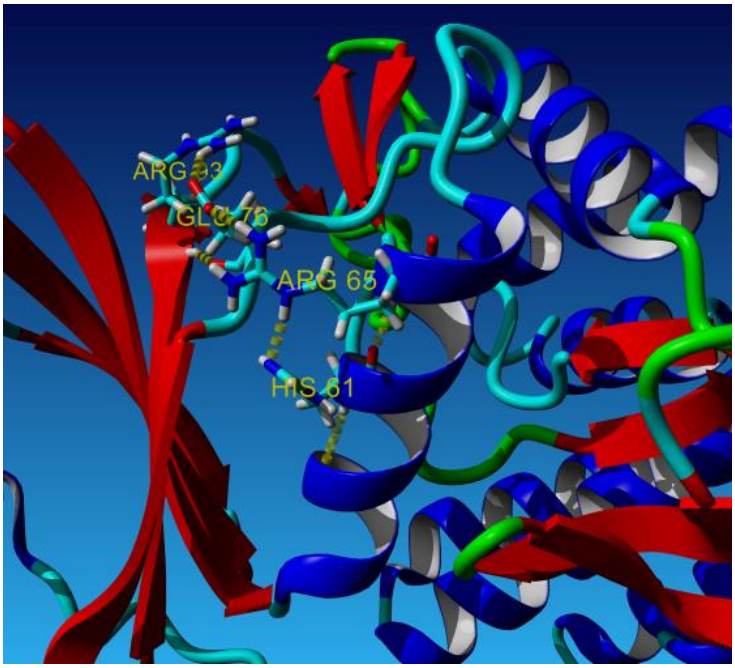


Figure 13. Section of the protein corresponding to the box. Formed bonds between Pro96 and its close amino acids for Hph17 (A) and Glu96 and its close amino acids for Hph5 (B).

A



B

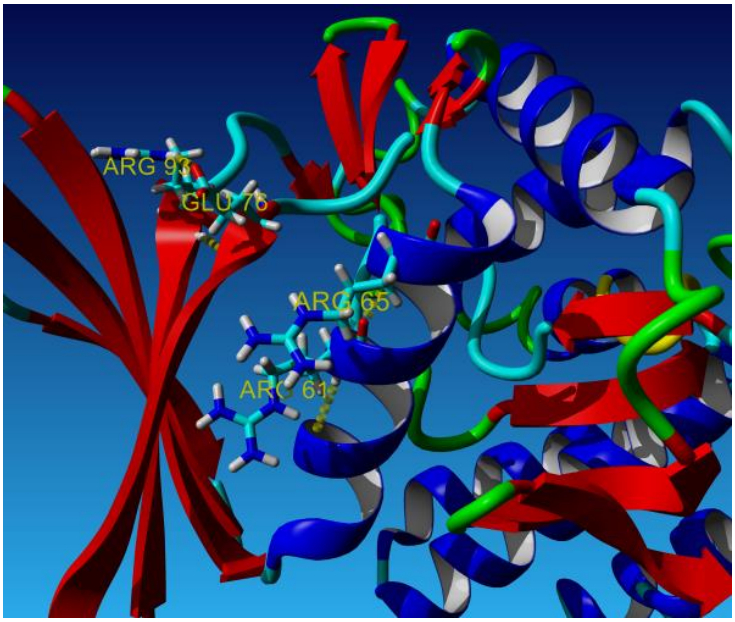


Figure 14. Hydrogen bonds formed between His61, Arg65, Glu76 and Arg93 for Hph17 (A) and hydrogens bonds formed between Arg61 and Arg65 and between Glu76 and Arg93 for Hph5(B).

Figure 14 shows the hydrogen bonds, the substitution R61H implies a new interaction between His61 and Arg65. This new bond causes a slight movement of Arg65, which permits a new hydrogen bond with Glu76 and finally this Glu76 binds with Arg93 with 3 bonds. While in Hph5, there is only one interaction Arg61 and Arg65, and thus there is no interaction between Arg65 and Glu76. Consequently, Glu76 only forms two bonds with Arg93.

His61 and Arg65 are in the same α-helix, but Glu76 is in a β-sheet and Arg93 in another β-sheet. The position of these amino acids in the protein and the new hydrogen bonds formed bring

closer the α-helix and the 2 β-sheets and, consequently, this substitution may stabilize the structure of Hph17.

3.2.2. Determination of melting temperature

Proteins Hph5 and Hph17, which have a molecular weight of 38 kDa, were expressed in *E. coli* and purified by IMAC *(Figure 15)*.

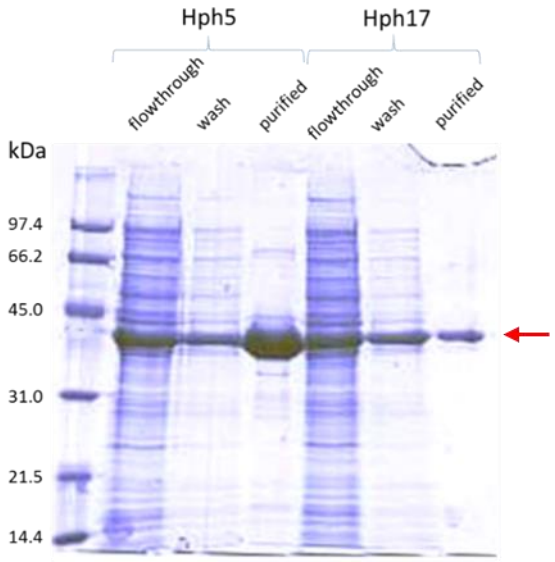


Figure 15. 12 % SDS-PAGE that shows flowthrough, wash and protein purified of both proteins, Hph5 and Hph17. The proteins were expressed in E. coli BL21. The molecular weight of Hph is 38 kDa. Band that might correspond to the expressed protein is marked by red arrow.

As shown in *Figure 15*, Hph17 had a high purity, but Hph5 did not obtain the same purity, as some other proteins were still present. Protein concentrations were determined using *Bio-Rad Protein Assay*. Hph5 had a concentration of 4.61 mg/mL and Hph17, 2.93 mg/mL.

The melting temperature (T_m) of both protein variants was determined by DSF (*Figure 16)*. This temperature was calculated to be 58.8 ± 1.52 °C for Hph5 and 59.2 ± 1.55 °C for Hph17. The T_m of Hph5 coincided exactly with the result obtained by Nakamura *et al.,* 2008, but the melting temperature difference is only 0.4 °C, having no significant differences between them. Although the apparent thermostability of Hph17 has not increased, it is obvious that the functionality at high temperature has increased, especially considering that the maximum temperature at which expressing transformants could grow is 67 and 74 °C for Hph5 and Hph17, respectively.

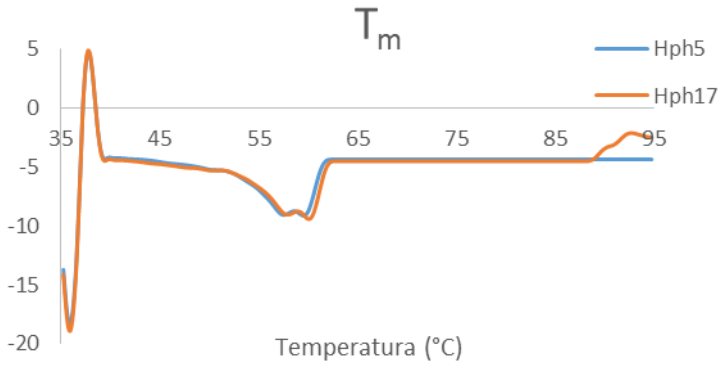


Figure 16. Values obtained at DSF at different temperatures. It is represented the average of the triplicates of both proteins.

The T_m of Hph17 variant has not showed a higher thermostability *in vitro* assays, however it has been observed that cells could grow up to 74 °C, suggesting that *in vivo* the thermostability of the protein can be influenced by other factors, such as compatible solutes that can protect the proteins (Faria *et al.,* 2008).

It also suggests that factors other than thermostability may be contributing to the enhanced survival of Hph17 expressing transformants. Although the substitutions are not in the catalytic centre, the conformational changes might confer an improvement in activity (Zhao *et al.,* 1997; Elleuche *et al.,* 2015), allowing a higher resistance to hygromycin B, thus increasing the growth.

It has been reported that selection for certain codons can help to thermostabilize the protein. Highly expressed genes often use a more restricted set of "preferred" synonymous codons than other, less highly expressed genes. The codon usage patterns have a functional significance. It has been shown that codon usage mirrors the distribution of tRNA abundances, indicating that the "preferred" codons are those that match the more abundant anticodons. It was shown significant changes in the relative frequencies between thermophiles and mesophiles: there are increases in 11 codons among the thermophiles, while there are significant decreases in 12 codons (Singer *et al.,* 2003). Reminding the nucleotide substitution, *hph17* gene includes one infrequent codon, but it also eliminates 2 uncommon codons, having a balance of losing one infrequent codon. It could be concluded that *hph17* changed the pattern of difference, but not enough to explain the fact that selection for survival at high temperature is linked to an organism's ability to produce sufficient amounts of functional proteins under conditions of thermal stress.

This result illustrates the rule 'you only get what you screen for' (You and Arnold, 1996), because even though the desired result was the screening for increased thermostability, the real

selection was for increased survival of the transformant, regardless of the mechanism used to achieve such survival. Ideally, selection methods to find enhanced enzymes should try to isolate a single feature of interest, in order to minimize unwanted false positives caused by e.g., higher expression levels or changes in the activity.

4. CONCLUSIONS

The method of folding interference at high temperature requires a reporter and a target protein. One reporter described is a thermostable hygromycin B phosphotransferase, Hph5, but this reporter was not thermostable enough, thus it did not provide selection. In this work, we have tried to obtain a more thermostable mutant, using *hph5* as a parental gene.

We have obtained a candidate to replace Hph5, it is Hph17, which has 5 more amino acid substitutions, R61H, S86G, Q96P, A185V and V322E. These changes implied differences in the protein model structure, as an interaction between an α-helix and a β-sheet or changes between amino acids interactions.

Hph17 permitted a cell growth of 74 °C, a difference of 6 °C comparing with Hph5, although it is important to mention that the melting temperature (T_m) has only increased 0.4 °C. This suggests that other factors, such activity or differences in the codon usage, would enhance survival of Hph17-expressing transformants.

However, the results demonstrate that Hph17 can be used as a selection marker, the following purpose would be to assess the potential for use of the Hph17 protein as reporter lies in measuring the thermostability with other methods, such circular dichroism or differential scanning calorimetry, or testing enzymatic analysis. Moreover, it should be verified that Hph17 is an optimum reporter in the method of folding interference providing an optimum correlation between the thermostability of the target protein and the host survival.

5. BIBLIOGRAPHY

Arnold, F. H., & Georgiou, G. (2003). *Directed evolution library creation. Methods in molecular biology*, volume 231, chapter 1.

Balkwill, D. L., Kieft, T. L., Tsukuda, T., Kostandarithes, H. M., Onstott, T. C., *et al.* (2004). Identification of iron-reducing *Thermus* strains as *Thermus scotoductus*. *Extremophiles*, 8(1), 37-44.

Becker, B., & Cooper, M. A. (2012). Aminoglycoside antibiotics in the 21st century. *ACS chemical biology*, 8(1), 105-115.

Beckman, R. A., Mildvan, A. S., & Loeb, L. A. (1985). On the fidelity of DNA replication: manganese mutagenesis in vitro. *Biochemistry*, 24(21), 5810-5817.

Borovinskaya, M. A., Shoji, S., Fredrick, K., & Cate, J. H. (2008). Structural basis for hygromycin B inhibition of protein biosynthesis. *Rna*, 14(8), 1590-1599.

Brakmann, S., & Schwienhorst, A. (Eds.). (2006). *Evolutionary methods in biotechnology: clever tricks for directed evolution.* John Wiley & Sons.

Bruins, M. E., Janssen, A. E., & Boom, R. M. (2001). Thermozymes and their applications. *Applied biochemistry and biotechnology*, 90(2), 155-186.

Cannio, R., Contursi, P., Rossi, M., & Bartolucci, S. (1998). An Autonomously Replicating Transforming Vector for *Sulfolobus solfataricus*. *Journal of bacteriology*, 180(12), 3237-3240.

Cava, F., Hidalgo, A., & Berenguer, J. (2009). *Thermus thermophilus* as biological model. *Extremophiles*, 13(2), 213-231.

Chautard, H., Blas-Galindo, E., Menguy, T., Grand'Mourselle, L., Cava, F., Berenguer, J., & Delcourt, M. (2007). An activity-independent selection system of thermostable protein variants. *Nature methods*, 4(11), 919-921.

Choi, J. M., Han, S. S., & Kim, H. S. (2015). Industrial applications of enzyme biocatalysis: Current status and future aspects. *Biotechnology advances*.

Eijsink, V. G., Gåseidnes, S., Borchert, T. V., & van den Burg, B. (2005). Directed evolution of enzyme stability. *Biomolecular engineering*, 22(1), 21-30.

Elleuche, S., Schäfers, C., Blank, S., Schröder, C., & Antranikian, G. (2015). Exploration of extremophiles for high temperature biotechnological processes. *Current opinion in microbiology*, 25, 113-119.

Faria, T. Q., Mingote, A., Siopa, F., Ventura, R., Maycock, C., & Santos, H. (2008). Design of new enzyme stabilizers inspired by glycosides of hyperthermophilic microorganisms. *Carbohydrate research*, 343(18), 3025-3033.

Friedrich, A., Prust, C., Hartsch, T., Henne, A., & Averhoff, B. (2002). Molecular analyses of the natural transformation machinery and identification of pilus structures in the extremely thermophilic bacterium *Thermus thermophilus* strain HB27. *Applied and environmental microbiology*, 68(2), 745-755.

Henne, A., Brüggemann, H., Raasch, C., Wiezer, A., Hartsch, *et al.* (2004). The genome sequence of the extreme thermophile *Thermus thermophilus*. *Nature biotechnology,* 22(5), 547-553.

López-López, O., Cerdán, M. E., & Siso, M. I. G. (2014). New extremophilic lipases and esterases from metagenomics. *Current protein & peptide science*, 15(5), 445.

Madigan, M. T., Clark, D. P., Stahl, D., & Martinko, J. M. (2010). *Brock Biology of Microorganisms 13th edition.* Benjamin Cummings. Chapter 5.

Markoulatos, P., Siafakas, N., & Moncany, M. (2002). Multiplex polymerase chain reaction: a practical approach. *Journal of clinical laboratory analysis*, (16), 47-51.

Nakamura, A., Takakura, Y., Kobayashi, H., & Hoshino, T. (2005). In vivo directed evolution for thermostabilization of *Escherichia coli* hygromycin B phosphotransferase and the use of the gene as a selection marker in the host-vector system of *Thermus thermophilus*. *Journal of bioscience and bioengineering*, 100(2), 158-163.

Nakamura, A., Takakura, Y., Sugimoto, N., Takaya, N., Shiraki, K., & Hoshino, T. (2008). Enzymatic analysis of a thermostabilized mutant of an *Escherichia coli* hygromycin B phosphotransferase. *Bioscience, biotechnology, and biochemistry*, 72(9), 2467-2471.

Niehaus, F., Bertoldo, C., Kähler, M., & Antranikian, G. (1999). Extremophiles as a source of novel enzymes for industrial application. *Applied microbiology and biotechnology*, 51(6), 711-729.

Ramirez, M. S., & Tolmasky, M. E. (2010). Aminoglycoside modifying enzymes. *Drug Resistance Updates*, 13(6), 151-171.

Rothschild, L. J., & Mancinelli, R. L. (2001). Life in extreme environments. *Nature*, 409(6823), 1092-1101.

Rozzell, J. D. (1999). Commercial scale biocatalysis: myths and realities. *Bioorganic & medicinal chemistry*, 7(10), 2253-2261.

Sadeghi, M., Naderi-Manesh, H., Zarrabi, M., & Ranjbar, B. (2006). Effective factors in thermostability of thermophilic proteins. *Biophysical chemistry*, 119(3), 256-270.

Singer, G. A., & Hickey, D. A. (2003). Thermophilic prokaryotes have characteristic patterns of codon usage, amino acid composition and nucleotide content. *Gene*, 317, 39-47.

Schwarzenlander, C., & Averhoff, B. (2006). Characterization of DNA transport in the thermophilic bacterium *Thermus thermophilus* HB27. *FEBS Journal*, 273(18), 4210-4218.

Shivange, A. V., Marienhagen, J., Mundhada, H., Schenk, A., & Schwaneberg, U. (2009). Advances in generating functional diversity for directed protein evolution. *Current opinion in chemical biology,* 13(1), 19-25.

Trivedi, S., Gehlot, H. S., & Rao, S. R. (2006). Protein thermostability in Archaea and Eubacteria. *Genetics and Molecular Research*, 5(4), 816-827.

Urbieta, M. S., Donati, E. R., Chan, K. G., Shahar, S., Sin, L. L., & Goh, K. M. (2015). Thermophiles in the genomic era: Biodiversity, science, and applications. *Biotechnology advances.*

Wilson, D. S., & Keefe, A. D. (2001). Random mutagenesis by PCR. *Current Protocols in Molecular Biology*, Chapter 8.3.

Wong, T. S., Roccatano, D., & Schwaneberg, U. (2007). Are transversion mutations better? A Mutagenesis Assistant Program analysis on P450 BM-3 heme domain. *Biotechnology journal*, 2(1), 133-142.

You, L., & Arnold, F. H. (1996). Directed evolution of subtilisin E in Bacillus subtilis to enhance total activity in aqueous dimethylformamide. *Protein Engineering*, 9(1), 77-83.

Zhao, H., & Arnold, F. H. (1997). Combinatorial protein design: strategies for screening protein libraries. *Current opinion in structural biology*, 7(4), 480-485.