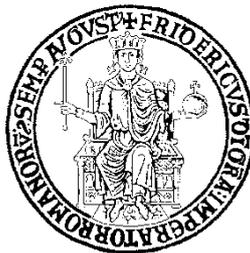


**UNIVERSITY OF NAPLES “FEDERICO II”**

DEPARTMENT OF BIOLOGY

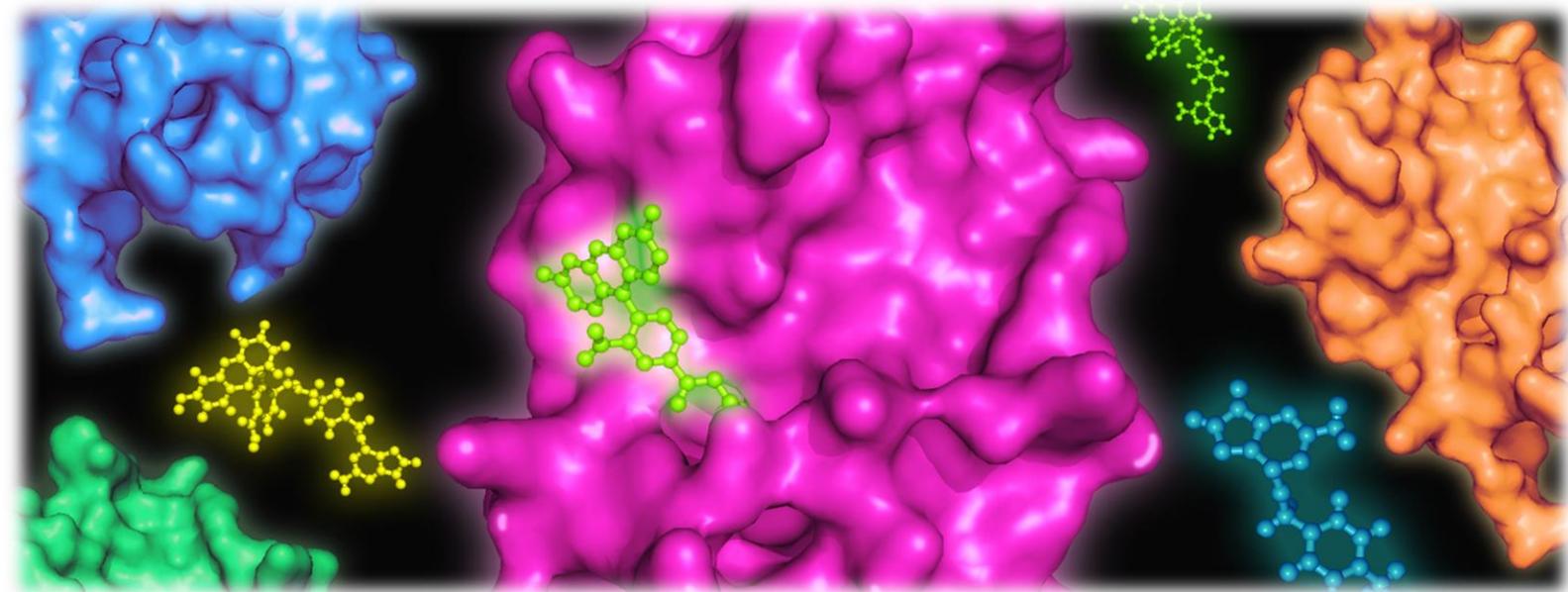


**DOCTORATE SCHOOL IN BIOLOGY**

Cycle XXXIII

# **Thermostable DNA repair enzymes for novel biotechnological applications**

ROSA MERLO



**UNIVERSITY OF NAPLES “FEDERICO II”**  
DEPARTMENT OF BIOLOGY



**DOCTORATE SCHOOL IN BIOLOGY**

Cycle XXXIII

**Thermostable DNA repair enzymes  
for novel biotechnological applications**

**Coordinator**

Prof. Salvatore Cozzolino

**PhD Student**

Rosa Merlo

**Supervisor (Unina)**

Prof. Ezio Ricca

**Supervisor (CNR)**

Dr Giuseppe Perugino

## **Index**

<i>Summary</i>	1
<b>Preface</b>	5
DNA direct repair activity upon alkylation damages	5
hMGMT as “biotech” tool	8
Extremophiles and thermophilic organisms: thermostable AGTs	9
<b>Aim of the thesis</b>	10
<b>Chapter I: (hyper)thermophilic <i>protein-tags</i></b>	11
<b>1. Introduction</b>	13
<i>1.1 The SNAP-tag<sup>®</sup> technology</i>	13
<i>1.2 SsOGT-H<sup>5</sup></i>	15
<i>1.3 Pyrococcus furiosus AGT</i>	17
<b>2. Results and Discussion</b>	19
<i>2.1 Identification and structure comparison of PfuAGT</i>	19
<i>2.2 Expression and purification</i>	20
<i>2.3 Biochemical characterization: activity assay, catalytic activity (progress curve) and thermal stability</i>	21
<b>3. Conclusions and perspectives</b>	24
<b>Chapter II: Evolution of the SNAP-tag<sup>®</sup> technology</b>	27
<b>1. Introduction</b>	29
<i>1.1 The Click Chemistry</i>	29
<i>1.2 The novel approach</i>	30
<b>2. Results and Discussion</b>	31
<i>2.1 Substrate specificity of AGTs on BG-based substrates</i>	31
<i>2.2 Molecular modelling on the H<sup>5</sup> with BG-azides</i>	32
<i>2.3 In vitro reaction of engineered AGTs with BG-azide substrates</i>	33
<i>2.4 Specificity and versatility of the chemo-enzymatic reaction</i>	36
<i>2.5 Application to the Bio-Layer Interferometry (BLI)</i>	38
<i>2.6 Permeability of eukaryotic and prokaryotic cells to BG-azides</i>	41
<b>3. Conclusions</b>	43
<b>Chapter III: Anchoring and self-labelling protein-tag system (ASLtag)</b>	44
<b>1. Introduction</b>	46
<i>1.1 Immobilization procedures</i>	46

<b>2. Results and Discussion</b>	48
2.1 <i>SsβGly</i>	51
2.2 <i>SspCA</i>	53
<b>3. Conclusions and perspectives</b>	59
<b>Materials and Methods</b>	62
<b>References</b>	74
<b>List of scientific publications</b>	85
<b>Meetings, Workshops and Webinar</b>	86
<b>Acknowledgements</b>	

## Summary

*O*<sup>6</sup>-alkylguanine-DNA-alkyltransferases (AGT, OGT or MGMT, EC 2.1.1.63) are small enzymes (< 25 kDa) involved in the DNA repair, by the removal of alkyl groups on *O*<sup>6</sup>-position of guanines, through an irreversible single-step reaction (Pegg, 2000). The peculiar mechanism led to the development of useful biotechnological tools for the specific labelling of proteins (Pegg et al., 2011), e.g. the so-called SNAP-*tag*<sup>®</sup> technology (Gautier et al., 2008). Thanks to their wide specificity, these enzymes are able to recognise and repair free nucleobases, even with bulky alkyl adducts on the *O*<sup>6</sup>-position (e.g. *O*<sup>6</sup>-benzyl guanine; BG) (Keppler et al., 2003): because of the retaining of the benzyl moiety in their active site, AGTs could be used as *protein-tag* for the covalent labelling with any chemical group of interest, if previously conjugated to the BG's benzylic group (Huber et al., 2004). However, despite the SNAP-*tag*<sup>®</sup> technology being a powerful tool, it has some limitations both in terms of the enzyme and of the relative BG-derivative substrates. In fact, given the mesophilic nature of this *protein-tag*, the general use of this approach is restricted to mild reaction conditions and mesophilic model systems. Furthermore, apart from commercially available BG-derivatives, the synthesis of *ad hoc* substrates has the disadvantage of tedious and complex purifications.

To overcome these limitations and in the context of expanding the biotechnological applications of SNAP-*tag*<sup>®</sup>, during my PhD I focused my activity on: *i*) the identification and characterization of a novel hyper-thermostable AGT from the archaeon *Pyrococcus furiosus*, in collaboration with the group of Prof. Michael Terns (University of Georgia) for the *in vivo* studies of proteins involved in the CRISPR Cas system (Chapter I); *ii*) the improvement and expansion of the SNAP-*tag*<sup>®</sup> technology, with an innovative *chemo-enzymatic* approach (Chapter II), in collaboration with Prof. Alberto Minassi (University of Piemonte Orientale); *iii*) the set up and characterization of a new one-step labelling and immobilization system (the *anchoring and self-labelling protein-tag-ASL*<sup>tag</sup>), in collaboration with Dr Clemente Capasso (National Research Council of Italy - IBBR) (Chapter III).

### Chapter I

The growing demand to apply a *protein-tag* to extreme conditions, especially in organisms living at very high temperatures, led us to look for new proteins from hot sources. Recently, the group of Dr

Perugino focused its attention on this aspect, aim to improving the SNAP-tag<sup>®</sup> technology, acting on the enzyme. Consequently, an AGT from the archaeon *Saccharolobus solfataricus* (*SsOGT*) was first characterized. Furthermore, while the enzyme has been demonstrated to remain stable at high temperatures and in physical and chemical denaturing agents (Perugino et al., 2012, Vettone et al., 2016), it is subsequently engineered to achieve a more active enzyme (*SsOGT-H<sup>5</sup>* variant).

In order to study *in vivo* CRISPR-Cas immune systems in the hyperthermophilic archaeon *Pyrococcus furiosus* (*Pfu*) (Terns and Terns 2013), with an optimal growth temperature of  $\geq 100$  °C, I identified, expressed and characterized the *PfuOGT*, following the same strategy used before for *SsOGT* (Perugino et al., 2012). *PfuOGT* displayed different enzymatic reaction rate and different stability from *SsOGT*. Apart from the DNA repair activity measured for this enzyme, a very high thermophilicity and thermostability was also confirmed, far exceeding the same measured biochemical parameters of the *Saccharolobus* counterpart (Perugino et al., 2012; Perugino et al., 2015). This data together proposed *PfuOGT* as a promising tool for the *in vivo* analysis and function of proteins of interest in (hyper)thermophilic model systems.

During the last part of my PhD program, I started working on the engineered version of *PfuOGT* protein, aimed to obtain a DNA binding-less mutant, by following the same approach used for *SsOGT-H<sup>5</sup>* (Gautier et al., 2008; Perugino et al., 2012). After the identification of the residues involved in the DNA binding activity, we prepared a synthetic gene to be replaced in the *PfuOGT* wild-type plasmid vector. From first analysis in terms of expression level, results encouraged us that the protein has been successfully expressed in *E. coli* cells and it was stable at high temperatures ( $> 80$  °C) in crude cell extract context.

The advent of Covid-19 emergency drastically slowed down our research activity. However, this protein will be characterized in the near future and tested by Prof. Terns and his group as *protein-tag*, to study the CRISPR-Cas system proteins *in vivo* in the already obtained *Pyrococcus furiosus ogt*- knock-out strain (KO).

## Chapter II

SNAP-tag<sup>®</sup> technology is a powerful technique for the covalent labelling of a protein of interest with a desired chemical group (BG-derivatives), without affecting its activity and stability. Although commercially available or *ad hoc* produced, the synthesis and purification of each BG substrate are necessary, increasing time and costs. Moreover, the risk of lower affinity and catalytic activity of these *tags* with new or customized BG-derivatives cannot be underestimated. To get over this issue, I proposed a modification of the SNAP-tag<sup>®</sup> technology, by developing a *chemo-*

*enzymatic* approach with a selected azide-based BG-substrate (BGSN), chemically synthesized by Prof. Minassi of University of Piemonte Orientale “Amedeo Avogadro” (UPO) and his group. *This approach is based on the SNAP-tag<sup>®</sup> enzymatic reaction as first step, followed by the extremely specific Huisgen-type Cu(I)-catalysed cycloaddition (Cu-AAC, or “click-chemistry” reaction).*

First, I tested the BGSN with the commercial SNAP-tag<sup>®</sup> and the thermostable SsOGT-H<sup>5</sup>: both enzymes were active and still correctly folded upon enzymatic reaction, allowing a sufficient exposure to the solvent of the azide-moiety covalently linked to the catalytic cysteine, and leading to the subsequent labelling of both proteins with different alkyne-based fluorophores by click reaction, as in the presence of *in vitro* and *in vivo* “perturbing” proteins, like in *E. coli* lysates as well as in eukaryotic cells. The sufficient solvent exposure of the azide group was also confirmed by the Molecular Dynamics analysis, in collaboration with Dr Miggiano of UPO. This new procedure allowed also the direct immobilization of these *protein-tags* on Bio-Layer Interferometry (BLI) solid surface, avoiding any steric hindrance. BLI is an optical analytical technique that allows the measure of the interactions between an immobilized (ligand) and a free molecule in solution (analyte). Since alkyne-derivative sensors are not currently commercially available, I necessarily had to functionalize the sensor with a *bi-functional* linker, in order to expose an alkyne group to the solvent. The sensorgrams obtained by the BLI showed that only the protein linked to the azide-substrate was covalently immobilized on the sensor, while the free protein and the only sensor had no obvious signal.

In order to set up a more efficient and safe copper-free reaction, I used Dibenzocyclooctyne (DBCO)-derivative chemical groups for the chemo-enzymatic approach: these compounds can covalently bond the azide-moiety conjugated to the protein substrate (BGSN) and tested on both SNAP-tag<sup>®</sup> and H<sup>5</sup> evaluating the enzymatic reaction which occurred between protein and BGSN, as well as the cycloaddition in terms of reaction rate. Moreover, in collaboration with the group of Prof. Leonardi (University of Naples “Federico II”), we tested the novel approach on an eukaryotic cell model for fluorescence imaging approach.

Finally, I demonstrated that splitting the SNAP-tag<sup>®</sup> reaction into two fast and highly specific steps does not affect the overall rate and efficiency of the protein labelling. Thus, this approach could offer the advantage to prepare only one universal substrate (taking also into account of a unique reaction rate) and by employing an infinite number of commercially available and inexpensive DBCO-based compounds, without the need of any other purification, reducing time and costs for the protein labelling and offer several possible biotechnological applications.

### Chapter III

It is known that the immobilization of enzymes is important to overcome the instability in harsh operational conditions and to improve their recycling and utilization in biotechnological applications (Zhou et al., 2013). The bacterial cell surface-display of enzymes and their direct immobilization represent one of the most interesting approaches (Samuelson et al., 2002; Daugherty, 2007).

During my first and second year, I focused my activity on the ASL<sup>tag</sup>, formed by the amino-terminal domain of the transmembrane ice nucleation protein (INP), from the Gram-negative bacterium *Pseudomonas syringae* (Cochet and Widehem 2000; Gurian-Sherman et al., 1993), and the thermostable *protein-tag* SsOGT-H<sup>5</sup>. I have characterized and tested this *innovative protein-tag* alone and in fusion with thermostable enzymes:  $\beta$ -glycoside hydrolase from *S. solfataricus* (Moracci et al., 1996) and the  $\alpha$ -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* (SspCA) (Del Prete et al., 2017; Capasso et al., 2012) in collaboration with Dr. Capasso of the Institute of Biosciences and BioResources (IBBR-CNR). I demonstrated that the presence of this novel tool does not affect the activity of both enzymes fused to the ASL<sup>tag</sup>, and the presence of this thermostable *tag* enhanced the SspCA thermostability, compared to its counterpart directly fused to the INPN.

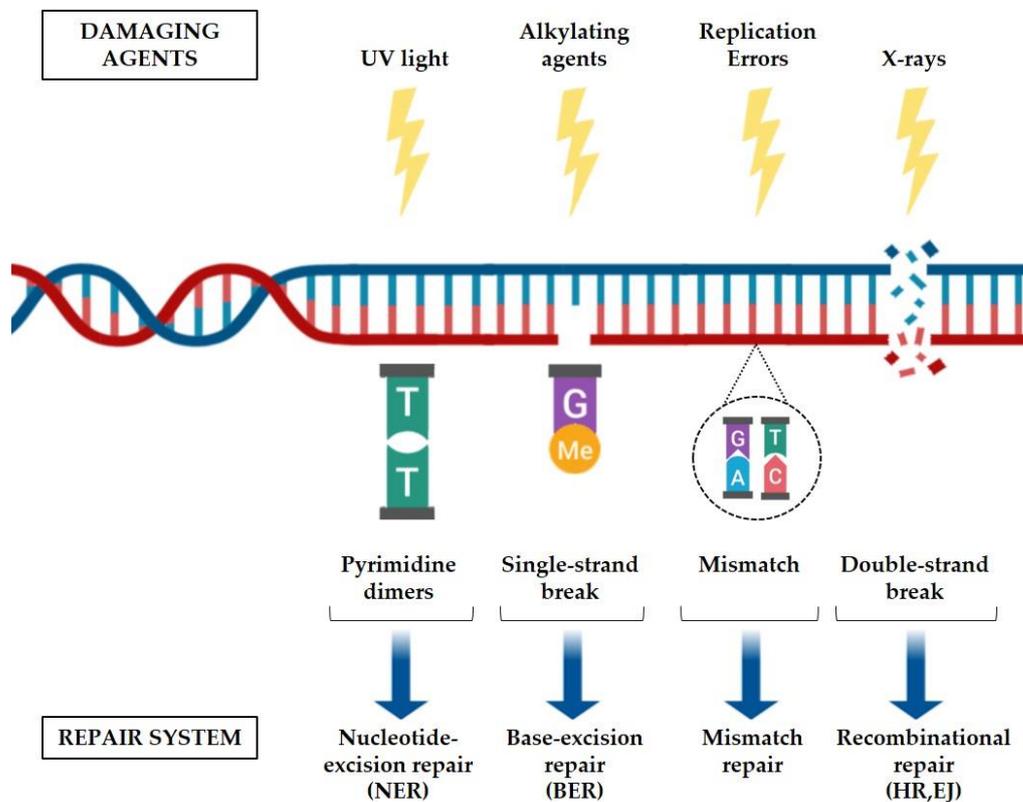
In the last part of my PhD, I worked on developing a novel ASL<sup>tag</sup> (hereinafter ASL<sup>tag2</sup>) system, based on the Halotag<sup>®</sup> protein, a haloalkane dehalogenase with a genetically modified active site, which reacts irreversibly with primary alkyl-halides (Los et al., 2008; England et al., 2015). Obtained data demonstrated that the new ASL<sup>tag</sup> also led to the expression and the immobilization of enzymes on the outer membrane of bacteria cells, and the presence of the *protein-tag* moiety, reacting with a fluorescent substrate, allowed the quantitative determination of the fused enzyme and its localization by *gel-imaging* techniques. Subsequently, ASL<sup>tag</sup> and ASL<sup>tag2</sup> have been successfully co-expressed in the same bacteria, presenting different orthogonal specificities, and allowing the contemporary labelling of *protein-tag* systems.

As part of my PhD activity, I was involved and I contributed to set up a novel AGTs assay using DNA triplex-based substrates (see paragraph 1.1 in Chapter I and List of scientific publications).

# Preface

## DNA direct repair activity upon alkylation damages

It is well known that the maintenance of genome integrity is essential for the survival of living organisms under different conditions. The nature of DNA and the processes to which it is subjected to, above all replication and recombination, continually expose it to various types of damage (Figure 1). As result, the cells set up different mechanisms for tolerating or repairing these events (Liu et al., 2006; Sharma et al., 2009). However, the malfunction of these systems and the maintenance of errors inside its structure may interfere with DNA metabolism and natural functions (Lindahl et al., 1993; Friedberg et al., 2003).



**Figure 1.** Schematization of different damages and repair systems. On top, the best-known examples of damaging agents and the relative effects on the DNA. At the bottom, the respective repair systems

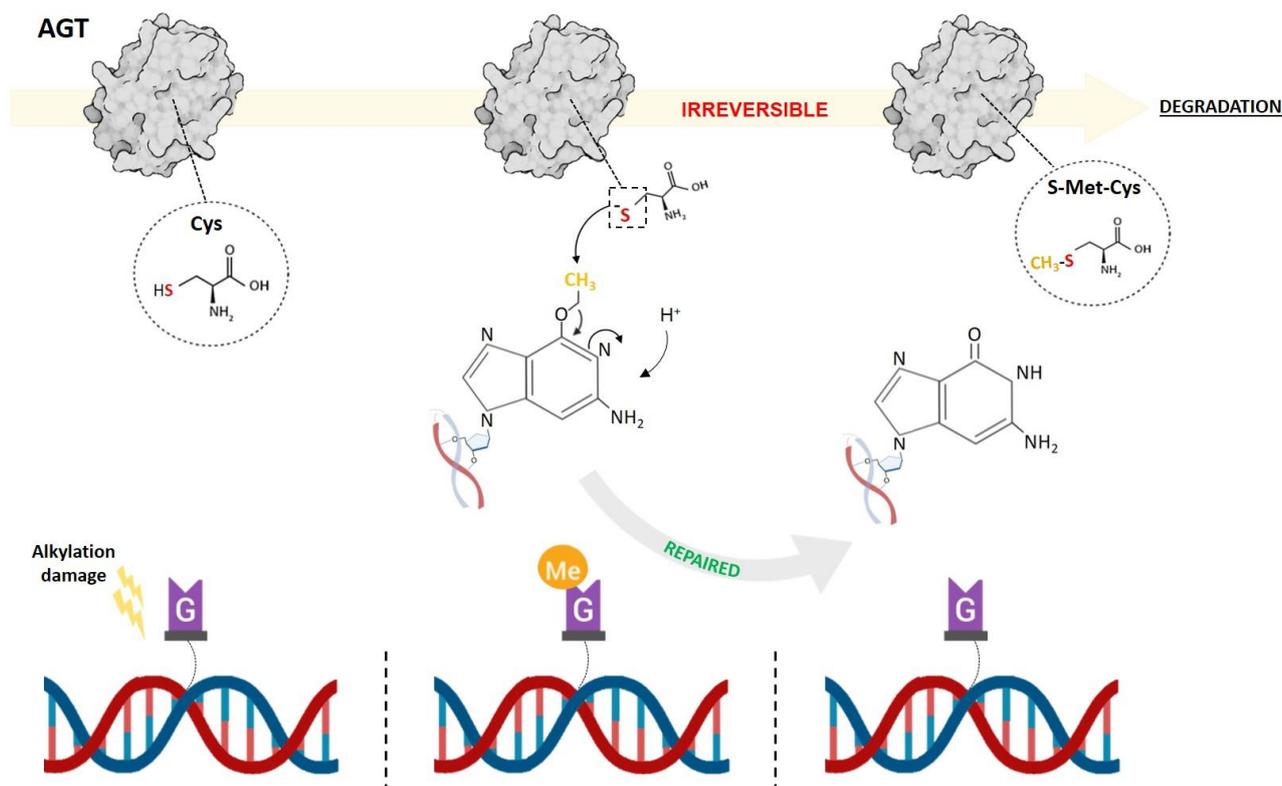
Because of its medical interest, the alkylation damage is still under study. In fact, alkylating agents are mutagenic molecules commonly used as therapeutic drugs in the treatment of cancer, which induce alkyl modification on DNA, but they are also present in the environment or produced inside cells as metabolic products. More specifically, they can induce DNA alkylation in several positions,

including the  $O^6$  of guanine ( $O^6$ -MG; 6% of adducts formed), the  $N^7$  of guanine ( $N^7$ -MG; 70%), and the  $N^3$  of adenine ( $N^3$ -MA; 9%) (Liu et al., 2006, Mattosovich et al., 2020), which can lead to replication and transcription arrest or apoptosis. In particular, the alkylation of guanine ( $O^6$ -AG) is particularly cytotoxic, because the  $O^6$ -AG incorrectly pairs with thymine generating a transition from G:C to A:T (Leonard et al., 1990, Singer et al., 1976) presenting potential injurious implications for the cells.

Differently from conventional multi-enzymatic DNA repair systems, alkylated DNA can be directly repaired by the activity of the  $O^6$ -alkyl-guanine-DNA-alkyl-transferases (AGTs or OGTs) or  $O^6$ -methyl-guanine-DNA-alkyl-transferases (MGMTs).

The AGTs are small enzymes (17-22 kDa) that are evolutionarily conserved in organisms of the three domains of life (bacteria, archaea, eukaryotes) (Pegg, 2011). Starting from the available 3D structures, AGTs consist of two domains connected by a large loop (Daniels et al., 2000): *i*) the highly conserved C-terminal domain contains the active site pocket where the catalytic cysteine (Cys145) is sited, and the DNA-binding domain, a *Helix-Turn-Helix* (HTH) motif, which binds the minor groove of the DNA double helix (Wibley et al., 2000); *ii*) the N-terminal domain, the most different moiety among species, which has a structural and regulatory role.

Despite the simplicity of the structure, what most characterizes AGTs is their peculiar reaction mechanism. First, the protein recognizes the single damaged base, with a still unclear mechanism, probably migrating along the DNA; then, AGT actively initiates the *flipping-out* of the injured nucleotide by sensing the distortion caused by an incorrect match (Duguid et al., 2005). After that, the alkyl group is correctly disposed in a hydrophobic pocket in the active site (Vora et al., 1998), the formation of a thiolate anion via proton transfer on the Cys145 occurs. At this point, the alkyl group is *irreversibly* transferred to a highly reactive cysteine by an irreversible one-step  $SN_2$ -like mechanism (Tubbs et al., 2007) (Figure 2, middle).



**Figure 1.** The AGT reaction mechanism. AGT protein is in grey (on top), with the catalytic cysteine (Cys) structure highlighted. The enzyme recognizes the methyl-guanine nucleobase and through the irreversible reaction, the methyl group is transferred to the Cys and the repaired guanine is reinserted into the DNA double helix. At the bottom is a schematic representation of alkylation damage and lesion repair on the DNA.

Upon reaction, whereas the repaired nucleobase is released and returns in the DNA double helix, the protein in its alkylated form is destabilized and rapidly degraded.

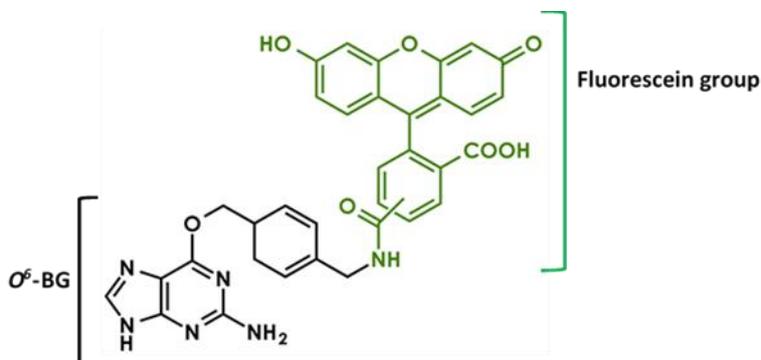
However, because of lack of data on the destabilized protein, which makes any biochemical studies very limited, the exact molecular mechanism of AGT destabilization is still under study (Brunk et al., 2013). What is known is that the AGTs can perform only one reaction before being degraded via the ubiquitin/proteasomal system (Srivenugopal et al., 1996, Xu-Welliver et al., 2002). For this reason, they are defined as *suicide* or *kamikaze* proteins, with a 1:1 stoichiometry of their reaction with the natural substrate (Gerson, 2004; Tubbs et al., 2007).

Since alkylating agents are the principle first-line chemotherapeutic agents used for the treatment of recurrent high-grade gliomas and have also been used to treat advanced malignant melanoma and other solid tumors (Fan et al., 2013), the importance of studying this class of proteins cannot be underestimated. In 1991, A.E. Pegg discovered the human AGT (hMGMT) (Pegg et al., 1991), which normally protects cells from cytotoxic effects, but on the other hand, it counteracts

chemotherapy treatments by preserving cancer cells from the killing effect of alkylating drugs (Gerson, 2002; Kaina et al., 2019). This is because, in a condition of cellular damage as in tumors under treatment, the expression level of the hAGT is extremely high, due also to the rapid turnover that the protein undergoes after completing the reaction. For all these reasons, many studies have been focused on the development of hMGMT inactivators to be used in combination with alkylating agents, resulting in a significant decrease of tumor resistance (Paranjpe et al., 2014). Particularly significant were the *O*<sup>6</sup>-benzylguanine (BG), an analog of the AGTs natural substrate, in which the methyl group is substituted with a benzyl group in the *O*<sup>6</sup> position, and the *O*<sup>6</sup>-[4-bromophenyl]-guanine (*O*<sup>6</sup>-BTG, also known as Lomeguatrib), one of the most powerful hMGMT inactivators found to date (Rabik et al., 2006; Kaina et al., 2010).

### hMGMT as “biotech” tool

The introduction of these inactivators led to the development of new biotechnological tool, by using the hMGMT as self-labelling *protein-tag*. In fact, given the peculiar and extremely specific reaction mechanism, Prof. K. Johnsson developed an engineered version called SNAP-*tag*<sup>®</sup> (Mollwitz et al., 2012; Gronemeyer et al., 2006) (as briefly described in Summary). All his technology is based on the reaction that SNAP-*tag*<sup>®</sup> performs with part of its substrate. Starting from the BG, several derivatives have been produced for this tool. These molecules are composed of a constant part, represented by the *O*<sup>6</sup>-benzyl-guanine, and a desired chemical group conjugated to the benzyl moiety. Most of these BG substrates are used for cell imaging applications, therefore they are characterized by the presence of a fluorescent or a quencher group bound to the benzyl ring of the substrate. A classic example is SNAP-Vista<sup>®</sup> Green (SVG) from the New England Biolabs, presenting a fluorescein group (Figure 3).



**Figure 2.** SNAP-Vista<sup>®</sup> Green substrate (NEB).

## Extremophiles and thermophilic organisms: thermostable AGTs

Organisms that live in conditions considered extreme for all life forms we know are defined “Extremophiles” (Rothschild and Mancinelli, 2001). These organisms dominated the evolutionary history of the planet, in fact, their origins probably date back to millions of years ago when the Earth was far from hosting the life forms we are used to thinking, and today they still populate those uncomfortable environments (Canganella and Wiegel, 2011). Most of the extremophiles belong to the Archaea domain (Woese CR et al., 1990), showing some traits that are peculiar and differ from those of Bacteria and Eukarya and others that are shared, as DNA replication and translation. Archaea are commonly classified based on their lifestyle and habitat, but of particular biotechnological interest are those who live at high temperatures: Thermophiles. As observed in organisms living under mesophilic conditions, the genomes of thermophilic and hyper-thermophilic organisms are attacked physically and chemically by environmental and endogenous alkylating agents. In addition, high temperatures may accelerate the process of alkylation, leading to DNA breaks (Valenti et al., 2006, Mattosovich et al., 2020). Thus, the presence of AGTs in (hyper)thermophilic organisms implies that these proteins play a crucial role in protecting genome stability and allow cell survival (Leclere et al., 1998; Skorvaga et al., 1998).

Among thermophilic archaea, the *Saccharolobus solfataricus* (previously known as *Sulfolobus solfataricus*) AGT (*SsOGT*) has been extensively characterized by Perugino and co-workers, even considering that differing from other thermophilic homologues, the protein heterologously expressed in *E. coli* is soluble, very stable and easy to purify (Perugino et al., 2012).

The open reading frame (ORF) n. SSO2487 encodes a protein of molecular mass 17 kDa, biochemically characterized and classified as *O*<sup>6</sup>-alkylguanine-DNA-alkyltransferase (Perugino et al., 2012). Structural highly similar in C-terminal domain to hMGMT and others, *SsOGT* has a high thermal stability and a maximum activity at 80 °C, but is also active even at lower temperatures. It is able to work under various reaction conditions such as pH extremes (5.0 - 8.0), high ionic strength (up to 2M NaCl) and in the presence of detergents (Perugino et al., 2012).

The biochemical characterization of *SsOGT* was carried out through the development of an innovative assay, based on the use the commercial fluorescein derivative SVG. Given the stoichiometric ratio of 1:1 between protein and substrate, the amount of covalently bound probe is a direct measure of protein activity. This made it possible to calculate the kinetic constants for this molecule and measure the activity of *SsOGT* on its natural substrate (methylated DNA) (Perugino et al., 2012).

## Aim of the thesis

SNAP-tag<sup>®</sup> is a powerful tool with wide applications; however, this technology has some limitations: this technology is suitable only for mesophilic model organisms and mild reaction conditions, and the synthesis of *ad hoc* substrates needs tedious and complex purification procedures.

This thesis is directed to get over these issues and to expand the biotechnological application potential of the SNAP-tag<sup>®</sup>. This aim was pursued by two strategies: acting on the enzyme, by the exploitation of homologous activities from extremophilic sources; and acting on the substrate, by the set-up of a novel approach.

### Acting on the enzyme

In part of my thesis (Chapter I and Chapter III), I addressed my work to the identification and characterization of a novel AGT from the hyperthermophilic organism *P. furiosus*, in order to expand the biotechnological potential of the SNAP-technology from “mesophilic” applications to *in vivo* and *in vitro* studies in extreme model organisms or in those processes that require non-moderate reaction conditions, and to the characterization of a novel immobilization and labelling tool.

### Acting on the substrate

In another part of my thesis (Chapter II), I focused my activity to the set-up of a novel approach to the SNAP-tag<sup>®</sup> reaction, by a *chemo-enzymatic approach*, which splits the reaction in two fast and extremely specific steps. The enzymatic reaction is performed by use of a selected azide-based BG-substrate, which allows the chemical step with a DBCO-based molecule via cycloaddition (or click-chemistry) reaction. I demonstrate that all the problems related to purification procedures can be bypassed by the utilization of a universal azide-derivative substrate, which is sufficiently exposed to the solvent to link any DBCO-molecules without affecting the reaction rate of the SNAP-tag<sup>®</sup>.

## **Chapter I: (hyper)thermophilic *protein-tags***

In this first chapter, the identification and biochemical characterization of an AGT from the hyperthermophilic archaeon *Pyrococcus furiosus* (*Pfu*) are described. This project is in collaboration with Professor M.P. Terns of the University of Georgia (USA) and it aims to find a (hyper)thermostable *protein-tag* to study *in vivo* the CRISPR-Cas immune systems in *P. furiosus*.

Briefly, here is illustrated a general introduction on the SNAP-*tag* technology and the first characterized “thermostable SNAP-*tag*” from *Saccharolobus solfataricus*. Successively, there are reported the obtained results from the identification and characterization of the *Pfu*OGT, describing its thermostability and repair activity on methylated DNA, as well as its activity on BG-derivative substrates.

These data have been published in *Extremophiles* journal (<https://doi.org/10.1007/s00792-019-01134-3>), together with another AGT identified with the same approach in *Thermotoga neapolitana* (*Tn*OGT).

Extremophiles (2020) 24:81-91

<https://doi.org/10.1007/s00792-019-01134-3>

SPECIAL FEATURE: ORIGINAL PAPER



12th International Congress on Extremophiles

## A journey down to hell: new thermostable protein-tags for biotechnology at high temperatures

Rosanna Mattossovich<sup>1\*</sup> · Rosa Merlo<sup>1\*</sup> · Angelo Fontana<sup>2</sup> · Giuliana d’Ippolito<sup>2</sup> · Michael P. Terns<sup>3</sup> · Elizabeth A. Watts<sup>3</sup> · Anna Valenti<sup>1</sup> · Giuseppe Perugino<sup>1</sup>

<sup>1</sup> Institute of Biosciences and BioResources, National Council of Research of Italy, Via P. Castellino 111, 80131 Naples, Italy

<sup>2</sup> Institute of Biomolecular Chemistry, National Council of Research of Italy, Via Campi Flegrei, 34, 80078 Pozzuoli, NA, Italy

<sup>3</sup> Departments of Biochemistry and Molecular Biology, Genetics, and Microbiology, University of Georgia, Athens, GA, USA

\* Rosanna Mattossovich and Rosa Merlo equally contributed to the present work.

# 1. Introduction

Monitoring and studying the behavior of proteins within a cellular context is essential to carry out their role within the cell. *Protein-tags* are short or long peptide sequences genetically fused to recombinant proteins for various purposes, as affinity purification, protein localization and general labelling procedures. The labelling of a protein of interest (POI) is possible thanks to the various available *tags*, such as peptides or proteins, which are generally fused with the POIs. Most of these not only label the protein of interest but determine its better solubilization, avoiding aggregation and misfolding, e.g. maltose-binding protein (MBP), and glutathione S-transferase (GST) (Mattosovich et al 2020). With the introduction of a plethora of fluorescent proteins, the *in vivo* cell biology scenario has really changed (Mattosovich et al., 2020; Chalfie et al. 1994; Tsien 1998; Aliye et al. 2015). However, although fluorescent proteins (FPs) are intrinsically fluorescent without the addition of any external substrate, they have some disadvantages: the relatively large dimensions and the sensitivity to environmental changes (pH or the absence of O<sub>2</sub>) affect the formation of the internal natural fluorophore (Ashby et al., 2004; Campbell and Choy 2000). On the contrary, the development of *self-labelling protein-tags* (SLPs) opened to new possibilities to analyse POIs, both *in vitro* with SDS-PAGE fluorescence scanning and *in vivo* (Hinner and Johnsson, 2010; Kosaka et al., 2009). SPLs are small enzymes that react covalently with their substrate and possess the advantages of high specificity and reaction speed, allowing the use of available commercial substrates or synthesized/customized ones. (Hinner and Johnsson, 2010). The best-known SLPs are the HaloTag<sup>®</sup>, derived from the haloalkane dehalogenase (Los et al., 2008), the SNAP-tag<sup>®</sup> (Keppler et al., 2003) and the CLIP-tag<sup>®</sup> (Gautier et al., 2008).

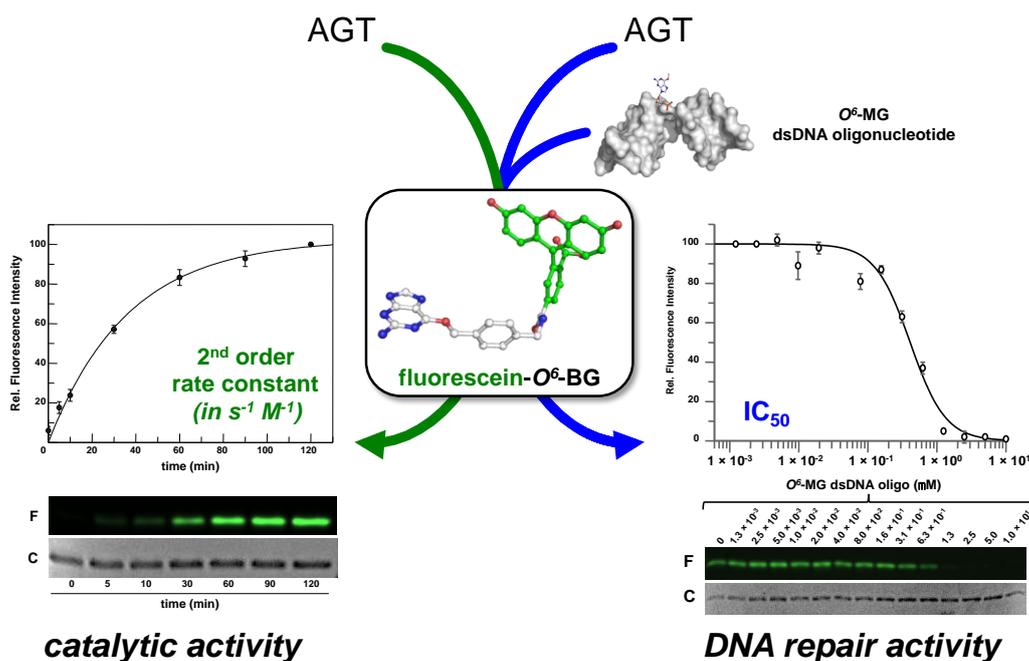
## 1.1 The SNAP-tag<sup>®</sup> technology

SNAP-tag<sup>®</sup> derives from hMGMT and was first introduced in 2003 by Prof. K. Johnsson (Mollwitz et al., 2012; Gronemeyer et al., 2006), and successively commercialized by New England Biolabs (NEB; <https://international.neb.com/tools-and-resources/feature-articles/snap-tag-technologies-novel-tools-to-study-protein-function>).

Due to the peculiar irreversible reaction mechanisms and to the low substrate specificity of the AGTs, this class of enzymes can be therefore exploited as biotechnological tool for the specific labelling of proteins (Pegg A.E. 2011). As above described, some AGTs can efficiently recognize free BG as substrate, as the case of hMGMT. In this way, any molecules conjugated to the 4-position of the benzyl ring can be irreversible transferred to the catalytic cysteine residue in the active site of the enzyme (see the Preface of this thesis and Figure 2.I for the description of the reaction mechanism).

Starting from this knowledge, Johnsson followed a directed-evolution approach to engineer the human protein and abolish any DNA binding activity, utilizing this variant as *protein-tag* for the *in vivo/in vitro* labelling of POIs (Keppler et al., 2003, Keppler et al., 2004; Juillerat et al., 2003; Gronemeyer et al., 2006; Mollwitz et al., 2012). The same strategy was also used by the same group to change the substrate specificity: from the SNAP-tag<sup>®</sup> they obtained the CLIP-tag<sup>®</sup> protein, which is active on *O*<sup>2</sup>-benzyl-cytosine derivatives, thus expanding the so-called *SNAP-tag technology* for multi-protein labelling (Gautier et al., 2008).

Despite the numerous assays extensively described in the literature (Perugino et al., 2012; Olsson et al., 1980; Wu et al., 1987; Klein and Oesch, 1992), the introduction of BG substrates allowed the measurement of the AGT activity in a safer and faster manner, not only *in vivo* but also *in vitro* in denaturing conditions, via SDS-PAGE gel-imaging analysis (Juillerat et al., 2003; Kindermann et al., 2003). In fact, if a fluorescent probe is conjugated to this molecule, it is irreversibly transferred to the AGT protein, labelling it in a 1:1 stoichiometry. As a consequence, the fluorescence intensity represents a direct measure of the enzyme activity and led to the determination of kinetic parameters for the reaction (See Figure 1.I, *left side*; Kindermann et al., 2003, Gronemeyer et al., 2006). According to this method, by adding a non-fluorescent alkylated DNA to the mixture containing the fluorescent substrate, it is possible to measure the protein affinity for its natural substrate and to determine the half-maximal inhibitory concentration (IC<sub>50</sub>), which can be converted to the dissociation constant for DNA (K<sub>DNA</sub>) and give an indirect measure of the methylation repair efficiency (Figure 1.I, *right side*) (Kindermann et al., 2003, Gronemeyer et al., 2006; Hinner and Johnsson, 2010). More recently, in collaboration with Prof. Ricci (University of Tor Vergata, Rome), the group of Dr Perugino has characterized a new substrate based on DNA-triplex, containing guanine nucleobases methylated in *O*<sup>6</sup>-position (*O*<sup>6</sup>-MeG), which undergo a conformational switch from a duplex to a triplex conformation upon enzymatic demethylation (Farag et al., 2021). The advantage of this new methodology is the possibility to measure in real-time the AGTs' activity directly on their natural substrate (methylated DNA).



**Figure 1.I** *Fluorescent assay for AGTs.* Fluorescein- $O^6$ -BG substrate could be used both for the determination of the AGT catalytic activity, and in combination with a competitive non-fluorescent substrate (in figure is represented a methylated double-strand DNA oligonucleotide) to indirectly measure the DNA repair activity (adapted from Mattosovich et al., 2020).

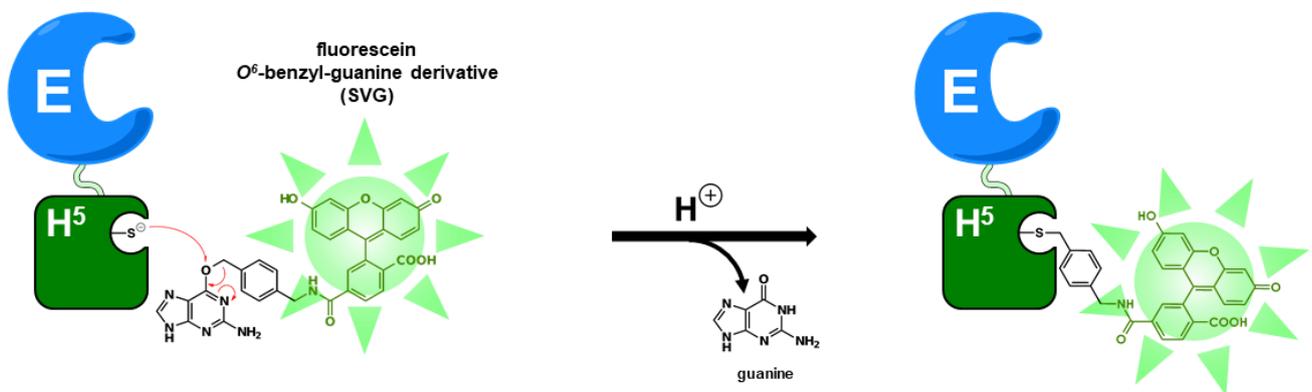
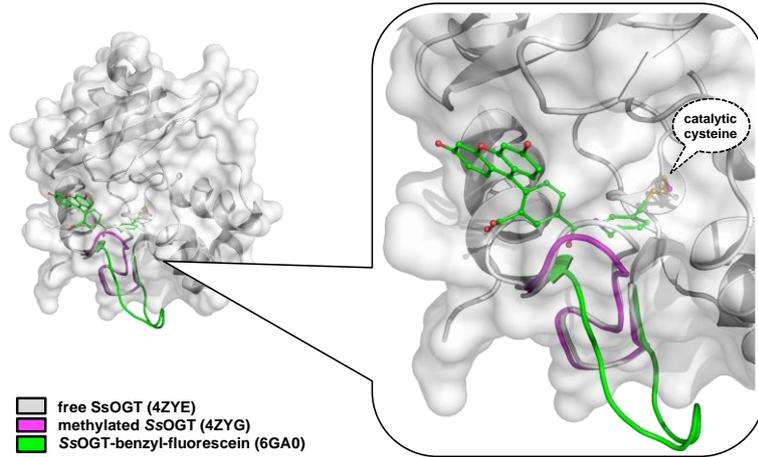
In addition to the hMGMT protein, this fluorescent assay can be applied to most prokaryotic orthologs, or in general to all those AGT sensitive to BG substrates. However, given the mesophilic nature and characteristics of the hMGMT, the *SNAP-tag technology* is strongly linked to mild reaction conditions and moderate temperatures. For this reason, following the same approach, Perugino and co-workers expanded this technology to extremophilic organisms and harsh reaction conditions. They developed a “thermo-*SNAP-tag*” by the production of a mutant from the *Saccharolobus solfataricus* OGT (*Ss*OGT- $H^5$ , hereinafter  $H^5$ ) (Perugino et al., 2012; Vettone et al., 2016).

## 1.2 *Ss*OGT- $H^5$

The variant  $H^5$  contains six mutations, five of which are in the helix-turn-helix domain abolishing the DNA binding activity. The sixth mutation is a substitution of a serine residue with a glutamic acid (S132E) in the so-called *active site loop*, which increases the catalytic activity of the protein as occurred in the molecular evolution of the *SNAP-tag*<sup>®</sup> (Juillerat et al., 2003).  $H^5$  was analysed in terms of thermal stability and compared to the *Ss*OGT wild-type, showing a moderate reduction of melting temperature ( $T_m$ ) at high temperature, but maintaining its resistance to denaturing and perturbing agents. But the most interesting data is that the protein activity is higher at low

temperatures (Perugino et al., 2012; Vettone et al., 2016), proposing this thermostable variant as a possible alternative to SNAP-tag<sup>®</sup> in biotechnological processes. For this purpose, H<sup>5</sup> mutant was first tested as *protein-tag* for two thermostable proteins derived from *S. solfataricus* and expressed in *E. coli*, without affecting their activity: a  $\beta$ -glycosidase (*Ss* $\beta$ gly) (Vettone et al., 2016) and the hyperthermophile-specific DNA topoisomerase reverse gyrase (Valenti et al., 2008; Perugino et al., 2009; Visone et al., 2017), both correctly folded. Successively, it was used as *protein-tag* in living thermophilic microorganisms, such as *Thermus thermophilus* HB27 and *Sulfolobus islandicus* E233S1, and as gene reporter in *in vitro* transcription/translation systems (Vettone et al., 2016; Visone et al., 2017; Lo Gullo et al., 2019). As a matter of fact, *Thermus thermophilus* is an *ogt*<sup>(-)</sup> species, showing only an alkyltransferase-like (ATL) activity (Morita et al., 2008). This class of proteins use a helix-turn-helix motif to bind the minor groove of the DNA, but they do not present repair activity, having the only task of recognising the lesion and recruiting proteins involved in the nucleotide excision repair system (NER) (Latypov et al., 2012; Lahiri et al., 2018). On the other hand, the *ogt* gene of *S. islandicus* was silenced by a CRISPR-based technique and tested as model organism (Visone et al., 2017). Through the fluorescent assay, it was shown that the H<sup>5</sup>-tag not only was correctly expressed and folded but maintained the activity *in vivo* on fluorescent substrates even at high temperatures.

Moreover, due to the great stability of this variant, it was possible to obtain for the first time the protein structure after the reaction with a substrate, revealing the initial destabilization of the *active site loop* (in green in Figure 2.I) after the alkylation of the catalytic cysteine (Rossi et al., 2018).



**Figure 2.I** SNAP-tag<sup>®</sup> technology with  $H^5$ . PDB structure of SsOGT- $H^5$  (PDB ID: 4ZYE) in complex with SNAP-Vista<sup>®</sup> Green (SVG) (PDB ID: 6GA0). In magenta, the conformational changes of the active site loop after methylation (PDB ID: 4ZYG) and in green after reaction with SVG. Below, a schematization of an enzyme of interest (in blue) genetically fused to  $H^5$  (in green), which can be labelled with any  $O^6$ -BG derivatives.

### 1.3 *Pyrococcus furiosus* AGT

The growing demand to apply a *protein-tag* to extreme conditions in thermophilic bacteria and archaea led to look for new protein from hot sources. Nowadays, there are no *protein-tags* for high temperatures (apart from His-tag). In order to study *in vivo* CRISPR-Cas immune systems in the hyperthermophilic archaeon *Pyrococcus furiosus* (*Pfu*) and expand the applicability of the SNAP-tag<sup>®</sup> near the point of boiling water, *P. furiosus* was chosen as reference microorganism.

Originally isolated from geothermal sediments in Vulcano Island (Italy) and described for the first time by Stetter and Fiala (Fiala and Stetter, 1986), *P. furiosus* is one of the best-studied extremophilic organisms (Kengen, 2017). With an optimum of growth temperature around 100 °C, it is one of the best sources of hyper-thermostable enzymes, most of them are currently used in molecular biology techniques (as an example the *Pfu*DNA polymerase I (Lundberg et al., 1991),

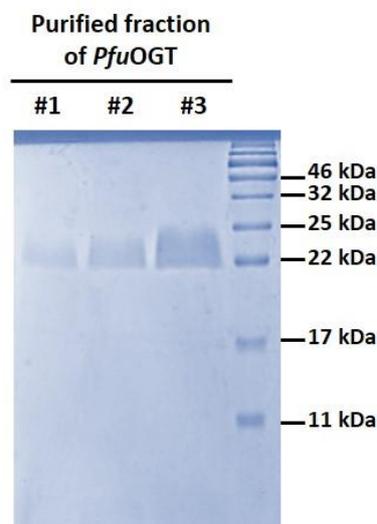
with an associated 3' – 5' exonuclease activity for proofreading). More recently, the discovery of a CRISPR-Cas system in *P. furiosus* offered new applications in biomedical and biotechnological fields (Hale et al., 2009; Terns and Terns, 2013). In 1998 was identified by Margison and co-workers a small protein that displays a potential AGT activity (Skorvaga et al., 1998). This was the starting point to investigate for a putative hyper-thermostable AGT to utilize as protein-tag *in vivo* at extreme temperature conditions.



## 2.2 Expression and purification

To verify and test the AGT activity, we produced and characterized the enzyme in recombinant form in *Escherichia coli* cells. For this purpose, we cloned the ORF PF1878 in the expression vector pHTP1 (NZYtech), obtaining the pHTP1/*Pfu*AGT plasmid. The positive clones were controlled by sequencing and the recombinant protein was successively purified by affinity chromatography and a further gel-filtration chromatography, using a Superdex 75 10/300 GL column, in order to improve the purity grade of the purification (as described in Experimental procedures).

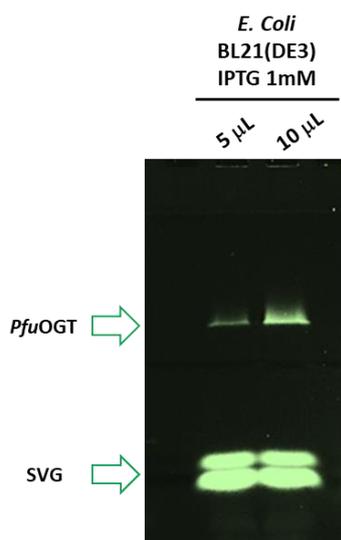
The protein was finally analysed by SDS-PAGE, in which a band of the expected molecular weight is present (Figure 4.I), stable in PBS 1×, pH 7.0.



**Figure 4.I** SDS-PAGE analysis of *Pfu*OGT. From the left, three fractions of purified *Pfu*OGT protein after desalting treatment. The non-canonical form of the bands is probably due to partially non-denatured protein population.

### 2.3 Biochemical characterization: activity assay, catalytic activity (progress curve) and thermal stability

The purified protein, named *Pfu*OGT, was tested on SVG commercial fluorescent substrate (SNAP-Vista® Green) (shown in Figure 3, in Preface), a BG conjugated with a fluorescein in  $O^6$  position. In particular, *Pfu*OGT was treated firstly at 50 °C up to 1h and showed a fluorescent band in gel-imaging analysis (Figure 5.I).



**Figure 5.I** *Pfu*OGT treated with SNAP-Vista® Green substrate. In the gel-imaging analysis it is possible to appreciate a fluorescent band corresponding to the protein of *Pfu*OGT, after treatment with SVG substrate performed at 50°C.

This demonstrated that *Pfu*OGT is active on BG derivative substrates, as other AGTs previously described in the literature. Furthermore, to verify its repair activity, an inhibition assay with methylated dsDNA (oligo sequences are listed in Table 1.IV in Materials and Methods section) was set up. However, due to its high thermophilicity, the repair activity of the *P. furiosus* protein was measurable only at 65 °C, instead of the standard procedures previously described (50 °C) (Mattosovich et al., 2020, Perugino et al., 2015, Morrone et al., 2017). Reaction mixtures were loaded on SDS-PAGE and competition between the SVG fluorescent substrate and the “dark” dsDNA were analysed *via* gel-imaging techniques.

We demonstrated the repair activity of *Pfu*OGT by the decrease of the fluorescent signal at crescent concentration of ds-oligo, as the methyl group initially present on the dsDNA has been transferred to the active site of the protein through an  $SN_2$ -like *one-step* reaction mechanism. Since the catalytic cysteine involved in the reaction has already covalently bound the methyl group, the active site is no longer available for the reaction with SVG (Table 1.I).

**Table 1.I.** Competitive inhibition assay ( $IC_{50}$ ) in the presence of double strands methylated oligonucleotides. *Ss*OGT and *Pfu*OGT proteins (5  $\mu$ M) were incubated with 5  $\mu$ M of fluorescent SVG substrate together with increasing concentrations (0–10  $\mu$ M) of methylated oligonucleotides. Here are reported the  $IC_{50}$  values obtained for both proteins.

	$IC_{50}$ (mM)	Notes
<b><i>Ss</i>OGT</b>	$1.1 \pm 0.08$	Perugino et al., 2015
<b><i>Pfu</i>OGT</b>	$0.88 \pm 0.10$	this work

The catalytic activity of *Pfu*OGT as a function of temperature was analysed by the fluorescent assay at various temperatures and timing conditions. Briefly, by using SVG s substrate, the *Pfu*OGT and *Ss*OGT (as control) activities were tested at 25, 50, 70, 80 and 90 °C. As expected, the former displayed a strong thermophilicity, with a measurable catalytic activity only at very high temperatures, as illustrated in the Table 2.I.

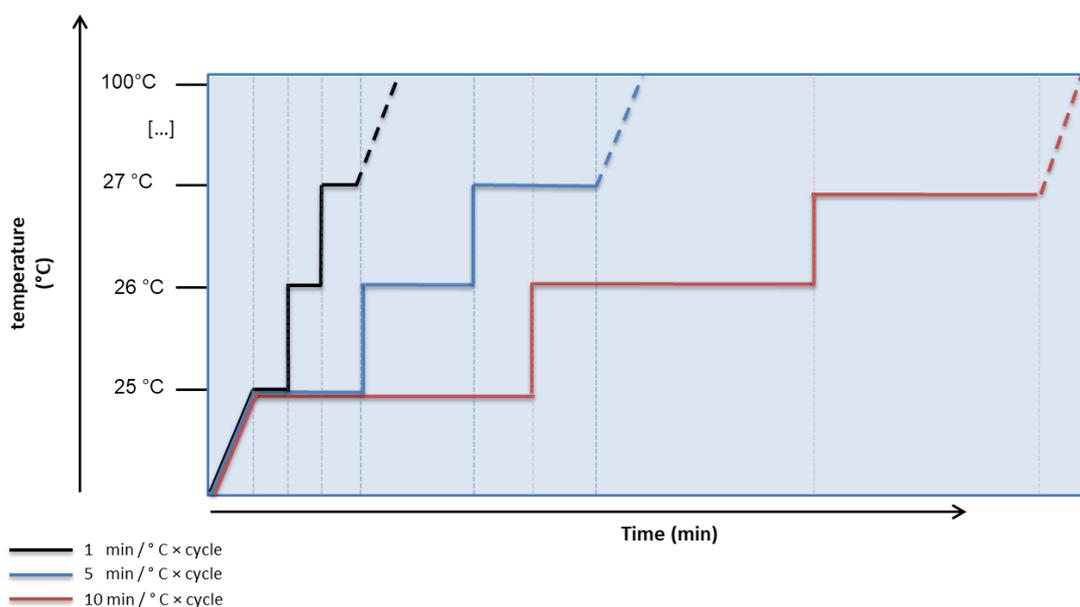
**Table 2.I** Catalytic activities as a function of temperature of thermostable AGTs, expressed as second-order rate constant values

Protein	Temperature (°C)	K ( $s^{-1} M^{-1}$ )	Notes
<i>Pfu</i> OGT	25	ND	this work
	50	$1.80 \times 10^3$	this work
	70	$2.83 \times 10^3$	this work
	80	$1.20 \times 10^4$	this work
	90	$1.50 \times 10^5$	this work
<i>Ss</i> OGT	25	$2.80 \times 10^3$	Perugino et al. (2012)
	50	$1.50 \times 10^3$	this work
	70	$5.33 \times 10^4$	Perugino et al. (2012)
	80	ND	this work

We investigated the thermostability of both *Ss*OGT wild-type and *Pfu*OGT proteins, tested in the same conditions by the (Table 3.I) by Differential Scanning Fluorimetry (DSF) or “thermofluor”. This technique is based on the measurement of the unfolding and aggregation of proteins in various

conditions, through an increase of fluorescence of specialized dyes, such as SYPRO Orange (Niesen et al. 2007). Protein samples are tested in a 96-well plate and subjected to a melting-curve protocol using a real-time thermal cycler. Due to test conditions, proteins unfold and the fluorogenic dye binds non-specifically their hydrophobic aminoacidic residues exposed to the solvent because of the denaturation, resulting in an increase in fluorescence.

By DSF analysis in PBS 1X pH 7.0 condition the melting temperature ( $T_m$ ) of *Pfu*OGT was 80 °C, while the *Ss*OGT  $T_m$  value was 68 °C. Due to its thermal characteristics and in order to obtain a sigmoidal melting curve for *Pfu*OGT, the heating rate of 10 min/°C × cycle was set up, whereas the  $T_m$  value measurements were usually performed at 1 min/°C × cycle (Niesen et al., 2007) (Figure 6.I).



**Figure 6.I** Schematization of Differential Scanning Fluorimetry programs. Different coloured lines show the incubation time per grade of the three protocols described in the main text. Respectively, in *black*, the protocol used by Niesen et al., 2007 with 1 min/ °C per cycle; in *blue*, the protocol used by Perugino et al., 2012 with 5 min of incubation/ °C per cycle, and in *red*, the protocols applied by Mattosovich et al., 2020, in which an incubation time of 10 min/ °C per cycle was needed to determine the  $T_m$  of *Pfu*OGT protein.

In addition, *Pfu*OGT and *Ss*OGT were tested in high ionic strength conditions and in presence of detergents. What we have observed is that the  $T_m$  value of *Pfu*OGT was negatively affected by the high salt concentration, whereas *Ss*OGT has shown an increase in its  $T_m$ . On the contrary, the  $T_m$  of *Pfu*OGT decreased only of 10 °C in the presence of denaturing agent (in this case SDS), while *Ss*OGT  $T_m$  dropped by 22 °C (Mattosovich et al., 2020).

**Table 3.I** Protein thermal stability by DSF method.

	Conditions	$T_m$ (°C)	Rate (min/°C × cycle)	Notes
<i>Ss</i> OGT	PBS 1 ×	80.0 ± 0.4	5	From Perugini et al. (2015)
	PBS 1 ×	67.9 ± 1.1	10	This study
	PBS 1 ×; NaCl 1.0 M	79.6 ± 0.3	10	This study
	PBS 1 ×; NaCl 4.0 M	82.9 ± 0.4	10	This study
<i>Pfu</i> OGT	PBS 1 ×; SDS 0.01%	46.5 ± 1.7	10	This study
	PBS 1 ×	78.8 ± 0.4	10	This study
	PBS 1 ×; NaCl 1.0 M	83.7 ± 0.3	10	This study
	PBS 1 ×; NaCl 4.0 M	50.4 ± 2.1	10	This study
	PBS 1 ×; SDS 0.01%	71.8 ± 0.3	10	This study

$T_m$  values were obtained by plotting the relative fluorescence intensity as a function of temperature. Data were achieved from three independent experiments

### 3. Conclusions and perspectives

The results described in the previous paragraph indicate that the ORF PF1878 of *P. furiosus* JFW02 strain actually encodes an AGT protein, which is sensitive to the BG substrates and presents a DNA repair activity on alkylating agents. Furthermore, its thermostability was confirmed by DSF analysis, showing also different stabilization strategies that occur in its structure, if compared to the *S. solfataricus* homologue, and underlining its potential role as hyper-thermostable *protein-tag* in living (hyper)thermophilic organisms.

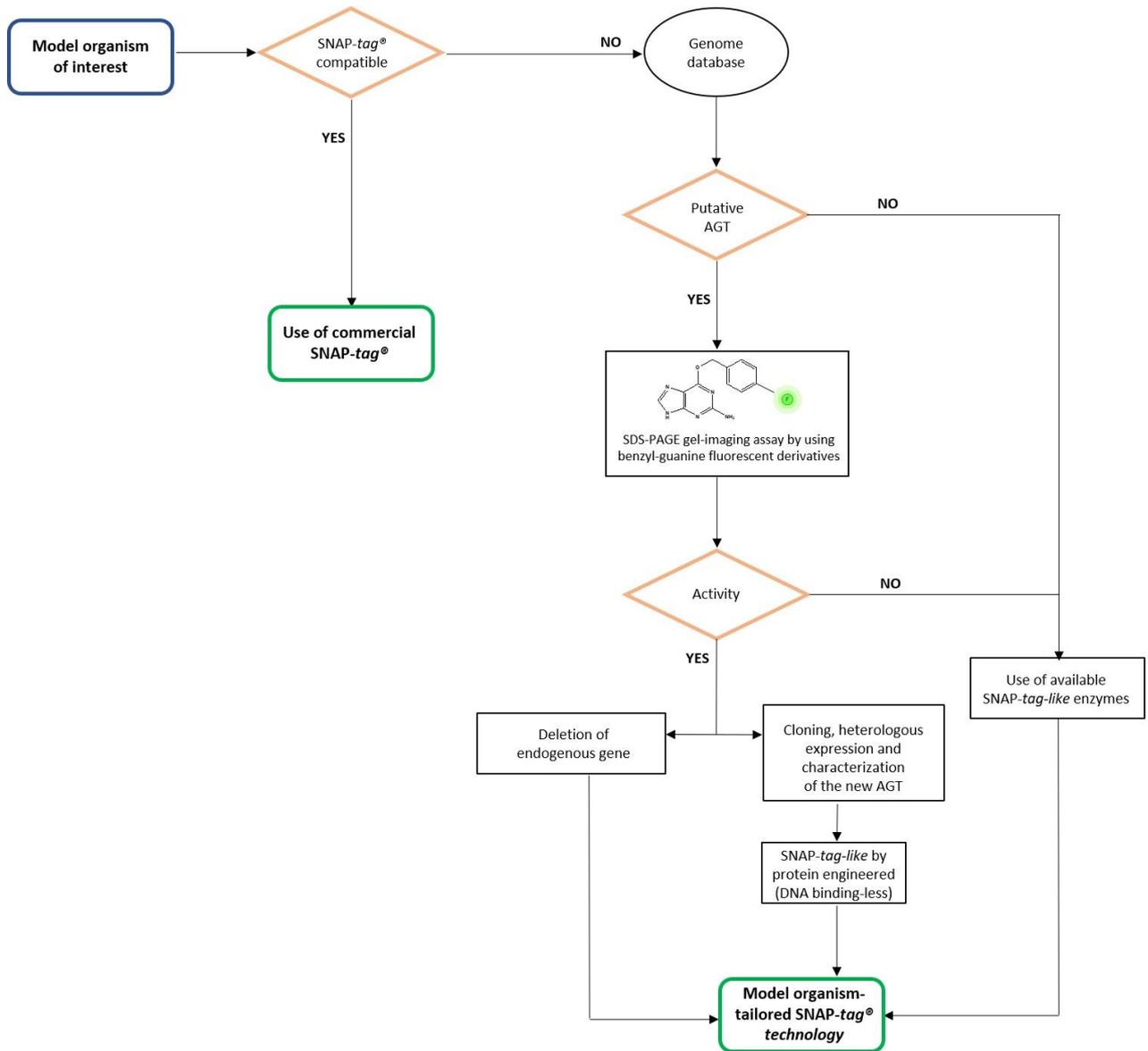
These results also opened remarkable avenues for the further development of a “hyper-thermoSNAP-*tag*”: from this wild-type version, the aim has been to proceed with an *in silico* approach to identify the residues to be engineered, as done before with *S. solfataricus* H<sup>5</sup>.

In the last part of my PhD, I compared sequences and 3D structures of *Ss*OGT wild-type, H<sup>5</sup> mutant, *Tkd*-MGMT and *Pfu*OGT proteins, identifying several aminoacidic residues involved in the DNA binding and in the catalytic activity. Successively, I expressed and purified the *Pfu*OGT mutant and, from preliminary data, I may affirm that this variant maintained the stability and activity of the wild-type. Moreover, from experiments with methylated ds-oligonucleotides, I

demonstrated the inability to bind the DNA. Therefore, the identified and selected mutations actually allowed the generation of a non DNA-binding version of the *Pfu*OGT protein.

The further steps will be the construction of an *ogt*-KO *P. furiosus* strain by the group of Prof. Terns, the *in vivo* validation of the *Pfu*OGT DNA binding less mutant and the consequent expression and purification the corresponding protein. Finally, a biochemical characterization in terms of activity and thermal stability will be necessary, as well as complementation experiments of the *Pfu*OGT mutant in the *ogt*-KO *P. furiosus* strain, to actually determine if the protein could be considered the first hyperthermophilic *protein-tag* hereto discovered and characterized.

The knowledge acquired in recent years on AGTs and their use in modern biotechnologies allows the development of a “workflow” aim to optimize this *protein-tag* for different needs (Figure 7.I). In the case of model organisms thriving in extreme conditions of temperature, salinity, pressure, etc., in which the commercial SNAP-*tag*<sup>®</sup> is not particularly suitable, it would be possible to imagine an identification of a relevant AGT and, through mutagenesis, obtain a “SNAP-*tag*-like” tool. This strategy has to necessarily take into account the fact that this identified AGT should be sensitive to *O*<sup>6</sup>-BG derivatives: some members of this family, such as *E. coli* AdaC, are resistant to the action of *O*<sup>6</sup>-BG (Elder et al., 1994; Goodtzova et al., 1997). In parallel, it is possible to perform an *ogt* gene deletion in the model organism, in order to eliminate any endogenous activity. This workflow was successfully realised in thermophilic microorganisms, such as the archaeon *S. islandicus* (Visone et al., 2017) and in the bacterium *T. thermophilus*. In the latter case, it was not necessary to eliminate the *ogt* gene, because this organism lacks the gene (Vettone et al., 2016).



**Figure 7.I** The SNAP-tag® workflow.

## **Chapter II: Evolution of the SNAP-*tag* technology**

In this chapter, the introduction and characterization of novel *chemo-enzymatic approach* for the SNAP-tag<sup>®</sup> technology is described. This project is in collaboration with the group of Prof. Alberto Minassi (University of Piemonte Orientale) and with Prof. Antonio Leonardi (University of Naples “Federico II”).

These data in have been published on the *Journal of Enzyme Inhibition and Medicinal Chemistry* ([https:// doi: 10.1080/14756366.2020.1841182.](https://doi.org/10.1080/14756366.2020.1841182))

JOURNAL OF ENZYME INHIBITION AND MEDICINAL CHEMISTRY  
2021, VOL. 36, NO. 1, 85–97  
<https://doi.org/10.1080/14756366.2020.1841182>



SHORT COMMUNICATION

OPEN ACCESS Check for updates

## The SNAP-tag technology revised: an effective *chemo-enzymatic approach* by using a universal azide-based substrate

Rosa Merlo<sup>a\*</sup>, Diego Caprioglio<sup>b\*</sup>, Michele Cillo<sup>c\*</sup>, Anna Valenti<sup>a</sup>, Rosanna Mattosovich<sup>a</sup>, Castrese Morrone<sup>b</sup>, Alberto Massarotti<sup>b,d</sup>, Franca Rossi<sup>b</sup>, Riccardo Miggiano<sup>b,d</sup>, Antonio Leonardi<sup>c</sup>, Alberto Minassi<sup>b</sup> and Giuseppe Perugino<sup>a</sup>

<sup>a</sup>Institute of Biosciences and BioResources, National Research Council of Italy, Naples, Italy; <sup>b</sup>Department of Pharmaceutical Sciences, University of Piemonte Orientale, Novara, Italy; <sup>c</sup>Department of Molecular Medicine and Medical Biotechnology, University of Naples “Federico II”, Naples, Italy; <sup>d</sup>IXTAL srl, Novara, Italy

\* These authors equally contributed to the present work.

# 1. Introduction

Although many BG-substrates for the SNAP-tag technology are commercially available, the possibility of conjugating infinite molecules to the 4-position on BGs leads to the synthesis of *ad hoc* substrates. This is generally possible through the reaction of “BG-building block” (e.g. the amine-reactive BG-NH<sub>2</sub>) with NHS-ester derivative compounds. However, this step is plagued with the main disadvantage to purify the final compounds to avoid the risk of contamination between chemical species in the mixture, increasing the times and costs of the experiments (Figure 1.II A). Furthermore, the presence itself of chemical groups conjugated to BG could affect the reaction efficiency (Figure 1.IIA) (Moschel et al., 1992; Terashima et al., 1997; Sun et al., 2016).

These limitations are overcome with a new *chemo-enzymatic approach* (combining an enzymatic reaction to a chemical one, each possessing a very high specificity), which proposes only one and universal BG-substrate associable to a unique enzyme reaction rate before proceeding to a highly specific click-chemistry reaction (Figure 1.II B).

## 1.1 The Click Chemistry

The term “Click Chemistry” was introduced by Kolb and co-workers in developing a chemical mechanism aimed at the rapid synthesis of useful new compounds. It describes the selective reaction that occurs between a pair of functional groups, such as terminal azide-alkyne (also known as azide-alkyne Huisgen cycloaddition reaction), in aqueous conditions (Kolb et al., 2001). The reaction they thought about is modular, wide in scope and generates high yields of products, moreover, it should have easily available reagents and at the end of the reaction the product should be simply isolated from the rest of the mixture (Kolb et al., 2001). To date, the concept of Click Chemistry is related to specific coupling procedures between two molecules (Sletten et al., 2009; Jewett et al., 2010), and it is used in different fields of biosciences (Xie et al., 2013; Su et al., 2013; Zeng et al., 2013). One of the most popular applications of Click Chemistry, apart from chemical synthesis, is the labelling of biomolecules, both *in vivo* and *in vitro*, for biomedical studies (Kim E. and Koo H., 2019). This type of reaction is well suited for this purpose, as it requires very mild reaction conditions, which can occur in the physiological environment of the biomolecule to be studied.

Among the plethora of possible reactions, the best known and used are:

- 1) Copper or Cu(I)- catalysed Azide-Alkyne reaction (CuAAC)
- 2) Strain-promoted Azide-Alkyne reaction (SPAAC)

### Cu(I)-catalyzed Azide-Alkyne reaction (CuAAC)

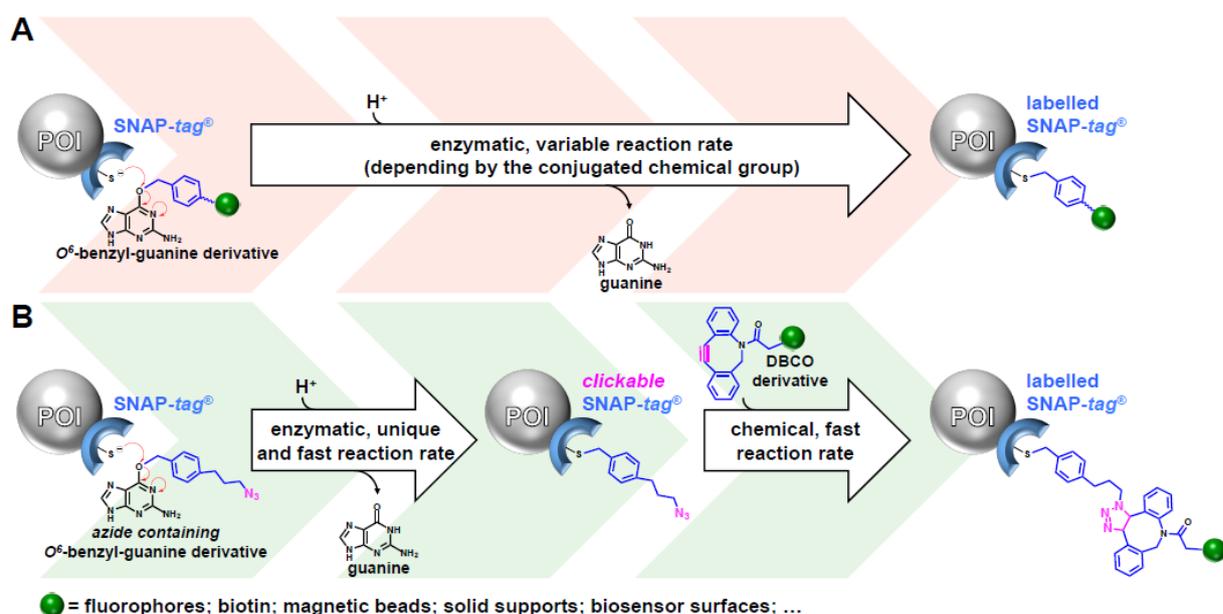
The reaction occurs between Azide- and Alkyne-functionalized molecules, forming a stable conjugate *via* a Triazole moiety. However, the low reactivity of Alkynes towards Azides makes the presence of a metal catalyst (such as copper) necessary (Presolski et al., 2011; Hong et al., 2009). Moreover, using CuAAC in live cells presents a big problem, because of toxicity related to copper or other metals. To overcome this limitation, the technique is evolved to SPAAC protocol (Jewett et al., 2010).

### Strain-promoted Azide-Alkyne reaction (SPAAC)

These reactions rely on the use of strained cyclo-octynes possessing lower activation energy and a faster second-order reaction rate, in contrast to terminal Alkynes under aqueous conditions without a catalyst (Debets et al., 2010; Debets et al., 2011; Ess et al., 2008). Various cyclo-octyne derivatives have been developed, but most of SPAAC reactions are actually based on Di-benzocyclo-octyne (DBCO), which combines high reactivity with hydrophilicity (Debets et al., 2010; Kuzmin et al., 2010). As the precedent class of click-chemistry reaction, Azide-DBCO is highly selective and suited for multiple labelling approaches, and because it is copper-free, it also suitable for *in vivo* cell imaging analysis (Yao et al., 2012) and drug target studies (Kim and Koo 2019).

## 1.2 The novel approach

Starting from this knowledge, part of my PhD research activity focused on the implementation of SNAP-tag<sup>®</sup> technology, by introducing a novel *chemo-enzymatic approach* that employs a universal azide BG-derivative substrate, that can be easily coupled with a potentially infinite number of commercially available DBCO-based molecules, through the Huisgen SPAAC reactions (Merlo et al., 2021).



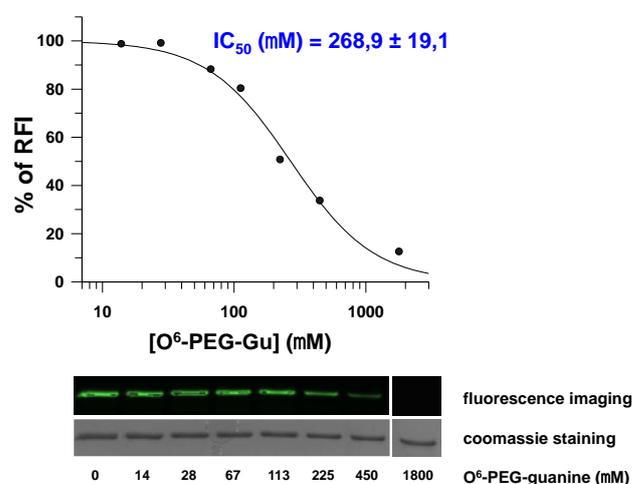
**Figure 1.II** Single-step reaction vs chemo-enzymatic approach. (A) The SNAP-tag<sup>®</sup> technology is based on BG-derivatives to synthesize and purify singularly, although it does not avoid that the conjugated chemical group (*green* sphere) could affect the enzymatic reaction rate. (B) The SNAP-technology revised uses a unique and universal azide BG-derivative (azide group in *magenta*), converting SNAP-tag<sup>®</sup> in a “clickable” form, ready to a fast and efficient SPAAC with DBCO-based chemical groups. POI, protein of interest genetically fused to the SNAP-tag<sup>®</sup>.

## 2. Results and Discussion

### 2.1 Substrate specificity of AGTs on BG-based substrates

To investigate the substrate specificity of the commercial SNAP-tag<sup>®</sup> and H<sup>5</sup> proteins, we first evaluated their activity on several *O*<sup>6</sup>-guanine-derivatives (Table 1.II). Because different substrates tested were non-fluorescent compounds, we performed a AGTs’ competitive inhibition assay by using the fluorescein-derivative SNAP-Vista<sup>®</sup> Green as substrate (SVG) (Perugino et al., 2012; Perugino et al., 2015; Vettone et al., 2016; Visone et al., 2017). As previously described, the reaction of an AGT with SVG led to the fluorescent labelling of the protein, which can be visualised as a fluorescent band in gel-imaging analysis. The increasing of amounts of the non-fluorescent competitor in the reaction leads to a decrease in the fluorescent signal, which can be measured and plotted for the IC<sub>50</sub> values determination (Perugino et al., 2015; Morrone et al., 2017). As shown in Table 1.II, the two enzymes presented different behaviours versus these competitors, irrespective of dimension and/or polarity of the conjugated chemical groups (Merlo et al., 2021). For example, SNAP-Cell<sup>®</sup> 430 (BG430) completely lost the competition with SVG, whereas both enzymes are extremely active on the SNAP Cell<sup>®</sup> Block (SCB), displaying the lowest IC<sub>50</sub> value measured. This

result can be explained because SCB has very similar structure to the Lomeguatrib, which is one of the most efficient inhibitors of the hMGMT protein and employed in the cancer treatment in combination with alkylating agents-based chemotherapeutics (Ranson et al., 2006). In general, all commercially used compounds (SVG, SCB, BG430, and BG-PEG-NH<sub>2</sub>, BGPA) are good substrates for SNAP-tag<sup>®</sup> and H<sup>5</sup> enzymes. However, some risks of lower reaction rate may occur in using customized substrates. This was the case of methyl-guanine-PEG-NH<sub>2</sub> (MGPA), an O<sup>6</sup>-methylguanine derivative used for the immobilisation of SNAP-tag<sup>®</sup> on nanoparticles (Colombo et al., 2012). This has not proved to be a good substrate, in fact, SNAP-tag<sup>®</sup> and H<sup>5</sup> completed the labelling only after an over-night incubation at 4 °C (Colombo et al., 2012) and 65 °C, respectively.

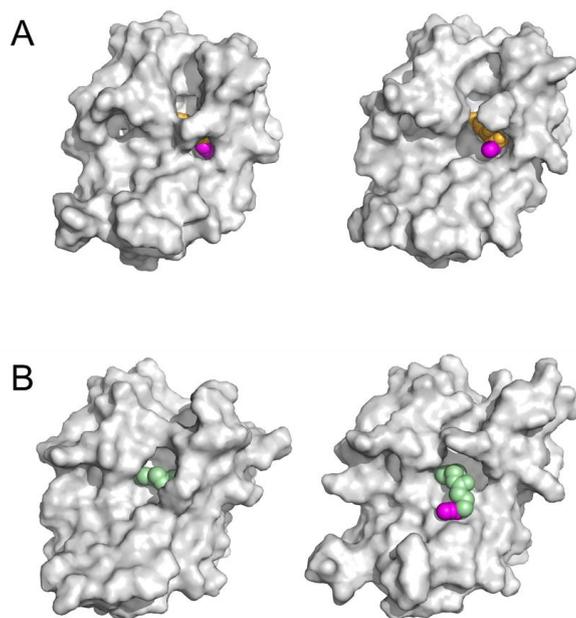


**Figure 2.II** IC<sub>50</sub> experiment on O<sup>6</sup>-PEG-Gu substrate.

## 2.2 Molecular modelling on the H<sup>5</sup> with BG-azides

Recent studies were focused on the synthesis of alternative “BG-building blocks” to produce SNAP substrates in an easier and faster way. Some of them are the alkyne substituted O<sup>6</sup>-BG, which was employed with azide-based fluorescent probes for the synthesis of compounds (Song et al., 2015) and the O<sup>6</sup>-BG-N<sub>3</sub> (BGN<sub>3</sub>) (Figure 4.II panel A, on top) for the conjugation with alkyne-based chemical groups (Zhang et al., 2011). Assuming that proteins are not static and the amino acids side-chain movements could prevent the click-chemistry reaction, BGN<sub>3</sub> and a new azide-based compound (BGSN<sub>3</sub>), that differs for the length of the chemical spacer, were analysed in a covalent complex with H<sup>5</sup> protein by the group of Prof. Alberto Massarotti (University of Piemonte Orientale) with Molecular Dynamics (MD) simulations using the Desmond package (see Materials and Methods), in order to determine which substrate was more prone to the labelling. The *S. solfataricus* protein was chosen as model because of its properties, which benefit crystallization, and the structures of wild-type and H<sup>5</sup> already registered in PDB databank (PDB IDs: 4ZYE and

6GA0, respectively). The complexes were simulated for 100 ns at 300 K using a standard protocol. The MD results were analysed in terms of Solvent Accessible Surface Area (SASA) of the compounds: more time the compounds are exposed to the solvent, the higher is the possibility to react (Ferraris et al., 2014). In Figure 3.II is reported the simulation models of the H<sup>5</sup> protein in complex with BGN3 (panel A) and BGSN3 (panel B). As expected, the BGN3 is less exposed to the solvent with a SASA value of  $32.967 \pm 18.573 \text{ \AA}^2$ , while BGSN3 shows a higher SASA value  $68.302 \pm 32.455 \text{ \AA}^2$ .



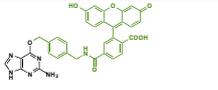
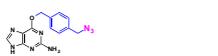
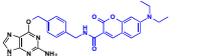
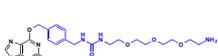
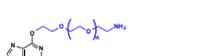
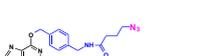
**Figure 3.II** Model simulation of H<sup>5</sup>-probe complexes. A) Structural docking from two different slants of H<sup>5</sup> protein with BGN3 substrate (orange). B) Structural docking of H<sup>5</sup> protein with BGSN3 substrate (green). In magenta, the azide group.

Thus, the simulation confirmed our biochemical data, proposing the BGSN3 as a better substrate for our chemo-enzymatic approach (Merlo et al., 2021).

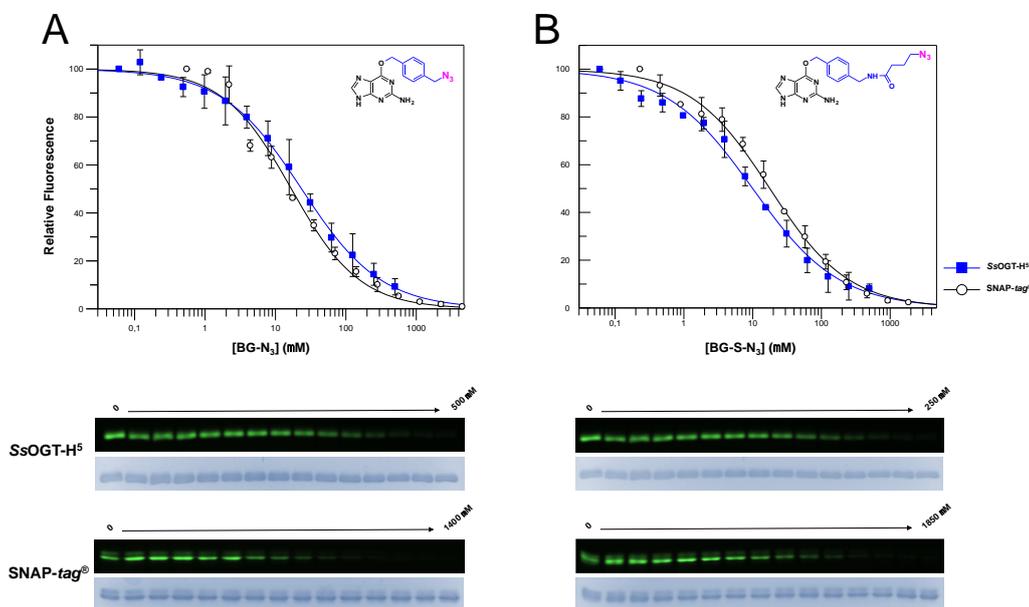
### 2.3 In vitro reaction of engineered AGTs with BG-azide substrates

To analyse the enzymatic reaction of SNAP-tag<sup>®</sup> and H<sup>5</sup>, we tested the proteins directly on BGN3 and a new synthesized BG-derivative containing a spacer between the benzyl-ring and the azide group (BGSN3) (Figure 4.II panel B, on top), and the resulting values were plotted for the determination of IC<sub>50</sub> values. As can be seen from the IC<sub>50</sub> values in Table 1.II and visually from Figure 4.II, the two substrates have an affinity for SNAP-tag<sup>®</sup> and H<sup>5</sup> proteins comparable to those commercially available.

**Table 1.II.** Substrate specificity of SNAP-tag<sup>®</sup> and H<sup>5</sup> on different BG-derivative substrates. Here are listed IC<sub>50</sub> values, obtained by competitive inhibition assay (IC<sub>50</sub>), and second-order rate constants of the enzymatic reaction of SNAP-tag<sup>®</sup> and H<sup>5</sup> proteins, only on SVG and BGSN3 substrates.

structure	name	SNAP-tag <sup>®</sup>		SsOGT-H <sup>5</sup>	
		IC <sub>50</sub> (mM)	k <sup>a</sup> (s <sup>-1</sup> M <sup>-1</sup> )	IC <sub>50</sub> (mM)	k (s <sup>-1</sup> M <sup>-1</sup> )
	SVG	-	2.8 × 10 <sup>4</sup> <sup>b</sup>	-	1.6 × 10 <sup>4</sup>
	BG	36.8 ± 5.6	-	10.1 ± 1.0	-
	SCB	2.1 ± 0.5	-	4.4 ± 0.8	-
	BGN3	15.6 ± 0.3	-	23.5 ± 1.0	-
	BG430	ND <sup>c</sup>	-	ND	-
	BGPA	86.0 ± 6.7	-	14.3 ± 1.9	-
	MGPA <sup>d</sup>	-	-	268.9 ± 19.1 <sup>e</sup>	-
	BGSN3	17.8 ± 1.1	4.64 ± 1.04 × 10 <sup>5</sup>	10.0 ± 0.7	1.40 ± 0.47 × 10 <sup>4</sup>

<sup>a</sup>reaction rates at 25 °C; <sup>b</sup>this value was obtained by using a BG-fluorescein substrate (BG-FL) very similar to SVG; <sup>c</sup>Not Determined; <sup>d</sup>this molecule is a O<sup>6</sup>-methyl-guanine derivative; <sup>e</sup>competitive assay for H<sup>5</sup> was performed at 65 °C.

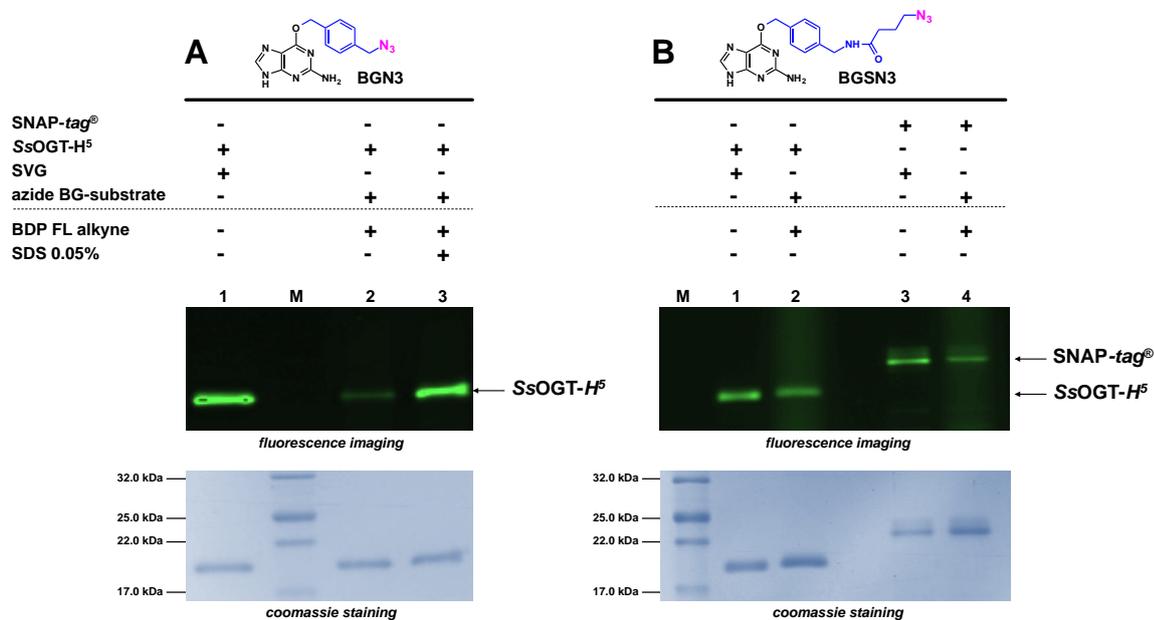


**Figure 4.II** IC<sub>50</sub> plots of the competitive fluorescent inhibition assay. SNAP-tag<sup>®</sup> and H<sup>5</sup> were incubated with BGN3 (A) and BGSN3 (B), using SVG as competitive substrate (see IC<sub>50</sub> values on Table 1.II). Values obtained from three independent experiments.

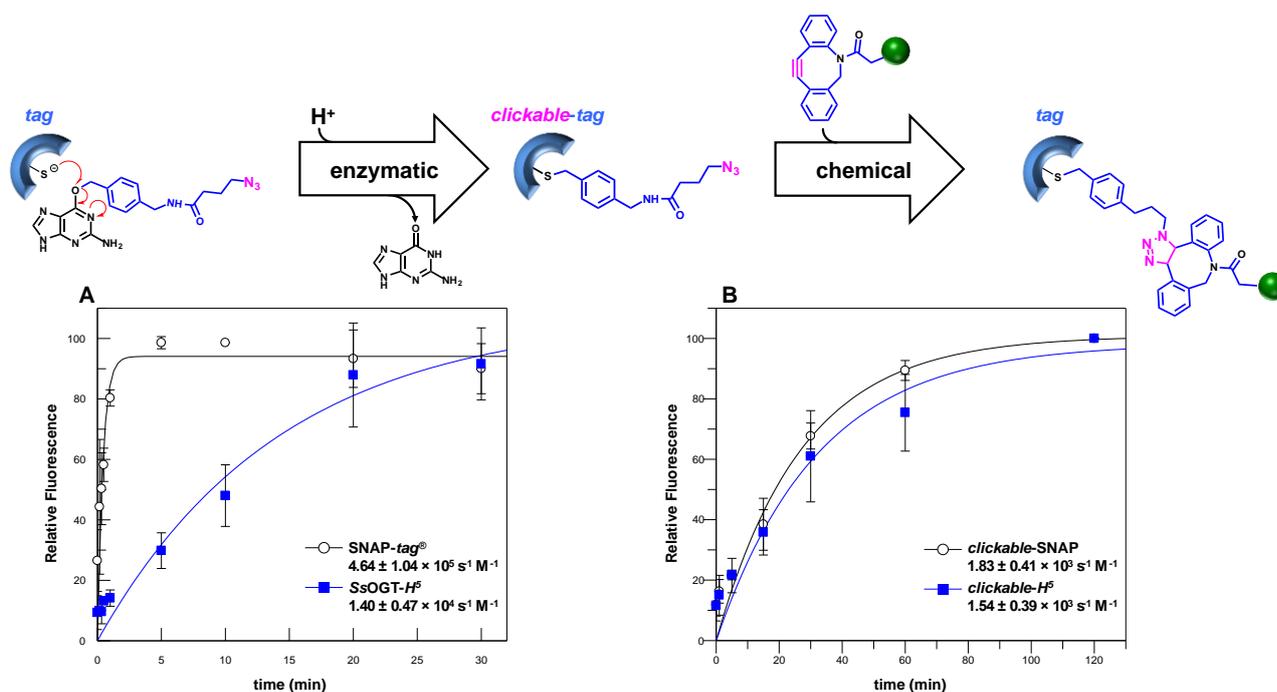
Upon the reaction with both the azide-substrates, an equimolar amount of SVG was added to the mixtures, followed by SDS-PAGE gel-imaging, where no fluorescent signal was obtained, thus confirming that the catalytic cysteine was irreversibly blocked by the benzyl-azide moiety.

After, to assay the click-chemistry, H<sup>5</sup> was tested with BGN3 and BGSN3 substrates and then the cycloaddition reaction using an alkyne-derivative of the fluorescein (BDP FL alkyne) was performed (Figure 5.II). However, the use of BGN3 led to a less efficient chemical reaction (Figure 5.II panel A, lane 2), in which the complete labelling of the protein was achieved only in the presence of SDS during the cycloaddition step (lane 3). This data suggested that the protein is still folded after the enzymatic reaction and the azide is hidden in the active site core, but becomes accessible only after the addition of denaturant, favouring a better exposure of the azide group to the solvent. On the contrary, using BGSN3 as substrate, the labelling of both the enzymes was comparable to the classical reaction with SVG without any denaturing agent (Figure 5.II panel B). In fact, the spacer present in BGSN3 sufficiently exposes the azide group for the CuAAC reaction (Figure 5.II panel B, lanes 2 and 4), confirming the MD simulations.

For this reason, we decided to perform further experiments by using the longer BG-azide (BGSN3) both for SNAP-tag<sup>®</sup> and H<sup>5</sup>, starting by the determination of the reaction rates, in order to compare the activity of commercial BG substrates and BGSN3 (Figure 6.II and Table 1.II).



**Figure 5.II** Gel-imaging SDS-PAGE of the chemo-enzymatic reaction with BG-azides. Proteins were first incubated in the presence of BGN3 (A) or BGSN3 (B). After the enzymatic reaction, the fluorescent BDP FL alkyne and all components for the click chemistry were added to the mixture. As control, each protein was incubated only with SVG.



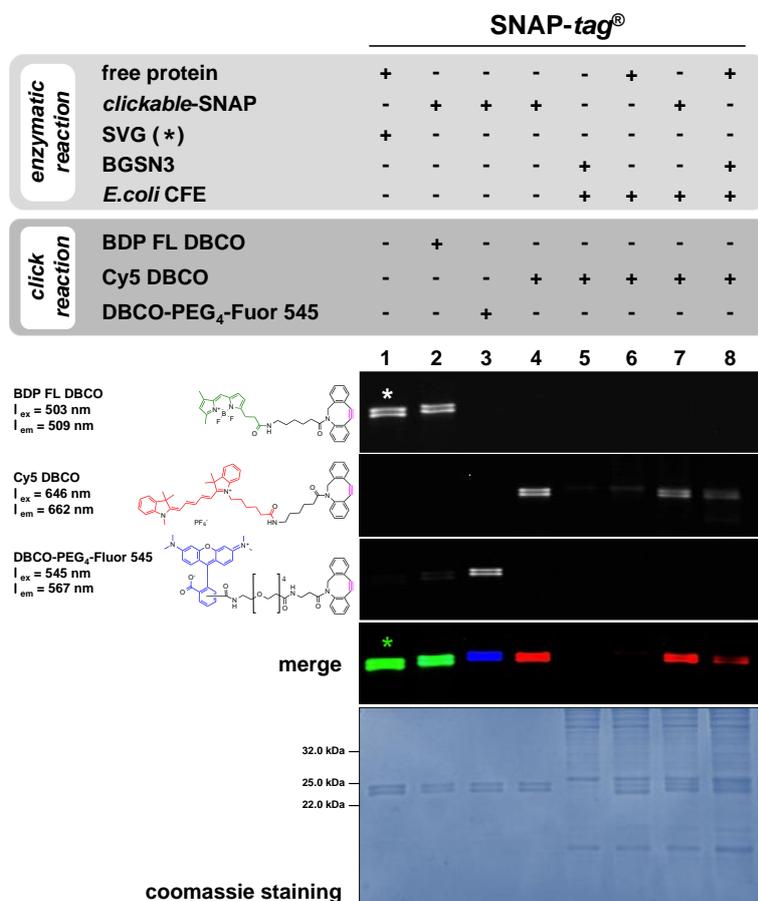
**Figure 6.II** Reaction rates of the chemo-enzymatic approach. Pseudo-first-order reaction of *protein-tags* for (A) the enzymatic reaction with BGSN3 (see *k* values also in Table 1.II), and of clickable-tags for (B) Huisgen reaction with DBCO-PEG4-Fluor 545. Values given are an average of three independent measurements and data are represented as mean  $\pm$ SEM.

The obtained results demonstrated that not only the novel substrate shows an activity comparable to the other commercially available substrates but also indicates that the complete protein labelling can be performed in less than an hour (Mollwitz et al., 2012; Gautier et al., 2008; Vettone et al., 2016).

## 2.4 Specificity and versatility of the chemo-enzymatic reaction

After establishing that BGSN3 was a good substrate for both *protein-tags*, the labelling efficiency of the *clickable-SNAP* and H<sup>5</sup> by using different DBCO-based fluorophores and the specificity of the “click” reaction were examined. We first quantitatively evaluated the rate (*k*) of the click reaction by using the DBCO-PEG4-Fluor 545 fluorophore: as expected, both the clickable-*tags* were labelled with the same efficiency ( $1.83 \pm 0.41 \times 10^3 \text{ s}^{-1} \text{ M}^{-1}$  for SNAP-tag<sup>®</sup>;  $1.54 \pm 0.39 \times 10^3 \text{ s}^{-1} \text{ M}^{-1}$  for H<sup>5</sup>), demonstrating that the chemical reaction is sufficiently fast and independent from the *tags* (Figure 6.II, panel B). Successively, upon the reaction with BGSN3, three cycloaddition reactions were performed with three different DBCO-based fluorophores in PBS 1 $\times$  buffer. All

click reactions were complete in ca. 30–45 min (Figure 7.II, lanes 2–4), with a protein labelling as efficient as the enzymatic reaction using the sole SVG (lane 1).

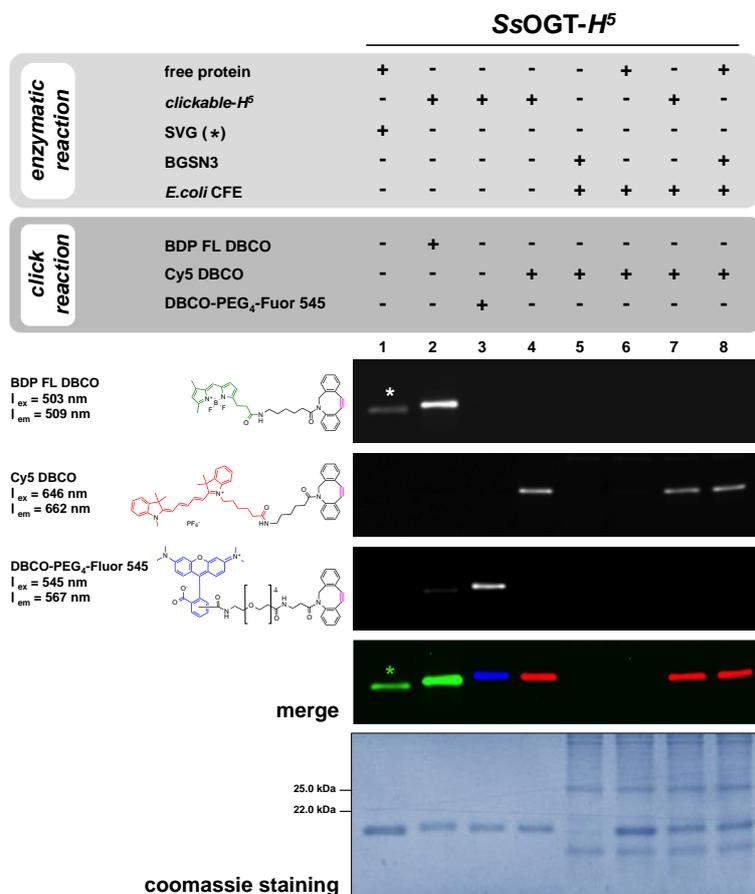


**Figure 7.II Specificity of the Huisgen reaction.** Gel-imaging analysis of SNAP-tag<sup>®</sup> labelling with BGSN3 and three different DBCO- fluorophores. Protein (5  $\mu$ M) was first incubated with 5  $\mu$ M of the BGSN3 for 60 min at 25 °C; then, an equimolar amount of DBCO-based substrate was added for the chemical click reaction. As control, SNAP-tag<sup>®</sup> was incubated only with SVG (lane 1, signal marked with an asterisk).

To test the specificity, we added the DBCO-fluorophore to a crude protein extract from *Escherichia coli* ABLE C (*EcCFE*) without any AGT activity (lane 5) and the sole free *protein-tag* in a context of *EcCFE* (lane 6), all resulting in no fluorescent signal at the gel-imaging analysis. On the contrary, previously purified *clickable*-SNAP (lane 7), as well as its free form in the presence of BGSN3 (lane 8), was specifically able to complete the chemo-enzymatic reaction with DBCO substrate, also in the presence of a perturbing environment (*EcCFE*).

The high specificity of the chemo-enzymatic approach was also confirmed by using the H<sup>5</sup> enzyme in the same and above-mentioned reaction conditions (Figure 8.II). These results clearly

demonstrated the high efficiency of our chemo-enzymatic approach for the labelling of both the *protein-tags* used.



**Figure 8.II Specificity of the Huisgen reaction.** Gel-imaging analysis of H<sup>5</sup> labelling by a chemo-enzymatic approach with BGSN3 and three different DBCO-derivative fluorophores.

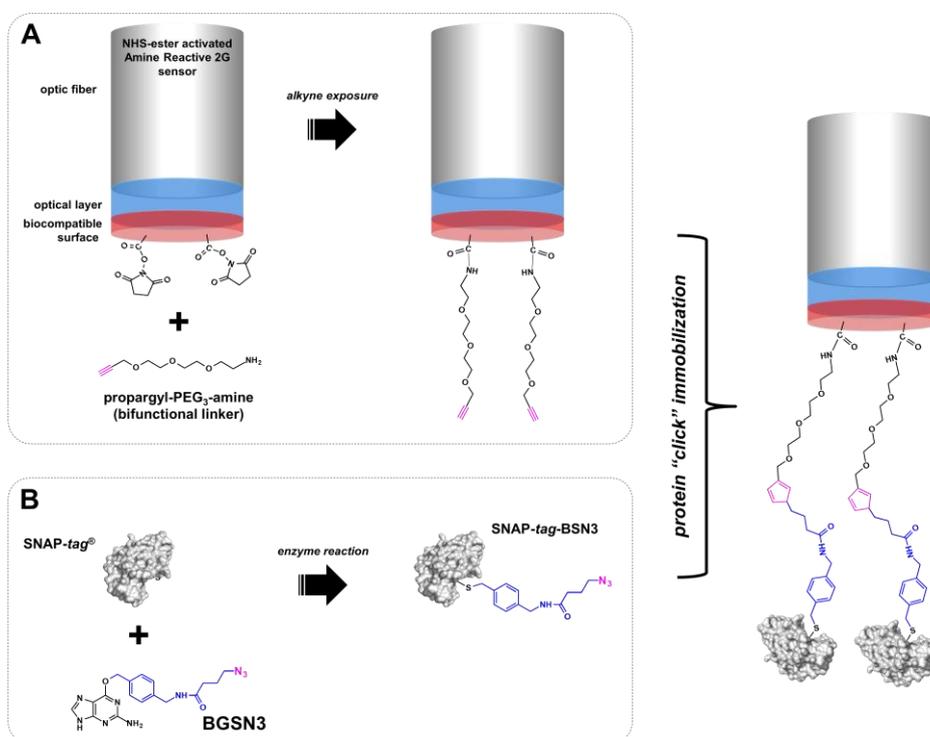
## 2.5 Application to the Bio-Layer Interferometry (BLI)

The possibility to apply the SNAP-tag<sup>®</sup> technology to the Surface Plasmon Resonance (SPR) for covalent immobilisation of a protein of interest was first explored by the group of Kai Johnsson and others (Kindermann et al., 2003; Huber et al., 2004; Niesen et al., 2016). However, their approaches still required the synthesis and the purification of the substrate to cover the sensor chip surface. On the contrary, by using the chemo-enzymatic approach with BGSN3 substrate it is possible to immobilize SNAP-tag<sup>®</sup> directly on an alkyne-derived sensor chip of the Bio-Layer Interferometry (BLI) equipment (Merlo et al., 2021). The SPR is based on the observation in real-time of binding interactions between an analyte in solution and a ligand immobilized on a biosensor through an optical phenomenon, and its applications range from the pharmaceutical or drug discovery to nanotechnology fields, generally based on the study of biomolecule complexes from the point of view of interactions. In SPR the sensor is covered with a functionalized polymer and reacts to

changes in the refractive index, these changes are directly proportional to the mass of the analysed molecule attached to the biosensor.

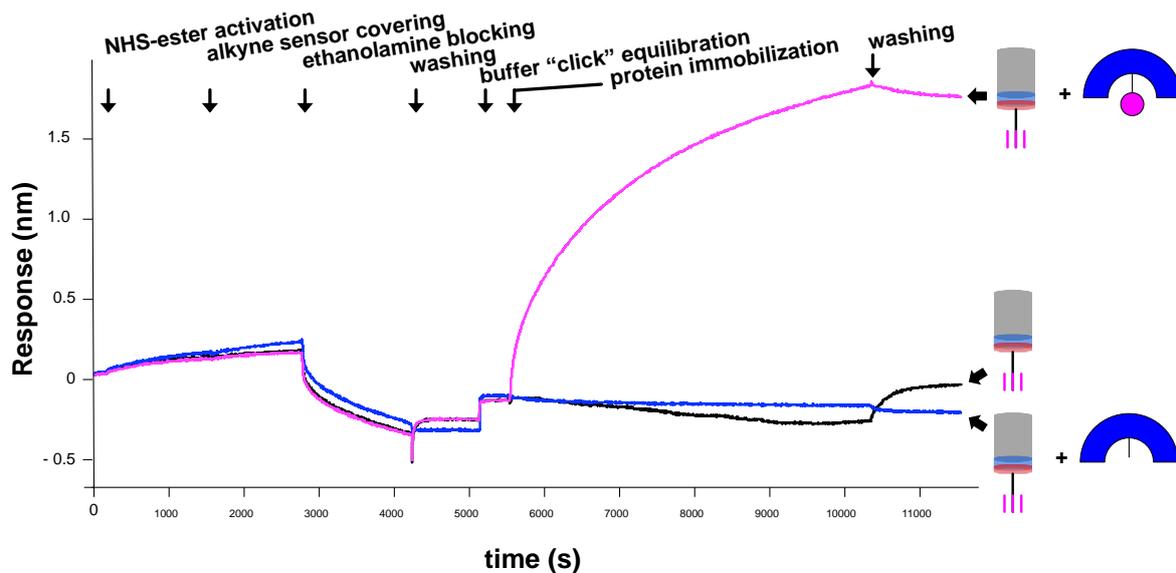
On the other hand, the Bio-Layer Interferometry (BLI) is based on the analysis of the interference profile generated by a white light reflected from two surfaces: a layer of molecules immobilized on the biosensor and an internal reference. Each change in the number of molecules bound at the sensor generates a shift in the interference pattern or a variation of the wavelength, that can be measured in real-time. This technique is more advantageous than the SPR because: (i) it needs a smaller amount of sample, making it more compatible to higher throughput (the capacity of running up to 96 samples in a parallel); (ii) the possibility to reuse samples, and (iii) of the total independency from any microfluidic issues.

Given the lack of any available BLI alkyne-derived sensors, we first activated the AR2G type sensor by a bi-functional linker (propargyl-PEG 3-amine) (see Materials and Methods section) in order to expose an alkyne group on the surface (Figure 9.II box A). This modified protocol provides the coating of the sensor tips with alkyne groups (approx. 80 min), occurring in the first wells of a plate. During this time, the reaction between the *protein-tag* and BGSN3 substrate can take place (Figure 9.II box B).



**Figure 9.II** Covalent immobilization of clickable-tags on the BLI sensor. (A) Covering of the BLI sensor with a bi-functional linker, exposing alkyne groups for the CuAAC reaction; (B) reaction of the SNAP-tag<sup>®</sup> with BGSN3.

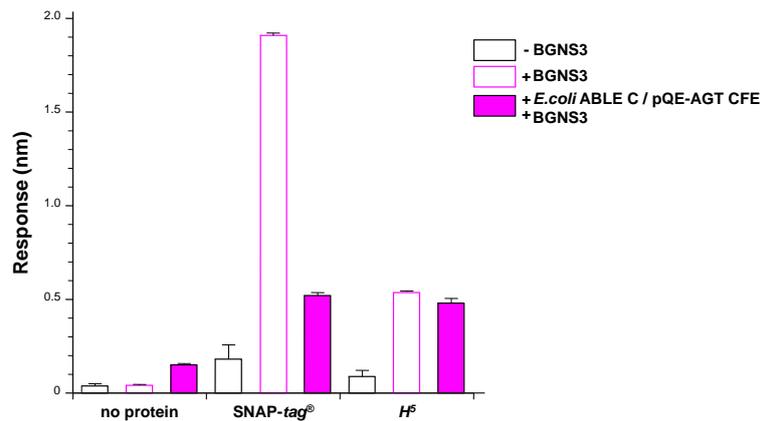
At the end of all analysis steps illustrated in Figure 10.II, it is possible to see that only the contemporary presence of the *clickable*-SNAP and the alkyne-coated sensor led to a measurable response. In fact, during the immobilization step, the signal of *clickable*-SNAP increases but does not drop down after washing procedures, given the covalent reaction between the protein and the sensor. Whereas, from the sensorgram no increasing signal has been measured for the only alkyne-coated sensor and the combination of sensor and free SNAP-*tag*<sup>®</sup> protein without BGSN3.



**Figure 10.II** Sensorgram of chemo-enzymatic SNAP-*tag*<sup>®</sup> immobilization on BLI. The alkyne-covered sensor (silver cylinder) was immersed in wells containing the buffer (in *black*), the free SNAP-*tag*<sup>®</sup> (in *blue*) and the clickable-SNAP (in *magenta*)

This result has been obtained also for H<sup>5</sup> protein (Figure 11.II), although temperature and times of the enzymatic reaction on BLI (30 °C) favoured the SNAP-*tag*<sup>®</sup> (Mollwitz et al., 2012; Vettone et al., 2016). Furthermore, in order to prove the specificity of the immobilization technique, EcCFEs, where both the enzymes were expressed, were analysed (Figure 11.II) (Merlo et al., 2021).

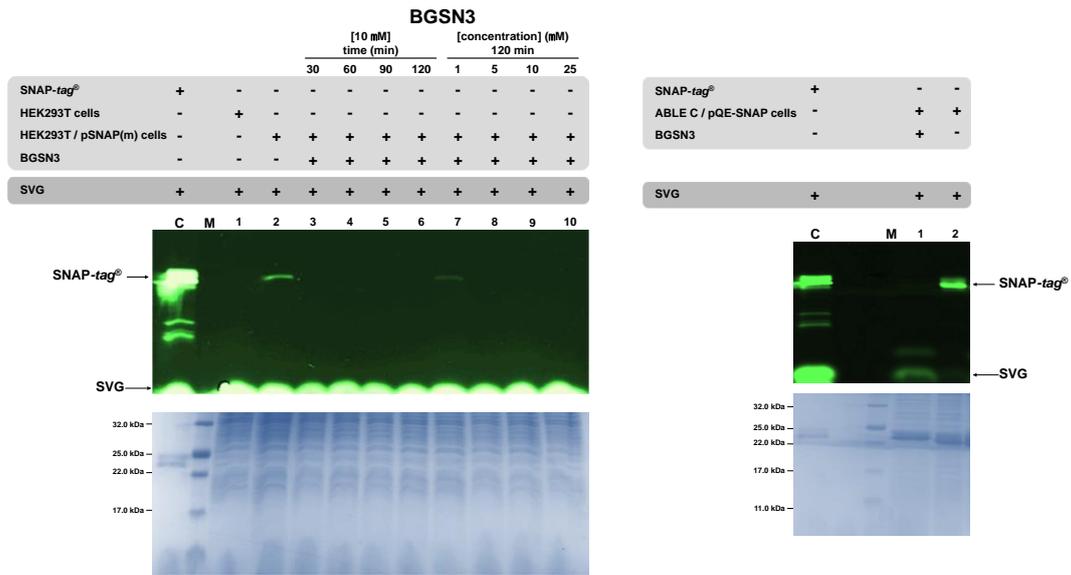
As shown in Figure 11.II, the specific immobilization of both proteins successfully occurred also in the context of a perturbing environment. However, the lower capacity to bound the sensor displayed by the SNAP-*tag*<sup>®</sup> in EcCFE context (filled *magenta* bar), compared to the purified protein (*magenta*-bordered bar), could be explained with a low expression level obtained in the preparation used.



**Figure 11.II** Column chart relative to the BLI immobilisation. Purified protein-tags alone (black-bordered bars), in the presence of BGSN3 (magenta-bordered bars) and using the EcCFE (filled magenta bars). Standard deviations were obtained from three independent experiments. Data are represented as mean  $\pm$  SEM.

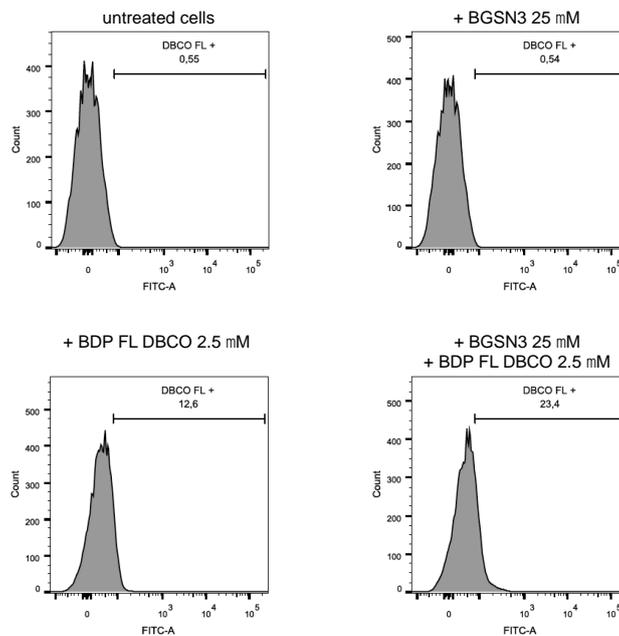
## 2.6 Permeability of eukaryotic and prokaryotic cells to BG-azides

Since one of the major applications of the SNAP-tag<sup>®</sup> technology concerns the field of cell biology to study protein functions and locations in living cells (Testa et al., 2013), we tested our chemo-enzymatic approach on eukaryotic cell model, by investigating the permeability of BGSN3 on HEK293T cells, in collaboration with the group of Prof. Leonardi (University of Naples “Federico II”). The cells were pre-treated by adding BGNS3 in culture medium and then lysed to obtain the protein extract. The samples are successively incubated with the SVG substrate and analysed by SDS-PAGE and gel-imaging: the absence of any fluorescent signal only in BG-azide treated lysates demonstrated that the internalisation of BGSN3 was fast (ca. 30 min; Figure 10.II, lane 3) and at concentrations comparable with commercial cell biology BG-substrates (in the range of  $<5 \mu\text{M}$ ; Figure 10.II, lane 8). The permeability was also tested and confirmed for *E. coli* bacterial cells (Figure 10.II) following the protocol reported in Materials and Methods section.



**Figure 10.II** Eukaryotic and prokaryotic permeability to BGSN3. SDS-PAGE analysis by gel-imaging and Coomassie staining of HEK293T and *E. coli* ABL C cell lysates. After BGSN3 addition in medium, lysates were incubated with SVG.

Preliminary experiments by FACS analysis confirmed that the *in vivo* cycloaddition between BGSN3 and the BDP-FL DBCO occurred (Figure 11.II).



**Figure 11.II** Eukaryotic permeability to BGSN3. FACS analysis of HeLa cells pre-treated with the BGSN3 and then incubated with the BDP-FL DBCO fluorophore.

### 3. Conclusions

These results demonstrate that the novel approach we developed led to an innovative modification of the SNAP-tag<sup>®</sup> technology, which could overcome the duration of the experiment and costs relative to the production and the utilisation of commercial or customised BG-derivatives. Although they are compatible in terms of catalytic activity as for the SNAP-tag<sup>®</sup>, as well as for the others AGTs (Perugino et al., 2012; Perugino et al., 2015; Vettone et al., 2016; Morrone et al., 2017; Kindermann et al., 2003; Miggiano et al., 2017) the risk of lowering the catalytic activity of these tags with customised BG-derivatives should not be underestimated. In addition, we proved that self-labelling *protein-tags* are still folded and enough stable in their *azidated* form after being bound by the BGSN3 and that the click chemistry is an extremely versatile, fast and specific reaction. In fact, the cycloaddition was recently used for the entrapment of catalytic activities by azide-based pseudo-substrates in *in vivo* condition, the so-called *activity-based protein profiling* method (ABPP) (Zweerink et al., 2017). Moreover, by using BGSN3 substrate, a reaction with a major number of DBCO-based molecules is possible, keeping high the specificity in the presence of *in vitro* “perturbing” proteins (like in cell lysates) and the *in vivo* labelling of expressed SNAP-tag<sup>®</sup> in eukaryotic cells.

In addition, the chemo-enzymatic approach allows the specific surface immobilization of SNAP-tag<sup>®</sup>, favouring a better orientation of tagged-POI for biological activities and allowing performing a directly *on-chip purification* from a crude lysate.

In conclusion, we demonstrated that splitting the SNAP-tag<sup>®</sup> reaction into two fast steps, as experimentally measured (Figure 6.II), does not affect the overall rate and efficiency of the protein labelling (Mollwitz et al., 2012; Vettone et al., 2016), opening new perspectives and widening the applications of this powerful biotechnology.

**Chapter III: *Anchoring and self-labelling***  
***protein-tag* system (ASL<sup>tag</sup>)**

In this chapter, the utilisation of the H<sup>5</sup> variant as part of a new enzyme immobilization tool (*the Anchoring and self-labelling protein-tag-ASL<sup>tag</sup>*) is described. This project is born in collaboration with the group of Dr Clemente Capasso (Institute of Bioscience and BioResources of the National Research Council of Italy - IBBR CNR).

Here are described several immobilization methods, with their own pros and cons, and the novel tool we developed, by using the “thermostable SNAP-tag” from *Saccharolobus solfataricus*. In particular, the obtained results from the construction of the first ASL<sup>tag</sup> are illustrated, describing its properties and advantages alone and in combination with POI genetically fused to it. These data have been published in the *Journal of Enzyme Inhibition and Medicinal Chemistry*.

(<https://doi.org/10.1080/14756366.2018.1559161>)

JOURNAL OF ENZYME INHIBITION AND MEDICINAL CHEMISTRY  
2019, VOL. 34, NO. 1, 490–499  
<https://doi.org/10.1080/14756366.2018.1559161>



RESEARCH PAPER

OPEN ACCESS

## An AGT-based *protein-tag* system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane

Rosa Merlo<sup>a\*</sup>, Sonia Del Prete<sup>a\*</sup>, Anna Valenti<sup>a</sup>, Rosanna Mattossovich<sup>a</sup>, Vincenzo Carginale<sup>a</sup>, Claudiu T. Supuran<sup>b</sup> , Clemente Capasso<sup>a</sup> and Giuseppe Perugino<sup>a</sup>

<sup>a</sup>Department of Biology Agriculture and Food Sciences, Institute of Bioscience and BioResources – National Research Council of Italy, Naples, Italy; <sup>b</sup>Neurofarba Department, University of Florence, Polo Scientifico, Sesto Fiorentino Firenze, Italy

(<https://doi.org/10.1080/14756366.2019.1605991>)

JOURNAL OF ENZYME INHIBITION AND MEDICINAL CHEMISTRY  
2019, VOL. 34, NO. 1, 946–954  
<https://doi.org/10.1080/14756366.2019.1605991>



RESEARCH ARTICLE

OPEN ACCESS

## Thermostability enhancement of the $\alpha$ -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* by using the anchoring-and-self-labelling-*protein-tag* system (ASL<sup>tag</sup>)

Sonia Del Prete<sup>a\*</sup>, Rosa Merlo<sup>a\*</sup>, Anna Valenti<sup>a</sup>, Rosanna Mattossovich<sup>a</sup>, Mosè Rossi<sup>a</sup>, Vincenzo Carginale<sup>a</sup>, Claudiu T. Supuran<sup>b</sup> , Giuseppe Perugino<sup>a</sup> and Clemente Capasso<sup>a</sup>

<sup>a</sup>Department of Biology Agriculture and Food Sciences, Institute of Bioscience and BioResources – National Research Council of Italy, Naples, Italy; <sup>b</sup>Neurofarba Department, University of Florence, Polo Scientifico, Sesto Fiorentino Firenze, Italy

\* These authors equally contributed to the present work.

# 1. Introduction

Nowadays, some of the principal demands of biotechnological industries are the enhancement of enzyme productivity and the increasing of their stability and life, to facilitate large-scale production and to minimize the costs (Datta et al., 2013). In fact, due to their relatively unstable nature, the utilisation of biocatalysts from mesophilic sources displays several disadvantages in bioprocesses, such as the incompatibility to harsh reaction condition or the recovery of the enzymes from the mixture after many reaction rounds (Gurung et al., 2013; Hiep Nguyen H. and Kim M., 2017). To overcome these limitations, two principal strategies can be applied: *i*) the research and study of biocatalysts from thermophilic sources: enzymes from organisms that live at high temperatures and/or harsher chemical-physical conditions have characteristics of stability; *ii*) immobilization of biocatalysts, which had particularly received the attention of the industrial sector (DiCosimo et al., 2013). When immobilized, an enzyme is “physically attached to specific solid supports, thus it can be used repeatedly and continuously while maintaining its catalytic activity” (Katchalski-Katzir E., 1993). In fact, a single batch of immobilized enzymes can be used multiple times, reducing production and purification costs, and the reaction can be controlled by removing the biocatalyst from the mixture, avoiding product contamination (Homaei et al., 2015). For these reasons, immobilized enzymes or whole cells (Kawaguti et al., 2006) are generally preferred over their free counterparts.

## 1.1 Immobilization procedures

Over the past decades, biochemical and biophysical studies have been performed to enhance the stability and activity of enzyme through their immobilization (Zhang et al., 2015) and several methods with natural/synthetic supports, such as inert polymers and inorganic materials, have been used (Datta et al., 2013). The best-known immobilization protocols are schematized in figure 1.III:

### **Adsorption**

This mechanism is based on weak bonds (e.g. hydrophobic interactions, Van der Waal's forces) (Sassolas et al., 2012) between biocatalyst and solid support. Briefly, the enzyme in solution is placed in contact with the support and the unbounded molecules are removed by washing with buffer solution. This method does not involve support functionalization, so it is simple and not expensive and generally it does not affect the enzyme activity, rather, adsorbed enzymes are preserved from aggregation and proteolysis (Spahn C., Minter S.D., 2008). However, since the bond between the support and enzymes is weak, the latter often detach due to temperature, pH or ionic strength changes (Mohamad et al., 2015). Furthermore, this method may suffer from non-

specific adsorption on solid surface of other molecules, causing contamination events (Hiep Nguyen H. and Kim M., 2017).

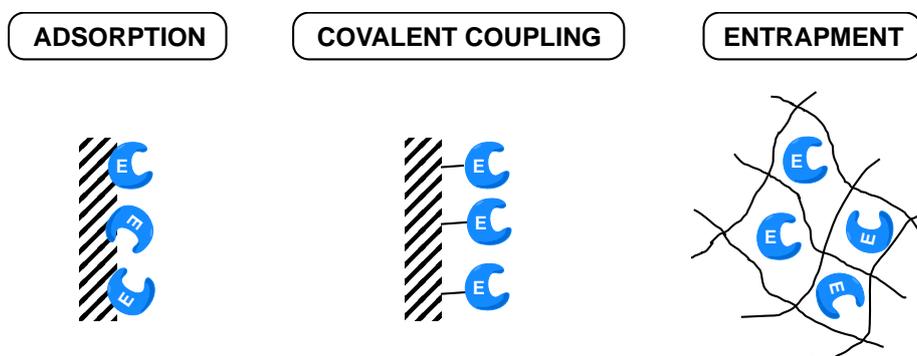
### **Covalent coupling**

This method is based on the covalent interactions between functional groups (e.g. the amino group, carboxylic group, sulfhydryl group, thiol group, imidazole group, hydroxyl group) present on the enzymes and the solid support (Novick and Rozzell, 2005). Of course, these groups must be non-essential for enzymatic activity; otherwise, the enzyme would be immobilized in an inactive or partially active form. This procedure involves a phase of activation of the support and the consequent stable binding with the enzyme. However, despite the strength of the bond, due to the high quantity of reagents to be used and the chemical changes that enzymes must undergo to acquire functional groups (which often cause their denaturation), it is possible to immobilize only a small amount of biocatalysts (a few milligrams of enzyme per gram of matrix) (Hiep Nguyen H. and Kim M., 2017).

### **Entrapment**

In this case, the enzyme is not directly attached to the surface but it is "entrapped" within a polymeric matrix that allows the only diffusion of substrate and products: first, the enzyme is mixed into a monomer solution, and then the polymerization of the monomer occurs.

This method seems to improve the stability of the enzymes and minimize denaturation phenomena since the enzyme does not chemically interact with the polymer. However, a major limitation of this system is that the thickness of the matrix tends to increase during polymerization. Thus, the substrate cannot diffuse easily and reach the enzyme (Hiep Nguyen H. and Kim M., 2017).



**Figure 1.III.** Best-known examples of Enzyme Immobilization. (a) Adsorption. Enzymes (E, in blue) are weakly attached on solid support; (b) Covalent coupling. After the activation of the support, enzymes are irreversibly bounded; (c) Entrapment. Enzymes are encapsulated into a polymeric matrix.

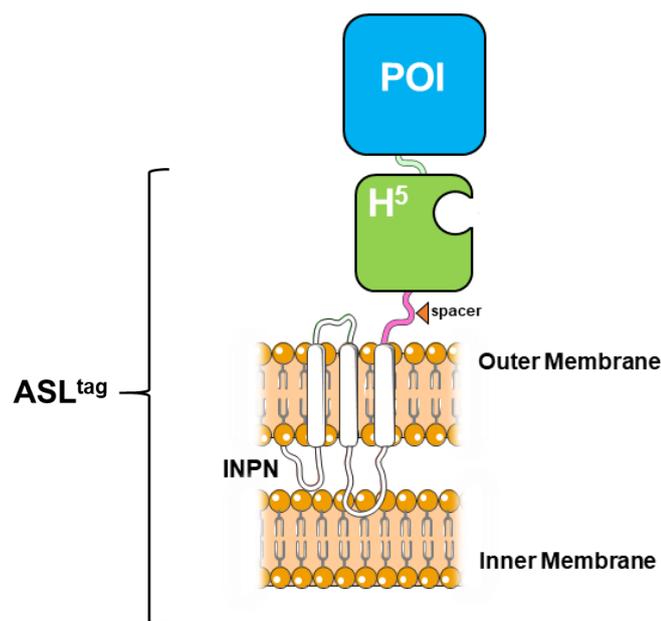
Nevertheless, although these three methods are well known and most used, they have the common problem of high costs because of the expensive production and purification procedures of biocatalysts. In addition, enzyme immobilization often involves conditions and chemical reagents that lead to conformational changes and loss of activity, due also to the “random” interaction and orientation that enzymes can achieve on solid support.

### **Enzyme display via Ice nucleation protein (INP)**

The introduction of natural systems, such as some of bacteria auto-transporter as the outer membrane protein A (OmpA) or phosphoprotein E (PhoE), have solved most of the limitations of immobilization procedures (Dautin et al., 2007; Henderson et al., 1998). In particular, the strategy of using the ice nucleation protein (INP) from *Pseudomonas syringae* offers the opportunity to display enzymes of interest on the outer membrane (OM) of Gram<sup>(-)</sup> bacteria, via an N-terminal signal peptide to transport enzymes into the periplasm, folding and insertion into the outer membrane and translocation onto the bacterial cell surface (Cochet et al., 2000; Binder et al., 2010; Samuelson et al., 2002; Lee et al., 2003). The INP protein is composed of an N-terminal domain (INPN) (N, 175 residues) and a structurally separated C-terminal domain (C, 49 residues) by a repetitive central motif (Graether et al., 2001). Both domains play a role in the anchoring of proteins to the OM. Specifically, the use of INPN as anchoring carrier is considered of great interest in biotechnological applications, ranging from the development of bacterial cell surface-display systems for vaccine delivery to the fabrication of whole-cell biocatalysts and biosensors (Samuelson et al., 2002; Lee et al., 2003; Daugherty PS, 2007).

## **2. Results and Discussion**

Starting from the knowledge of enzyme display via INPN, part of my PhD activity was focused on the development and characterization of a novel tool for the immobilization of biocatalysts, the *Anchoring-and-Self-Labeling-protein-tag* (hereinafter ASL<sup>tag</sup>), which includes the thermostable protein H<sup>5</sup> as *protein-tag* and important part of the construct. In fact, this system is composed of two moieties: *i*) the N-terminal domain of INP (INPN), which immobilizes a protein or an enzyme of interest on *E. coli* outer membrane, allowing its exposition to the solvent and its correct orientation for biological activities (Del Prete et al., 2017); and *ii*) H<sup>5</sup> protein variant from *S. solfataricus* which can be labeled with any desired chemical groups (opportunistically conjugated to the benzyl-guanine) (Figure 2.III).

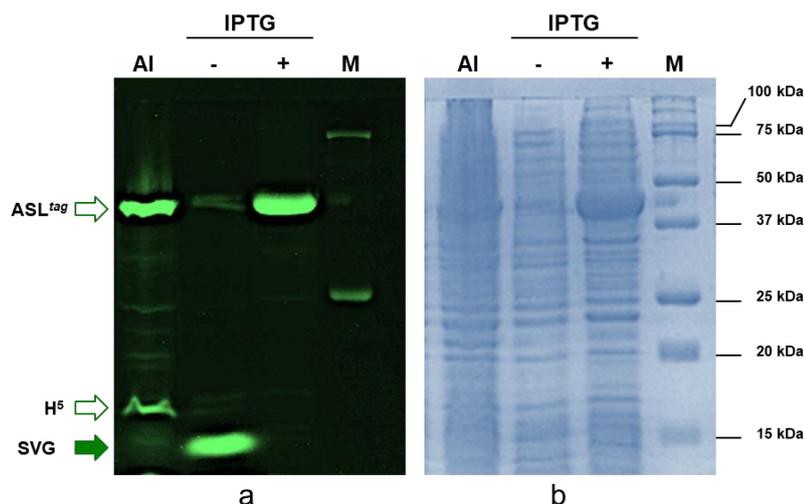


**Figure 2.III** The *ASL<sup>tag</sup>*. This construct allows the *in vivo* protein of interest (POI) (*in light blue*) immobilization on the bacterial Gram<sup>(-)</sup> outer membrane (OM) by the INPN domain (*in white*) and at the same time the possibility to label it or combine its biological activities with desired chemical groups by the covalent labelling operated by the H<sup>5</sup> enzyme (*in green*).

As already mentioned in Chapter I, H<sup>5</sup> is the first thermostable SNAP-*tag* described. Due to its peculiar reaction mechanism and to the available fluorescent assay, it has been possible to analyse the expression of the *ASL<sup>tag</sup>* construct, both *in vivo* and *in vitro* in *E. coli* BL21(DE3) transformed cells (Merlo et al., 2019; Del Prete et al., 2019).

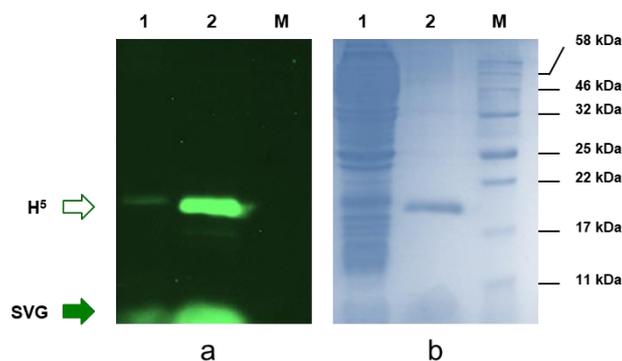
First, to evaluate the expression level of the *ASL<sup>tag</sup>*, the cells were grown in LB medium in different induction conditions; some batches were induced with IPTG, while others were grown in Auto-Induction medium (AI) (as described in Materials and Methods section). Then, the fluorescence assay was performed and samples were analysed by SDS-PAGE and gel-imaging technique (Figure 3.III).

Although the fusion protein was satisfactorily expressed in both conditions, only the samples grown in AI medium present the signal of the H<sup>5</sup> protein alone in gel-imaging analysis. This suggests that interruptions or failure events may occur in the translocation process on the OM of bacteria cells, which lead to breaking points or cleavage of the fusion protein, especially concentrated in the spacer region between the two moieties. On the other hand, no fragments but a defined signal of expected molecular weight is visible in samples after the IPTG induction.



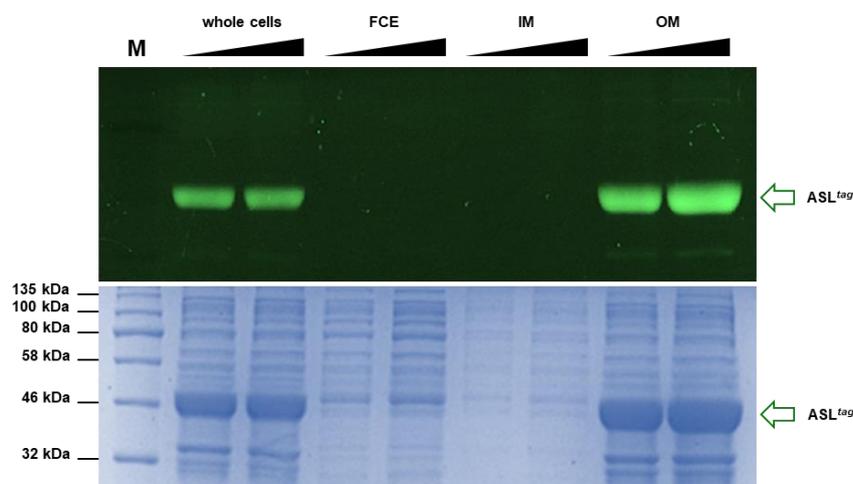
**Figure 3.III** *ASL<sup>tag</sup>* expression in *E. coli*. BL21(DE3) strain transformed with the pET-*ASL<sup>tag</sup>* plasmid was grown in IPTG-induced or in auto-induction medium (AI). After the *in vivo* AGT assay, a defined amount in micrograms of whole cells at OD<sub>600nm</sub> of 1.0 was directly loaded on SDS-PAGE, followed by gel-imaging fluorescence (a) and Coomassie staining (b) analyses. Open and closed green arrows indicate fluorescent signals of H<sup>5</sup> or *ASL<sup>tag</sup>* and the free SVG substrate, respectively. M is the molecular weight marker.

This result not only highlighted the optimal growth conditions for *ASL<sup>tag</sup>* expression but also demonstrated that the extracellular translocation of the H<sup>5</sup> protein made its expression possible in BL21 (DE3) strain. In fact, to date, the heterologous expression of *SsOGT* has been performed in *E. coli* ABLE C strain, which keeps the number of *ogt*-plasmid copies low (Perugino et al., 2012; Vettone et al., 2016). To corroborate this result, pQE*ogtH<sup>5</sup>* plasmid was used to transform BL21(DE3) cells (Vettone et al., 2016), showing a very low expression level (Merlo et al., 2019) (Figure 4.III, lane 1).



**Figure 4.III** The expression of free H<sup>5</sup> protein in *E. coli* BL21 (DE3). In lane 1, BL21 cells transformed with pQE*ogtH<sup>5</sup>* plasmid display a very low fluorescent signal and corresponding low amount of protein expressed. In lane 2, as control, purified H<sup>5</sup>

The fluorescent assay on H<sup>5</sup> was used to confirm also the anchoring of the ASL<sup>tag</sup> through the INPN trans-membrane domain (Merlo et al., 2019). After been fractionated with several cycles of centrifugation, both the whole cells and the steps of membrane purification, in order cytoplasmatic fraction, inner membrane and outer membrane, were tested with SVG and analysed by SDS-PAGE. The fluorescence of H<sup>5</sup> containing samples was visualized in gel-imaging (as described in Materials and methods section). As result, only whole cells and the OM fraction displayed a fluorescent band of the expected molecular weight, corresponding to the ASL<sup>tag</sup> fusion protein, whereas the signal was missed in the lanes relative to the cytoplasmic and IM fractions (Figure 5.III).



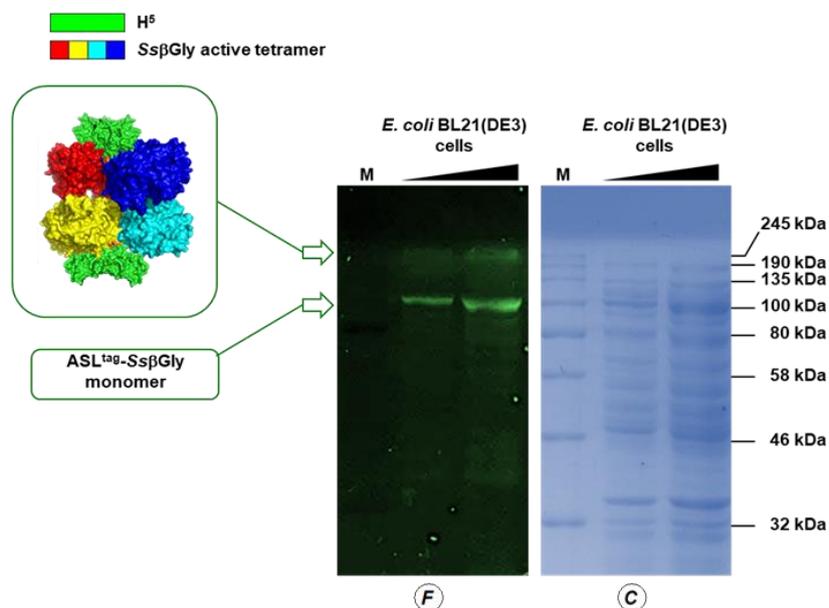
**Figure 5.III** Localisation of ASL<sup>tag</sup> in *E. coli*. (a) Gel-imaging and Coomassie staining analyses after SDS-PAGE of different loaded amounts of the whole cells, the relative cytoplasmic fraction (FCE), the inner (IM) and the outer membrane (OM) fractions.

After confirming that the ASL<sup>tag</sup> system is sited only on the outer surface of the bacteria cell, it was tested in fusion with two thermostable enzymes: the  $\beta$ -glycoside hydrolase from the thermophilic archaeon *S. solfataricus* (*Ss* $\beta$ Gly) (Moracci et al., 1996) and a carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* (*SspCA*) (Di Fiore et al., 2013).

## 2.1 *Ss* $\beta$ Gly

In our laboratory, it was previously demonstrated that the cytoplasmic H<sup>5</sup>-*Ss* $\beta$ Gly fusion protein is stable and active for both the enzymatic assays, suggesting that the two enzymes do not interfere with each other, in terms of folding and catalytic activity (Vettone et al., 2016). However, the idea of anchoring the ASL<sup>tag</sup>-*Ss* $\beta$ Gly protein to the bacterial OM was particularly challenging, because it

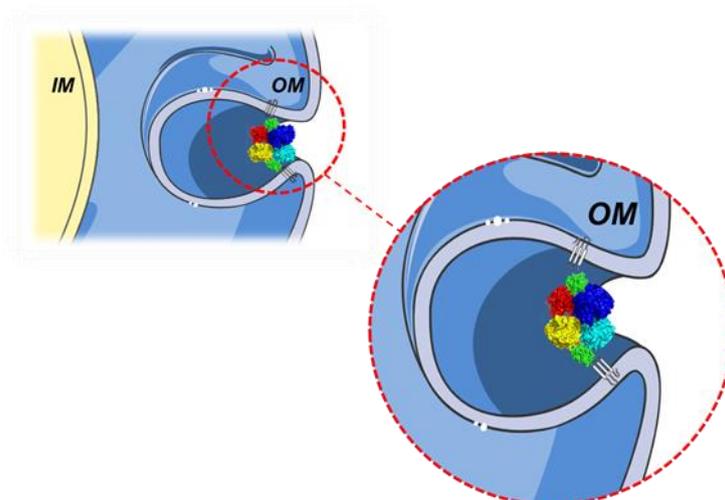
has been demonstrated that *SsβGly* is active only in its tetrameric form (Aguilar et al., 1997, Pouwels et al., 2000). The first evidence of the  $\beta$ -glycoside hydrolase activity has been obtained after transforming BL21(DE3) cells with ASL<sup>tag</sup>-*SsβGly* plasmid and plated on LB agar in presence of a glucoside chromogenic derivative (5-Bromo-4-chloro-3-indolyl  $\beta$ -D-glucopyranoside, also known as X-glucoside or X-Glc), which is a preferred substrate of *SsβGly* (Perugino et al., 2003) but not of the *E. coli* LacZ (a  $\beta$ -galactosidase enzyme). The presence of blue colonies only for the *E. coli* BL21(DE3) transformed with ASL<sup>tag</sup>-*SsβGly*-containing plasmid was a convincing indication that the oligomerization of this thermostable enzyme occurred. Subsequently, we tested the expression level of the fusion protein by SDS-PAGE analysis (Figure 6.III). In this case, IPTG-induced *E. coli* BL21(DE3)/pET-ASL<sup>tag</sup>-*SsβGly* cells displayed a fluorescent signal of expected molecular weight (98.8 kDa, which corresponds to one monomer of fusion protein), indicating the successful expression of ASL<sup>tag</sup>-*SsβGly*. However, with this construct, we have obtained a total amount of fusion protein lower than the sole ASL<sup>tag</sup>, probably because of the complexity of the enzymatic structure. In addition, a higher band is visible in the fluorescence analysis, out of the molecular weight marker range used. This band could correspond to the partially denatured tetrameric form of ASL<sup>tag</sup>-*SsβGly* (ca. 400 kDa), which is particularly resistant to thermal denaturation (Vettone et al., 2016, Moracci et al., 1996, Aguilar et al., 1997) (Figure 6.III).



**Figure 6.III** Heterologous expression of thermostable  $\beta$ Gly enzyme fused to ASL<sup>tag</sup>. SDS-PAGE analysis of the ASL<sup>tag</sup>-*SsβGly* expression. Fluorescent signals (on the F panel) shows bands of molecular weights corresponding to the protein fusions and a very high fluorescent band, presumably corresponding to the tetrameric form of *SsβGly* (represented in the green box)

To demonstrate the correct folding of the *Ss*βGly tetramer in non-denaturing condition, an amount of 0.24 μg/mg of immobilized *Ss*βGly (based on calculated H<sup>5</sup> pmol, see Figure 11.III and Table 1.III) was assayed on 2 and 4-Nitrophenyl β-D-glucopyranoside (2 Np- and 4 Np-Glc) at three different temperatures, 50-60 and 70 °C. The resulting data show an activity of 12.6 ± 0.7 and 8.8 ± 0.4 (50 °C), 30.3 ± 0.4 and 20.3 ± 0.9 (60 °C), 51.5 ± 0.9 and 30.8 ± 1.5 U/mg (70 °C), respectively, whereas OM fraction containing the sole ASL<sup>tag</sup> did not show any β-glucosidase activity. These values are correctly related to the activity of the free form of *Ss*βGly (Moracci et al., 1998). Thus, since the activity of *Ss*βGly depends on its homotetrameric structure, these data confirm that the formation of the quaternary structure of *Ss*βGly on the *E. coli* OM occurred.

However, since the 3D structure of this glycoside hydrolase indicates that it is not a planar structure (Aguilar et al., 1997), we hypothesized an invagination of the *E. coli* OM to allow the assembly of all four units of the ASL<sup>tag</sup>-*Ss*βGly (Figure 7.III), resulting in the active form of this thermostable enzyme.



**Figure 7.III** The tetrameric form of the ASL<sup>tag</sup>-*Ss*βGly fusion protein. Schematic representation of the possible spatial disposition of the tetrameric *Ss*βGly (PDB ID: 4GOW) linked to four H<sup>5</sup> units (PDB ID: 6GA0). The hypothesized invagination of the external membrane of *E. coli* would make it possible the assembling of the tetrameric form and the consequent measured catalytic activity of the *Ss*βGly.

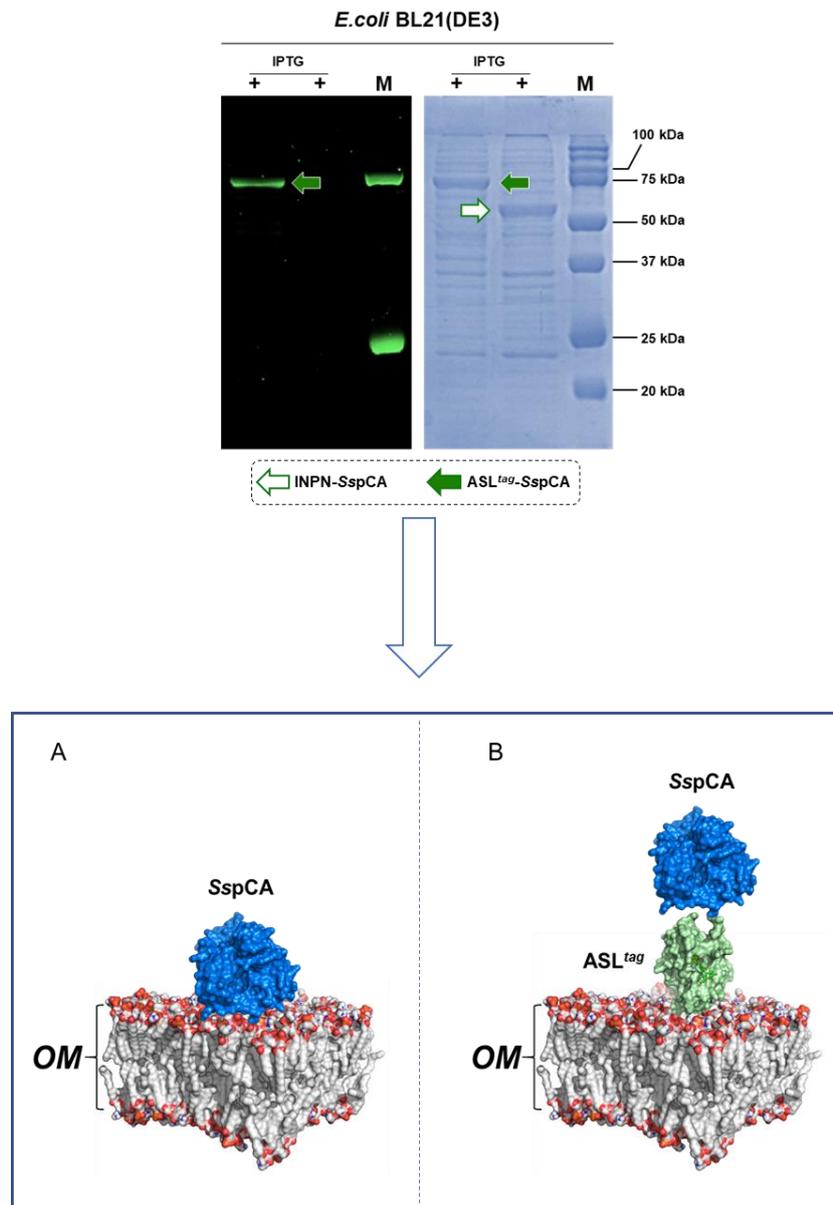
## 2.2 *SspCA*

Several studies were carried out on Carbonic anhydrase enzymes (CAs, EC 4.2.1.1), a superfamily of ubiquitous metalloenzymes present in all living organisms on the planet. The bacterial CAs received particular attention in medical fields as new drug target, due to their inhibition effects on the growth or virulence of many pathogens (Annunziato et al., 2016; Capasso and Supuran, 2015;

Ozensoy Guler et al., 2016), and in biotechnological applications (De Simone et al., 2002; Di Fiore et al., 2015) as post-combustion carbon capture process, artificial lungs, and biosensors (Alterio et al., 2015).

However, these processes are often characterized by harsh conditions, thus, especially in Biotech fields, there is great interest in searching for more resistant biocatalysts. In this context, the group of Dr Capasso and his collaborators identified and immobilized in a one-step procedure (using INPN strategy) a CA from the extreme thermophile *S. yellowstonense* (*SspCA*) (Del Prete et al., 2017; Capasso et al., 2012). This enzyme showed to be highly thermostable and capable of retaining its catalytic activity even when heated for a prolonged time at a high temperature (Vullo et al., 2013; Akdemir et al., 2013; De Luca et al., 2012; Vullo et al., 2012).

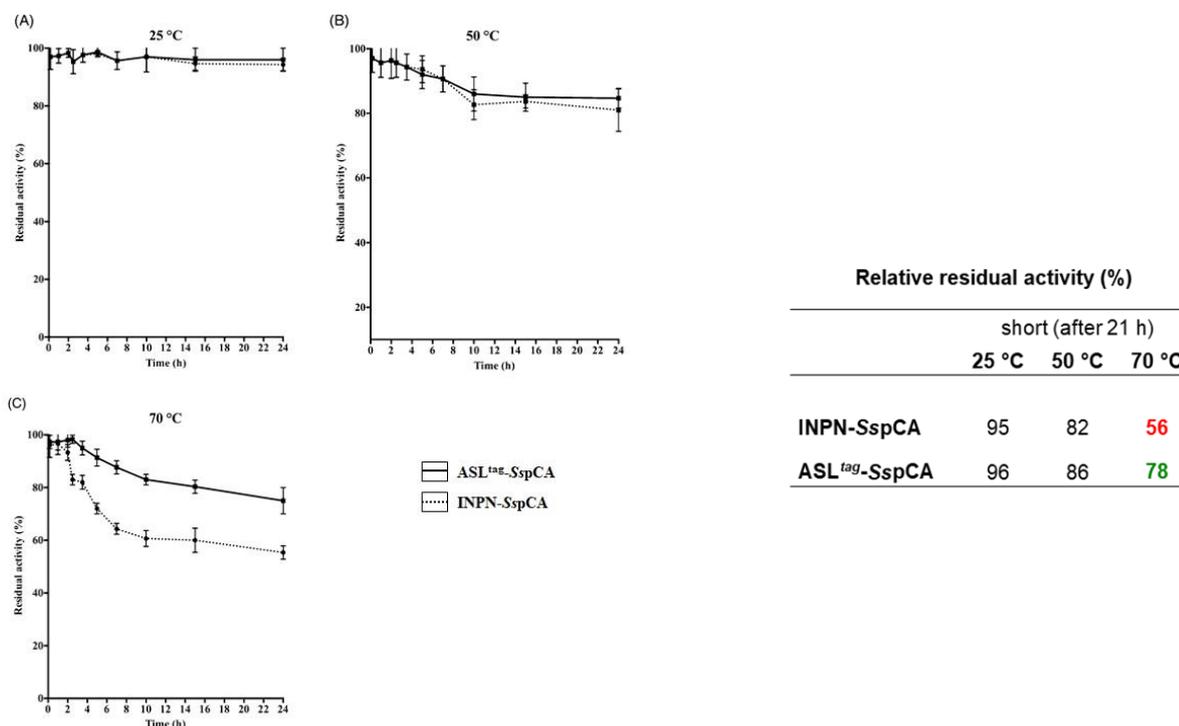
Starting from this point, part of my research activity was focused on the utilization of the ASL<sup>tag</sup> system to express the CA, trying to implement its activity. First, we realized a ASL<sup>tag</sup>-*SspCA* construct by inserting the *ogtH*<sup>5</sup> gene between the INPN and *SspCA*. Successively, both the ASL<sup>tag</sup>-*SspCA* and the INPN-*SspCA* plasmids were used to transform BL21(DE3) cells (as described in Materials and Methods) and the relative protein expressions were compared. As result, by SDS-PAGE and gel-imaging analysis, a fluorescent signal of expected molecular weight is present only in the first lane, representing the full-length ASL<sup>tag</sup>-*SspCA* (70.2 kDa), whereas the band corresponding to the INPN-*SspCA* fusion protein (53 kDa) can be appreciated only in Coomassie Blue staining in the second lane, because missing of any activity on BG derivative substrates (Figure 8.III) (Merlo et al., 2019; Del Prete et al., 2019).



**Figure 8.III** *ASL<sup>tag</sup>-SspCA* expression in *E. coli*. On top, fluorescence gel-imaging and Coomassie staining of INPN-SspCA and ASL<sup>tag</sup>-SspCA (white and filled green arrows respectively). Below, a model representation of an OM (PDB from Tieleman and Berendsen) describing the *in vivo* immobilization of INPN-SspCA (in blue; PDB ID: 4G7A; panel A) and in fusion with H<sup>5</sup> (in green; PDB ID: 6GA0; panel B). The INPN domain is omitted because inserted in the OM.

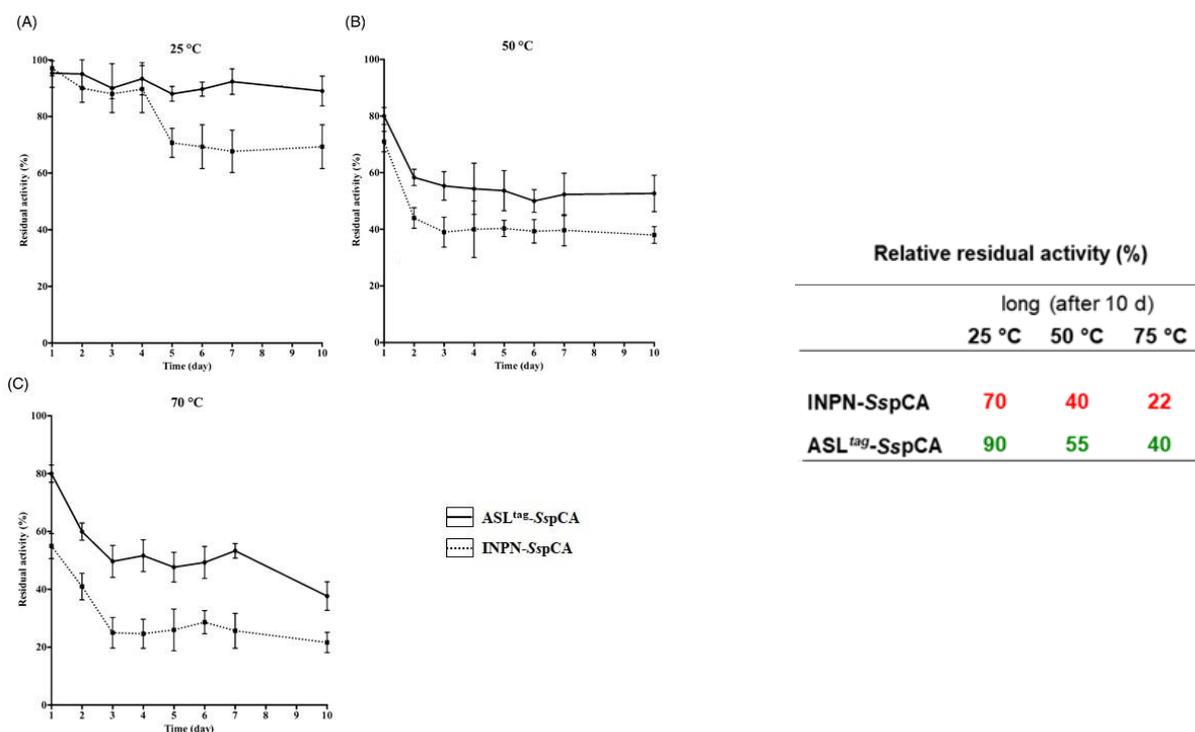
Furthermore, we proved that the ASL<sup>tag</sup> moiety does not affect the activity of SspCA; on the contrary it resulted to be more stable when treated at high temperatures for a prolonged time, if compared to the INPN-SspCA (Del Prete et al., 2019). Briefly, together with our collaborators, we investigated the residual CA activity of whole bacterial cells expressing the enzyme on the external surface, incubating them at three different temperatures (25-50-70 °C). What emerged is that, while both ASL<sup>tag</sup>-SspCA and the INPN-SspCA show a comparable residual hydratase activity when incubated at 25 °C and 50 °C for up to 24 hours (Figure 9.III, panels A and B), the ASL<sup>tag</sup>-SspCA

fusion protein maintains an activity of 80% when treated at 70 °C (Figure 9.III, panel C), whereas the INPN-*SspCA* drops its activity to 60% after 2 h, when treated in the same condition (Figure 9.III, panel C). These results demonstrated that the anchoring ASL<sup>tag</sup> system, enhanced the *SspCA* stability of about 20%. This aspect is crucial in the context of the post-combustion carbon capture process, since it requires temperatures ranging from 40 and 60 °C (Russo et al., 2013).



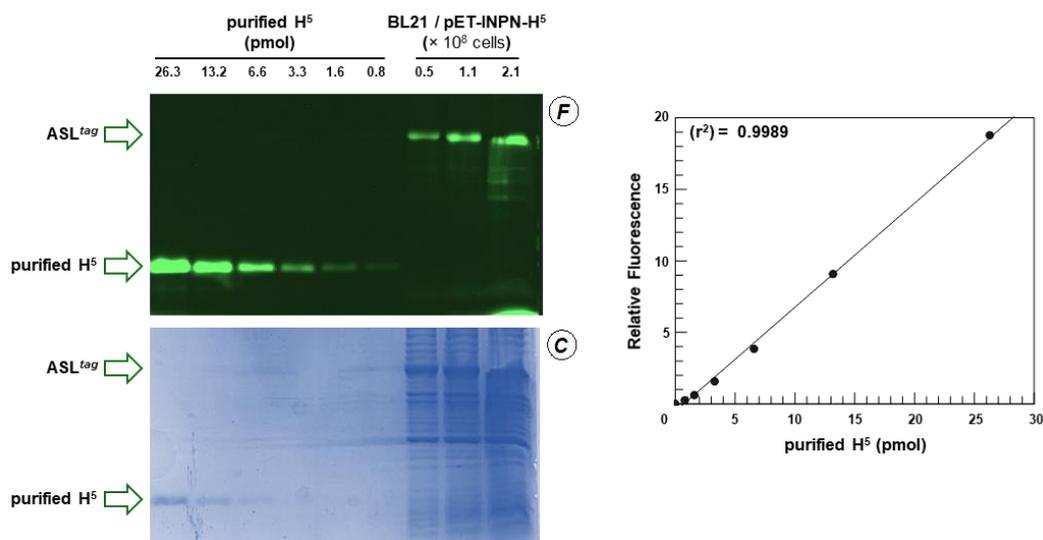
**Figure 9.III** The thermostability of immobilized *SspCA* and ASL<sup>tag</sup>-*SspCA* on the bacterial surface. Measures were carried out at indicated temperatures, by using aliquots of the whole cells incubated up to 24 h. Continuous line, membrane-bound ASL<sup>tag</sup>-*SspCA*; dashed line, membrane-bound *SspCA*. Each point is the mean of three independent determinations.

In addition, the residual activity was analysed and monitored also for up to 10 days. In this case, the differences between the two forms of anchored CAs become more evident. At 25 °C, the INPN-*SspCA* residual activity started to decrease after 4 days and reached a value of about 70% after 10 days, whereas at 50 and 70 °C, its activity decreased up to 40 and 20%, respectively (Figure 10.III, panel B and C). The activity of ASL<sup>tag</sup>-*SspCA* remained almost constant at 90% at 25 °C (Figure 10.III, panel A) and showed a residual activity of about 60 and 40% when tested at 50 and 70 °C, respectively (panel B and C). These data informed us that the presence of a thermostable *protein-tag* between the INPN anchoring domain and the *SspCA* significantly improved the long-term stability of this enzyme (Del Prete et al., 2019).



**Figure 10.III** The long-term stability of immobilised *SspCA*. Measures were carried out at indicated temperatures up to 10 days, using aliquots of whole bacterial cells.

The presence of defined amount of free H<sup>5</sup> enzyme and different volumes of expressed ASL<sup>tag</sup>/ASL<sup>tag</sup>-fused proteins allowed us to quantitatively measure the heterologous expression of our construct as pmol/mg of the whole wet cells (Figure 11.III). Briefly, free H<sup>5</sup> in decreasing quantities and ASL<sup>tag</sup> constructs were tested with SVG fluorescent substrate and then analysed by SDS-PAGE and gel-imaging. Considering the INPN:H<sup>5</sup>:POI ratio as 1:1:1, the amount of the whole fusion proteins expressed was estimated (see Materials and Methods section) and listed in Table 1.III.

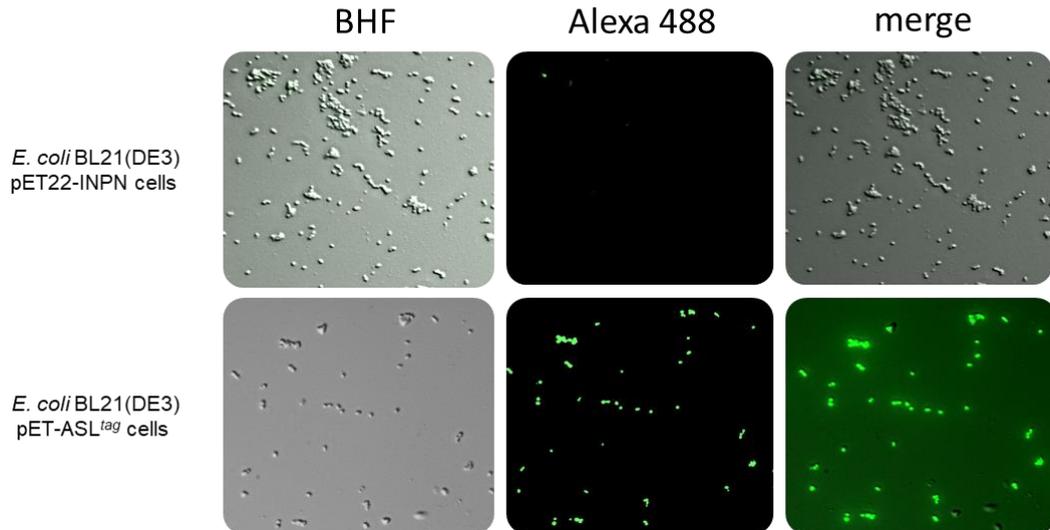


**Figure 11.III** *Quantitative estimation of the ASL<sup>tag</sup> expression.* Defined amount of cells and purified H<sup>5</sup> protein (in pmols) were loaded and analyzed on a SDS-PAGE (on the left). Fluorescent values obtained from H<sup>5</sup> were fitted in a linear plot (on the right), as described in Materials and Methods.

**Table 1.III** *Quantization of ASL<sup>tag</sup> and ASL<sup>tag</sup>-fused protein*

	MW (kDa)	ASL <sup>tag</sup> - Enzyme ratio	Fusion yield (pmol H <sup>5</sup> /mg)	Enzyme yield (μg/mg)	(r <sup>2</sup> )
ASL <sup>tag</sup>	42	-	152.5 ± 15.7	-	0.9977
ASL <sup>tag</sup> -SspCA	70.2	1:1	58.4 ± 11.6	1.54 ± 0.30	0.9986
ASL <sup>tag</sup> -SβGly	98.8	4:1	30.6 ± 13.2	7.2 ± 3.1	0.9980

Finally, as anticipated at the beginning of this chapter, the activity of ASL<sup>tag</sup> was observed also in living cells by microscope analysis. Upon the labelling with the fluorescent substrate, images of living *E. coli* BL21(DE3) cells had shown specific fluorescent signals only in those transformed with the ASL<sup>tag</sup>-containing plasmid (Figure 12.III). These data suggest that ASL<sup>tag</sup> is suitable for localization and analysis of membrane proteins and provide an opportunity for *in vivo* analyses under physiological conditions of proteins of interest fused to the ASL<sup>tag</sup>.

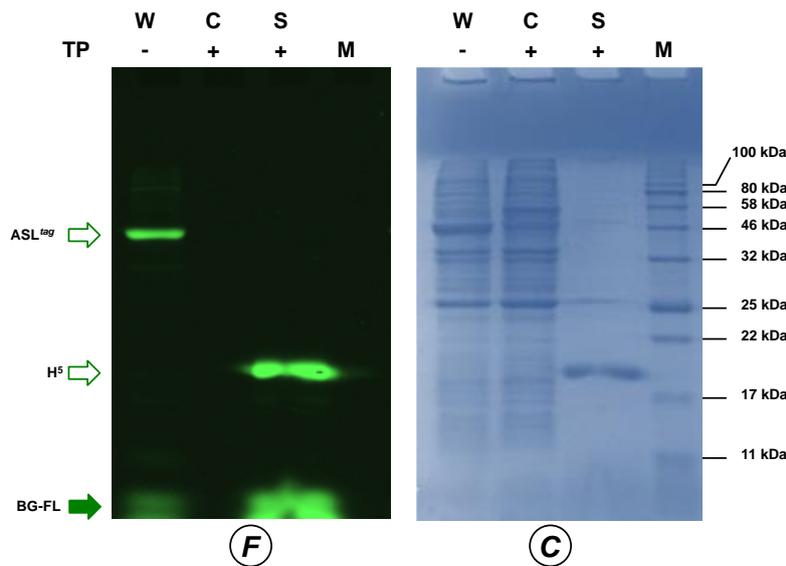


**Figure 12.III** *ASL<sup>tag</sup> in Fluorescence Microscopy.* *E. coli* BL21(DE3) cells transformed with pET-22b/INPN plasmid (Top) or with pET-ASL<sup>tag</sup> (bottom) were incubated with SVG and then analysed at fluorescence microscopy. Images show bright field (BHF), AlexaFluor488 (*green*) and merged images

### 3. Conclusions and perspectives

In conclusion, the ASL<sup>tag</sup> system can be considered as a novel and functional strategy to easily express and immobilize in one-step proteins/enzymes of interest putting them in a correct orientation, and to further increase the thermostability of proteins to be used in biotechnological applications, in which a highly effective and thermostable catalyst is needed.

Moreover, it can lay the foundations for new purification protocols, without going through traditional and expensive procedures. In fact, a cleavage site for thrombin protease was localized between the INPN domain and the H<sup>5</sup> moiety (Merlo et al., 2019). After treating at the same time the whole cells expressing ASL<sup>tag</sup> with this protease and SVG substrate, the fluorescent signal corresponding to the MW of the H<sup>5</sup> protein was present only in the supernatant fraction; indicating that a soluble form of free enzyme can be recovered (Figure 13.III).

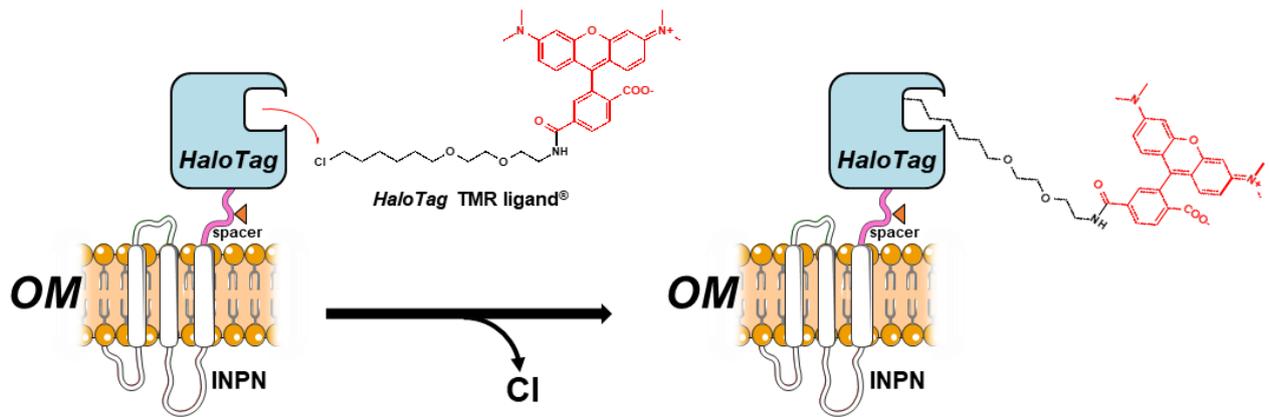


**Figure 13.III** Cleavage of ASL<sup>tag</sup> by the Thrombin protease (TP) on whole cells (W). After the H<sup>5</sup> reaction and simultaneous protease treatment, cells were centrifuged and the supernatant (S) was separated from the intact cells (C). the samples were then analysed by fluorescence imaging and Coomassie staining

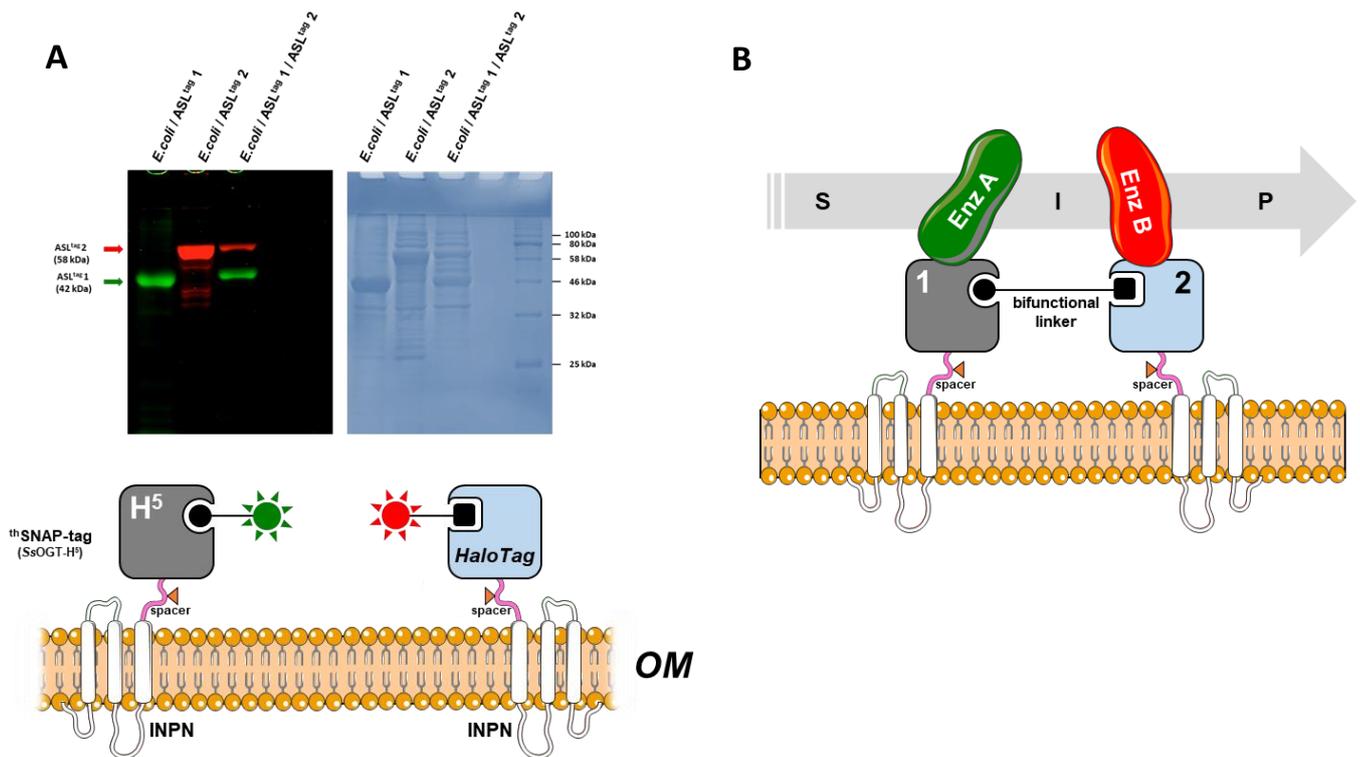
### Anchoring and self-labelling protein-tag 2 (a new ASL<sup>tag</sup>)

Following the same approach, I started to set-up and characterize a novel ASL<sup>tag</sup> composed by the transmembrane INPN domain and the *self-labelling protein-tag* HaloTag (<https://ita.promega.com/resources/technologies/halotag/>) (Figure 14.III) (Los et al., 2008). As the previously described SNAP-tag<sup>®</sup>, this enzyme is an engineered version of a haloalkane dehalogenase which recognizes as substrate a reactive chloro-alkane linker bound to a functional group of interest, with an irreversible reaction. Preliminary data demonstrated the successful expression of the new construct alone or in co-expression with the ASL<sup>tag</sup> on the same *E. coli* cell (Figure 15.III A).

These results can lead to the development of a powerful biotechnological tool for the implementation of enzyme cascade reactions, both in terms of yield and for obtaining a product with high added value, enzymatically and without dispersion of reaction intermediates (Figure 15.III B).



**Figure 14.III.** The  $ASL^{tag} 2$  system on external bacterial surface. In figure is also represented the reaction mechanism of HaloTag protein.



**Figure 15.III.** The co-expression. In panel A, the co-expression of both  $ASL^{tag} 1$  and 2 on the same bacteria cell is represented. In panel B, a possible utilization of this system for biotechnological applications.

## IV. Materials and Methods

### **Bacterial Strains (*E. coli*)**

**ABLE C** (C lac(LacZ $\omega$ -) [Kanr McrA-McrCB-McrF-Mrr-

**DH5 $\alpha$**  (F- 80dlacZ M15 (lacZYA-argF) U169 recA1 endA1hsdR17(rk-, mk+) phoAsupE44 -thi-1 gyrA96 relA1

**KRX** [F', traD36,  $\Delta$ ompP, proA+B+, lacIq,  $\Delta$  (lacZ) M15]  $\Delta$ ompT, endA1, recA1, gyrA96 (Nalr), thi-1, hsdR17 (rk-, mk+), e14- (McrA-), relA1, supE44,  $\Delta$ (lacproAB),  $\Delta$ (rhaBAD)::T7 RNA polymerase.

**BL21 Rosetta (DE3)** F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)

### **Culture media**

#### **LB (Lysogeny Broth) (1 liter):**

10 g NaCl,

5 g yeast extract,

5 g tryptone.

#### **Auto-induced media (1 liter):**

1M MgSO<sub>4</sub>,

20 mL 50x 5052 [25% glycerol, 2.5% glucose, 10%  $\alpha$ -lactose],

50 mL 20x NPS [(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 0.5 M, KH<sub>2</sub>PO<sub>4</sub> 0.5 M, Na<sub>2</sub>PO<sub>4</sub> 1M; pH 6.75],

LB to volume

### **Reagents**

Commercially available substrates for AGT activity, pSNAP-tag(m) plasmid, DNA restriction and modification enzymes were purchased from New England Biolabs (USA); whereas *O*<sup>6</sup>-Benzylguanine was from Activate Scientific GmbH (UK). BDP FL alkyne, BDP FL DBCO, Cy5 DBCO were purchased from Lumiprobe GmbH (Germany). DBCO-PEG4-Fluor 545, Tris(2-carboxyethyl)phosphin (TCEP), Tris [(1-benzyl-1H-1,2,3-triazol-4-yl)-methyl]amine (TBTA) were from Sigma-Aldrich (USA). SYPRO™ Orange for Protein Gel Stain (5,000X Concentrate in

DMSO), Pierce™ Premium Grade Sulfo-NHS (N-hydroxy-sulfo-succinimide) and Pierce™ Premium Grade 1-ethyl-3-(3-dimethyl-amino-propyl)-carbodiimide hydrochloride (EDC) were from Thermo Fischer Scientific (USA). All molecular biology kits for the plasmid preparations were from Macherey-Nagel GmbH (Germany); NZYProof 2x Colourless Master Mix, containing NZYProof DNA polymerase, was from NZYTech (Portugal). Lyophilised Thrombin Protease was purchased from GE Healthcare (USA).

To determine the protein concentration the Bio-Rad protein assay kit (from Bio-Rad Pacific) was used, using purified BSA as standard. Synthesis of the oligonucleotides listed in Table 1.IV and the DNA sequencing service were performed by Eurofins Genomics (Germany)

Methyl-guanine-PEG-NH<sub>2</sub> substrate was a gift of Prof Prosperi (University of Milano-Bicocca, Milan, Italy)

**Table 1.IV.** Oligonucleotides used in this study

Name	Sequence	Note
INPN-Fwd	5'-TAATACGACTCACTATAGGG-3'	-
INPN-Rev	5'-GGTGATGGTGAGATCCTCTCGGAACCAGAGATCCATAGGCTTCAATCAGATCGC-3'	-
H <sup>5</sup> -Fwd	5'-GCGATCTGATTGAAGCCTATGGATCTCTGGTTCCGAGAGGATCTCACCATCACC-3'	-
H <sup>5</sup> -Rev	5'-TCACCTTCATATGACCATTCATGTTTCGCTACCATAAGCTTCTGTGCGACGGTACCTCGAGTTCTGG-3'	-
H <sup>5</sup> -Rev2	5'-ATTGAGCAACTGACTGAAATGCC-3'	-
SspCA-Fwd	5'-CCAGAACTCGAGGTACCGTCGACAGAGGCTTATGGTAGCGAACTGAATGGTCATATGAAGGTGA-3'	-
SspCA-Rev	5'-CTAGTTATTGCTCAGCGGT-3'	-
Lig5- <i>Pfu</i> AGT	5'-TCAGCAAGGGCTGAGGCCATGGTATTGGAAGTTAGG-3'	Nco I site underlined
Lig3- <i>Pfu</i> AGT	5'-CCTCAGCGGAAGCTGAGGTTAGCTTGCCATCCTTCC-3'	
NZY-His Fwd2	5'-CATGGCACACCATCACCATCACCATACGGG-3'	For
NZY-His Rev2	5'-CATGCCCGTATGGTGATGGTGATGGTGTGC-3'	insertion of an His <sub>6</sub> -tag
Fwd <sup>m4</sup>	5'-ggcMgtaggcctagcatgacaatctgattggtgatcacgg-3'	<sup>a</sup>
Rev4	5'-ccgtgatcaccaatgcagattgcatgctaggcctaccgc-3'	<sup>a</sup>
QE-SNAP-Fwd	5'-ATGGCAGGATCCAATGGACAAAGACTGCGAAATG-3'	BamH I and
QE-SNAP-Rev	5'-CTATCAAAGCTTAAACCCAGCCAGGCTT GCCAG-3'	Hind III sites underlined

<sup>a</sup> From Perugino et al., 2015; M=O<sup>6</sup>-methyl guanine

## DNA constructs

### Cloning of SNAP-tag<sup>®</sup> and H<sup>5</sup> proteins

The cloning procedures in pQE31 expression vector (Qiagen, Germany) were similar for both proteins. In particular, the pSNAP-tag(m) Vector was used as template to amplify the DNA fragment relative to the SNAP-tag<sup>®</sup> gene, by using QE-SNAP-Fwd/Rev oligonucleotides pairs. Afterwards, the resulting fragment and the pQE31 vector were digested with BamH I and Hind III restriction enzymes and ligated, leading to the final pQE-SNAP plasmid. The final SNAP-tag<sup>®</sup> protein was expressed with an extra N-terminal aminoacidic sequence, comprising a His<sub>6</sub>-tag (MRGSHHHHHTDP-). H<sup>5</sup> was cloned as described (Perugino et al., 2012).

### Cloning of pET-ASL<sup>tag</sup>

The cloning procedures for pET-ASL<sup>tag</sup> constructs were obtained as described in Merlo et al., 2019. Briefly, pET-22b/INPN-SspCA and pQE-*ogtH*<sup>5</sup> vectors was used as template to obtain the pET-ASL<sup>tag</sup> recombinant plasmid. More precisely, the original plasmid vectors were used to recover the DNA fragments relative to INPN and H<sup>5</sup> genes respectively, which were amplified by PCR, by using INPN and H<sup>5</sup>-Fwd/Rev synthetic oligonucleotides (listed in Table 1.IV). The amplification reaction was performed under the following conditions: an initial denaturation at 95 °C for 5 min, 30 cycles of 30 s at 95 °C, 30 s at 50 °C and 30 s at 72 °C, followed by a final extension of 5 min at 72 °C. The resulting DNA fragments were first purified by using PCR Kleen<sup>™</sup> Purification Spin Columns (Bio-Rad) and then fused to each other in a further amplification reaction in the conditions mentioned above, using INPN-Fwd/H<sup>5</sup>-Rev oligo pairs. This was possible due to the total complementarity of INPN-Rev to the H<sup>5</sup>-Fwd oligonucleotide.

To obtain the final ASL<sup>tag</sup> construct, pET-22b/INPN-SspCA plasmid was used as accepting vector for the PCR product, by removing the INPN-SspCA fusion gene. This vector and the amplified DNA were both digested with Hind III and Xba I restriction endonucleases, then gel-purified using NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) and ligated. The resulting ligation mixture was cloned in *E. coli* DH5α cells and positive colonies were screened by colony PCR and confirmed by DNA sequencing. The same strategy was used to obtain the pET-ASL<sup>tag</sup>-SspCA plasmid, this time by using the H<sup>5</sup>-Rev2 oligo to fuse the ASL<sup>tag</sup> expressing gene to SspCA one, previously amplified use SspCA Fwd/Rev oligo pairs listed in Table 1.IV.

For the pET-ASL<sup>tag</sup>-lacS plasmid preparation, pQE-*ogtH*<sup>5</sup>-lacS vector was digested with Hind III/BamH I endonucleases to obtain the DNA fragment relative to the *ogt-lacS* chimera (*ogtH*<sup>5</sup> gene fused to the  $\beta$ -glycoside hydrolase from *Saccharolobus solfataricus*); likewise the pET-ASL<sup>tag</sup> plasmid. The pET22 resulting recipient was subjected to the first cloning round, where positive blue colonies were selected by the hydrolase activity on Ampicillin selective-Lysogeny Broth (LB-Amp) Agar plates supplemented with 5-bromo-4-chloro-3-indolyl-b-D-glucopyranoside (X-Glc). Successively, the plasmid was digested with BamH I and treated with Alkaline Phosphatase Calf Intestinal (CIP), then ligated to the INPN DNA fragment, derived from the BamH I/BamH I digestion of the pET-ASL<sup>tag</sup> plasmid. Positive blue colonies on LB-Amp-X-Glc agar plates were confirmed by univocal Pst I digestion pattern analysis.

### Cloning of *Pfu*OGT

The ORF PF1878, encoding a putative AGT, was amplified from genomic DNA from *Pyrococcus furiosus* JFW02 strain, by using Lig5/Lig3-*Pfu*AGT oligonucleotides and cloned into the expression vector pHTP1 (NZYtech) following the NZYEasy Cloning and Expression kit I (NZYTech) manual, as described in Mattosovich et al., 2020. Subsequently, commercially available *E. coli* DH5 $\alpha$  cells (NZY5 $\alpha$  Competent Cells-NZYTech) were transformed with the ligation mixture and positive colonies were confirmed by sequencing. In addition, in order to insert a DNA sequence expressing a His<sub>6</sub>-tag sequence (MAHHHHHHTG-) similar to those of *Ss*OGT protein from Perugino et al., 2012, pHTP1-*Pfu*AGT plasmid was digested with Nco I enzyme and ligated to a double-stranded oligonucleotides pairs (NZY-His Fwd2/Rev2, Table 1.IV). The final construct was used to transform *E. coli* KRX cells and the sequence was confirmed by sequencing.

### **Protein expression and purification**

#### SNAP-tag<sup>®</sup> and H<sup>5</sup>

Both proteins were expressed in *E. coli* ABLE C strain, incubated at 37 °C in Lysogeny Broth (LB) (Bertani G. 1951; Bertani G. 2004; Ezraty et al., 2014) medium supplemented with 50 mg/L kanamycin and 100 mg/L ampicillin. The protein expression was induced with 1 mM isopropyl-thio- $\beta$ -D-galactoside (IPTG) when an OD<sub>600nm</sub> of 0.5–0.6 was reached. After an overnight growth, cells were collected and resuspended 1:3 (w/v) in purification buffer A (50 mM phosphate, 300 mM NaCl; pH 8.0) supplemented with 1% Triton X-100 and stored overnight at –20 °C. Subsequently,

the biomass was treated in ice with lysozyme and DNase for 60 min and then sonicated as described (ref). After a centrifugation of 30 min at  $60,000 \times g$ , the cell extract was recovered and applied to a Protino Ni-NTA Column 1 mL (Macherey-Nagel) for His<sub>6</sub>-tag affinity chromatography. After two washing steps of 10 column volumes of Buffer A and 10 column volumes of 10% buffer A supplemented with 250 mM imidazole, the elution was performed in 20 column volumes of buffer A, by applying a linear gradient of 0–250 mM imidazole. The eluted fractions containing the protein were collected and dialysed against phosphate-buffered saline (PBS 1×, 20 mM phosphate buffer, NaCl 150 mM, pH 7.3) and the protein purification was confirmed by SDS-PAGE analysis.

### ASL<sup>tag</sup>

All DNA constructs were used to transform *E. coli* BL21(DE3) cells, grown at 37 °C in LB or in the ZY auto-induction medium (AI) selective medium supplemented with 100 mg/L ampicillin and 30 mg/L chloramphenicol and proteins expression was induced with 1 mM IPTG at OD<sub>600nm</sub> value of 0.6. After an overnight incubation, whole cells expressing the *anchoring and self-labelling* construct were collected and tested by the activity of H<sup>5</sup> moiety for qualitative measurement of expression level.

### PfuOGT

*PfuOGT* was expressed and purified as described for SNAP-tag<sup>®</sup> and H<sup>5</sup> proteins. In this case, *E. coli* BL21 (DE3) cells and LB medium supplemented with 50 mg/L kanamycin and 30 mg/L chloramphenicol were used. In addition to the purification steps, in order to remove *E. coli* contaminants, the cell extract was incubated 20 min at 70 °C, centrifugated at  $13,000 \times g$  at 4 °C and the supernatant was diluted 1:2 (v/v) in purification buffer for a His<sub>6</sub>-tag affinity chromatography. The fractions containing the *PfuOGT* protein were pooled, concentrated and subjected to a further gel-filtration chromatography, using a Superdex 75 10/300 GL column (GE Healthcare Life Sciences). Finally, the protein was loaded on 15% SDS-PAGE gel to confirm its purity and stored at – 20 °C.

### ***In vitro* fluorescent alkyltransferase assay**

The fluorescent substrate SNAP-Vista<sup>®</sup> Green (SVG) was used to assay the catalytic activity of all AGTs analysed, both free proteins and bacterial cells expressing the ASL<sup>tag</sup>/ASL<sup>tag</sup>-POI on cell

surface, as described (Perugino et al., 2012, Perugino et al., 2015; Miggiano et al., 2013; Vettone et al., 2016; Visone et al., 2017; Merlo et al., 2019; Del Prete et al., 2019, Mattossovich et al., 2020). In sum, 5  $\mu\text{M}$  of protein/whole cells were incubated with 10  $\mu\text{M}$  of SVG in Fluo Buffer 1 $\times$  (50 mM phosphate, 100 mM NaCl, 1 mM DTT; pH 6.5) at different temperatures and times, based on the thermophilic or mesophilic nature of each protein; the reactions were stopped by adding Laemmli buffer 1 $\times$  (formamide 95%; EDTA 20 mM; bromophenol 0.05%) and loaded on SDS-PAGE. The resulting gels were analysed by fluorescence imaging on a VersaDoc 4000<sup>TM</sup> system (Bio-Rad), by applying as excitation/emission parameters a blue LED/530 bandpass filter, and by Coomassie Blue staining.

For a quantitative determination of the ASL<sup>tag</sup>/ASL<sup>tag</sup>-POI expression, whole cells were diluted to an OD<sub>600nm</sub> of 1 and, after the reaction with SVG, different volumes of bacterial cells were loaded in the same gel with defined amounts (in the range of 0.8–50 pmols) of purified free H<sup>5</sup> protein. The values of the relative fluorescence were fitted by a linear equation. Whose parameters were successively used for the estimation of the amount of expressed H<sup>5</sup>-POI. Given that the concentration of *E. coli* cell cultures of 1 OD<sub>600nm</sub> is 8  $\times 10^8$  cells/mL, and the amount of wet cells is 1.7 g/L, it has been also possible to calculate the yield of expressed proteins in terms of pmol/mg of cells (Glazyrina et al., 2010; Merlo et al., 2019).

## Competitive assay and IC<sub>50</sub> calculation

Competitive inhibition assay was performed as described (Perugino et al., 2015). As reported, an increasing amount of several guanine-derivatives (0–2 mM) (Table 1.II, chapter II) was added to fixed concentration of the fluorescent SVG (5  $\mu\text{M}$ ) and enzymes (5  $\mu\text{M}$ ). The reactions were incubated for 30 min at 25 °C and 50 °C for SNAP-tag<sup>®</sup> and H<sup>5</sup> respectively and loaded on SDS-PAGE. The resulting fluorescent bands were analysed by gel-imaging and the data were plotted by equation (1),

$$RF = \frac{100\%}{1 + \left(\frac{[I]}{IC_{50}}\right)^{[S]}} \quad (1)$$

where RF is the obtained Relative Fluorescence, [I] and [S] are the concentration of the inhibitor and the substrate, respectively, and the IC<sub>50</sub> is the concentration needed to reduce by 50% the fluorescence intensity of the protein band.

The activity of SNAP-tag<sup>®</sup> and H<sup>5</sup> on BGN3 and BGSN3 substrates was analysed by using the above-mentioned method (Table 1.II, chapter II).

Regarding the characterization of novel hyperthermophilic wild-type AGT, the competitive assay was performed in the presence of double strands (ds) oligonucleotides (Fwdm4:Rev4; Table 1.IV), which contain a single *O*<sup>6</sup>-methyl-guanine, as described (Perugino et al., 2015). The reaction was incubated for 10 min at 65 °C with increasing concentrations (0–10 μM) of ds-Fwdm4, keeping constant SVG concentration (5 μM). Finally, corrected data of fluorescence intensity were fitted with equation (1), as previously described (Perugino et al., 2015; Morrone et al., 2017; Mattosovich et al., 2020).

### **β-glycoside hydrolase assay**

The β-glycoside hydrolase assay was performed at different temperatures in sodium phosphate buffer 50 mM at pH 6.5 with 5 mM final concentration of 2 Np- and 4 Np-Glc substrates, as previously described (Perugino et al., 2003). OM fractions, containing 1 to 5 μg of ASL<sup>tag</sup> and relative fusion POI, were used in each assay. The enzymatic activity was calculated based on the molar extinction coefficient ( $\epsilon_M$ ) values of 2- and 4-nitrophenol in mentioned buffer and reaction conditions, as previously reported (Perugino et al., 2003). By definition, one unit of enzyme activity is the amount of enzyme, which hydrolyses 1 μmol of substrate in 1 min, under the above described conditions. All the reactants except the enzyme (blank mixture), were considered for the correction of the spontaneous hydrolysis of the substrates.

### ***Pfu*OGT: Protein stability analysis**

Differential scan fluorimetry analysis (DSF) was used to determine the stability of the thermostable AGTs at different conditions, adapting the protocol previously described for *Ss*OGT and relative mutants (Niesen et al., 2007; Perugino et al., 2015; Vettone et al., 2016; Morrone et al., 2017). Triplicates of each condition, containing 25 μM of enzyme (ca. 0.5 mg/mL) in PBS 1× buffer and SYPRO Orange dye 1×, were subjected to a scan analysis of 70 cycles at temperatures from 25 to 94 °C for 10 min/°C × cycle in a Real-Time Light Cycler<sup>™</sup> (Bio-Rad). Relative fluorescence data have been normalized to the maximum fluorescence value in each scan. The resulting plots of fluorescence intensity vs temperature displayed sigmoidal curves (typical of a two-state transition), which allowed the determination of the inflection points ( $T_m$  values) by fitting the Boltzmann equation (Niesen et al., 2007; Perugino et al., 2015; Vettone et al., 2016; Morrone et al., 2017).

### **ASL<sup>tag</sup>: Outer Membranes preparation**

The *E. coli* outer (OM) and inner (IM) membranes were purified by following a procedure previously described (Del Prete et al., 2017). Briefly, harvested cells were resuspended 1:20 (w/v) in 25 mM Tris/HCl buffer, pH 8.0 and disrupted by sonication on ice (10 cycles of 10 s / 50 s on:off treatment). Cell extract was centrifuged at 120,000 ×g for 1 h, and the supernatant containing the cytoplasmic fraction was discarded. The pellet containing membranes fractions were resuspended in 20 mL of PBS 1×, containing 0.01 mM MgCl<sub>2</sub> and 2% Triton X-100. After incubation at room temperature for 30 min, the solution was centrifuged as described above. The OM fraction obtained was resuspended in PBS 1× and used for further experiments. Moreover, to analyse of the total amount of fusion protein expressed, the OM fraction was assayed for the H<sup>5</sup> activity as previously described.

### **ASL<sup>tag</sup>: Thrombin assay**

To separate the H<sup>5</sup> moiety from the transmembrane domain, the ASL<sup>tag</sup> on the OM was cleaved with the Thrombin Protease. A suspension of bacterial whole cells was gently centrifuged at 3000 ×g for 10 min at 4 °C and resuspended in PBS 1× buffer. The sample was incubated with 30 U of Thrombin Protease at 25 °C overnight under gentle agitation with 5 μM of SVG, then centrifuged under the above-mentioned conditions. Cells and supernatants were separately loaded on SDS-PAGE and analysed by gel-imaging and Coomassie staining as previously described.

### **ASL<sup>tag</sup>: *in vivo* microscopy analysis of *E. coli* cells**

*E. coli* BL21(DE3) cells transformed with pET-22b/INPN-SspCA or pET-ASL<sup>tag</sup> plasmids were IPTG-induced and grown overnight at 37 °C. After been diluted until 1 OD<sub>600nm</sub>, 1 mL of cell culture was washed twice in PBS 1× and finally resuspended in 50 μL of same buffer supplemented with 5 μM of the SVG. Subsequently, the reaction mixture was incubated at 37 °C for 30 min, washed twice and incubated again for 30 min at 37 °C, to allow the external diffusion of the unreacted substrate. For *in vivo* imaging, the mixture was spotted on poly-L-lysine coated slides for microscopy analysis. Images were collected using a DM6 fluorescence microscope and Hamamatsu camera under the control of Leica LAS AF 6000 software; excitation and emission wavelengths used suitably for AlexaFluor488 dye were  $\lambda_{exc}$  490 nm;  $\lambda_{em}$  525 nm, respectively.

## ***In vitro* Huisgen Cu(I)-catalysed cycloaddition reaction**

The Huisgen chemical reaction was evaluated on SNAP-tag<sup>®</sup> and H<sup>5</sup> previously incubated with BGN3 and BGSN3. An opportune amount of purified proteins was incubated with in an equimolar ratio of these substrates for 60-120 min at 25 °C and 37 °C respectively, to ensure the complete enzymatic labelling reaction. Later, we performed the subsequent cycloaddition using 5 μM of an alkyne-derivative of the fluorescein (BDP FL alkyne), in the presence of copper (1 mM), TCEP (1 mM), TBTA (0.1 mM) and, where indicated, of SDS (0.05 %). Finally, mixtures were loaded on SDS-PAGE and analysed as described in previous sections.

## **Determination of the rate constants of the chemo-enzymatic labelling approach**

Rate constants of the enzymatic reactions with the only BGSN3 were determined by the method of (Gautier et al., 2008). In this case, purified proteins (5 μM) were incubated with the substrate (5 μM) in PBS 1× buffer at 25 °C. Aliquots were taken at different times, the reactions were immediately stopped in Leamli Buffer 1× in addition with 10 μM of Cy5 DBCO fluorophore and placing tubes on ice.

Rate constants for the chemical reaction needed of the preliminary achievement of the *clickable*-SNAP and *clickable*-H<sup>5</sup> with BGSN3, which was obtained by the above-described protocol (Section 5), in order to get the complete labelling. Then, to each aliquot of 5 μM of *clickable* proteins, 20 μM of DBCO-PEG4-Fluor 545 fluorophore was added. At different times, an excess of sodium azide (NaN<sub>3</sub>, 300 mM) was immediately added to each aliquot and then placing tubes on ice, in order to stop the click reaction between the azide group on the BGSN3 and the DBCO-PEG4-Fluor 545 molecule.

Finally, for both the experiments, all aliquots were boiled in SDS buffer for 5 min, and immediately loaded on SDS-PAGE, for the gel-imaging and Coomassie staining analyses. Data were fitted to a pseudo-first-order reaction model using the GraFit 5.0 software package (Erithacus Software Ltd.). Second-order rate constants  $k$  (in s<sup>-1</sup> M<sup>-1</sup>) were then obtained by dividing the pseudo-first-order constant by the concentration of substrate. Values given are an average of at least three independent measurements.

## ***In vitro* Huisgen copper-free cycloaddition reaction with different DBCO-fluorophores**

For the SPAAC reaction, aliquots of 5  $\mu$ M of each *clickable*-protein were incubated for 60 min at room temperature in the dark with 5  $\mu$ M of fluorescent DBCO-derivative substrates (BDP FL DBCO, Cy5 DBCO and DBCO-PEG4-Fluor 545) in a total volume of 10  $\mu$ L of PBS 1 $\times$  buffer. The reactions were finally stopped in Leammi Buffer 3 $\times$ , loaded on SDS-PAGE and analysed as described in Section 4, by applying a blue LED/530 bandpass filter, red LED/695 bandpass filter and green LED/605 bandpass filter as excitation/emission parameters for each DBCO-fluorophores, respectively. The click reaction was also performed on 5  $\mu$ M of both the enzymes, but in the presence of an *EcCFE* diluted in PBS 1 $\times$  buffer.

## **Procedure for protein immobilization on Bio Layer Interferometry (BLI)**

OctetRED96<sup>TM</sup> (ForteBio, CA, USA) was used to immobilize specifically SNAP-tag<sup>®</sup> and H<sup>5</sup> with the chemo-enzymatic approach. Samples and reaction buffers were located in black 96-well plates (OptiPlate-96 Black, Black Opaque 96-well Microplate, PerkinElmer, USA) in a maximum reaction volume of 300  $\mu$ L per well with 800 rpm shaking for each step. For the immobilization procedure, AR2G sensors were first wetted in 200  $\mu$ L of pure water for at least 15 min, followed by an equilibration step (3 min) in acetate buffer 0.1 M, pH 5.0. Afterwards, they were activated with 20 mM 1-ethyl-3-(3-dimethyl-amino-propyl) carbodiimide hydrochloride (EDC)/20 mM N-hydroxy-sulfo-succinimide (sulfo-NHS) mixture in acetate buffer (60 min) and covered with 2 mM propargyl-PEG3-amine bifunctional linker (BroadPharm, San Diego, CA, USA) in Loading step (20 min). To avoid the presence of any free amine groups on the biosensors, a Blocking step with Ethanolamine 1 M (30 min) was performed. Subsequently, a Washing step (15 min) with water and an Equilibration step in click-reaction buffer (15 min) are followed.

During the above-described procedure, proteins were labelled with BGSN3. Finally, the immobilization step for each sample via Huisgen reaction was carried out at 30  $^{\circ}$ C for 80 min, followed by a Washing step (20 min), in order to remove all the unbound molecules. This procedure was the same in the presence of the *EcCFE*. All measurements were performed in triplicates.

## **Permeability of eukaryotic and prokaryotic cells to BGSN3 and DBCO-fluorophores**

HEK293T cells were maintained at 37 °C with 5% CO<sub>2</sub> in Dulbecco's Modified Essential Medium (Invitrogen, USA) supplemented with 10% Fetal Bovine Serum (FBS) (Invitrogen, USA) and 100 U/mL Penicillin/Streptomycin (Roche, Switzerland). HEK293T cells were transfected with SNAP-tag<sup>®</sup> plasmid by using Lipofectamine 2000 (Invitrogen, USA) following manufacturer's protocol. The treatment with BGSN3 were performed, at the concentration and time indicated for each experiment. Twenty-four hours after transfection, we treated cells with BGSN3 for 2 hours at different concentrations ranging (from 1 to 25 µM), directly dissolving the compound in complete culture medium. Then cells were harvested, washed with PBS 1× buffer and lysed with 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5 mM EDTA, 0.1% Triton X-100 supplemented with complete protease (Roche, Switzerland) and phosphatase (SERVA Electrophoresis, Germany) inhibitors. Afterwards, transfected cells were treated with a fixed concentration of BGSN3 (10 µM) at different time points (from 30 to 120 min). Again, HEK293T cells were washed and lysed as described above. To confirm the reaction with BGSN3, the same amount of protein extract (0,91 µg/µL for each sample) was incubated for 30 min at 25 °C with SVG. Subsequently, proteins were loaded on SDS-PAGE and analysed by gel-imaging on a VersaDoc 4000<sup>™</sup> system (Bio-Rad), by applying a blue LED/530 bandpass filter (Figure 4 in the main text).

For flow cytometry analysis, HeLa cells were seeded in 24-well plates and transfected with SNAP-tag<sup>®</sup> plasmid by using Lipofectamine 2000 (Invitrogen, USA) following manufacturer's protocol. Twenty-four hours after the transfection, cells were treated with 25 µM BGSN3 for one hour and the excess of the substrate was washed out by 2 x 15 min, followed by 1 x 30 min washes. Cells were then treated with 2.5 µM BDP FL DBCO for 30 min and unbound fluorophore was removed by following the same procedure performed for the BGSN3. All treatments and washes were performed at 37 °C in complete culture medium. Lastly, cells were harvested by trypsinization and fluorescence was measured using FACS CANTO II instrument. Analysis was performed on live singlet cells using FlowJo software (Fig S11).

*E. coli* ABLE C strain was transformed with SNAP-tag<sup>®</sup> plasmid and protein expressed as previously described. After an overnight growth, samples of 2 mL were treated with 100 µM of BGSN3 for 2 hours at 25 °C and then collected by centrifugation at 2000 ×g. Cell pellets of 0,05 g were resuspended 1:3 (w/v) in PBS 1× supplemented with 1% Triton X-100 and subjected to cell lysis, by applying 5 cycles of freeze-thawing. After a centrifugation at 13000 ×g, the supernatants containing the protein extract were incubated 30 min at 25 °C with SVG and proteins were loaded on SDS-PAGE. Finally, fluorescent bands were analysed by gel-imaging techniques (Figure S12).

## **Images were collected using a DM6 fluorescence**

microscope and Hamamatsu camera under the control of Leica LAS AF 6000 software; excitation and emission wavelengths used suitably for AlexaFluor488 dye were  $\lambda_{exc}$  490 nm;  $\lambda_{em}$  525 nm, respectively.

## **Data analysis and software**

Prism Software Package (GraphPad Software) and GraFit 5.0 Data Analysis Software (Erithacus Software) were used for the data analysis from activity, competitive inhibition and stability assays.

## References

- Aliye N.**, Fabbretti A., Lupidi G., Tsekoa T., Spurio R. Engineering color variants of green fluorescent protein (GFP) for thermostability, pH-sensitivity, and improved folding kinetics. *Appl Microbiol Biotechnol* 2015; 99:1205–1216
- Alterio V.**, Langella E., De Simone G., Monti SM. Cadmium-containing carbonic anhydrase CDCA1 in marine diatom *Thalassiosira weissflogii*. *Mar Drugs* 2015; 13:1688–97.
- Aguilar C.**, Sanderson I., Moracci M., et al. Crystal structure of the beta-glycosidase from the hyperthermophilic archeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability. *J Mol Biol* 1997; 271:789–802.
- Akdemir A.**, Vullo D., De Luca V., et al. The extreme-alpha-carbonic anhydrase (CA) from *Sulfurihydrogenibium azorense*, the fastest CA known, is highly activated by amino acids and amines. *Bioorg Med Chem Lett* 2013; 23:1087–90.
- Annunziato G.**, Angeli A., D’Alba F., et al. Discovery of new potential anti-infective compounds based on carbonic anhydrase inhibitors by rational target-focused repurposing approaches. *Chem Med Chem* 2016; 11:1904–14.
- Ashby M.C.**, Ibaraki K., Henley J.M. It’s green outside: tracking cell surface proteins with pH-sensitive GFP. *Trends Neurosci* 2004; 27:257–261
- Bertani G.** Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* 1951; 62:293–300.
- Bertani G.** Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *J. Bacteriol.* 2004; 186:595–600.
- Binder U.**, Matschiner G., Theobald I. and Skerra. J.A. High-throughput Sorting of an Anticalin Library via EspP-mediated Functional Display on the *Escherichia coli* Cell Surface. *Mol. Biol.* 2010; 400, 783–802
- Brunk, E.**, Mollwitz, B., Rothlisberger, U. Mechanism to Trigger Unfolding in O6-Alkylguanine-DNA Alkyltransferase. *ChemBioChem* 2013; 14, 703–710.
- Campbell T.N.**, Choy F.Y.M The effect of pH on green fluorescent protein: a brief review. *Mol Biol Today* 2000; 2:1–4
- Canganella F.**, Wiegel J. Extremophiles: from abyssal to terrestrial ecosystems and possibly beyond. *Naturwissenschaften.* 2011; 98:253-79.
- Capasso C.**, De Luca V., Carginale V., et al. Biochemical properties of a novel and highly thermostable bacterial alphacarboxylic anhydrase from *Sulfurihydrogenibium yellowstonense* YO3AOP1. *J Enzyme Inhib Med Chem* 2012; 27:892–7.
- Capasso C.**, Supuran C.T. Bacterial, fungal and protozoan carbonic anhydrases as drug targets. *Expert Opin Ther Targets* 2015; 19:1689–704.

**Chalfie M.**, Tu Y., Euskirchen G. et al. Green fluorescent protein as a marker for gene expression. *Science* 1994; 263, 802–805.

**Chudakov D.M.**, Matz M.V., Lukyanov S., Lukyanov K.A. Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol Rev* 2010; 90:1103–63.

**Cochet N.**, Widehem P. Ice crystallization by *Pseudomonas syringae*. *Appl Microbiol Biotechnol* 2000; 54:153–61.

**Colombo M.**, Mazzucchelli S., Montenegro J.M., et al. Protein oriented ligation on nanoparticles exploiting O<sup>6</sup>-alkylguanine-DNA transferase (SNAP) genetically encoded fusion. *Small* 2012; 8:1492–7.

**Daniels D.S.**, Mol C.D., Arvai A.S. et al. Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. *EMBO J* 2000; 19:1719–1730

**Datta S.**, Christena L.R. and Rajaram Y.R.S. Enzyme immobilization: an overview on techniques and support materials. *3 Biotech* 2013; 3, 1–9.

**Daugherty PS.** Protein engineering with bacterial display. *Curr Opin Struct Biol* 2007; 17:474–80.

**Dautin N.** and Bernstein H.D. Protein secretion in Gram-negative bacteria via the autotransporter pathway. *Annu. Rev. Microbiol.* 2007; 61, 89–112.

**De Luca V.**, Vullo D, Scozzafava A, et al. Anion inhibition studies of an alpha-carbonic anhydrase from the thermophilic bacterium *Sulfurihydrogenibium yellowstonense* YO3AOP1. *Bioorg Med Chem Lett* 2012; 22:5630–4.

**De Simone G.**, Monti S.M., Alterio V., et al. Crystal structure of the most catalytically effective carbonic anhydrase enzyme known, SazCA from the thermophilic bacterium *Sulfurihydrogenibium azorense*. *Bioorg Med Chem Lett* 2015; 25:2002–6.

**Del Prete S.**, Perfetto R., Rossi M. et al. A one-step procedure for immobilising the thermostable carbonic anhydrase (SspCA) on the surface membrane of *Escherichia coli*. *J. Enz. Inhib. Med. Chem.* 2017; 32: 1120-1128.

**Del Prete S.**, Merlo R., Valenti A. et al. Thermostability enhancement of the  $\alpha$ -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* by using the anchoring-and-self labelling-protein-tag system (ASLtag). *J. Enz. Inhib. Med. Chem.* 2019; 34: 946-954

**Debets M.F.**, van Berkel S.S., Schoffelen S. et al. Aza-dibenzocyclooctynes for fast and efficient enzyme PEGylation via copper-free (3+2) cycloaddition. *Chem. Commun.* 2010; 46: 97-9.

**Debets M.F.**, van Berkel S. S., Dommerholt J. et al. Bioconjugation with strained alkenes and alkynes. *Acc Chem Res.* 2011; 44, 805–815.

**DiCosimo R.**, McAuliffe J, Poulouse AJ and Bohlmann G. Industrial use of immobilized enzymes. *Chem Soc Rev.* 2013;42:6437-74.

- Di Fiore A.**, Capasso C., De Luca V. et al. X-ray structure of the first “extremo-alpha-carbonic anhydrase”, a dimeric enzyme from the thermophilic bacterium *Sulfurihydrogenibium yellowstonense* YO3AOP1. *Acta Crystallogr D Biol Crystallogr* 2013; 69:1150–9.
- Di Fiore A.**, Alterio V., Monti S.M. et al. Thermostable carbonic anhydrases in biotechnological applications. *Int J Mol Sci* 2015; 16:15456–80.
- Duguid E.M.**, Rice P.A., He C. The structure of the human AGT protein bound to DNA and its implications for damage detection. *J. Mol. Biol.* 2005; 350, 657–666.
- Elder R.H.**, Margison G.P., Rafferty J.A. Differential inactivation of mammalian and *Escherichia coli* O6-alkylguanine-DNA alkyltransferases by O6-benzylguanine. *Biochem. J.* 1994; 298: 231-235.
- England C.G.**, Luo H., Cai W. HaloTag technology: a versatile platform for biomedical applications. *Bioconjug Chem* 2015; 26:975–86.
- Ess D.H.**, Jones G.O., Houk K.N. Transition states of strain-promoted metal-free click chemistry: 1,3-dipolar cycloadditions of phenyl azide and cyclooctynes. *Org Lett.* 2008; 10:1633-6.
- Eunha K.** and Heebeom K. Biomedical applications of copper-free click chemistry: *in vitro*, *in vivo*, and *ex vivo*. *Chem. Sci.* 2019; 10, 7835-7851.
- Ezraty B.**, Henry C., Hérisse M., Denamur E., Barras F. Commercial Lysogeny Broth culture media and oxidative stress: a cautious tale. *Free Radic Biol Med.* 2014; 74:245-51.
- Fan C.H.**, Liu W.L., Cao H., Wen C. Chen L., Jiang G. O<sup>6</sup>-methylguanine DNA methyltransferase as a promising target for the treatment of temozolomide-resistant gliomas. *Cell Death Dis.* 2013; 4(10):e876.
- Farag N.**, Mattosovich R., Merlo R. et al. Folding-upon-repair DNA nanoswitches for monitoring DNA repair enzymes activity. *Angew Chem Int Ed Engl.* 2021; 60:7283-7289.
- Ferraris D.M.**, Spallek R., Oehlmann W. et al. Structure of *Thermococcus litoralis* D1-pyrroline-2-carboxylate reductase in complex with NADH and L-proline. *Proteins* 2014; 82:2268–74.
- Fiala G.**, Stetter K.O. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaeobacteria growing optimally at 100 °C. *Arch Microb* 1986; 145:56–61
- Friedberg E.C.** DNA damage and repair. *Nature* 2003; 421, 436-440.
- Gautier A.**, Juillerat A., Heinis C. et al. An engineered *protein tag* for multiprotein labeling in living cells. *Chem Biol* 2008; 15:128–136.
- Gerson S.L.** Clinical Relevance of MGMT in the Treatment of Cancer. *J Clin Oncol.* 2002; 20:2388-99.
- Gerson S.L.** MGMT: Its role in cancer aetiology and cancer therapeutics. *Nat. Rev. Cancer.* 2004; 4, 296–307.
- Glazyrina J.**, Materne EM, Dreher T, et al. High cell density cultivation and recombinant protein production with *Escherichia coli* in a rocking-motion-type bioreactor. *Microb Cell Fact* 2010; 9:42–52.

**Goodtzova K.**, Kanugula S., Edara S. et al. Repair of O6-benzylguanine by the Escherichia coli Ada and Ogt and the human O6-alkylguanine-DNA alkyltransferase. *J. Biol. Chem.* 1997; 272: 8332-8339.

**Graether S.P.** and Jia Z. Modeling *Pseudomonas syringae* ice-nucleation protein as a beta-helical protein. *Biophys J* 2001; 80:1169–73.

**Gronemeyer T.**, Chidley C., Juillerat A. et al. Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Eng. Des. Sel.* 2006; 19, 309–316.

**Gurian-Sherman D.** and Lindow S.E. Bacterial ice nucleation: Significance and molecular basis. *FASEB J.* 1993; 7: 1338-1343.

**Gurung N.**, Ray S., Bose S., and Rai V. A broader view: microbial enzymes and their relevance in industries, medicine, and beyond. *Biomed Res Int.* 2013; 2013:329121.

**Hale C.R.**, Zhao P., Olson S. et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 2009; 139, 945–956.

**Hashimoto H.**, Inouel T., Nishioka M. et al. Hyperthermostable protein structure maintained by intra and inter-helix ion-pairs in archaeal O6-methylguanine-DNA methyltransferase. *J Mol Biol* 1999; 292:707–716

**Henderson I.R.**, Navarro-Garcia F. and Nataro J.P. The great escape: structure and function of the autotransporter proteins. *Trends Microbiol.* 1998; 6, 370–378.

**Henderson I.R.** and Nataro J.P. Virulence functions of autotransporter proteins. *Infect. Immun.* 2001; 69, 1231–1243.

**Hiep Nguyen H.**, Kim, M. An Overview of Techniques in Enzyme Immobilization. *Appl. Sci. Conver. Technol* 2017; 26, 157–163.

**Hinner M.J.** and Johnsson K. How to obtain labeled proteins and what to do with them *Curr Opin Biotechnol.* 2010; 21:766-76.

**Hoelzel C.A.** and Zhang X., Visualizing and Manipulating Biological Processes by Using HaloTag and SNAP-Tag Technologies *ChemBioChem* 2020; 21, 1–13

**Homaei A.** Enzyme immobilization and its application in the food industry *Adv. Food Biotechnol.* 2015, 10.1002/9781118864463.ch09

**Hong V.**, Presolski S.I., Ma C., Finn M.G. Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew Chem Int Ed Engl.* 2009; 48:9879-83.

**Huber W.**, Perspicace S., Kohler J., et al. SPR-based interaction studies with small molecular weight ligands using hAGT fusion proteins. *Anal Biochem* 2004; 333:280–8.

**Jewett J.C.** and Bertozzi C.R. Cu-free click cycloaddition reactions in chemical biology. *Chem. Soc. Rev.* 2010; 39:1272.

**Juillerat A.**, Gronemeyer T., Keppler A. et al. Directed evolution of O<sup>6</sup>-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules *in vivo*. *Chem Biol* 2003; 10:313–7.

**Kaina B.**, Margison G.P., Christmann M. Targeting O<sup>6</sup>-methylguanine-DNA methyltransferase with specific inhibitors as a strategy in cancer therapy. *Cell Mol Life Sci.* 2010; 67:3663-81.

**Kaina B.**; Christmann M. DNA repair in personalized brain cancer therapy with temozolomide and nitrosoureas. *DNA Repair* 2019; 78: 128-141.

**Katchalski-Katzir E.** Immobilized enzymes--learning from past successes and failures. *Trends Biotechnol.* 1993; 11:471-8.

**Kawaguti H.Y.**, Manrich E., Sato H.H. Production of isomaltulose using *Erwinia* sp. D12 cells: culture medium optimization and cell immobilization in alginate. *Biochem Eng J* 2006; 29:270–277

**Kengen S.W.M.** 'Pyrococcus furiosus, 30 years on'. *Microb Biotechnol.* 2017;10:1441-1444.

**Keppler A.**, Gendreizig S., Gronemeyer T. et al. A general method for the covalent labeling of fusion proteins with small molecules *in vivo*. *Nat Biotechnol* 2003; 21:86–9.

**Keppler A.**, Pick H., Arrivoli C. et al. Labeling of fusion proteins with synthetic fluorophores in live cells. *Proc Natl Acad Sci USA* 2004; 10:9955–9.

**Kim E.**, Koo H. Biomedical applications of copper-free click chemistry: *in vitro*, *in vivo*, and *ex vivo*. *Chem Sci.* 2019; 10:7835-7851.

**Kindermann M.**, George N., Johnsson N., Johnsson, K. Covalent and selective immobilization of fusion proteins. *J. Am. Chem. Soc.* 2003; 125, 7810–7811.

**Klein S.**, Oesch F. Assay for O<sup>6</sup>-alkylguanine-DNA-alkyltransferase using oligonucleotides containing O<sup>6</sup>-methylguanine in a BamHI recognition site as substrate. *Anal. Biochem.* 1992; 205, 294–299.

**Kolb H.C.**, Finn M.G., Sharpless K.B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew Chem Int Ed Engl.* 2001; 40:2004-2021.

**Kosaka N.**, Ogawa M., Choyke P.L. et al. *In vivo* stable tumor-specific painting in various colors using dehalogenase-based protein-tag fluorescent ligands. *Bioconjug Chem* 2009; 20:1367-1374

**Kuzmin A.**, Poloukhtin A., Margreet A. et al. Surface Functionalization Using Catalyst-Free Azide-Alkyne Cycloaddition. *Bioconjugate Chem.* 2010; 21: 2076.

**Lahiri, S.**, Rizzi M., Rossi F., Miggianno R. *Mycobacterium tuberculosis* UvrB forms dimers in solution and interacts with UvrA in the absence of ligands. *Proteins* 2018; 86, 98–109.

**Latypov V.F.**, Tubbs J.L., Watson, A.J. et al. AtI1 regulates choice between global genome and transcription-coupled repair of O(6)-alkylguanines. *Mol Cell.* 2012; 47, 50–60.

- Leclere M.M.**, Nishioka M., Yuasa, T. et al. The O<sup>6</sup>-methylguanine-DNA methyltransferase from the hyperthermophilic archaeon *Pyrococcus* sp. KOD1: A thermostable repair enzyme. *Mol Gen Genet* 1998; 258, 69–77.
- Lee S.Y.**, Choi J.H., Xu Z. Microbial cell-surface display. *Trends Biotechnol.* 2003; 21:45-52.
- Leonard G.A.**, Thomson J., Watson W.P., Brown T. High-resolution structure of a mutagenic lesion in DNA. *Proc. Natl. Acad. Sci. USA* 1990; 87, 9573–9576.
- Lindahl T.** Instability and decay of the primary structure of DNA. *Nature* 1993; 362, 709-715.
- Liss V.**, Barlag B., Nietschke M., Hensel M. Self-labelling enzymes as universal tags for fluorescence microscopy, super-resolution microscopy and electron microscopy. *Sci Rep* 2015; 5:17740–52.
- Liu L.**, Gerson S.L. Targeted modulation of MGMT: Clinical implications. *Clin. Cancer Res.* 2006; 12, 328–331.
- Lo Gullo G.**, Mattosovich R., Perugino G. et al. Optimization of an *in vitro* transcription/translation system based on *Sulfolobus solfataricus* cell lysate. *Archaea* 2019; 2019:9848253.
- Los G.V.**, Encell L.P., McDougall M.G. et al. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* 2008; 3:373-382
- Lundberg K.S.**, Shoemaker D.D., Adams M.W.W., Short J.M., Sorge J.A., Mathur E.J. High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* 1991; 108: 1-6.
- Mattosovich R.**, Merlo R., Miggiano R. et al. O<sup>6</sup>-alkylguanine-DNA Alkyltransferases in Microbes Living on the Edge: From Stability to Applicability. *Int. J. Mol. Sci.* 2020; 21: 2878.
- Mattosovich R.**, Merlo R., Fontana A. et al. A journey down to hell: new thermostable protein-tags for biotechnology at high temperatures. *Extremophiles* 2020; 24: 81-91.
- Merlo R.**, Del Prete S., Valenti A. et al. An AGT-based *protein tag* system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane. *J Enzyme Inhib Med Chem* 2019; 34:490–9.
- Merlo R.**, Caprioglio D., Cillo M. et al. The SNAP-tag technology revised: an effective *chemo-enzymatic* approach by using a universal azide-based substrate. *J. Enz. Inhib. Med. Chem.* 2020; 36: 85-97.
- Miggiano R.**, Casazza, V., Garavaglia, S. et al. Biochemical and structural studies of the *Mycobacterium tuberculosis* O<sup>6</sup>-methylguanine methyltransferase and mutated variants. *J. Bacteriol.* 2013; 195, 2728–2736.
- Miggiano R.**, Valenti A., Rossi F. et al. Every OGT is illuminated...by fluorescent and synchrotron lights. *Int J Mol Sci* 2017; 18:2613–31.

- Mohamad** N.R., Marzuki N.H., Buang N.A. et al. An overview of technologies for immobilization of enzymes and surface analysis techniques for immobilized enzymes. *Biotechnol Biotechnol Equip.* 2015; 29:205-220.
- Mollwitz** B., Brunk E., Schmitt S. et al. Directed evolution of the suicide protein O6-alkylguanine-DNA alkyltransferase for Increased reactivity results in an alkylated protein with exceptional stability. *Biochemistry* 2012; 51:986–94.
- Morita**, R., Nakagawa, N., Kuramitsu, S., Masui, R. An O6-methylguanine-DNA methyltransferase-like protein from *Thermus thermophilus* interacts with a nucleotide excision repair protein. *J. Biochem.* 2008; 144, 267–277.
- Morrone** C., Miggiano R., Serpe M. et al. Interdomain interactions rearrangements control the reaction steps of a thermostable DNA alkyltransferase. *Biochim Biophys Acta Gen Subj* 2017; 1861:86–96.
- Moracci** M., Capalbo L., Ciaramella M., Rossi M. Identification of two glutamic acid residues essential for catalysis in the beta-glycosidase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Protein Eng* 1996; 9:1191–5.
- Moracci** M., Trincone A., Perugino G. et al. Restoration of the activity of active-site mutants of the hyperthermophilic-glycosidase from *Sulfolobus solfataricus*: dependence of the mechanism on the action of external nucleophiles. *Biochemistry* 1998; 37:17262–70.
- Moschel** R.C., McDougall M.G., Dolan M.E., et al. Structural features of substituted purine derivatives compatible with depletion of human O6-alkylguanine-DNA alkyltransferase. *J Med Chem* 1992; 35:4486–91.
- Niesen** F.H., Berglund H., Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* 2007; 2:2212–2221
- Niesen** J., Sack M., Seidel M. et al. SNAP-Tag technology: a useful tool to determine affinity constants and other functional parameters of novel antibody fragments. *Bioconjug Chem* 2016; 27:1931–41.
- Nishikori** S., Shiraki K., Fujiwara S. et al. Unfolding mechanism of a hyperthermophilic protein O6-methylguanine-DNA methyltransferase. *Biophys Chem* 2005; 116:97–104
- Novick**, S.J., Rozzell, J.D. Immobilization of enzymes by covalent attachment. *Microbial Enzymes and Biotransformations* 2005; pp. 247–272.
- Olsson** M., Lindahl T. Repair of alkylated DNA in *Escherichia coli*. Methyl group transfer from O6-methylguanine to a protein cysteine residue. *J. Biol. Chem.* 1980; 255, 10569–10571.
- Ozensoy Guler** O., Capasso C., Supuran C.T. A magnificent enzyme superfamily: carbonic anhydrases, their purification and characterization. *J Enzyme Inhib Med Chem* 2016;31: 689–94.
- Paranjpe** A., Zhang R., Ali-Osman F. et al. Disulfiram is a direct and potent inhibitor of human O6-methylguanine-DNA methyltransferase (MGMT) in brain tumor cells and mouse brain and markedly increases the alkylating DNA damage. *Carcinogenesis* 2014; 35, 692–702.

**Pegg** A.E., Wiest L., Mummert C. and Dolan M.E. Production of antibodies to peptide sequences present in human O<sup>6</sup>-alkylguanine-DNA-alkyltransferase and their use to detect this protein in cell extracts. *Carcinogenesis* 1991; 12, 1671-7.

**Pegg** A.E. Repair of O(6)-alkylguanine by alkyltransferases. *Mutat. Res.* 2000; 462: 83-100.

**Pegg**, A.E. Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA damage, resistance to chemotherapy, and research tools. *Chem. Res. Toxicol.* 2011; 24, 618–639.

**Perugino** G., Trincone A, Giordano A. et al. Activity of hyper-thermophilic glycosynthases is significantly enhanced at acidic pH. *Biochemistry* 2003; 42:8484–93.

**Perugino** G., Valenti A., D’Amaro A., Rossi M., Ciaramella M. Reverse gyrase and genome stability in hyperthermophilic organisms. *Biochem. Soc. Trans.* 2009; 37, 69–73.

**Perugino** G., Vettone A., Illiano G. et al. Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage. *J Biol Chem* 2012; 287:4222–31.

**Perugino** G., Miggianno R., Serpe M. et al. Structure-function relationships governing activity and stability of a DNA alkylation damage repair thermostable protein. *Nucleic Acids Res* 2015; 43:8801–16.

**Pouwels** J., Moracci M., Cobucci-Ponzano B. et al. Activity and stability of hyperthermophilic enzymes: a comparative study on two archaeal-glycosidases. *Extremophiles* 2000; 4:157–64.

**Presolski** S.I., Hong V.P. and Finn M.G. Copper-Catalyzed Azide-Alkyne Click Chemistry for Bioconjugation. *Curr Protoc Chem Biol.* 2011; 3:153.

**Rabik** C.A., Njoku M.C., Dolan ME Inactivation of O<sup>6</sup>-alkylguanine DNA alkyltransferase as a means to enhance chemotherapy. *Cancer Treat Rev* 2006; 32:261–276

**Ranson** M., Middleton M.R., Bridgewater J. et al. Lomeguatrib, a potent inhibitor of O<sup>6</sup>-alkylguanine-DNA-alkyltransferase: phase I safety, pharmacodynamic, and pharmacokinetic trial and evaluation in combination with temozolomide in patients with advanced solid tumors. *Clin Cancer Res* 2006; 12:1577–84.

**Rossi** F., Morrone C., Massarotti A. et al. Crystal structure of a thermophilic O<sup>6</sup>-alkylguanine-DNA alkyltransferase-derived self-labeling *protein-tag* in covalent complex with a fluorescent probe. *Biochem. Biophys. Res. Comm.* 2018; 500, 698–703.

**Rothschild** L.J., Mancinelli R.L. Life in extreme environments. *Nature* 2001; 409:1092-101.

**Russo** M.E., Olivieri G., Capasso C. et al. Kinetic study of a novel thermo-stable alpha-carbonic anhydrase for biomimetic CO<sub>2</sub> capture. *Enzyme Microb Technol* 2013; 53:271–7.

**Samuelson** P., Gunneriusson E., Nygren P.A., Ståhl S. Display of proteins on bacteria. *J Biotechnol* 2002; 96:129–54.

**Sassolas** A., Blum L.J., Leca-Bouvier BD. Immobilization strategies to develop enzymatic biosensors. *Biotechnol Adv.* 2012; 30:489-511.

**Sharma S.**, Salehi F., Scheithauert B.W. et al., Role of MGMT in Tumor Development, Progression, Diagnosis, Treatment and Prognosis. *Anticancer Research* 2009; 29, 3759-3768.

**Singer B.** All oxygens in nucleic acids react with carcinogenic ethylating agents. *Nature* 1976; 264, 333-9.

**Skorvaga M.**, Raven N.D., Margison G.P. Thermostable archaeal O<sup>6</sup>-alkylguanine-DNA alkyltransferases. *Proc. Natl. Acad. Sci. USA* 1998; 95, 6711–6715.

**Sletten E.M.**, Bertozzi C.R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angew. Chem. Int. Ed.* 2009; 48:6998.

**Song X.**, Wang C., Han Z., et al. Terminal alkyne substituted O<sup>6</sup>-benzylguanine for versatile and effective syntheses of fluorescent labels to genetically encoded SNAP-tags. *RSC Adv* 2015; 5:23646–9.

**Spahn C.** and Minter S.D. Enzyme immobilization in biotechnology. *Recent Pat Eng* 2008; 2:195–200

**Srivenugopal K.S.**, Yuan X.H., Friedman H.S., Ali-Osman F. Ubiquitination-dependent proteolysis of O<sup>6</sup>-methylguanine-DNA methyltransferase in human and murine tumor cells following inactivation with O<sup>6</sup>-benzylguanine or 1,3-bis-(2-chloroethyl)-1-nitrosourea. *Biochemistry* 1996; 35, 1328-1334.

**Su Y.**, Ge J., Zhu B. et al. Target identification of biologically active small molecules via in situ methods. *Curr Opin Chem Biol.* 2013; 17:768.

**Sun G.**, Fan T., Zhang N. et al. Identification of the structural features of guanine derivatives as MGMT inhibitors using 3D-QSAR modeling combined with molecular docking. *Molecules* 2016; 21:823–44.

**Terashima I.**, Kawate H., Sakumi K. et al. Substrate specificity of human O<sup>6</sup>-methylguanine-DNA methyltransferase for O<sup>6</sup>-benzylguanine derivatives in oligodeoxynucleotides. *Chem Res Toxicol* 1997; 10:1234–9.

**Terns, R.M.**, Terns, M.P. The RNA- and DNA-targeting CRISPR-Cas immune systems of *Pyrococcus furiosus*. *Biochem. Soc. Trans.* 2013; 41, 1416–1421.

**Testa I.**, Wurm C.A., Medda R. et al. Multicolor fluorescence nanoscopy in fixed and living cells by exciting conventional fluorophores with a single wavelength. *Biophys J* 2010; 99: 2686–94.

**Tieleman D.P.** and Berendsen H.J. A molecular dynamics study of the pores formed by *Escherichia coli* OmpF porin in a fully hydrated palmitoylcholine bilayer. *Biophys J* 1998; 74:2786–801.

**Trevan M.** Techniques of Immobilization *Immobilized Enzymes: an introduction and applications in biotechnology* 1980; Wiley, Chichester, New York, pp 1-9

**Tsien, R.Y.** The green fluorescent protein. *Annu. Rev. Biochem.* 1998; 67, 509–554.

**Tubbs** J.L., Pegg A.E., Tainer J.A. DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair* 2007; 6:1100–15.

**Valenti** A., Napoli A., Ferrara M.C. et al. Selective degradation of reverse gyrase and DNA fragmentation induced by alkylating agent in the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res.* 2006; 34, 2098–2108.

**Valenti** A., Perugino G., D’Amaro A. et al. Dissection of reverse gyrase activities: Insight into the evolution of a thermostable molecular machine. *Nucleic Acids Res.* 2008; 36, 4587–4597.

**Valenti** A., Perugino G., Nohmi T. et al. Inhibition of translesion DNA polymerase by archaeal reverse gyrase. *Nucleic Acids Res.* 2009; 37, 4287–4295.

**Vettone** A., Serpe M., Hidalgo A. et al. A novel thermostable *protein-tag*: optimization of the *Sulfolobus solfataricus* DNAalkyl-transferase by protein engineering. *Extremophiles* 2016; 20:1–13.

**Visone** V., Han W., Perugino G. et al. *In vivo* and *in vitro* protein imaging in thermophilic archaea by exploiting a novel *protein tag*. *PloS One* 2017;12:e0185791.

**Vora** R.A., Pegg A.E. and Ealick S.E. A new model for how O6-methylguanine-DNA methyltransferase binds DNA. *Proteins* 1998; 32, 3-6.

**Vullo** D., De Luca V., Scozzafava A. et al. The first activation study of a bacterial carbonic anhydrase (CA). The thermostable alpha-CA from *Sulfurihydrogenibium yellowstonense* YO3AOP1 is highly activated by amino acids and amines. *Bioorg Med Chem Lett* 2012; 22:6324–7.

**Vullo** D., Luca V.D., Scozzafava A. et al. The alpha-carbonic anhydrase from the thermophilic bacterium *Sulfurihydrogenibium yellowstonense* YO3AOP1 is highly susceptible to inhibition by sulfonamides. *Bioorg Med Chem* 2013; 21:1534–8.

**Wibley** J.E.A., Pegg A.E., Moody P.C.E. Crystal structure of the human O6-alkylguanine-DNA alkyltransferase. *Nucleic Acids Res.* 2000; 28, 393-401.

**Woese** C.R., Kandler O., Wheelis M.L. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A.* 1990; 87:4576-9.

**Wu** R.S., Hurst-Calderone S., Kohn, K.W. Measurement of O6-alkylguanine-DNA-alkyltransferase activity in human cells and tumor tissues by restriction endonuclease inhibition. *Cancer Res.* 1987; 47, 6229–6235.

**Xie** R., Hong S., Chen X. Cell-selective metabolic labeling of biomolecules with bioorthogonal functionalities. *Curr Opin Chem Biol.* 2013; 17:747.

**Xu-Welliver** M. and Pegg A.E. Degradation of the alkylated form of the DNA repair protein, O6-alkylguanine-DNA alkyltransferase. *Carcinogenesis* 2002; 23, 823–830.

**Yao** J.Z., Uttamapinant C., Poloukhine A. et al. Fluorophore Targeting to Cellular Proteins via Enzyme-Mediated Azide Ligation and Strain-Promoted Cycloaddition. *J. Am. Chem. Soc.* 2012; 134: 3720.

**Zeng D.**, Zeglis B.M., Lewis J.S., Anderson C.J. The Growing Impact of Bioorthogonal Click Chemistry on the Development of Radiopharmaceuticals. *J Nucl Med* 2013; 54:829.

**Zhang C.J.**, Li L., Chen G.Y.J. et al. One- and two-photon live cell imaging using a mutant SNAP-tag protein and its FRET substrate pairs. *Org Lett* 2011; 13:4160–3.

**Zhang Y.**, Ge J. and Liu Z. Enhanced Activity of Immobilized or Chemically Modified Enzymes. *ACS Catal.* 2015; 5, 4503–4513

**Zhou Z.**, Hartmann M. Progress in enzyme immobilization in ordered mesoporous materials and related applications. *Chem Soc Rev* 2013; 42:3894–912.

**Zweerink S.**, Kallnik V., Ninck S. et al. Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea. *Nat Commun* 2017; 8:15352.

## List of scientific publications

Farag N., Mattossovich R., **Merlo R.**, Nierzwicki Ł., Palermo G., Porchetta A., Perugino G., Ricci F. Folding-upon-repair DNA nanoswitches for monitoring DNA repair enzymes activity. *Angew Chem Int Ed Engl.* 2021, 60:7283-7289. doi: [10.1002/anie.202016223](https://doi.org/10.1002/anie.202016223).

**Merlo R.**, Caprioglio D., Cillo M., Valenti A., Mattossovich R., Morrone C., Massarotti A., Rossi F., Miggiano R., Leonardi A., Minassi A., Perugino G. The SNAP- tag technology revised: an effective chemo-enzymatic approach by using a universal azide-based substrate. *J Enzyme Inhib Med Chem.* 2021, 36(1):85-97. doi: [10.1080/14756366.2020.1841182](https://doi.org/10.1080/14756366.2020.1841182).

Mattossovich R.\*, **Merlo R.\***, Miggiano R., Valenti A., Perugino G. O<sup>6</sup>-alkylguanine-DNA Alkyltransferases in Microbes Living on the Edge: From Stability to Applicability. *Int. J. Mol. Sci.* 2020, 21(8), 2878. doi.org/[10.3390/ijms21082878](https://doi.org/10.3390/ijms21082878).

Mattossovich R.\*, **Merlo R.\***, Fontana A., d'Ippolito G., Terns M.P., Watts E.A., Valenti A., Perugino G. A journey down to hell: new thermostable protein-tags for biotechnology at high temperatures. *Extremophiles* 2020, 24(1):81-91. doi:[10.1007/s00792-019-01134-3](https://doi.org/10.1007/s00792-019-01134-3).

Del Prete S.\*, **Merlo R.\***, Valenti A., Mattossovich R., Rossi M., Carginale V., Supuran C.T., Perugino G., Capasso C. Thermostability enhancement of the  $\alpha$ -carbonic anhydrase from Sulfurihydrogenibium yellowstonense by using the anchoring-and-self-labelling-protein-tag system (ASLtag). *J Enzyme Inhib Med Chem.* 2019, 34(1):946-954. doi:[10.1080/14756366.2019.1605991](https://doi.org/10.1080/14756366.2019.1605991).

**Merlo R.**, Del Prete S.\*, Valenti A., Mattossovich R., Carginale V., Supuran C.T., Capasso C., Perugino G. An AGT-based protein-tag system for the labelling and surface immobilization of enzymes on E. coli outer membrane. *J Enzyme Inhib Med Chem.* 2019, 34(1):490-499. doi:[10.1080/14756366.2018.1559161](https://doi.org/10.1080/14756366.2018.1559161).

\*These authors contributed equally

### **Books chapter:**

**Merlo R.**, Mattossovich R., Valenti A. and Perugino G. Recent advances in the use of the SNAP-tag<sup>®</sup> in the modern biotechnology. *Recent Research Advances in Biology*. Accepted for publication (Ref. no. 2020/BP/8226D)

## Meetings, Workshops and Webinar

- Speaker for “CNR-Institute of Biosciences and BioResources (IBBR) Webinar” on February 17th, 2021.  
Oral Presentation title: “*O*<sup>6</sup>-alkylguanine-DNA Alkyltransferase (AGT): when Nature meets Biotechnology”  
Author: **Rosa Merlo**
- Speaker at “1st IBBR Memorial Workshop – Maria Ciaramella”, held in Naples (Italy) on January 28th, 2020.  
Oral Presentation title: “*O*<sup>6</sup>-alchilguanina DNA-alchiltransferasi: da ambienti estremi ad applicazioni biotecnologiche”  
Author: **Rosa Merlo**
- Has attended the “II Industrial Biotechnology Congress: BioID&A Biotechnology Identity and Application” held in Naples (Italy) on October 28th, 2019
- Selected speaker at the “15th International Congress on Thermophiles”, Fukuoka (Japan), September 2-6, 2019.  
Oral Presentation title: “From Hot Sources to Biotechnological Processes: A Novel Tool for an In Vivo Enzyme Labelling and Immobilization”  
Authors: **Rosa Merlo**, Sonia Del Prete, Anna Valenti, Rosanna Mattosovich, Vincenzo Carginale, Claudiu T. Supuran, Clemente Capasso, Giuseppe Perugino
- Selected speaker and Poster presenter at the “2° Workshop BIO/10 di Docenti e Ricercatori di Biochimica della Campania”, held in Naples (Italy) on May 17th, 2019.  
Oral Presentation and Poster title: “A new biotechnological tool for an in vivo enzyme labelling and immobilization”  
Authors: **Rosa Merlo**, Sonia Del Prete, Anna Valenti, Rosanna Mattosovich, Vincenzo Carginale, Claudiu T. Supuran, Clemente Capasso, Giuseppe Perugino
- Speaker at “VI edizione dell’European Biotech Week” held in Naples (Italy) on September 24 – 30, 2018.  
Oral Presentation title: “Dagli ambienti estremi le biotecnologie del futuro”  
Author: **Rosa Merlo**
- Poster presenter at the 12th edition of “International Congress on Extremophiles”, Ischia, Naples (Italy), September 16-20, 2018.  
Poster title: “AGTs from hot sources: from stability to applicability”

Authors: Rosanna Mattosovich\*, **Rosa Merlo\***, Sonia Del Prete, Vincenzo Carginale,  
Clemente Capasso, Anna Valenti, Andreas Jaekel, Barbara Saccà, Giuliana Ippolito, Angelo  
Fontana, Maria Ciaramella, Giuseppe Perugino



## An AGT-based *protein-tag* system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane

Rosa Merlo, Sonia Del Prete, Anna Valenti, Rosanna Mattosovich, Vincenzo Carginale, Claudiu T. Supuran, Clemente Capasso & Giuseppe Perugino

To cite this article: Rosa Merlo, Sonia Del Prete, Anna Valenti, Rosanna Mattosovich, Vincenzo Carginale, Claudiu T. Supuran, Clemente Capasso & Giuseppe Perugino (2019) An AGT-based *protein-tag* system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane, Journal of Enzyme Inhibition and Medicinal Chemistry, 34:1, 490-499, DOI: [10.1080/14756366.2018.1559161](https://doi.org/10.1080/14756366.2018.1559161)

To link to this article: <https://doi.org/10.1080/14756366.2018.1559161>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 06 Feb 2019.



[Submit your article to this journal](#)



Article views: 29



[View Crossmark data](#)

RESEARCH PAPER



## An AGT-based *protein-tag* system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane

Rosa Merlo<sup>a\*</sup>, Sonia Del Prete<sup>a\*</sup>, Anna Valenti<sup>a</sup>, Rosanna Mattosovich<sup>a</sup>, Vincenzo Carginale<sup>a</sup>, Claudiu T. Supuran<sup>b</sup> , Clemente Capasso<sup>a</sup>  and Giuseppe Perugino<sup>a</sup>

<sup>a</sup>Department of Biology Agriculture and Food Sciences, Institute of Bioscience and BioResources – National Research Council of Italy, Naples, Italy; <sup>b</sup>Neurofarba Department, University of Florence, Polo Scientifico, Sesto Fiorentino Firenze, Italy

### ABSTRACT

The use of natural systems, such as outer membrane protein A (OmpA), phosphoprotein E (PhoE), ice nucleation protein (INP), etc., has been proved very useful for the surface exposure of proteins on the outer membrane of Gram-negative bacteria. These strategies have the clear advantage of unifying in a one-step the production, the purification and the *in vivo* immobilisation of proteins/biocatalysts onto a specific biological support. Here, we introduce the novel *Anchoring-and-Self-Labeling-protein-tag* (ASL<sup>tag</sup>), which allows the *in vivo* immobilisation of enzymes on *E. coli* surface and the labelling of the neosynthesised proteins with the engineered alkylguanine-DNA-alkyl-transferase (H<sup>5</sup>) from *Sulfolobus solfataricus*. Our results demonstrated that this *tag* enhanced the overexpression of thermostable enzymes, such as the carbonic anhydrase (*SspCA*) from *Sulfurihydrogenibium yellowstonense* and the  $\beta$ -glycoside hydrolase (*SsBGly*) from *S. solfataricus*, without affecting their folding and catalytic activity, proposing a new tool for the improvement in the utilisation of biocatalysts of biotechnological interest.

### ARTICLE HISTORY

Received 30 October 2018  
Revised 28 November 2018  
Accepted 4 December 2018

### KEYWORDS

Carbonic anhydrase;  
 $\beta$ -glycoside hydrolase;  
thermostable *protein-tag*;  
ice nucleation protein;  
enzyme immobilisation

## 1. Introduction

The term *immobilised enzymes* refers to 'enzymes physically confined or localised in a certain defined region of space with retention of their catalytic activities, and which can be used repeatedly and continuously'<sup>1</sup>. The immobilisation of enzymes on solid supports is historically very important for overcoming their general instability in harsh operational conditions and their low shelf-life, as well as the need for their recycling more times<sup>2</sup>. Furthermore, the physical separation of the biocatalyst from the reaction mixture avoids the protein contamination of the products. Although a reduction in reaction rates sometimes occurs, because the enzyme cannot mix freely with the substrate or a particular conformational change is needed for the biocatalyst efficiency, there are many examples of increased activity and stability of immobilised enzymes<sup>3</sup>. Many chemical or physical methods for the enzyme immobilisation are currently available, from the physical adsorption to the covalent coupling on supports (Figure 1(a–c,e–f))<sup>3–6</sup>. Recently, the use of *protein-tags* based on an engineered version of the human O<sup>6</sup>-alkyl-guanine-DNA-alkyl-transferases (hAGT) is an effective alternative for the covalent immobilisation of proteins and enzymes (Figure 1(d))<sup>7–9</sup>. AGTs (or OGTs, MGMTs; E.C.: 2.1.1.63) are DNA repair enzymes, which *irreversibly* transfer the alkyl group from the damaged DNA containing O<sup>6</sup>-alkyl-guanines to their cysteine residue in the active site<sup>10–13</sup>. In 2003, Johnsson and his group demonstrated that most enzymes of this class display relatively low substrate specificity, making them reactive also with free O<sup>6</sup>-benzyl-guanines (O<sup>6</sup>-BG)

nucleobases<sup>14</sup>. This led to the development of the so-called SNAP-tag<sup>TM</sup> technology, which uses derivatives of O<sup>6</sup>-BG potentially conjugated with an unlimited number of chemical groups<sup>15–18</sup>. This system allows the immobilisation on O<sup>6</sup>-BG-derivatised surface of the protein expressed in fusion with the SNAP-tag<sup>18</sup> (Figure 1(d)). However, all these approaches mainly depend on the high costs due to the isolation and purification of the biocatalysts. This limitation can be easily overcome by the heterologous expression of enzymes and their *in vivo* direct immobilisation on the surface of bacterial hosts, by the utilisation of transmembrane protein domains, as the ice nucleation protein (INP) of the Gram-negative bacterium *Pseudomonas syringae* (Figure 1(g))<sup>19,20</sup>. This protein is composed of an N-terminal domain (N, 175 residues) structurally separated from a C-terminal domain (C, 49 residues) by a repetitive central domain<sup>21</sup>. Both domains play a role in the anchoring of proteins to the outer membrane<sup>21</sup>. The use of INP as *anchoring carrier* is considered of great interest in biotechnological applications, ranging from the development of bacterial cell surface-display systems for vaccine delivery to the fabrication of whole-cell biocatalysts and biosensors<sup>22–24</sup>. The N-terminal domain of INP (INPN) was recently and successfully used for the one-step procedure immobilisation (Figure 1(g))<sup>22–29</sup>. Moreover, Capasso *et al.* demonstrated that the amount of a thermostable carbonic anhydrase<sup>30–33</sup> fused to the INPN domain and expressed on the bacterial cell surface had a hydratase activity similar to that of the enzyme covalently immobilised onto magnetic nanoparticles<sup>30,34</sup>. Here, we introduce a novel *protein-tag* system, (hereinafter

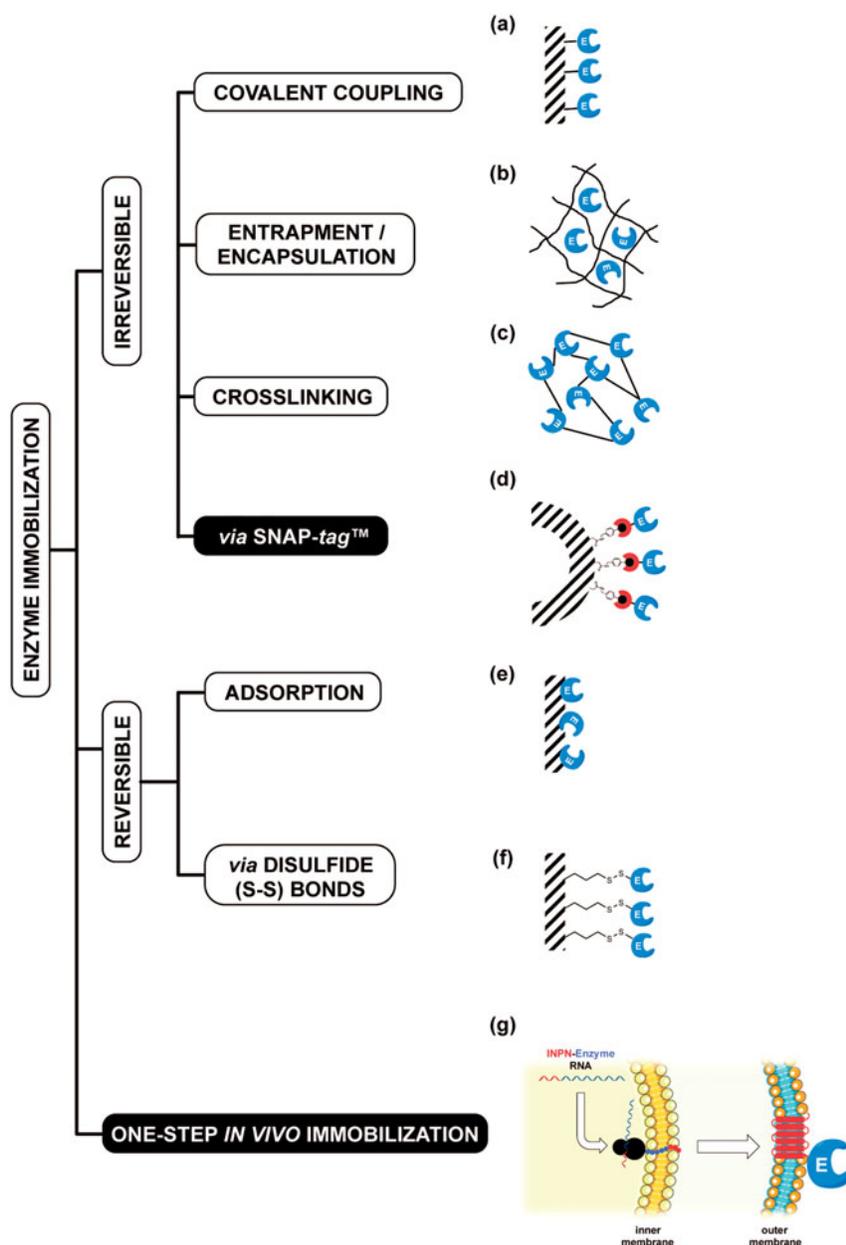
**CONTACT** Clemente Capasso  [clemente.capasso@ibbr.cnr.it](mailto:clemente.capasso@ibbr.cnr.it); Giuseppe Perugino  [giuseppe.perugino@ibbr.cnr.it](mailto:giuseppe.perugino@ibbr.cnr.it)  Department of Biology Agriculture and Food Sciences, Institute of Bioscience and BioResources – National Research Council of Italy, Via Pietro Castellino 111, Naples 80131, Italy

\*These authors equally contributed to this work.

 Supplemental data for this article can be accessed on the publisher's website.

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** Examples of enzyme immobilisation methods. Among the traditional methods for the irreversible immobilisation (a–c, e–f) of an enzyme (E, in blue), the recent introduction of the SNAP-tag™ technology (d, red semicircle) allowed an indirect immobilisation of protein of interest. The one step *in vivo* immobilisation of an enzyme (g) is possible when it is expressed as fusion protein with the N-terminal domain of the ice nucleation protein (INPN, in red). Because of the presence of a peptide leader upstream the coding sequence, the nascent polypeptide is translocated to the cell OM by the anchoring transmembrane INPN domain, leading to the immobilisation and exposition of the biocatalyst outside the cell.

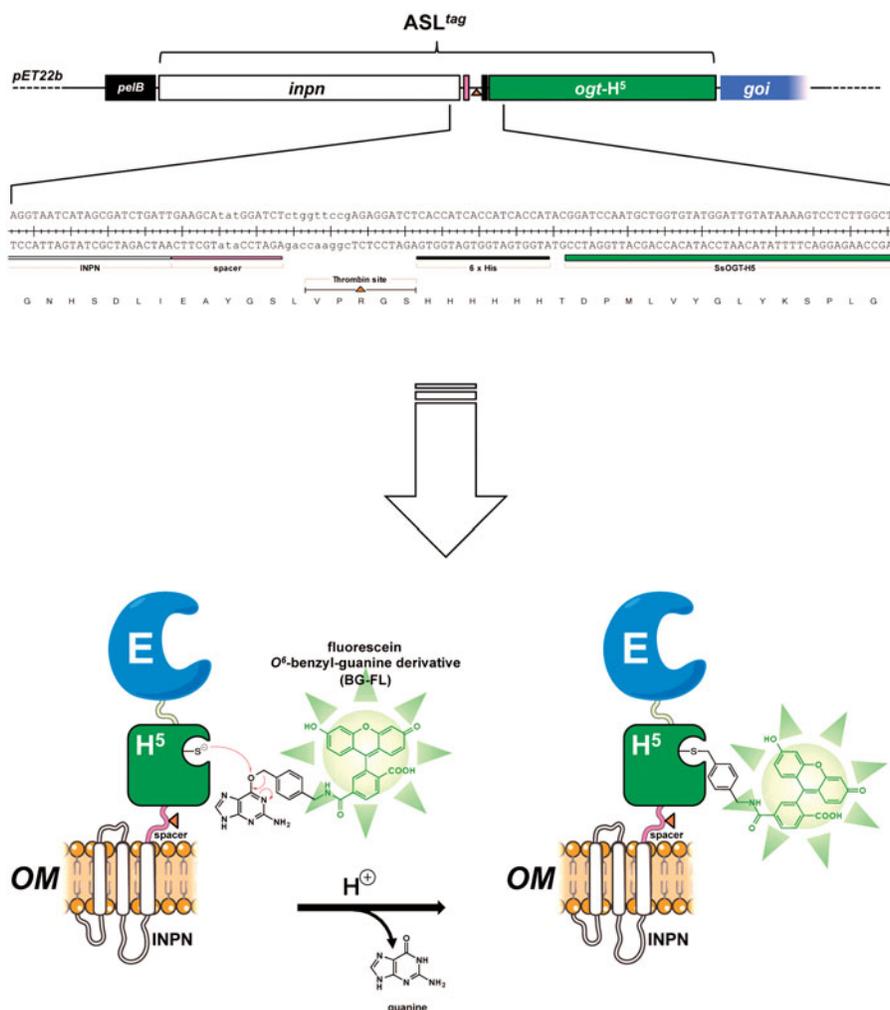
Anchoring-and-Self-Labeling-protein-tag or ASL<sup>tag</sup>), which simultaneously allows the *in vivo* immobilisation of the enzyme of interest on the *E. coli* surface and its quantitative determination (Figure 2). The ASL<sup>tag</sup> system is formed by the INPN domain fused to an engineered and thermostable variant of the alkylguanine-DNA-alkyl-transferase (H<sup>5</sup>) from the hyperthermophilic archaeon *Sulfolobus solfataricus*<sup>35,36</sup>. This enzyme was extensively characterised, suggesting its biotechnological role as thermostable alternative to the commercial SNAP-tag™ and its utilisation as *protein-tag* for heterologous expression of proteins of interest in *E. coli* and, for the first time, in thermophilic organisms as *Thermus thermophilus* and *Sulfolobus islandicus*<sup>35–37</sup>. Thus, using the substrate of H<sup>5</sup>, a fluorescein derivative of the O<sup>6</sup>-BG (Figure 2), we successfully estimated the expression of the ASL<sup>tag</sup> in *E. coli* cells, by *in vitro gel-imaging* techniques, as well as by *in vivo* fluorescent

microscopy. Furthermore, we demonstrated that the activity and the stability of the enzymes of interest (*SspCA*, the  $\alpha$ -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense*; and *S $\beta$ Gly*, the  $\beta$ -glycoside hydrolase from *S. solfataricus*) fused to the ASL<sup>tag</sup> and exposed on the surface of *E. coli* cells were not affected by the presence of this novel *protein-tag*.

## 2. Materials and methods

### 2.1. Reagents

All DNA restriction and modification enzymes and the fluorescent substrate for the OGT activity (SNAP-Vista Green™, hereinafter BG-FL) were purchased from New England Biolabs (Ipswich, MA); molecular biology kits for the plasmid preparations and DNA gel



**Figure 2.** The  $ASL^{tag}$  protein. The  $ASL^{tag}$  gene is composed by the *inpn* ORF (in white) in frame fused to the *ogtH<sup>5</sup>* gene (in green) in the pET22b expression vector. This *tag* can be further fused to a gene of interest (*goi*, in blue), for a *one step procedure* of the expression and immobilisation of an enzyme (E). The presence of the  $H^5$  moiety allows the quantitative estimation of the yield of E by the irreversible alkyl-transferase assay using a fluorescent  $O^6$ -benzyl-guanine derivative (BG-FL). Between *inpn* and *ogtH<sup>5</sup>*, a spacer (in pink), a thrombin cleavage site (shown as an orange triangle) and a  $6 \times$  His-tag (in black) were inserted, for the easy separation and purification of any  $H^5$ -E fusion protein.

extractions (NucleoSpin<sup>®</sup> Gel and PCR Clean-up<sup>®</sup>) were from Macherey-Nagel GmbH (Germany); Lyophilised Thrombin Protease from GE Healthcare (Illinois, US). Eurofins Genomics (Germany) performed the oligonucleotides synthesis and the DNA sequencing service.

## 2.2. DNA constructs

To obtain the pET- $ASL^{tag}$  construct, we replaced the  $\alpha$ -carbonic anhydrase (*SspCA*) gene with the *ogtH<sup>5</sup>* gene in the previously described vector pET-22b/INPN-*SspCA*<sup>30</sup>. By the latter and the pQE-*ogtH<sup>5</sup>*<sup>35</sup> plasmid as template, the DNA fragments relative to the INPN domain and  $H^5$  were respectively amplified with the INPN- and  $H^5$ -Fwd/Rev oligonucleotides pairs (listed in Table 1), under the following conditions: an initial denaturation at 95.0 °C for 5 min, 30 cycles of 30 s at 95.0 °C, 30 s at 50.0 °C and 30 s at 72.0 °C, followed by a final extension of 5 min at 72.0 °C. DNA products were fused to each other in a further PCR amplification, taking advantage of the total complementarity of the INPN-Rev to the  $H^5$ -Fwd oligonucleotide, obtaining the final  $ASL^{tag}$  DNA fragment. Subsequently, this fragment and the pET-22b/INPN-*SspCA* vector were digested with Hind III and Xba I restriction endonucleases, gel-purified, and ligated. The ligation mixture was used to

transform the *E. coli* DH5 $\alpha$  strain, and positive colonies were confirmed by colony PCR and DNA sequencing.

The same cloning strategy was used to achieve the pET- $ASL^{tag}$ -*SspCA* construct: the  $ASL^{tag}$  DNA fragment obtained this time using the  $H^5$ -Rev2 oligonucleotide was used as template to further fuse to the *SspCA* gene (obtained by amplification with *SspCA* Fwd/Rev oligonucleotide pairs, Table 1). Again, positive colonies after ligation and transformation were confirmed as above.

Finally, the pET- $ASL^{tag}$ -*lacS* plasmid preparation started with the achievement of the DNA fragment relative to the *ogtH<sup>5</sup>* gene fused to the  $\beta$ -glycoside hydrolase from the thermophilic archaea *Sulfolobus solfataricus* (*lacS*), obtained by the Hind III/BamH I digestion from the pQE-*ogtH<sup>5</sup>*-*lacS* plasmid<sup>36</sup>. The pET- $ASL^{tag}$  plasmid was similarly digested to obtain the pET22 recipient for the first ligation/transformation round. Positive blue colonies were selected by the hydrolase activity of the *lacS* gene product (*Ss* $\beta$ Gly) on the Ampicillin selective Luria-Bertani (LB) Agar plates supplemented with 5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucopyranoside (X-Glc), and the insertion confirmed using PCR analysis. This intermediate vector was further digested with BamH I, treated with Alkaline Phosphatase Calf Intestinal (CIP) and ligated to the BamH I/BamH I INPN DNA fragment, derived from the digestion of the pET- $ASL^{tag}$  plasmid. In this case, positive blue colonies on

**Table 1.** Oligonucleotides used in this study.

Oligonucleotide	Sequence
INPN-Fwd	5'-TAATACGACTCACTATAGGG-3'
INPN-Rev	5'-GGTGATGGTGAGATCCTCTCGGAACAGAGATCCATAGGCTTCAATCAGATCGC-3'
H <sup>5</sup> -Fwd	5'-GCGATCTGATTGAAGCCTATGGATCTCTGGTCCGAGAGGATCTCACCATACC-3'
H <sup>5</sup> -Rev	5'-TCACCTTCATATGACCATTATGTTCTGCTACCATAAGCTTCTGTCGACGGTACCTCGAGTTCTGG-3'
H <sup>5</sup> -Rev2	5'-ATTGAGCAACTGACTGAAATGCC-3'
SspCA-Fwd	5'-CCAGAACTCGAGGTACCGTCGACAGAGGCTTATGGTAGCGAACTGAATGGTCATGAAGGTGA-3'
SspCA-Rev	5'-CTAGTTATTGCTCAGCGGT-3'

LB-Amp-X-Glc agar plates were confirmed by Pst I restriction enzyme digestion analysis.

### 2.3. Determination of the protein expression by a fluorescent assay based on the H<sup>5</sup> activity

All constructs were used to transform *E. coli* BL21(DE3) cells. Cultures were grown at 37.0 °C in LB selective medium supplemented with 100.0 mg/L ampicillin and 30.0 mg/L chloramphenicol; expression was induced with 1.0 mM isopropyl-thio- $\beta$ -D-galactoside (IPTG) when an absorbance value of 0.5–0.6 A<sub>600nm</sub> was reached, or in the ZY auto-induction medium (AI)<sup>38</sup>, supplemented with the same selective antibiotics. After an overnight incubation, whole cells were collected and assayed by using the BG-FL fluorescent H<sup>5</sup> substrate previously described<sup>35–37,39</sup> for a qualitative measurement of the protein expression, an aliquot of 1.0 mL of cells was centrifuged at 4000  $\times$  g and the pellet was resuspended in 50.0  $\mu$ L of 5.0  $\mu$ M BG-FL in phosphate-buffered saline (PBS 1 $\times$ , 20.0 mM phosphate buffer, NaCl 150.0 mM, pH 7.3). After an incubation for 2.0 h at 37.0 °C, reactions were stopped by denaturing the samples in Leamml Buffer 1 $\times$  and directly loaded on SDS-PAGE, followed by *gel-imaging* on a VersaDoc 4000<sup>TM</sup> system (Bio-Rad), by applying a blue LED/530 bandpass filter as excitation/emission parameters, respectively. Finally, the fluorescence intensity of each band was normalised to the intensity of the signal obtained from the Coomassie Blue staining analysis.

For a quantitative determination of the expression, whole cells were opportunely diluted to achieve an OD<sub>600nm</sub> of 1.0. By following the same above-mentioned assay, three different volumes of whole cells were loaded *in the same gel* with defined amounts (in the range of 0.8–50.0 pmols) of purified free H<sup>5</sup> protein, after the reaction on BG-FL in the same conditions. The obtained values of the relative fluorescence as a function of the purified loaded H<sup>5</sup> were fitted by a linear equation, whose parameters were then used for the estimation of the amount of expressed H<sup>5</sup>-derivated fusion proteins, assuming that the activity of the H<sup>5</sup> moiety in the fusions is not affected by the presence of the other protein partner(s). Given that the concentration of *E. coli* cell cultures of 1.0 OD<sub>600nm</sub> is ca. 8.0  $\times$  10<sup>8</sup> cells/mL, and the amount of wet cells is 1.7 g/L, it is possible to calculate the yield of expressed proteins in terms of pmol/mg of cells<sup>40</sup>.

### 2.4. Membranes fractionation

The *E. coli* outer (OM) and inner (IM) membranes were purified by following a procedure previously described<sup>27</sup>. Briefly, harvested bacterial cells were resuspended 1:20 (g/mL) in 25.0 mM Tris/HCl buffer, pH 8.0 and disrupted by sonication on ice (10 cycles of 10s: 50s on:off treatment). Cell extract was centrifuged at 120,000  $\times$  g for 1.0 h, and the supernatant containing the cytoplasmic fraction was discarded. Both IM and OM fractions were recovered in the pellet and resuspended in 20.0 mL of PBS 1 $\times$ ,

containing 0.01 mM MgCl<sub>2</sub> and 2.0% Triton X-100. After incubation at room temperature for 30.0 min, the solution was centrifuged as described above. Then, a defined amount of the OM fraction obtained was assayed for the H<sup>5</sup> activity, leading to the quantitative determination of the total amount of fusion protein, as previously described.

### 2.5. Thrombin assay

The ASL<sup>tag</sup> on the OM was cleaved with the Thrombin Protease, in order to separate the H<sup>5</sup> moiety from the INPN transmembrane domain. A suspension of *E. coli* whole cells was gently centrifuged at 3000  $\times$  g for 10.0 min at 4.0 °C and then resuspended in a half volume of PBS 1 $\times$ . The sample was incubated with 30.0 U of Thrombin Protease at 25.0 °C overnight under gentle agitation in the presence of 5.0  $\mu$ M of BG-FL, followed by the same centrifugation. Cells and supernatants were separately loaded on SDS-PAGE and analyzed by *gel-imaging* and Coomassie staining as described.

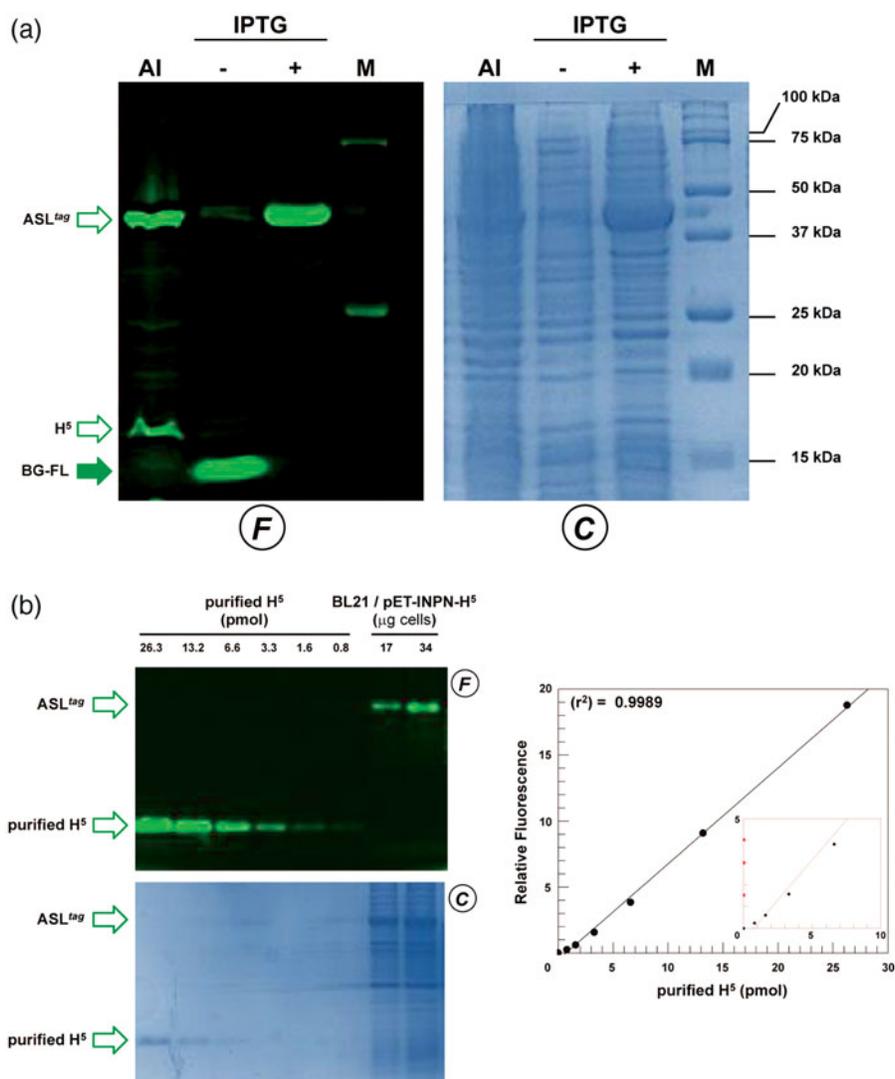
### 2.6. Microscopy analysis of *E. coli* cells

For *in vivo* imaging, *E. coli* BL21(DE3) cells transformed with pET-22b/INPN-SspCA or pET-ASL<sup>tag</sup> plasmids were IPTG-induced, grown overnight at 37.0 °C and finally diluted until OD<sub>600nm</sub> of 1.0. An amount of cells corresponding to a volume of 1.0 mL of the culture was washed twice in PBS 1 $\times$  and finally resuspended in 50.0  $\mu$ L of the same buffer supplemented with 5.0  $\mu$ M of the BG-FL substrate. After an incubation at 37.0 °C for 30.0 min, cells were washed twice, resuspended and again incubated for 30.0 min at 37.0 °C, to allow the external diffusion of the unreacted substrate. Labelling was first verified by fluorescence *gel-imaging* on SDS-PAGE and then spotted on poly-L-lysine coated slides for microscopy analysis.

Images were collected using a DM6 fluorescence microscope and Hamamatsu camera under the control of Leica LAS AF 6000 software; excitation and emission wavelengths used suitably for AlexaFluor488 dye were  $\lambda_{ex}$  = 490 nm;  $\lambda_{em}$  = 525 nm, respectively.

### 2.7. $\beta$ -glycoside hydrolase assay

The  $\beta$ -glycoside hydrolase assay was performed as previously described<sup>41</sup> at different temperatures in 50 mM sodium phosphate buffer at pH 6.5, in the presence of 2 Np- and 4 Np-Glc substrates at 5.0 mM final concentration. OM fractions containing ASL<sup>tag</sup> and relative fusion proteins amounts ranging from 1.0 to 5.0  $\mu$ g were used in each assay. For the correction of the spontaneous hydrolysis of the substrates, all the reactants except the enzyme (blank mixture), was taken into account. The enzymatic activity was calculated on the basis of the molar extinction coefficient ( $\epsilon_M$ ) values of 2- and 4-nitrophenol in 50 mM sodium phosphate buffer pH 6.5 at different temperatures, as previously reported<sup>41</sup>. We defined as one unit of enzyme activity the amount of enzyme hydrolyzing 1.0  $\mu$ mol of substrate in 1.0 min, under the above-described conditions.



**Figure 3.** The ASL<sup>tag</sup> expression in *E. coli*. (a) *E. coli* BL21(DE3) strain transformed with the pET-ASL<sup>tag</sup> plasmid was grown in IPTG-induced or in auto-induction medium (AI). After the *in vivo* OGT assay, a defined amount in micrograms of whole cells at OD<sub>600nm</sub> of 1.0<sup>40</sup> was directly loaded on SDS-PAGE, followed by *gel-imaging* fluorescence (F) and Coomassie staining (C) analyses. Open and closed green arrows indicate fluorescent signals of free H<sup>5</sup> or H<sup>5</sup>-based fusion proteins, and the free BG-FL substrate, respectively. M: molecular weight marker. (b) Quantitative estimation of the ASL<sup>tag</sup> expression: defined amount of cells and purified H<sup>5</sup> protein (in pmols) were loaded and analyzed on a SDS-PAGE (on the left). Fluorescent values obtained from H<sup>5</sup> were fitted in a linear plot (on the right), as described in Materials and Methods. Obtained parameters allowed the quantitative determination of the amount of ASL<sup>tag</sup> in *E. coli* cells and shown in Table 2.

**Table 2.** Quantitative estimation of the heterologous expression of ASL<sup>tag</sup> and relative fusion proteins in *E. coli* BL21(DE3) whole wet cells.

	MW <sup>a</sup> (kDa)	ASL <sup>tag</sup> -E ratio	Fusion yield <sup>b</sup> (pmol H <sup>5</sup> /mg)	E yield <sup>c</sup> (μg/mg)	(r <sup>2</sup> ) <sup>d</sup>
ASL <sup>tag</sup>	42.0	–	152.5 ± 15.7	–	0.9977
ASL <sup>tag</sup> -SspCA	70.2	1:1	58.4 ± 11.6	1.54 ± 0.30	0.9986
ASL <sup>tag</sup> -SsβGly	98.8	4:1	30.6 ± 13.2	7.2 ± 3.1 <sup>e</sup>	0.9980

<sup>a</sup>Calculated from the primary structure.

<sup>b</sup>On the basis of the H<sup>5</sup> activity on fluorescent *gel-imaging* analysis (see Materials and Methods).

<sup>c</sup>The amount of the immobilised enzyme without the ASL<sup>tag</sup>.

<sup>d</sup>Correlation coefficient of the linear curve obtained from the H<sup>5</sup> values of fluorescence as a function of the amount of the loaded protein.

<sup>e</sup>The tetrameric form of the catalytically active SsβGly enzyme.

## 2.8. Carbonic anhydrase assay

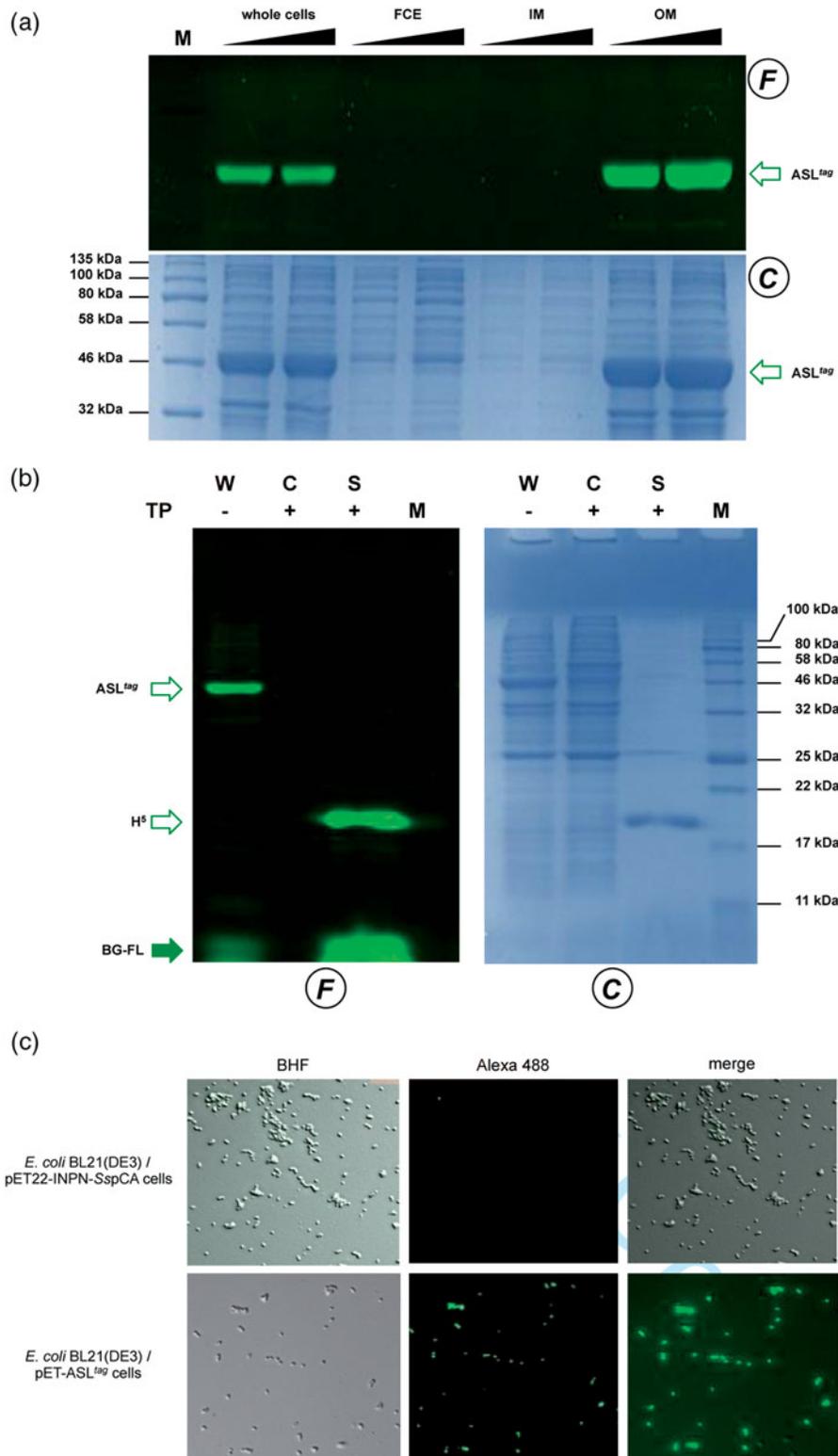
CA activity assay was a modification of the procedure described by Capasso *et al.*<sup>33</sup>. Briefly, the hydratase assay was performed at 0°C using CO<sub>2</sub> as substrate following the pH variation due to the catalyzed conversion of CO<sub>2</sub> to bicarbonate. Bromothymol blue was used as pH indicator. The production of hydrogen ions during

the CO<sub>2</sub> hydration reaction lowers the pH of the solution leading to a colour transition of the dye. The time required for the colour change is inversely proportional to the amount of CA present in the sample. The Wilbur–Anderson units (WAU) were calculated according to the following definition: one WAU of CA activity is defined as the ratio  $(T_0 - T)/T$ , where  $T_0$  (the time needed for the pH indicator color change for the uncatalyzed reaction) and  $T$  (the time needed for the pH indicator color change for the catalyzed reaction) are recorded as the time (in seconds) required for the pH to drop from 8.3 to the transition point of the dye (pH 6.8) in a control buffer and the presence of enzyme, respectively.

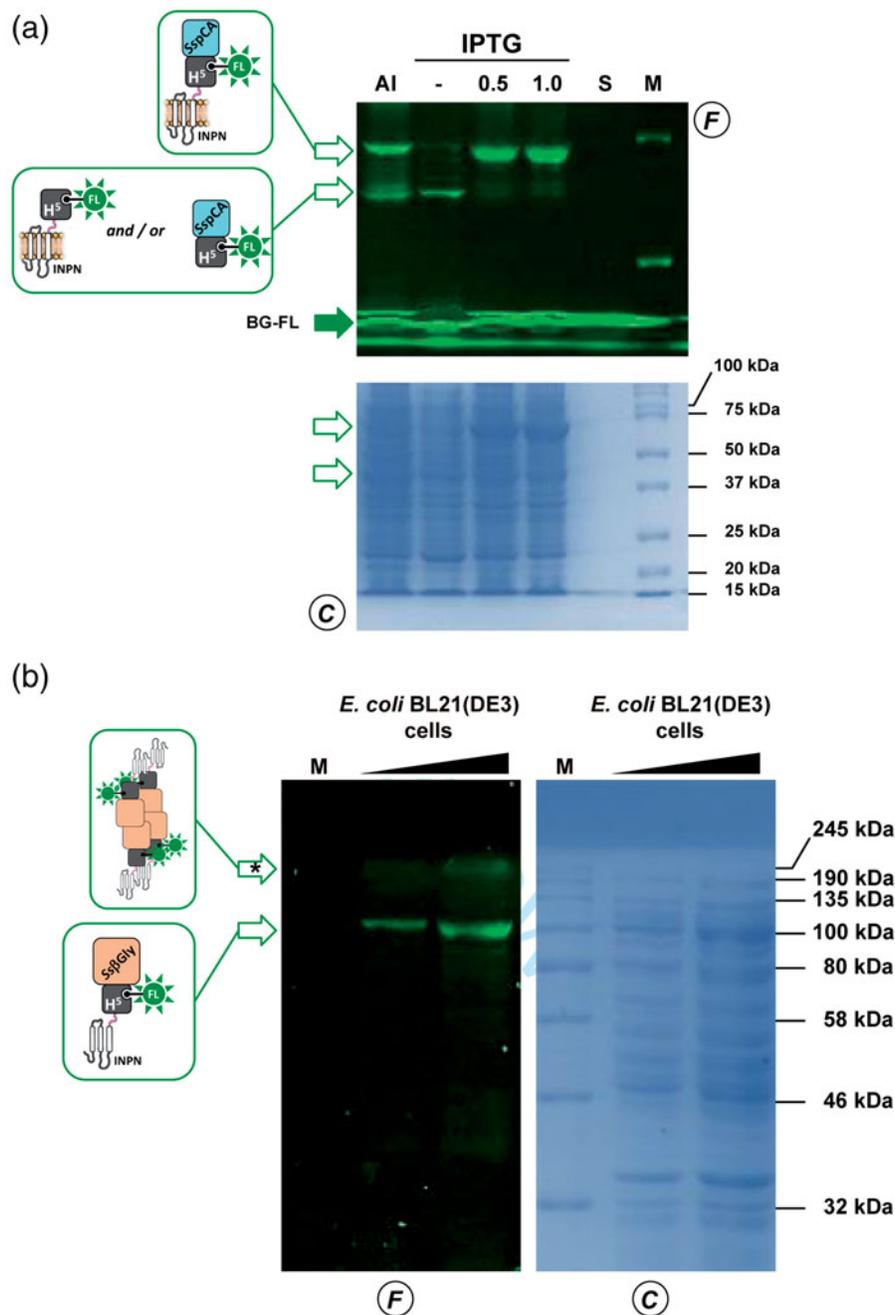
## 3. Results and discussion

### 3.1. Expression analysis and localisation of ASL<sup>tag</sup> in *E. coli*

The *in vivo* alkyl-transferase fluorescent assay of H<sup>5</sup> on whole bacterial cells is a useful method to determine the heterologous expression of this protein and/or relative fusion proteins in the mesophilic *E. coli* and in the thermophilic species *T. thermophilus*



**Figure 4.** Localisation of ASL<sup>tag</sup> in *E. coli*. (a) gel-imaging and Coomassie staining analyses after SDS-PAGE of different loaded amount of the whole cells, the relative cytoplasmic fraction (FCE), the inner (IM) and the outer membrane (OM) fractions. (b) Cleavage of ASL<sup>tag</sup> by the Thrombin protease (T) on whole cells (W). After the H<sup>2</sup> reaction during the protease treatment, cells were centrifuged, separating the supernatant (S) from the intact cells (C). (c) *E. coli* BL21(DE3) cells transformed with pET-22b/INPN-SspCA (Top) or with pET-ASL<sup>tag</sup> (bottom) were incubated with BG-FL and then analyzed at fluorescence microscopy. Images show bright field (BHF), AlexaFluor488 (green) and merged images. All used symbols are described in Figure 3(a).



**Figure 5.** Heterologous expression of thermostable enzymes fused to ASL<sup>tag</sup>. SDS-PAGE analysis of the ASL<sup>tag</sup>-SspCA (a) and ASL<sup>tag</sup>-SsβGly (b) expression. (a) Fluorescent signals on the F panel correspond to bands whose molecular weights are attributable to the protein fusions shown in the schemes. (b) 1.0 mM IPTG-induced cells gave a very high fluorescent band (marked by an asterisk), presumably corresponding to the tetrameric form of SsβGly linked to 4 × ASL<sup>tag</sup> units, as shown. All used symbols are described in Figure 3(a).

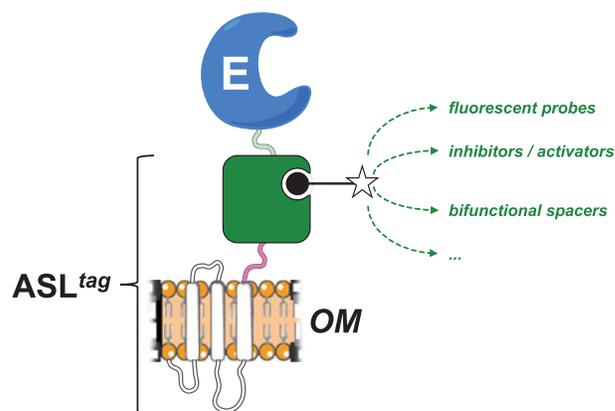
and *S. islandicus*, without any purification procedure<sup>35–37</sup>. The *in vivo* assay is possible since the bacterial cells are permeable to OGT fluorescent substrates (as the commercially available SNAP-Vista Green<sup>TM</sup> or SNAP-Cell TMR<sup>TM</sup>); and the catalytic activity at mesophilic temperatures (25–37 °C) of H<sup>5</sup> mutant is one order of magnitude higher than the SsOGT wild-type counterpart and comparable to the commercial SNAP-tag<sup>TM36</sup>. Despite GFPs utilisation, the covalent conjugation of H<sup>5</sup> with the benzyl-fluorophore moiety of the substrate (Figure 2) allows the denaturation of the whole cells and the direct loading of the samples on SDS-PAGE for the *gel-imaging* analysis, as described in the Materials and Methods<sup>35–37,39</sup>.

To evaluate the expression of the ASL<sup>tag</sup>, *E. coli* BL21(DE3)/pET-ASL<sup>tag</sup> cells were grown in LB medium and induced with IPTG or in the AI. In the latter case, although the expression level of the fusion protein was satisfactory, the presence of only the H<sup>5</sup> signal (ca. 20% of the total fluorescence intensity) in the *gel-imaging* analysis is difficult to rationalise (Figure 3(a)). Probably, during the expression of the protein (e.g., in the advanced stage of the growth) could occur interruptions or failures in the translocation process of the ASL<sup>tag</sup> on the outer membrane of *E. coli* with following breaking/cleavage events, especially in the spacer region between INPN and H<sup>5</sup> (Figure 2). On the other hand, after the IPTG induction, a strong and clear signal at the expected

molecular weight is visible, without any fragmentation (Figure 3(a)). This is an important result since, to date, the heterologous expression of SsOGT wild-type and relative variants in *E. coli* has been generally performed in the ABLE C strain, which keeps low the number of copies of the *ogt*-containing plasmids<sup>35,36</sup>. As proof, transformed *E. coli* BL21(DE3) strain with the pQE-*ogtH*<sup>5</sup> plasmid<sup>36</sup> showed a very low expression of free H<sup>5</sup> (Supplementary Figure S1), whereas the fusion with the INPN domain and the consequent translocation on the outer cell membrane made the H<sup>5</sup> expression possible in this strain. The comparative fluorescent analysis with a defined amount of free H<sup>5</sup> enzyme, allowed us to quantitatively measure the heterologous expression of ASL<sup>tag</sup> as pmol/mg of the whole wet cells (Figure 3(b) and Table 2). The assay on H<sup>5</sup> confirmed the anchoring function of the INPN trans-membrane protein: only the whole cells and the fraction containing the outer membrane displayed a fluorescent band corresponding to the ASL<sup>tag</sup> fusion protein, whereas the signal was missed in the lanes relative to the cytoplasmic and the inner membrane fractions (Figure 4(a)). Besides, the evidence of the INPN on the bacterial outer membrane was confirmed by treating the whole bacterial cells with the thrombin, too (Figure 2). A cleavage site for this protease was localised between the INPN domain and the H<sup>5</sup> moiety (Figure 2). The fluorescent signal corresponding to the MW of the H<sup>5</sup> protein was present only in the supernatant fraction when the protease was added at the same time with the BG-FL substrate (Figure 4(b)). Finally, we checked the OGT activity of ASL<sup>tag</sup> in living cells by microscope analysis, upon the labelling with the fluorescent substrate. The obtained images of living *E. coli* BL21(DE3) cells showed a strong and specific fluorescent signal only in those transformed with the ASL<sup>tag</sup>-containing plasmid (Figure 4(c)), indicating that the fusion protein is stable and proficient to labelling. This data suggest that ASL<sup>tag</sup> is suitable for localisation and analysis of membrane proteins, and provide an opportunity for further *in vivo* analyses of ASL<sup>tag</sup>-tagged proteins of interest under physiological conditions.

### 3.2. *In vivo* immobilisation of thermostable enzymes by ASL<sup>tag</sup>

As described in the literature, it has been demonstrated that the monomeric  $\alpha$ -carbonic anhydrase (SspCA) from the thermophilic bacterium *S. yellowstonense* can be actively expressed onto the outer membrane of *E. coli*<sup>30</sup>. Following this strategy, we realised a plasmid expressing the ASL<sup>tag</sup>-SspCA construct by inserting the *ogtH*<sup>5</sup> gene between the INPN and SspCA. The expression profile analyzed by following the H<sup>5</sup> activity on BG-FL confirmed that in the AI multiple fluorescent signals are present (Figure 5(a)), mainly represented by the full-length ASL<sup>tag</sup>-SspCA (70.2 kDa; ca. 65% of the total fluorescence intensity) and lower band (ca. 35%) closer to 37 kDa than 50 kDa (Figure 5(a)). This signal is compatible to the ASL<sup>tag</sup> (42.0 kDa) as well as the H<sup>5</sup>-SspCA moiety (46.0 kDa), suggesting that both translation interruption and translocation failure events in this growth conditions can be not excluded. Again, we detected an optimal achievement of the full-length fusion protein under IPTG-based expression in LB medium (ca. 95%; Figure 5(a)). In this condition and considering the INPN:H<sup>5</sup>:SspCA ratio as 1:1:1, the amount of the whole fusion protein expressed was estimated as ca. 60.0 pmol/g cells, which corresponds to ca. 1.5  $\mu$ g of the sole immobilised carbonic anhydrase per mg of cells (Table 2). Preliminary assays on ASL<sup>tag</sup>-SspCA indicated that the presence of H<sup>5</sup> does not hamper the hydratase activity of the



**Figure 6.** The biotechnological potential of the ASL<sup>tag</sup>. Any chemical group of interest (open star) conjugated to the BG-substrate (black closed circle) could be covalently bound to the H<sup>5</sup> moiety (in green) of the ASL<sup>tag</sup>. This enhances the potential use of an immobilised enzyme (E) on the *E. coli* surface (OM), making available to it a series of molecules, e.g., fluorescent probes and enzymatic activity modulators, or bi-functional groups for cascade reactions with other biocatalysts.

SspCA, if compared with that of the previously expressed INPN-anchored enzyme<sup>30</sup> (data not shown).

ASL<sup>tag</sup> system was also tested with another thermostable enzyme, the  $\beta$ -glycoside hydrolase (Ss $\beta$ Gly) from the thermophilic archaea *S. solfataricus*<sup>42</sup>. We previously demonstrated that the cytoplasmic H<sup>5</sup>-Ss $\beta$ Gly fusion protein is stable and active for both the OGT and the  $\beta$ -glycosidase assays, suggesting that the presence of one enzyme does not interfere with the folding and activity of the other<sup>36</sup>. Interesting to note that anchoring this protein fusion to the bacterial outer membrane by the INPN domain is particularly challenging because Ss $\beta$ Gly is active only in its tetrameric form<sup>43,44</sup>. The presence of blue colonies on LB agar plate in the presence of a glucoside chromogenic derivative (X-Glc), which is a preferred substrate of Ss $\beta$ Gly<sup>41</sup> but not of the *E. coli* LacZ (a  $\beta$ -galactosidase enzyme) was a first and convincing indication of the oligomerisation of this thermostable enzyme. Although in this case the amount of the expressed fusion protein was lower than the above examples (Figure 5(b) and Table 2), IPTG-induced *E. coli* BL21(DE3)/pET-ASL<sup>tag</sup>-Ss $\beta$ Gly cells displayed a fluorescent signal of expected molecular weight (98.8 kDa), corresponding to one monomer of Ss $\beta$ Gly fused to one ASL<sup>tag</sup> unit. However, a higher band is clearly visible in the fluorescence analysis (marked with an asterisk in Figure 5(b)), out of the molecular weight marker range used: as for the cytoplasmic H<sup>5</sup>-Ss $\beta$ Gly fusion, it could suggest that it corresponds to a partially denatured part of the tetrameric form of ASL<sup>tag</sup>-Ss $\beta$ Gly (ca. 400.0 kDa), which is particularly resistant to thermal denaturation<sup>36,42,43,45</sup>. Finally, an amount of 0.24  $\mu$ g/mg of immobilised Ss $\beta$ Gly on the OM fraction (on the basis of the calculated H<sup>5</sup> pmol) was assayed on 2Np- and 4Np-Glc at three different temperatures. The results show an activity of  $12.6 \pm 0.7$  and  $8.8 \pm 0.4$  (50 °C),  $30.3 \pm 0.4$  and  $20.3 \pm 0.9$  (60 °C),  $51.5 \pm 0.9$  and  $30.8 \pm 1.5$  U/mg (70 °C), respectively, whereas OM fraction containing the sole ASL<sup>tag</sup> did not result in any  $\beta$ -glucosidase activity, as expected (data not shown). These values are correctly related to the activity of Ss $\beta$ Gly<sup>45</sup>, clearly indicating that the formation of the quaternary structure on the *E. coli* OM occurs. However, since the 3D structure of this thermostable glycoside hydrolase showed that it is not laying on a surface<sup>43</sup>, we hypothesised an invagination of the *E. coli* outer membrane to allow the assembly of four units of the ASL<sup>tag</sup>-Ss $\beta$ Gly (Supplementary Figure S2).

#### 4. Conclusions and perspectives

In the present work, we introduced and demonstrated the utility of a novel *protein-tag*, composed by the N-terminal domain of the INP protein fused to a DNA repair enzyme. From our results, it is readily apparent that ASL<sup>tag</sup> offers: (i) an easy expression and *in vivo one-step procedure* of enzyme immobilisation on biological supports (e.g., *E. coli* outer membrane); (ii) significantly reduces the costs of the enzyme purification and those of the immobilisation support, allowing a direct exposition of the enzyme to the solvent<sup>20,30</sup>; (iii) an indirect labeling, by the reaction of a thermostable variant of the SNAP-tag<sup>TM</sup> (H<sup>5</sup>)<sup>46,47</sup>, which covalently links desired chemical groups conjugated to its benzyl-guanine substrate<sup>14,35</sup>. ASL<sup>tag</sup> favoured the expression of a monomeric protein (e.g., the thermostable SspCA) and an enzyme having a complex quaternary structure (e.g., the thermophilic SsβGly), without compromising their overall folding and enzymatic activity. Moreover, we showed that the utilisation of a fluorescein-derivative of the BG led to the localisation and the quantitative determination of the yield of the expressed ASL<sup>tag</sup> and the relative fusion proteins (Table 2). On the other hand, despite the GFPs utilisation limited only to all fluorescence-based applications, the possibility to conjugate different groups to the BG- for the H<sup>5</sup> reaction<sup>36</sup> dramatically expands the biotechnological potential of this novel *protein-tag*. For example, it will be possible to modulate the activity of biocatalysts (by introducing inhibitors/activators), or connecting them with other proteins for the improvement of cascade reactions in the presence of bi-functional chemical groups (Figure 6).

#### Acknowledgements

We are grateful to Maria Ciaramella for the constant and fruitful discussions during the execution of the experiments and the preparation of the manuscript. We are also grateful to Giovanni Del Monaco for technical assistance.

#### Disclosure statement

No potential conflict of interest was reported by the authors.

#### Funding

This work is supported by FIRB-Futuro in Ricerca RBF12001G\_002 “Nematic” and by the grant “SMART GENERATION-Sistemi e tecnologie sostenibili per la generazione di energia-PON03PE\_00157\_1, OR3-Bio-sistemi di cattura ed utilizzazione della CO<sub>2</sub>”.

#### ORCID

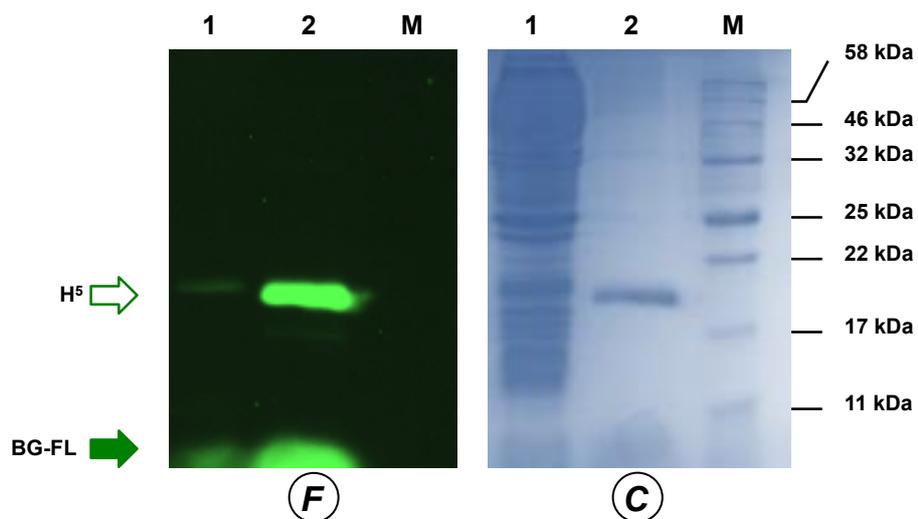
Claudiu T. Supuran  <http://orcid.org/0000-0003-4262-0323>

Clemente Capasso  <http://orcid.org/0000-0003-3314-2411>

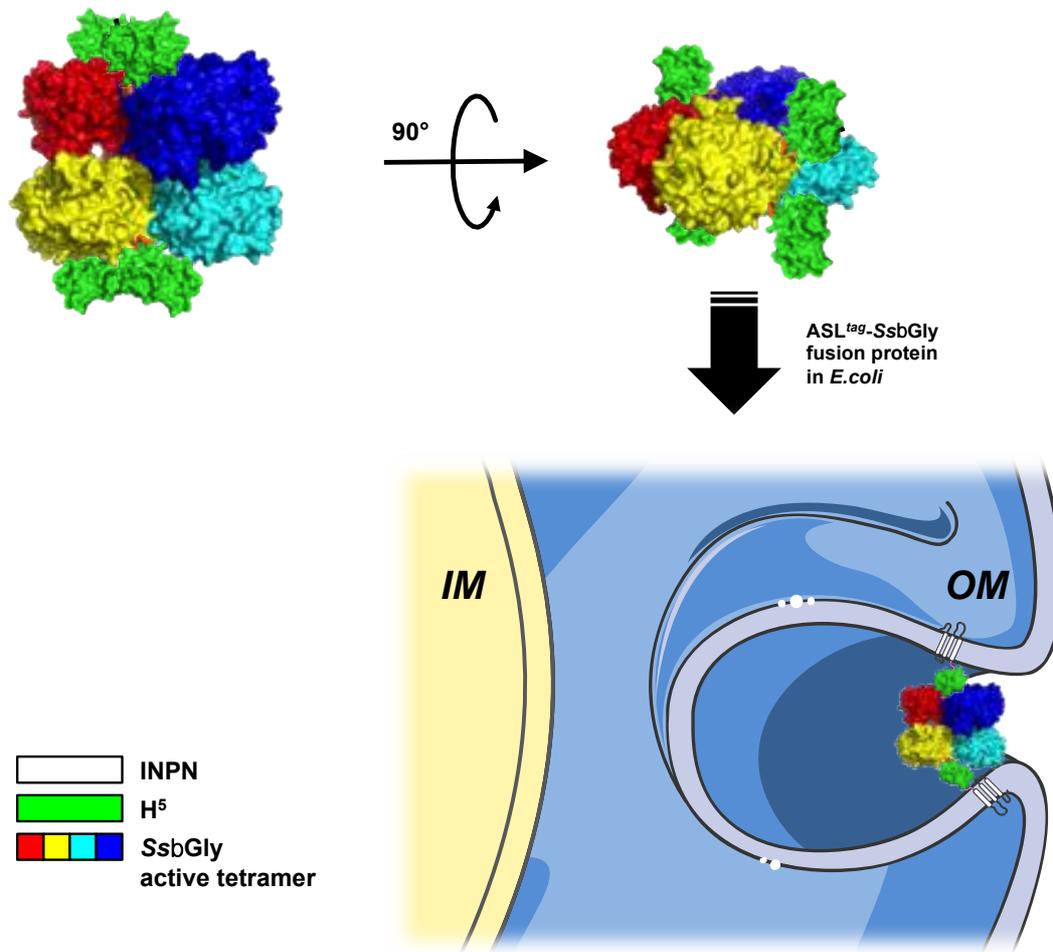
#### References

1. Tosa T, Mori T, Fuse N, Chibata I. Studies on continuous enzyme reactions. I. Screening of carriers for preparation of water-insoluble aminoacylase. *Enzymologia* 1966;31:214–24.
2. Zhou Z, Hartmann M. Progress in enzyme immobilization in ordered mesoporous materials and related applications. *Chem Soc Rev* 2013;42:3894–912.
3. Mohamad NR, Che Marzuki NH, Buang NA, et al. An overview of technologies for immobilization of enzymes and surface analysis techniques for immobilized enzymes. *Biotech Biotech Equip* 2015;29:205–20.
4. Nguyen HH, Kima M. An overview of techniques in enzyme immobilization. *App Sci Conv Tech* 2017;26:157–63.
5. Jakub ZJ, Meyer AS, Jesionowski T, Pinelo M. A general overview of support materials for enzyme immobilization: characteristics, properties, practical utility. *Catalysts* 2018;8:92.
6. Sirisha VL, Jain A, Chapter JA. Nine – enzyme immobilization: an overview on methods, support material, and applications of immobilized enzymes. *Adv Food Nut Res* 2016;79: 179–211.
7. Engin S, Trouillet V, Franz CM. Benzylguanine thiol self-assembled monolayers for the immobilization of SNAP-tag proteins on microcontact-printed surface structures. *Langmuir* 2010;26:6097–101.
8. Engin S, Fichtner D, Wedlich D, Fruk L. SNAP-tag as a tool for surface immobilization. *Curr Pharm Des* 2013;19:5443–8.
9. Meldal M, Schoffelen S. Recent advances in covalent, site-specific protein immobilization. *F1000Res*. 2016;5. pii: F1000 Faculty Rev-2303. eCollection 2016.
10. Pegg AE. Repair of O6-alkylguanine by alkyltransferases. *Mutat Res* 2000;462:83–100.
11. Daniels DS, Mol CD, Arvai AS, et al. Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. *Embo J* 2000;19:1719–30.
12. Daniels DS, Woo TT, Luu KX, et al. DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat Struct Mol Biol* 2004;11:714–20.
13. Tubbs JL, Pegg AE, Tainer JA. DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair* 2007;6:1100–15.
14. Keppler A, Gendreizig S, Gronemeyer T, et al. A general method for the covalent labeling of fusion proteins with small molecules *in vivo*. *Nat Biotech* 2003;21:86–9.
15. Gautier A, Juillerat A, Heinis C, et al. An engineered protein-tag for multi-protein labeling in living cells. *Chem Biol* 2008; 15:128–36.
16. Gronemeyer T, Chidley C, Juillerat A, et al. Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Prot Eng Des Sel* 2006;19:309–16.
17. Yu Q, Griss R, Schena A, Johnsson K. Highly modular bioluminescent sensors for small molecules and proteins. *Met Enz* 2017;589:365–82.
18. Huber W, Perspicace S, Kohler J, et al. SPR-based interaction studies with small molecular weight ligands using hAGT fusion proteins. *Anal Biochem* 2004;333:280–8.
19. Cochet N, Widehem P. Ice crystallization by *Pseudomonas syringae*. *Appl Microbiol Biotechnol* 2000;54:153–61.
20. Gurian-Sherman D, Lindow SE. Bacterial ice nucleation: significance and molecular basis. *FASEB J* 1993;7:1338–43.
21. Graether SP, Jia Z. Modeling *Pseudomonas syringae* ice-nucleation protein as a beta-helical protein. *Biophys J* 2001; 80:1169–73.
22. Samuelson P, Gunneriusson E, Nygren PA, Ståhl S. Display of proteins on bacteria. *J Biotechnol* 2002;96:129–54.
23. Lee SY, Choi JH, Xu Z. Microbial cell-surface display. *Trends Biotechnol* 2003;21:45–52.
24. Daugherty PS. Protein engineering with bacterial display. *Curr Opin Struct Biol* 2007;17:474–80.

25. Jung H-C, Park J-H, Park S-H, et al. Expression of carboxymethylcellulase on the surface of *Escherichia coli* using *Pseudomonas syringae* ice nucleation protein. *Enzyme Microb Technol* 1998;22:348–54.
26. Li L, Kang DG, Cha HJ. Functional display of foreign protein on surface of *Escherichia coli* using N-terminal domain of ice nucleation protein. *Biotechnol Bioeng* 2004;85:214–21.
27. Li Q, Yu Z, Shao X, et al. Improved phosphate biosorption by bacterial surface display of phosphate-binding protein utilizing ice nucleation protein. *FEMS Microbiol Lett* 2009;299:44–52.
28. Li Q, Yan Q, Chen J, et al. Molecular characterization of an ice nucleation protein variant (inaQ) from *Pseudomonas syringae* and the analysis of its transmembrane transport activity in *Escherichia coli*. *Int J Biol Sci* 2012;8:1097–108.
29. Xu Y, Liu Q, Zhou L, et al. Surface display of GFP by *Pseudomonas syringae* truncated ice nucleation protein in attenuated *Vibrio anguillarum* strain. *Mar Biotechnol* 2008;10:701–8.
30. Del Prete S, Peretto R, Rossi M, et al. A one-step procedure for immobilising the thermostable carbonic anhydrase (SspCA) on the surface membrane of *Escherichia coli*. *J Enz Inhib Med Chem* 2017;32:1120–8.
31. Nakagawa S, Shtaih Z, Banta A, et al. *Sulfurihydrogenibium yellowstonense* sp. nov., an extremely thermophilic, facultatively heterotrophic, sulfur-oxidizing bacterium from Yellowstone National Park, and emended descriptions of the genus *Sulfurihydrogenibium*, *Sulfurihydrogenibium subterraneum* and *Sulfurihydrogenibium azorense*. *Int J Syst Evol Microbiol* 2005;55:2263–8.
32. Akdemir A, Vullo D, De Luca V, et al. The extremo-alpha-carbonic anhydrase (CA) from *Sulfurihydrogenibium azorense*, the fastest CA known, is highly activated by amino acids and amines. *Bioorg Med Chem Lett* 2013;23:1087–90.
33. Capasso C, De Luca V, Carginale V, et al. Biochemical properties of a novel and highly thermostable bacterial alpha-carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* YO3AOP1. *J Enzyme Inhib Med Chem* 2012;27:892–789.
34. Peretto R, Del Prete S, Vullo D, et al. Production and covalent immobilisation of the recombinant bacterial carbonic anhydrase (SspCA) onto magnetic nanoparticles. *J Enzyme Inhib Med Chem* 2017;32:759–66.
35. Perugino G, Vettone A, Illiano G, et al. Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage. *J Biol Chem* 2012;287:4222–31.
36. Vettone A, Serpe M, Hidalgo A, et al. A novel thermostable protein-tag: optimization of the *Sulfolobus solfataricus* DNA-alkyl-transferase by protein engineering. *Extremophiles* 2016;20:13.
37. Visone V, Han W, Perugino G, et al. In vivo and in vitro protein imaging in thermophilic archaea by exploiting a novel protein tag. *PLoS One* 2017;12:e0185791.
38. Studier FW. Protein production by auto-induction in high density shaking cultures. *Prot Expr Purif* 2005;41:207–34.
39. Miggiano R, Valenti A, Rossi F, et al. Every OGT is illuminated ... by fluorescent and synchrotron lights. *Int J Mol Sci* 2017;18:2613–31.
40. Glazyrina J, Materne EM, Dreher T, et al. High cell density cultivation and recombinant protein production with *Escherichia coli* in a rocking-motion-type bioreactor. *Microb Cell Fact* 2010;9:42–52.
41. Perugino G, Trincon A, Giordano A, et al. Activity of hyperthermophilic glycosynthases is significantly enhanced at acidic pH. *Biochemistry* 2003;42:8484–93.
42. Moracci M, Capalbo L, Ciaramella M, Rossi M. Identification of two glutamic acid residues essential for catalysis in the beta-glycosidase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Protein Eng* 1996;9:1191–5.
43. Aguilar C, Sanderson I, Moracci M, et al. Crystal structure of the beta-glycosidase from the hyperthermophilic archeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability. *J Mol Biol* 1997;271:789–802.
44. Pouwels J, Moracci M, Cobucci-Ponzano B, et al. Activity and stability of hyperthermophilic enzymes: a comparative study on two archaeal-glycosidases. *Extremophiles* 2000; 4:157–64.
45. Moracci M, Trincon A, Perugino G, et al. Restoration of the activity of active-site mutants of the hyperthermophilic-glycosidase from *Sulfolobus solfataricus*: dependence of the mechanism on the action of external nucleophiles. *Biochemistry* 1998; 37:17262–70.
46. Perugino G, Miggiano R, Serpe M, et al. Structure-function relationships governing activity and stability of a DNA alkylation damage repair thermostable protein. *Nuc Ac Res* 2015; 43:8801–16.
47. Rossi F, Morrone C, Massarotti A, et al. Crystal structure of a thermophilic O6-alkylguanine-DNA alkyltransferase-derived self-labeling protein-tag in covalent complex with a fluorescent probe. *Bioch Biophys Res Comm* 2018;500:698–703.



**Figure S1.** *Expression of free H<sup>5</sup> in BL21(DE3) strain.* Whole IPTG-induced cells transformed with pQE-*ogtH<sup>5</sup>* plasmid (lane 1)<sup>36</sup> were incubated and loaded on SDS-PAGE for the *gel-imaging* and Coomassie staining analyses, as described in the Materials and Methods. Lane 2 corresponds to the 1.0  $\mu\text{g}$  of the purified free H<sup>5</sup> enzyme. M, molecular weight marker. All used symbols are described in Figure 3a.



**Figure S2.** *The tetrameric form of the ASL<sup>tag</sup>-SsbGly fusion protein.* Schematic representation of the possible spatial disposition of the tetrameric SsbGly (PDB ID: 4GOW<sup>43</sup>) linked to four H<sup>5</sup> units (PDB ID: 6GA0<sup>47</sup>), taking into account of the exposed first methionine residue of each SsbGly monomer (*in orange*). The hypothesized invagination of the external membrane of *E. coli* would makes it possible the assembling of the tetrameric form and the consequent measured catalytic activity of the SsbGly.



## Thermostability enhancement of the $\alpha$ -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* by using the anchoring-and-self-labelling-*protein-tag* system (ASL<sup>tag</sup>)

Sonia Del Prete, Rosa Merlo, Anna Valenti, Rosanna Mattosovich, Mosè Rossi, Vincenzo Carginale, Claudiu T. Supuran, Giuseppe Perugino & Clemente Capasso

To cite this article: Sonia Del Prete, Rosa Merlo, Anna Valenti, Rosanna Mattosovich, Mosè Rossi, Vincenzo Carginale, Claudiu T. Supuran, Giuseppe Perugino & Clemente Capasso (2019) Thermostability enhancement of the  $\alpha$ -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* by using the anchoring-and-self-labelling-*protein-tag* system (ASL<sup>tag</sup>), *Journal of Enzyme Inhibition and Medicinal Chemistry*, 34:1, 946-954, DOI: [10.1080/14756366.2019.1605991](https://doi.org/10.1080/14756366.2019.1605991)

To link to this article: <https://doi.org/10.1080/14756366.2019.1605991>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 30 Apr 2019.



Submit your article to this journal [↗](#)



View Crossmark data [↗](#)

RESEARCH ARTICLE

 OPEN ACCESS 

## Thermostability enhancement of the $\alpha$ -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* by using the anchoring-and-self-labelling-protein-tag system (ASL<sup>tag</sup>)

Sonia Del Prete<sup>a\*</sup>, Rosa Merlo<sup>a\*</sup>, Anna Valenti<sup>a</sup>, Rosanna Mattosovich<sup>a</sup>, Mosè Rossi<sup>a</sup>, Vincenzo Carginale<sup>a</sup>, Claudiu T. Supuran<sup>b</sup> , Giuseppe Perugino<sup>a</sup> and Clemente Capasso<sup>a</sup> 

<sup>a</sup>Department of Biology Agriculture and Food Sciences, Institute of Bioscience and BioResources – National Research Council of Italy, Naples, Italy; <sup>b</sup>Neurofarba Department, University of Florence, Polo Scientifico, Sesto Fiorentino Firenze, Italy

### ABSTRACT

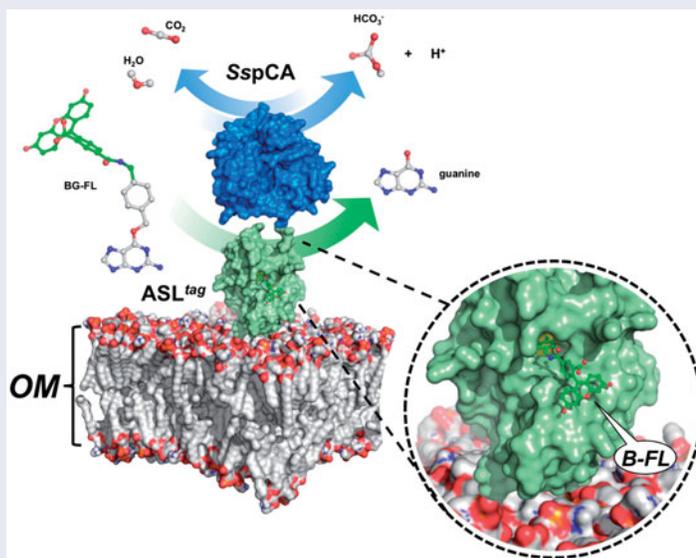
Carbonic anhydrases (CAs, EC 4.2.1.1) are a superfamily of ubiquitous metalloenzymes present in all living organisms on the planet. They are classified into seven genetically distinct families and catalyse the hydration reaction of carbon dioxide to bicarbonate and protons, as well as the opposite reaction. CAs were proposed to be used for biotechnological applications, such as the post-combustion carbon capture processes. In this context, there is a great interest in searching CAs with robust chemical and physical properties. Here, we describe the enhancement of thermostability of the  $\alpha$ -CA from *Sulfurihydrogenibium yellowstonense* (SspCA) by using the anchoring-and-self-labelling-protein-tag system (ASL<sup>tag</sup>). The anchored chimeric H<sup>5</sup>-SspCA was active for the CO<sub>2</sub> hydration reaction and its thermostability increased when the cells were heated for a prolonged period at high temperatures (e.g. 70 °C). The ASL<sup>tag</sup> can be considered as a useful method for enhancing the thermostability of a protein useful for biotechnological applications, which often need harsh operating conditions.

### ARTICLE HISTORY

Received 25 February 2019  
Revised 3 April 2019  
Accepted 5 April 2019

### KEYWORDS

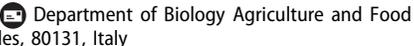
Carbonic anhydrase; thermostability; *Sulfurihydrogenibium yellowstonense*; alkylguanine-DNA-alkyl-transferase; ASLtag



## 1. Introduction

The hydration/dehydration reaction involving carbon dioxide, water, bicarbonate, and protons ( $\text{CO}_2 + \text{H}_2\text{O} \rightleftharpoons \text{HCO}_3^- + \text{H}^+$ ) is a fundamentally important process for the planet and all its associated forms of life<sup>1–9</sup>. The dissolution of CO<sub>2</sub> in the aqueous phase develops carbonic acid (H<sub>2</sub>CO<sub>2</sub>), which is subject to an ionisation

reaction producing bicarbonate (HCO<sub>3</sub><sup>−</sup>), whereas this last species then generates carbonate (CO<sub>2</sub>) through a second dissociation reaction. These species are disseminated in the fluids of the all living organisms and are involved in a large number of physiological processes, such as some biosynthetic pathways, photosynthesis, respiration, pH homeostasis, secretion of electrolytes, etc.<sup>9–11</sup>. At

**CONTACT** Clemente Capasso  [clemente.capasso@ibbr.cnr.it](mailto:clemente.capasso@ibbr.cnr.it); Giuseppe Perugino  [giuseppe.perugino@ibbr.cnr.it](mailto:giuseppe.perugino@ibbr.cnr.it) 

\*These authors equally contributed to this work.

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

physiological pH, the naturally uncatalysed CO<sub>2</sub> hydration reaction has a catalytic constant ( $k_{\text{cat}}$ ) of 0.15 s<sup>-1</sup>, whereas the uncatalysed dehydration shows a  $k_{\text{cat}}$  of 50.0 s<sup>-1</sup><sup>12,13</sup>. These values are typical of slow reactions and are not sufficient for accomplishing fast cellular physiological processes which support metabolic activities dependent on the dissolved inorganic carbon species (CO<sub>2</sub>, HCO<sub>3</sub><sup>-</sup>, CO<sub>3</sub><sup>2-</sup>)<sup>9</sup>. Probably this is the reason why living organisms evolved a superfamily of ubiquitous metalloenzymes, the carbonic anhydrases (CAs, EC 4.2.1.1), which catalyse, and highly accelerate, the above-mentioned reactions, at a very high rate with respect to the non-catalysed reaction. CAs show kinetic constants  $k_{\text{cat}}$  varying from 10<sup>4</sup> to 10<sup>6</sup> s<sup>-1</sup> for the hydration reaction<sup>12,13</sup>.

Up to date, CA superfamily contains seven genetically distinct families (or classes), named  $\alpha$ -,  $\beta$ -,  $\gamma$ -,  $\delta$ -,  $\zeta$ -,  $\eta$ -, and  $\theta$ -CAs<sup>9,14,15</sup>, characterised by multiple transcript variants and protein isoforms, with different biochemical properties and specific tissue/organ and sub-cellular localisations<sup>7,9,12,16-19</sup>. Generally, only  $\alpha$ -class enzymes are present in the animals<sup>20,21</sup>, whereas  $\alpha$ -,  $\beta$ -,  $\gamma$ -,  $\delta$ - and  $\theta$ -classes are found in plants and algae;  $\alpha$ - and  $\beta$ -CAs in fungi;  $\alpha$ -,  $\beta$ -, and/or  $\eta$ -CAs in protozoans;  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CA classes in bacteria<sup>7,19,15,22-25</sup>.

Studies carried out on the bacterial CAs concern two main aspects. They are considered an attractive and rather new drug target, because their inhibition affects the growth or virulence of many pathogens<sup>4,7,26-28</sup>. Furthermore, they are biocatalysts often used in biotechnological applications<sup>29,30</sup>, such as the post-combustion carbon capture process, artificial lungs, and biosensors<sup>31,32</sup>. Many such processes are characterised by conditions, which may be deleterious to an enzyme belonging to the mesophilic organisms<sup>25,33-45</sup>. In the field of biotechnology, there is a great interest in searching proteins with robust chemical and physical properties, which resist the hard conditions of industrial processes. In this context, our groups identified in the genome of the extreme thermophiles *Sulfurihydrogenibium yellowstonense* and *Sulfurihydrogenibium azorense* two CAs, indicated with the acronyms SspCA and SazCA, respectively. It has been demonstrated that these two CAs belong to the  $\alpha$ -CA class and showed an excellent activity as catalysts for the CO<sub>2</sub> hydratase reaction ( $k_{\text{cat}} = 10^5 - 10^6 \text{ s}^{-1}$ )<sup>30,46-52</sup>. Interestingly, the two extreme enzymes resulted to be highly thermostable, retaining an excellent catalytic activity when heated for a prolonged period at a temperature higher than 80 °C<sup>30,46-52</sup>. The X-ray tridimensional structures of the two proteins demonstrated that the high compactness of the dimeric structure, the higher content of secondary-structural elements, the increased number of charged residues on the protein surface, and the vast number of ionic networks with respect to the mesophilic counterparts, are the main structural elements responsible for the protein thermostability<sup>29,30</sup>. Moreover, Russo et al. reported the use of free SspCA in experiment of CO<sub>2</sub> absorption<sup>53</sup> demonstrating that it is an excellent candidate for the biomimetic capture of CO<sub>2</sub>. Subsequently, the necessity to use this biocatalyst repeatedly and continuously, led to the immobilisation of the recombinant SspCA on polyurethane foam (PU), a pre-polymer of polyethylene glycol<sup>54</sup>; onto supported ionic liquid membranes (SMLs), in order to realise a system able to selectively separate and transform CO<sub>2</sub><sup>55</sup>. Furthermore, the immobilisation onto magnetic support for recovering the biocatalyst from the bioreactor effortlessly and practically, for example through the use of a magnet, was also proposed for these thermostable CAs<sup>56</sup>. Unfortunately, these strategies may discourage the wide utilisation of enzymes in industrial applications because of the high costs connected to the biocatalyst production and purification, and the expenses for the preparation of the immobilisation support.

Thus, to overcome this limitation, a *one-step immobilisation procedure* has been proposed, which consists in the overexpression of SspCA directly onto the surface of bacterial hosts, by using the ice nucleation protein (INP) from the Gram-negative bacterium *Pseudomonas syringae*<sup>57</sup>.

In this article, we describe the improvement of the thermostability of SspCA by using a novel *protein-tag* system, the ASL<sup>tag</sup><sup>58</sup>. The anchored SspCA was fused to the thermostable variant of the alkylguanine-DNA-alkyl-transferase (H<sup>5</sup>) from the hyperthermophilic archaeon *Sulfolobus solfataricus*. The chimeric H<sup>5</sup>-SspCA was efficiently overexpressed on the bacterial surface of *Escherichia coli*. The protonography technique showed that the neosynthesised H<sup>5</sup>-SspCA was active for the CO<sub>2</sub> hydration reaction. Even more intriguing, the chimeric H<sup>5</sup>-SspCA expressed onto the bacterial surface resulted to be more stable with respect to the non-chimeric SspCA, when treated at high temperatures (50.0 and 70.0 °C) for a prolonged time. The ASL<sup>tag</sup> system may thus be considered as a brilliant strategy to further increase the thermostability of proteins to be used in biotechnological applications, in which a highly effective and thermostable catalyst is needed.

## 2. Materials and methods

### 2.1 Construction of vectors for surface fusion and H<sup>5</sup>-SspCA overexpression

The vector pET-22b/INPN-SspCA was used to produce the pET-ASL<sup>tag</sup>-SspCA vector, which overexpressed onto the bacterial surface the chimeric H<sup>5</sup>-SspCA. The pET-22b/INPN-SspCA and pET-ASL<sup>tag</sup>-SspCA vectors were prepared as described previously<sup>57,58</sup>. For overexpressing the chimeric H<sup>5</sup>-SspCA or SspCA on the bacterial cell surface, competent *E. coli* BL21 (DE3) cells were transformed with the above-mentioned constructs. They were grown at 37.0 °C and induced with 1.0 mM isopropyl-thio- $\beta$ -D-galactoside (IPTG) and 0.5 mM ZnSO<sub>4</sub> at an OD<sub>600</sub> of 0.6–0.7. After additional growth for 6 h, the cells were harvested by centrifugation and washed three times with PBS 1×. Aliquots of cells were resuspended in 25 mM Tris/HCl and used to determine the enzyme activity and for the preparation of the outer membrane fraction.

### 2.2 Carbonic anhydrase assay and SDS-PAGE

CA activity assay was a modification of the procedure described by Capasso et al.<sup>59</sup>. Briefly, the assay was performed at 0 °C using CO<sub>2</sub> as substrate and following the pH variation due to the catalysed conversion of CO<sub>2</sub> to bicarbonate. Bromothymol blue was used as the indicator of pH variation. The production of hydrogen ions during the CO<sub>2</sub> hydration reaction lowers the pH of the solution until the colour transition point of the dye is reached. The time required for the colour change is inversely related to the quantity of CA present in the sample. Wilbur-Anderson units (WAU) were calculated according to the following definition: one WAU of activity is defined as (T<sub>0</sub>–T)/T, where T<sub>0</sub> (uncatalysed reaction) and T (catalysed reaction) are recorded as the time (in seconds) required for the pH to drop from 8.3 to the transition point of the dye in a control buffer and in the presence of enzyme, respectively. Assay of the membrane-bound enzyme (H<sup>5</sup>-SspCA or SspCA) was carried out using an amount of whole cells or outer membranes ranging from 1.0 to 5.0 mg. Sodium dodecyl sulphate (SDS)-polyacrylamide gel electrophoresis (PAGE) was performed as described by Laemmli using 12%

gels.<sup>60</sup> Samples were dissolved in buffer with 5%  $\beta$ -mercaptoethanol. The gel was stained with Coomassie blue and protein concentration was determined by Bio-Rad assay kit (Bio-Rad, Hercules, CA).

### 2.3 Protonography and his-tag Western blotting

To perform the protonography, wells of 12% SDS-gel were loaded with solubilised outer membranes having on their surface  $H^5$ -SspCA or SspCA, and a solution of free SspCA (enzyme overexpressed as cytoplasmic protein and purified as described previously<sup>59</sup>). Samples were mixed with loading buffer without 2-mercaptoethanol and without boiling the samples, to solubilise cells and avoid protein denaturation. The gel was run at 180 V until the dye front moved off the gel. Following the electrophoresis, the 12% SDS-gel was subject to protonography to detect the cytoplasmic SspCA, the surface-SspCA, and surface- $H^5$ -SspCA hydratase activity on the gel as described by Del Prete et al.<sup>61,62</sup> and De Luca et al.<sup>63</sup>. To perform the Western-Blot, after a 12% (w/v) SDS-PAGE, the overexpressed cytoplasmic SspCA and the membrane-bound enzymes (SspCA and  $H^5$ -SspCA) were also electrophoretic transferred to a PVDF membrane with transfer buffer (25 mM Tris, solubilised whole cells 192 mM glycine, 20% methanol) by using Trans-Plot SD Cell (Bio-Rad, Hercules, CA). His-tag Western blot was carried out using the Pierce Fast Western Blot Kit (Thermo Scientific, Waltham, MA). The blotted membrane has been placed in the wash blot solution Fast Western 1 $\times$  Wash Buffer to remove transfer buffer. Primary Antibody Working Dilution was added to the blot and incubated for 30.0 min at room temperature (RT) with shaking. After, the blot was removed from the primary antibody solution and incubated for 10.0 min with the Fast Western Optimized HRP Reagent Working Dilution. Subsequently, the membrane was washed two times in about 20 ml of Fast Western 1 $\times$  Wash Buffer. Finally, the membrane was incubated with the Detection Reagent Working Solution and incubated for 3.0 min at RT and then developed with X-ray film.

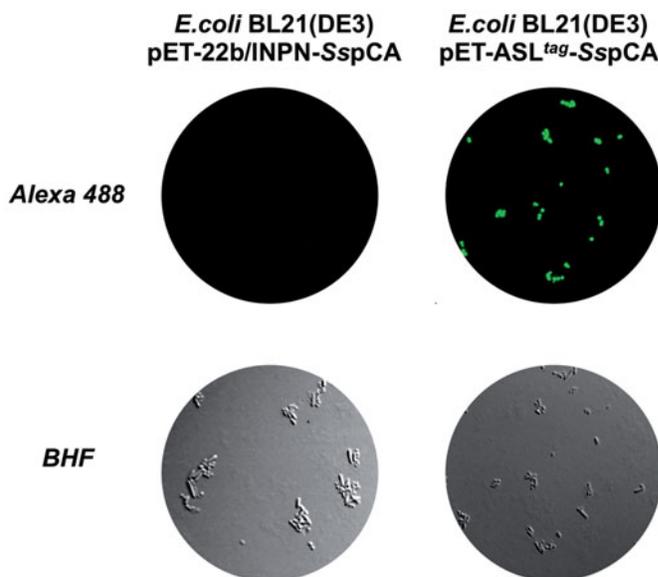
### 2.4 Determination of the $H^5$ activity by an *in vitro* and *in vivo* fluorescent assay

Whole overnight induced *E. coli* BL21(DE3) cells were collected and the expression of the  $H^5$ -derived fusion proteins was analysed by an *in vitro* assay with the fluorescent SNAP-Vista Green<sup>TM</sup> substrate (New England Biolabs, Ipswich, MA; hereinafter BG-FL), as previously described<sup>58,64</sup>. The *in vivo* imaging was carried out as described by Merlo et al.<sup>58</sup>. Briefly, bacterial cells expressing the  $H^5$ -SspCA onto cell surface were washed twice in PBS 1 $\times$  and resuspended in 50.0  $\mu$ l of the same buffer supplemented with 5.0  $\mu$ M of the BG-FL. After incubation at 37.0 °C for 30.0 min, cells were washed twice, resuspended, and again incubated for 30.0 min at 37.0 °C, to allow the external diffusion of the unreacted substrate. Images were collected using a DM6 fluorescence microscope and Hamamatsu camera under the control of Leica LAS AF 6000 software; excitation and emission wavelengths used suitably for AlexaFluor488 dye were  $\lambda_{ex}$  = 490 nm;  $\lambda_{em}$  = 525 nm, respectively.

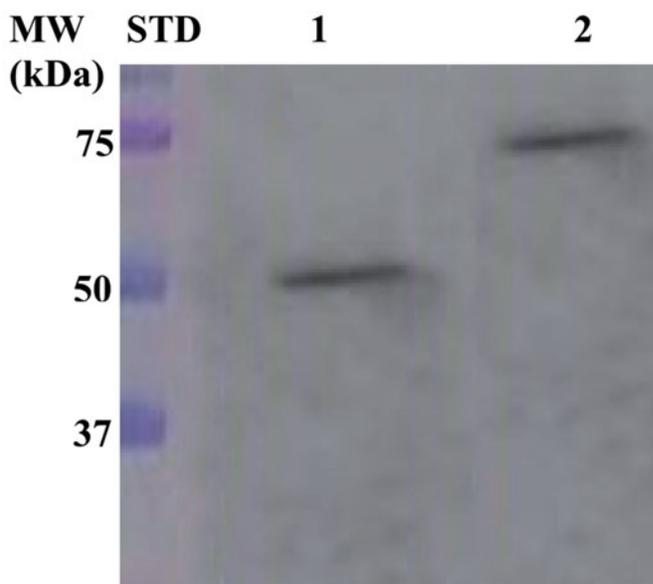
### 2.5 Outer membrane preparation

The bacterial outer membranes were fractioned by inner membranes as described previously by Del Prete et al.<sup>57</sup>. Briefly, 2.0 g

of harvested bacterial cells were resuspended and disrupted by sonication on the ice. Cell extract was ultracentrifuged to recover the total membrane fraction. The outer membrane fraction was purified resuspending the pellet in phosphate-buffered saline (PBS 1 $\times$ ) containing 0.01 mM  $MgCl_2$  and 2% Triton X-100 and incubated at RT for 30.0 min to solubilise the inner membrane. The outer membrane fraction was then pelleted by ultracentrifugation at 120,000 $\times g$  and used for further experiments.



**Figure 1.** Fluorescence microscopy of *E. coli* BL21(DE3) cells transformed with pET-22b/INPN-SspCA (left) or with pET-ASL<sup>tag</sup>-SspCA (right). The cells were incubated with BG-FL and then analysed by fluorescence microscopy. Images show bright field (BHF) and AlexaFluor488 (green). As expected, the fluorescence is only evidenced for the bacterial cell transformed with the ASL<sup>tag</sup> system.



**Figure 2.** Western Blot performed on the outer membrane purified from the whole bacterial cells. The anti His-tag antibody was raised against the C-terminus of His-tagged SspCA. Legend: Lane Std, molecular markers, M.W. starting from the top: 75.0, 50.0, and 37.0 kDa; Lane 1, anchored SspCA; Lane 2, anchored  $H^5$ -SspCA.

## 2.6 Temperature stability studies

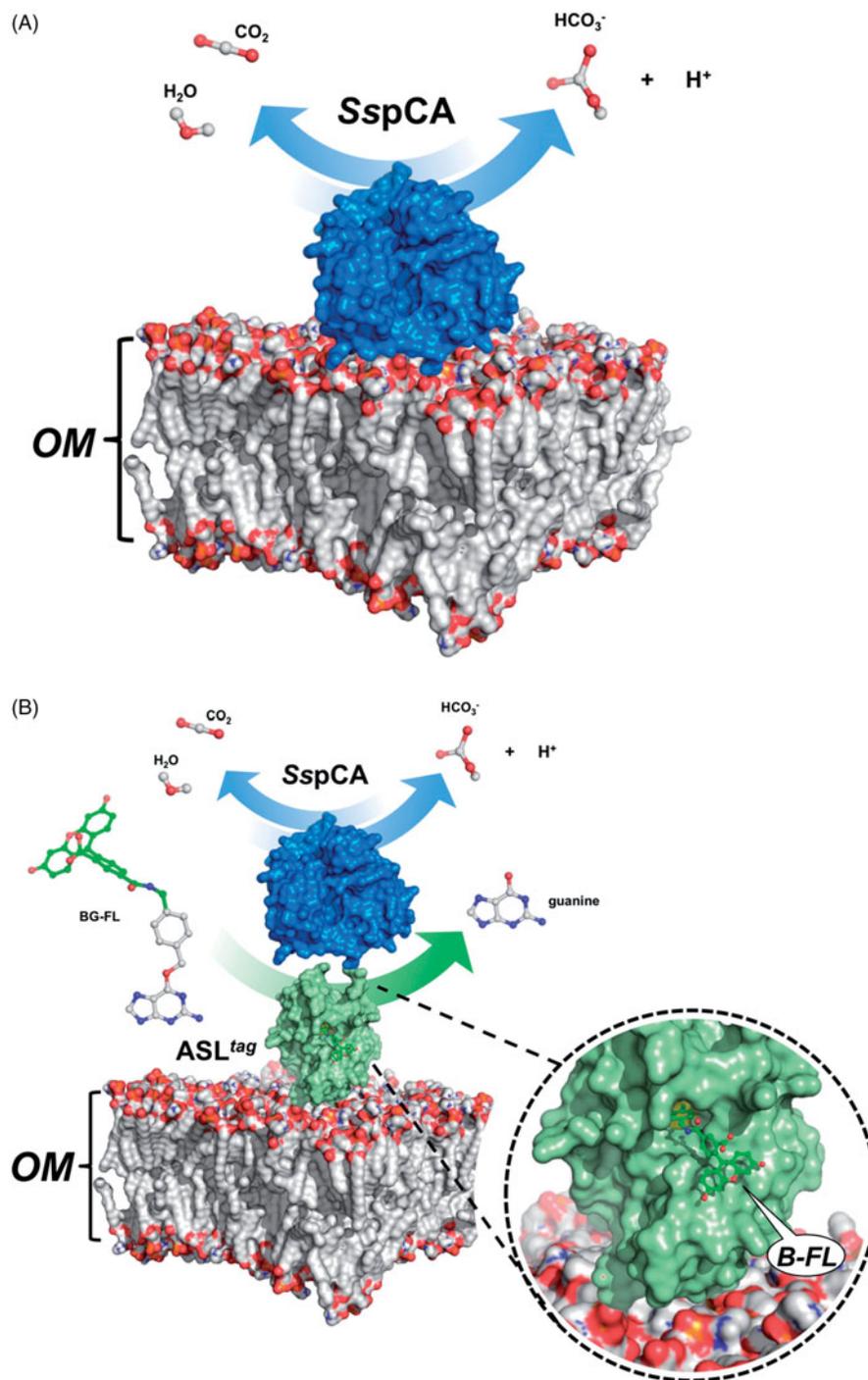
### 2.6.1 Thermostability

Bacterial cells (2.0 g/20 ml) were incubated at 25.0, 50.0, and 70.0 °C for different time up to 24 h to compare the stability of the membrane-bound enzymes (SspCA and H<sup>5</sup>-SspCA) at the above-indicated temperatures. Cell membrane-bound enzymes aliquots were withdrawn at appropriate times and the residual activity was measured using CO<sub>2</sub> as the substrate. All data have been analysed using GraphPad Prism version 5.0 software (GraphPad

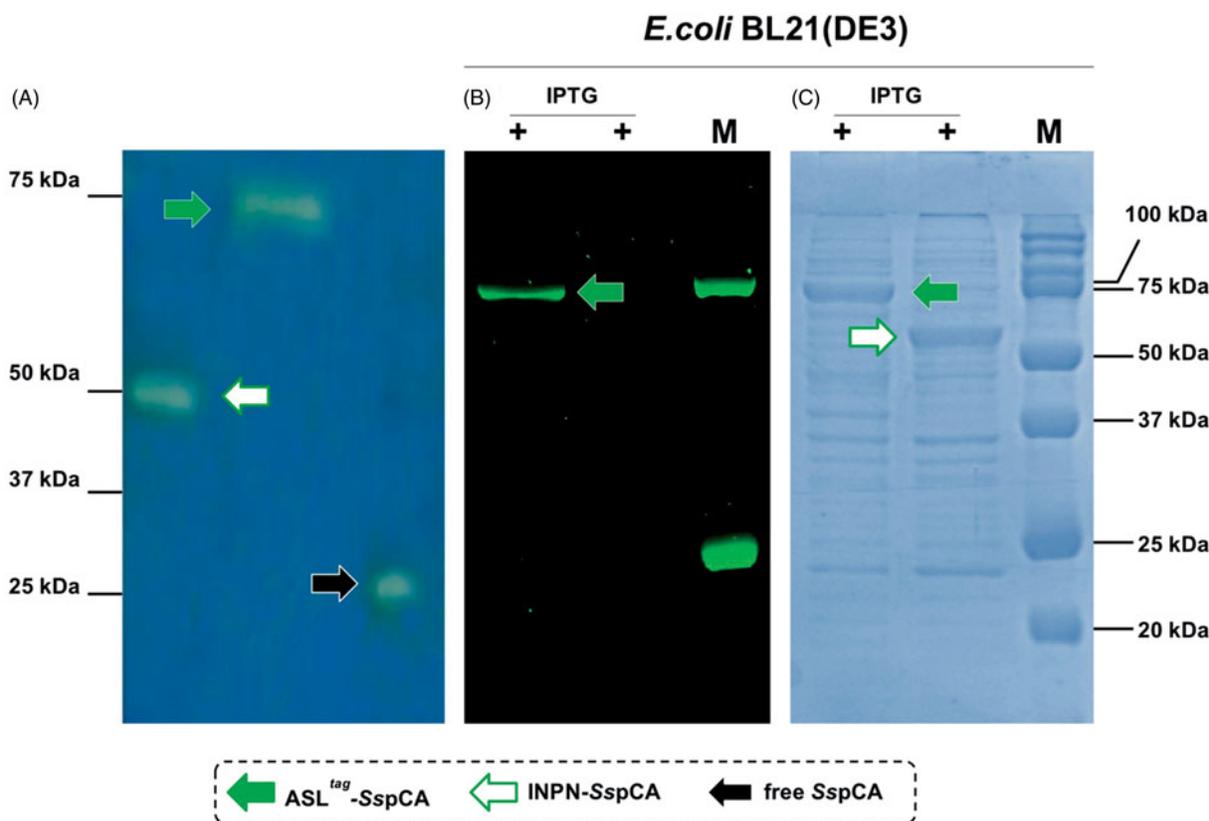
Software, San Diego, CA). Curves were obtained by the mean of three independent determinations.

### 2.6.2 Long-term stability

Membrane-bound SspCA or H<sup>5</sup>-SspCA were investigated for their long-term stability at different temperatures (25.0, 50.0, and 70.0 °C) by assaying their hydratase residual activities using CO<sub>2</sub> as substrate and withdrawing aliquots of cell surface SspCA or



**Figure 3.** Model representation of an outer membrane fraction (OM; pdb from Tieleman and Berendsen<sup>65</sup>) describing the *in vivo* immobilisation of SspCA (*in blue*; PDB ID: 4G7A; panel A) and in fusion with H<sup>5</sup> (*in green*; PDB ID: 6GA0; panel B). The INPN domain is omitted because inserted in the OM. The catalytic reaction of SspCA (the hydration/dehydration of CO<sub>2</sub>) and H<sup>5</sup> (the conversion of BG-FL in the free guanine and the fluorescent benzyl-guanine derivative, B-FL, covalently linked to the active site of H<sup>5</sup>) are also shown.



**Figure 4.** Protonography (Panel A), fluorescence *gel-imaging* (Panel B) and Coomassie staining (Panel C) of SspCA and H<sup>5</sup>-SspCA carried out with different amounts of the whole *E. coli* cells (see Materials and Methods). Filled green, white and black arrows represent the ASL<sup>tag</sup>-SspCA, INPN-SspCA and the free SspCA, respectively.

H<sup>5</sup>-SspCA at appropriate times. All the buffers were sterilised by using a sterile 0.22  $\mu\text{m}$  filter, while samples containing the membrane-bound enzymes were treated with a diluted solution of NaN<sub>3</sub> to avoid contamination. All data were obtained by the mean of 3 independent determinations.

### 3. Results and discussion

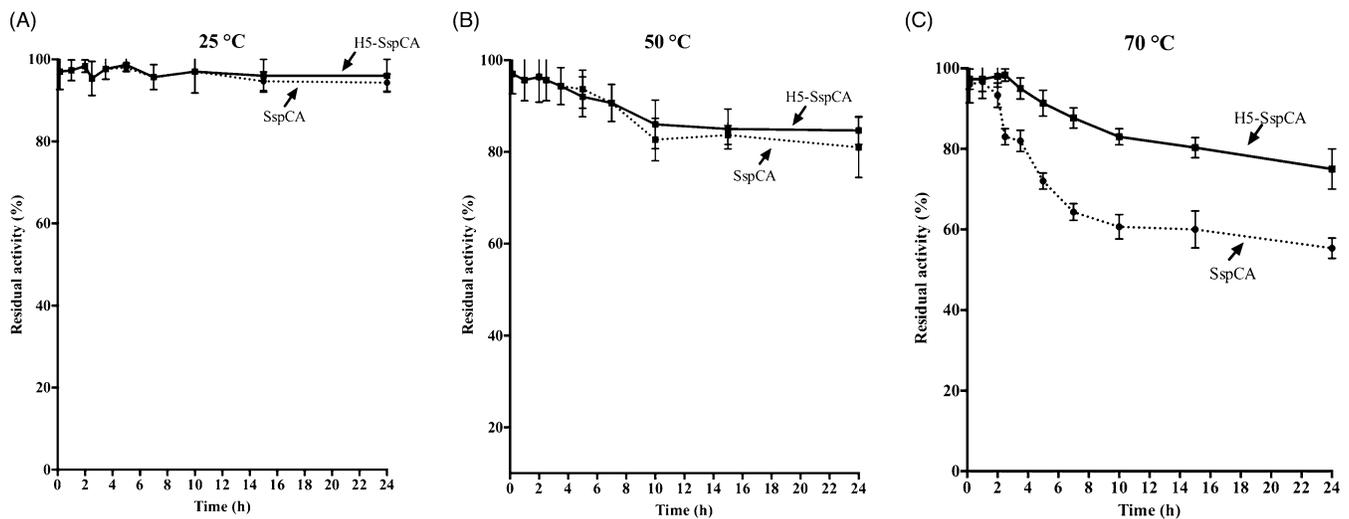
#### 3.1 Expression and surface localisation of SspCA and H<sup>5</sup>-SspCA

Expression of the SspCA and H<sup>5</sup>-SspCA was realised through the *one-step procedure*, by transforming the *E. coli* cells with the construct expressing a gene composed of a signal peptide (necessary for the periplasmic translocation of the protein), the *P. syringae* INPN domain (fundamental for displaying the overexpressed protein onto the bacterial surface), and the protein of interest (SspCA or H<sup>5</sup>-SspCA). This strategy has the advantage to overexpress and directly immobilise *in vivo* the  $\alpha$ -CA or other proteins on the bacterial cell surface. Besides, the system expressing the H<sup>5</sup>-SspCA, named anchoring-and-self-labelling-protein-tag (ASL<sup>tag</sup>), allowed the labelling of the neosynthesised protein fused to H<sup>5</sup> through the use of the fluorescein derivative of the O<sup>6</sup>-BG (BG-FL), which is the substrate of H<sup>5</sup>. As reported in Figure 1, the expression of the chimeric H<sup>5</sup>-SspCA on the bacterial surface has been confirmed using the H<sup>5</sup> substrate and analysing the whole bacterial cells with fluorescent microscopy. The irreversible reaction of the ASL<sup>tag</sup> system with a fluorescent substrate allowed the quantitative determination of the immobilised bacterial  $\alpha$ -CA or of other proteins fused to H<sup>5</sup>, by *in vitro* gel-imaging techniques as described by Del Prete et al.<sup>57</sup> and Merlo et al.<sup>58</sup>. Diversely from H<sup>5</sup>-SspCA, the expression of the anchored His-tagged SspCA

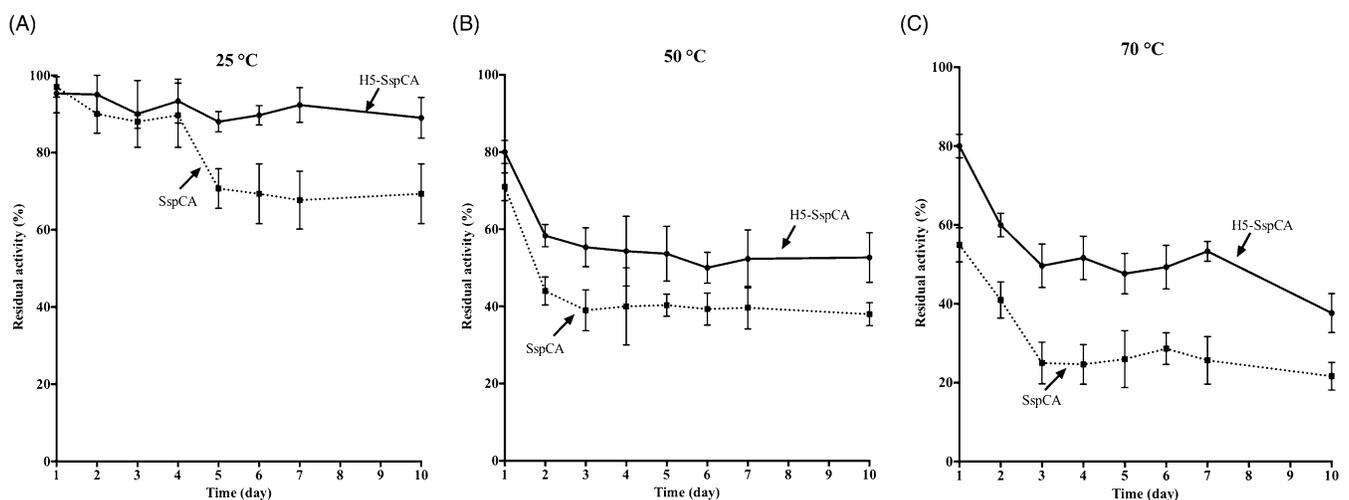
(without the H<sup>5</sup>) has been confirmed only by the Western Blot analysis using an anti-His-tag antibody (Figure 2), indicating an expected molecular weight of 50.0 kDa (the sum of the INPN and SspCA polypeptide chains produced with the construct pET-22b/INPN-SspCA; see the experimental section). Anchored His-tagged H<sup>5</sup>-SspCA showed a higher molecular weight (70.0 kDa) with respect to the non-chimeric protein because of the presence of the H<sup>5</sup> protein (158 amino acid residues). The H<sup>5</sup>-SspCA Western-Blot fully supports the fluorescence microscopy results. Thus, using this *one-step procedure*, the thermostable proteins  $\alpha$ -CA (SspCA) and the chimeric ASL<sup>tag</sup>-SspCA<sup>56</sup> were efficiently expressed on the external side of the bacterial outer membrane. Figure 3 reports a model representing the *in vivo* immobilisation of SspCA (Panel A) and chimeric H<sup>5</sup>-SspCA (Panel B) on the bacterial external cell surface. Moreover, Figure 3 shows the reactions catalysed by both the biocatalysts (the hydration/dehydration of CO<sub>2</sub> and the conversion of BG-FL in the free guanine and the fluorescent benzyl-guanine derivative covalently linked to the active site of H<sup>5</sup>).

#### 3.2 Hydratase activity of the membrane anchored SspCA and H<sup>5</sup>-SspCA

Using CO<sub>2</sub> as the substrate, the hydratase activity of all the forms of SspCA has been investigated in solution<sup>57</sup>. The results showed that the membrane-bound SspCA with and without H<sup>5</sup> was an active enzyme, when immobilised on the bacterial surface. The CO<sub>2</sub> hydratase activity of SspCA and H<sup>5</sup>-SspCA did not show any differences. The results also evidenced that 1.0  $\mu\text{g}$  of bacterial cells had a CO<sub>2</sub> hydratase activity corresponding to that of 10.0 ng of



**Figure 5.** The thermostability of immobilised SspCA and H<sup>5</sup>-SspCA on the bacterial surface. Measures were carried out at indicated temperatures, by using aliquots of the whole cells incubated up to 24 h. Legend: continuous line, membrane-bound H<sup>5</sup>-SspCA; dashed line, membrane-bound SspCA. Each point is the mean of three independent determinations.



**Figure 6.** The long-term stability of immobilised SspCA and H<sup>5</sup>-SspCA on the bacterial surface. Measures were carried out at indicated temperatures up to 10 d, using aliquots of whole bacterial cells. Continuous line: free SspCA; Dashed line: membrane-bound SspCA. Each point is the mean of three independent determinations.

free SspCA. Probably, anchored SspCA or H<sup>5</sup>-SspCA is subjected to various phenomena, which influence the enzymatic reaction, e.g. a reduction of the structural conformational changes (this is typical of an immobilised enzyme); a different substrate access to the active site with respect to the free biocatalyst due to the bacterial cell surface microenvironment, and, finally, an aggregation of the cells or outer membranes used in the assay. Otherwise, the activity of SspCAs was compared by using the protonography, which is a technique able to reveal the hydrogen ions produced by the hydratase activity reaction as a yellow band on the SDS-PAGE. The protonography results showed that the all the forms of SspCA (the two membrane-bound ones and the free enzyme) had a comparable enzyme activity and a different molecular weight on SDS-PAGE (Figure 4, panel A and C). Protonography corroborated the results obtained with the fluorescent microscopy (Figure 1) and Western Blot (Figure 2). Interestingly, H<sup>5</sup>-SspCA fluorescent band at a molecular weight of about 70.0 kDa (Figure 4, panel B)

indicated that the presence of SspCA does not affect the activity of the thermostable H<sup>5</sup> enzyme on the BG-FL substrate.

### 3.3 Stability of SspCA and H<sup>5</sup>-SspCA linked to the bacterial cell surface

Using the whole bacterial cells expressing on the external surface SspCA or H<sup>5</sup>-SspCA, the effect of the CO<sub>2</sub> hydratase reaction as a function of temperature has been investigated. In Figure 5, the residual activity of the SspCA and H<sup>5</sup>-SspCA remained almost constant at 25.0 and 50.0 °C, retaining their residual activity at 100% up to 24 h (panel A) and at 70% up to 6 h of incubation (panel B), respectively. In contrast, it is readily apparent that at higher temperatures (70.0 °C) SspCA and H<sup>5</sup>-SspCA behave differently (Figure 5, panel C): the residual activity of SspCA started to decline rapidly after 2 h, getting a value of about 60% after 14 h of incubation;

whereas the stabilising effect of H<sup>5</sup> on the SspCA showed a residual activity of about 85% and remained almost constant for the rest of the time indicated in the figure (panel C). These results demonstrated that the anchoring ASL<sup>tag</sup> system, enhanced the SspCA stability of about 20%. On the other hand, it is important to highlight that both anchored enzymes continued to work for several hours at temperatures considered prohibitive for the free enzymes, as SspCA, which Russo et al. demonstrated to show a residual activity of 20% when heated at 70.0 °C for 15 min<sup>57</sup>. This aspect is crucial in the context of the post-combustion carbon capture process, which requires temperatures ranging from 40.0 and 60.0 °C<sup>53</sup>. Figure 6 shows the residual activity for the CO<sub>2</sub> hydration reaction for SspCA and H<sup>5</sup>-SspCA when the whole cells were treated at different temperatures for a very long period (up to 10 d). At 25.0 °C, the SspCA residual activity started to decrease after 4 d and reached a value of about 70% after 10 d, while H<sup>5</sup>-SspCA remained almost constant (panel A). At 50.0 and 70.0 °C, the residual activity of SspCA decreased up to 40 and 20%, respectively (panel B and C), whereas H<sup>5</sup>-SspCA showed a residual activity of about 60 and 40%, respectively (panel B and C). All these data confirmed that the presence of a thermostable *protein-tag* between the INPN anchoring domain and the SspCA significantly improved the long-term stability and the storage of this CA.

#### 4. Conclusions

The ASL<sup>tag</sup> system efficiently overexpressed the chimeric H<sup>5</sup>-SspCA onto to the bacterial cell surface, as demonstrated by fluorescence microscopy and Western-Blot. As expected, the CO<sub>2</sub> hydratase assay and the protonography showed that SspCA was still very active, even linked on the bacterial surface and the H<sup>5</sup> moiety, showing a CO<sub>2</sub> hydratase activity similar to that of its anchored counterpart without H<sup>5</sup>. Furthermore, by investigating the behaviour of the whole bacterial cells expressing on the external surface SspCA or H<sup>5</sup>-SspCA at different temperatures, we demonstrated an enhancement in terms of thermal stability of the chimeric protein. In conclusion, the H<sup>5</sup>-SspCA obtained by the ASL<sup>tag</sup> system constitutes a valid strategy for further increasing the thermostability of proteins, for processes in which a highly effective, thermostable catalyst is needed.

#### Acknowledgements

We are grateful to Giovanni Del Monaco for technical assistance.

#### Disclosure statement

The authors state no conflict of interests.

#### Funding

This work was supported by FIRB-Futuro in Ricerca RBF12001G\_002 "Nematic" and by the grant "SMART GENERATION – Sistemi e tecnologie sostenibili per la generazione di energia-PON03PE\_00157\_1, OR3-Bio-sistemi di cattura ed utilizzazione della CO<sub>2</sub>".

#### ORCID

Claudiu T. Supuran  <http://orcid.org/0000-0003-4262-0323>  
Clemente Capasso  <http://orcid.org/0000-0003-3314-2411>

#### References

1. Capasso C, Supuran CT. Inhibition of bacterial carbonic anhydrases as a novel approach to escape drug resistance. *Curr Top Med Chem* 2017;17:1237–48.
2. Del Prete S, Vullo D, De Luca V, et al. Cloning, expression, purification and sulfonamide inhibition profile of the complete domain of the eta-carbonic anhydrase from *Plasmodium falciparum*. *Bioorg Med Chem Lett* 2016;26:4184–90.
3. Del Prete S, Vullo D, De Luca V, et al. Anion inhibition profiles of the complete domain of the eta-carbonic anhydrase from *Plasmodium falciparum*. *Bioorg Med Chem* 2016;24:4410–4.
4. Annunziato G, Angeli A, D'Alba F, et al. Discovery of new potential anti-infective compounds based on carbonic anhydrase inhibitors by rational target-focused repurposing approaches. *Chem Med Chem* 2016;11:1904–14.
5. Del Prete S, Vullo D, De Luca V, et al. Anion inhibition profiles of alpha-, beta- and gamma-carbonic anhydrases from the pathogenic bacterium *Vibrio cholerae*. *Bioorg Med Chem* 2016;24:3413–7.
6. Abdel Gawad NM, Amin NH, Elsaadi MT, et al. Synthesis of 4-(thiazol-2-ylamino)-benzenesulfonamides with carbonic anhydrase I, II and IX inhibitory activity and cytotoxic effects against breast cancer cell lines. *Bioorg Med Chem* 2016;24:3043–51.
7. Capasso C, Supuran CT. An overview of the carbonic anhydrases from two pathogens of the oral cavity: streptococcus mutans and porphyromonas gingivalis. *Curr Top Med Chem* 2016;16:2359–68.
8. Del Prete S, Vullo D, De Luca V, et al. Comparison of the sulfonamide inhibition profiles of the alpha-, beta- and gamma-carbonic anhydrases from the pathogenic bacterium *Vibrio cholerae*. *Bioorg Med Chem Lett* 2016;26:1941–6.
9. Supuran CT, Capasso C. Biomedical applications of prokaryotic carbonic anhydrases. *Expert Opin Ther Pat* 2018;28:745–54.
10. Nishimori I, Onishi S, Takeuchi H, Supuran CT. The alpha and beta classes carbonic anhydrases from *Helicobacter pylori* as novel drug targets. *Curr Pharm Des* 2008;14:622–30.
11. Morishita S, Nishimori I, Minakuchi T, et al. Cloning, polymorphism, and inhibition of beta-carbonic anhydrase of *Helicobacter pylori*. *J Gastroenterol* 2008;43:849–57.
12. Supuran CT. Advances in structure-based drug discovery of carbonic anhydrase inhibitors. *Expert Opin Drug Discov* 2017;12:61–88.
13. Supuran CT. Structure and function of carbonic anhydrases. *Biochem J* 2016;473:2023–32.
14. Supuran CT, Capasso C. An overview of the bacterial carbonic anhydrases. *Metabolites* 2017;7:p11: E56.
15. Capasso C, Supuran CT. An overview of the alpha-, beta- and gamma-carbonic anhydrases from bacteria: can bacterial carbonic anhydrases shed new light on evolution of bacteria? *J Enzyme Inhib Med Chem* 2015;30:325–32.
16. Supuran CT. Carbonic anhydrase activators. *Future Med Chem* 2018;10:561–73.
17. Supuran CT, Capasso C. An overview of the bacterial carbonic anhydrases. *Metabolites* 2017;7:56–74.

18. Supuran CT, Capasso C. Carbonic anhydrase from *Porphyromonas gingivalis* as a drug target. *Pathogens* 2017; 6:30–42.
19. Capasso C, Supuran CT. Bacterial, fungal and protozoan carbonic anhydrases as drug targets. *Expert Opin Ther Targets* 2015;19:1689–704.
20. Aspatwar A, Tolvanen ME, Ortutay C, Parkkila S. Carbonic anhydrase related proteins: molecular biology and evolution. *Subcell Biochem* 2014;75:135–56.
21. Supuran CT. Carbonic anhydrases as drug targets—an overview. *Curr Top Med Chem* 2007;7:825–33.
22. Supuran CT, Capasso C. The  $\eta$ -class carbonic anhydrases as drug targets for antimalarial agents. *Expert Opin Ther Targets* 2015;19:551–63.
23. Capasso C, Supuran CT. An overview of the selectivity and efficiency of the bacterial carbonic anhydrase inhibitors. *Curr Med Chem* 2015;22:2130–9.
24. Capasso C, Supuran CT. Sulfa and trimethoprim-like drugs - antimetabolites acting as carbonic anhydrase, dihydropteroate synthase and dihydrofolate reductase inhibitors. *J Enzyme Inhib Med Chem* 2014;29:379–87.
25. Capasso C, Supuran CT. Anti-infective carbonic anhydrase inhibitors: a patent and literature review. *Expert Opin Ther Pat* 2013;23:693–704.
26. Ozensoy Guler O, Capasso C, Supuran CT. A magnificent enzyme superfamily: carbonic anhydrases, their purification and characterization. *J Enzyme Inhib Med Chem* 2016;31: 689–94.
27. Del Prete S, Vullo D, De Luca V, et al. Sulfonamide inhibition studies of the beta-carbonic anhydrase from the pathogenic bacterium *Vibrio cholerae*. *Bioorg Med Chem* 2016;24: 1115–20.
28. Del Prete S, De Luca V, De Simone G, et al. Cloning, expression and purification of the complete domain of the eta-carbonic anhydrase from *Plasmodium falciparum*. *J Enzyme Inhib Med Chem* 2016;31:54–59.
29. De Simone G, Monti SM, Alterio V, et al. Crystal structure of the most catalytically effective carbonic anhydrase enzyme known, SazCA from the thermophilic bacterium *Sulfurihydrogenibium azorense*. *Bioorg Med Chem Lett* 2015; 25:2002–6.
30. Di Fiore A, Capasso C, De Luca V, et al. X-ray structure of the first 'extremo-alpha-carbonic anhydrase', a dimeric enzyme from the thermophilic bacterium *Sulfurihydrogenibium yellowstonense* YO3AOP1. *Acta Crystallogr D Biol Crystallogr* 2013; 69:1150–9.
31. Alterio V, Langella E, De Simone G, Monti SM. Cadmium-containing carbonic anhydrase CDCA1 in marine diatom *Thalassiosira weissflogii*. *Mar Drugs* 2015;13:1688–97.
32. Di Fiore A, Alterio V, Monti SM, et al. Thermostable carbonic anhydrases in biotechnological applications. *Int J Mol Sci* 2015;16:15456–80.
33. Supuran CT. CA IX stratification based on cancer treatment: a patent evaluation of US2016/0002350. *Expert Opin Ther Pat* 2016;26:1105–9.
34. Lomelino C, McKenna R. Carbonic anhydrase inhibitors: a review on the progress of patent literature (2011–2016). *Expert Opin Ther Pat* 2016;26:947–56.
35. Monti SM, Supuran CT, De Simone G. Anticancer carbonic anhydrase inhibitors: a patent review (2008–2013). *Expert Opin Ther Pat* 2013;23:737–49.
36. Masini E, Carta F, Scozzafava A, Supuran CT. Antiglaucoma carbonic anhydrase inhibitors: a patent review. *Expert Opin Ther Pat* 2013;23:705–16.
37. Scozzafava A, Supuran CT, Carta F. Antiobesity carbonic anhydrase inhibitors: a literature and patent review. *Expert Opin Ther Pat* 2013;23:725–35.
38. Aggarwal M, Kondeti B, McKenna R. Anticonvulsant/antiepileptic carbonic anhydrase inhibitors: a patent review. *Expert Opin Ther Pat* 2013;23:717–24.
39. Carta F, Supuran CT. Diuretics with carbonic anhydrase inhibitory action: a patent and literature review (2005–2013). *Expert Opin Ther Pat* 2013;23:681–91.
40. Winum JY, Capasso C. Novel antibody to a carbonic anhydrase: patent evaluation of WO2011138279A1. *Expert Opin Ther Pat* 2013;23:757–60.
41. Aggarwal M, McKenna R. Update on carbonic anhydrase inhibitors: a patent review (2008–2011). *Expert Opin Ther Pat* 2012;22:903–15.
42. Carta F, Scozzafava A, Supuran CT. Sulfonamides: a patent review (2008–2012). *Expert Opin Ther Pat* 2012;22:747–58.
43. Carta F, Supuran CT, Scozzafava A. Novel therapies for glaucoma: a patent review 2007–2011. *Expert Opin Ther Pat* 2012;22:79–88.
44. Poulsen SA. Carbonic anhydrase inhibition as a cancer therapy: a review of patent literature, 2007–2009. *Expert Opin Ther Pat* 2010;20:795–806.
45. Boone CD, Habibzadegan A, Gill S, McKenna R. Carbonic anhydrases and their biotechnological applications. *Biomolecules* 2013;3:553–62.
46. Alafeefy AM, Abdel-Aziz HA, Vullo D, et al. Inhibition of carbonic anhydrases from the extremophilic bacteria *Sulfurihydrogenibium yellowstonense* (SspCA) and *Sulfurihydrogenibium azorense* (SazCA) with a new series of sulfonamides incorporating aroylhydrazone-, [1,2,4]triazolo[3,4-b][1,3,4]thiadiazinyl- or 2-(cyanophenylmethylene)-1,3,4-thiadiazol-3(2H)-yl moieties. *Bioorg Med Chem* 2014;22: 141–7.
47. Vullo D, De Luca V, Scozzafava A, et al. The extremo-alpha-carbonic anhydrase from the thermophilic bacterium *Sulfurihydrogenibium azorense* is highly inhibited by sulfonamides. *Bioorg Med Chem* 2013;21:4521–5.
48. Vullo D, Luca VD, Scozzafava A, et al. The alpha-carbonic anhydrase from the thermophilic bacterium *Sulfurihydrogenibium yellowstonense* YO3AOP1 is highly susceptible to inhibition by sulfonamides. *Bioorg Med Chem* 2013;21:1534–8.
49. Akdemir A, Vullo D, De Luca V, et al. The extremo-alpha-carbonic anhydrase (CA) from *Sulfurihydrogenibium azorense*, the fastest CA known, is highly activated by amino acids and amines. *Bioorg Med Chem Lett* 2013;23:1087–90.
50. De Luca V, Vullo D, Scozzafava A, et al. Anion inhibition studies of an alpha-carbonic anhydrase from the thermophilic bacterium *Sulfurihydrogenibium yellowstonense* YO3AOP1. *Bioorg Med Chem Lett* 2012;22:5630–4.
51. Vullo D, De Luca V, Scozzafava A, et al. Anion inhibition studies of the fastest carbonic anhydrase (CA) known, the extremo-CA from the bacterium *Sulfurihydrogenibium azorense*. *Bioorg Med Chem Lett* 2012;22:7142–5.
52. Vullo D, De Luca V, Scozzafava A, et al. The first activation study of a bacterial carbonic anhydrase (CA). The thermostable alpha-CA from *Sulfurihydrogenibium yellowstonense* YO3AOP1 is highly activated by amino acids and amines. *Bioorg Med Chem Lett* 2012;22:6324–7.

53. Russo ME, Olivieri G, Capasso C, et al. Kinetic study of a novel thermo-stable alpha-carbonic anhydrase for biomimetic CO<sub>2</sub> capture. *Enzyme Microb Technol* 2013;53:271–7.
54. Migliardini F, De Luca V, Carginale V, et al. Biomimetic CO<sub>2</sub> capture using a highly thermostable bacterial alpha-carbonic anhydrase immobilized on a polyurethane foam. *J Enzyme Inhib Med Chem* 2014;29:146–50.
55. Abdelrahim MYM, Martins CF, Neves LA, et al. Supported ionic liquid membranes immobilized with carbonic anhydrases for CO<sub>2</sub> transport at high temperatures. *J Membr Sci* 2017;528:225–30.
56. Perfetto R, Del Prete S, Vullo D, et al. Production and covalent immobilisation of the recombinant bacterial carbonic anhydrase (SspCA) onto magnetic nanoparticles. *J Enzyme Inhib Med Chem* 2017;32:759–66.
57. Del Prete S, Perfetto R, Rossi M, et al. A one-step procedure for immobilising the thermostable carbonic anhydrase (SspCA) on the surface membrane of *Escherichia coli*. *J Enzyme Inhib Med Chem* 2017;32:1120–8.
58. Merlo R, Del Prete S, Valenti A, et al. An AGT-based protein-tag system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane. *J Enzyme Inhib Med Chem* 2019;34:490–9.
59. Capasso C, De Luca V, Carginale V, et al. Biochemical properties of a novel and highly thermostable bacterial alpha-carbonic anhydrase from *Sulfurihydrogenibium yellowstone* YO3AOP1. *J Enzyme Inhib Med Chem* 2012;27:892–7.
60. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970;227:680–5.
61. Del Prete S, De Luca V, Supuran CT, Capasso C. Protonography, a technique applicable for the analysis of eta-carbonic anhydrase activity. *J Enzyme Inhib Med Chem* 2015;30:920–4.
62. Del Prete S, De Luca V, Iandolo E, et al. Protonography, a powerful tool for analyzing the activity and the oligomeric state of the gamma-carbonic anhydrase identified in the genome of *Porphyromonas gingivalis*. *Bioorg Med Chem* 2015;23:3747–50.
63. De Luca V, Del Prete S, Supuran CT, Capasso C. Protonography, a new technique for the analysis of carbonic anhydrase activity. *J Enzyme Inhib Med Chem* 2015;30:277–82.
64. Vettone A, Serpe M, Hidalgo A, et al. A novel thermostable protein-tag: optimization of the *Sulfolobus solfataricus* DNAalkyl-transferase by protein engineering. *Extremophiles* 2016;20:13.
65. Tieleman DP, Berendsen HJ. A molecular dynamics study of the pores formed by *Escherichia coli* OmpF porin in a fully hydrated palmitoyl-oleoylphosphatidylcholine bilayer. *Biophys J* 1998;74:2786–801.



# A journey down to hell: new thermostable protein-tags for biotechnology at high temperatures

Rosanna Mattossovich<sup>1</sup> · Rosa Merlo<sup>1</sup> · Angelo Fontana<sup>2</sup> · Giuliana d'Ippolito<sup>2</sup> · Michael P. Terns<sup>3</sup> · Elizabeth A. Watts<sup>3</sup> · Anna Valenti<sup>1</sup> · Giuseppe Perugino<sup>1</sup>

Received: 20 June 2019 / Accepted: 13 September 2019 / Published online: 25 September 2019  
© Springer Japan KK, part of Springer Nature 2019

## Abstract

The specific labelling of proteins in recent years has made use of self-labelling proteins, such as the SNAP-tag<sup>®</sup> and the Halotag<sup>®</sup>. These enzymes, by their nature or suitably engineered, have the ability to specifically react with their respective substrates, but covalently retaining a part of them in the catalytic site upon reaction. This led to the synthesis of substrates conjugated with, e.g., fluorophores (proposing them as alternatives to fluorescent proteins), but also with others chemical groups, for numerous biotechnological applications. Recently, a mutant of the OGT from *Saccharolobus solfataricus* (H<sup>5</sup>) very stable to high temperatures and in the presence of physical and chemical denaturing agents has been proposed as a thermostable SNAP-tag<sup>®</sup> for in vivo and in vitro harsh reaction conditions. Here, we show two new thermostable OGTs from *Thermotoga neapolitana* and *Pyrococcus furiosus*, which, respectively, display a higher catalytic activity and thermostability respect to H<sup>5</sup>, proposing them as alternatives for in vivo studies in these extreme model organisms.

**Keywords** (Hyper)thermophiles · Thermostable proteins · Protein-tag · Biotechnology

---

Communicated by M. Moracci.

---

This manuscript is part of a special issue of Extremophiles journal for the 12th International Congress of Extremophiles (Extremophiles 2018) that was held on 16–20 September 2018 in Ischia, Naples, Italy.

---

Rosanna Mattossovich and Rosa Merlo equally contributed to the present work.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00792-019-01134-3>) contains supplementary material, which is available to authorized users.

---

✉ Giuseppe Perugino  
giuseppe.perugino@ibbr.cnr.it

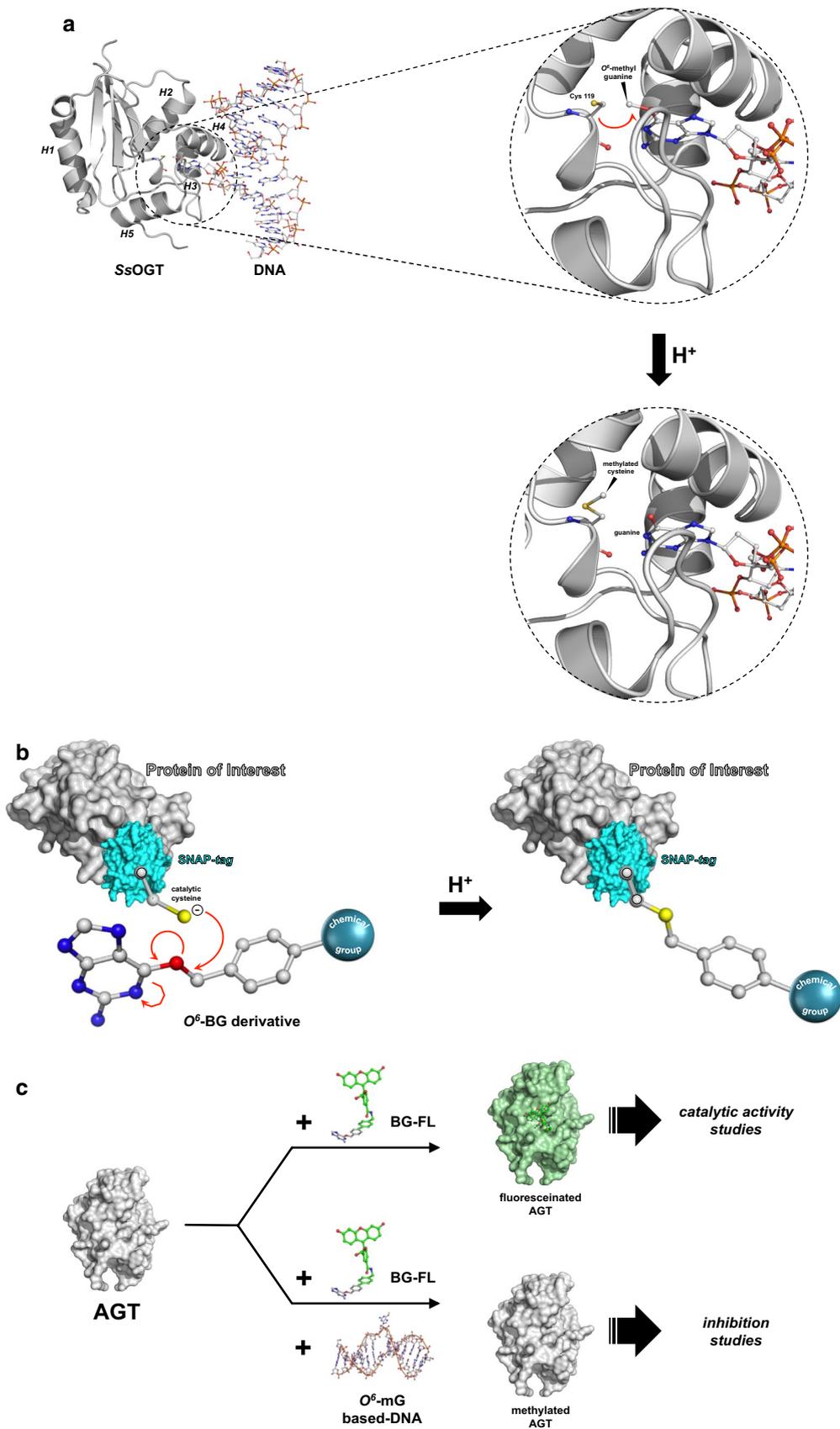
<sup>1</sup> Institute of Biosciences and BioResources, National Council of Research of Italy, Via P. Castellino 111, 80131 Naples, Italy

<sup>2</sup> Institute of Biomolecular Chemistry, National Council of Research of Italy, Via Campi Flegrei, 34, 80078 Pozzuoli, NA, Italy

<sup>3</sup> Departments of Biochemistry and Molecular Biology, Genetics, and Microbiology, University of Georgia, Athens, GA, USA

## Introduction

Protein-tags are short or long peptide sequences genetically fused to recombinant proteins for various purposes, as affinity purification, protein localization and general labelling procedures. In some cases, the presence of these tags enhances the solubilization of proteins and enzymes expressed in chaperone-deficient species such as *Escherichia coli*, to assist in the proper folding of proteins and avoid precipitation. These include commercially available thioredoxin (TRX), poly(NANP), maltose-binding protein (MBP), and glutathione *S*-transferase (GST). The discovery of Fluorescent Proteins (FPs) has revolutionized the world of cell and molecular biology, allowing several applications as reporter gene in fluorescence microscopy (Chalfie et al. 1994; Tsien 1998; Aliye et al. 2015). Although FPs are intrinsically fluorescent without the addition of any external substrate, they have some disadvantages. Indeed, their relatively large dimensions and the sensitivity to environmental changes (pH or the absence of O<sub>2</sub>) affect the formation of the internal natural fluorophore (Ashby et al. 2004; Campbell and Choy 2000).



**Fig. 1 a** A cartoon of the AGTs irreversible reaction mechanism is shown as example: the *Ss*OGT enzyme (PDB ID: 4ZYE) recognizes the methylated guanine on DNA (from PDB ID: 4ZYD) and proceeds to the irreversible transfer of the methyl group from the guanine to the catalytic cysteine, inactivating the protein (from PDB ID: 4ZYG) (Perugino et al. 2015). **b** The SNAP-tag<sup>®</sup> technology is based of an engineered variant of the hAGT, which is able to recognize BG-derivative substrates, and irreversibly transferring in its active site a desired chemical group conjugated to the benzyl moiety. Consequently, this allows the labelling of a protein of interest expressed as fusion protein to the SNAP-tag<sup>®</sup>. **c** The use of commercially available fluorescent AGT substrates led to the setting up of new safety AGT's assay, for the determination of the catalytic activities, or inhibition studies in the presence of natural substrates (e.g., alkylated DNA) or AGT's inhibitors

Recently, studies on *O*<sup>6</sup>-alkylguanine-DNA-alkyl-transferases (AGTs, OGTs or MGMTs; EC: 2.1.1.63) led to the proposal of alternative protein-tags with new behaviours, useful for several applications in many fields. In nature, these ubiquitous small proteins have a crucial role in the direct repair of DNA by alkylating agents (Mishina et al. 2006; Serpe et al. 2019) and the human homolog (hAGT) is target of chemotherapeutic drugs for its crucial role in many types of tumours (Sun et al. 2018; Aoki and Natsume 2019). Their catalytic activity is based on the recognition of the damage on DNA (an *O*<sup>6</sup>-alkylguanine or an *O*<sup>4</sup>-alkylthymine) and, by a *one-step* reaction mechanism (via SN<sub>2</sub> type), the alkyl group from the damaged base is irreversibly transferred to a cysteine residue in their own active site (Daniels et al. 2000, 2004; Fang et al. 2005; Tubbs et al. 2007; Pegg 2011; Yang et al. 2009) (Fig. 1a). For these reasons, they are called suicide or kamikaze proteins, with a 1:1 stoichiometry of their reaction with the natural substrate. However, some AGTs resulted very reactive with a strong inhibitor, the *O*<sup>6</sup>-benzyl-guanine (*O*<sup>6</sup>-BG), which was used in combination with chemotherapeutic drugs (Pegg et al. 2001; Luu et al. 2002; Coulter et al. 2007). This information led professor Kai Johnsson and his group to put their effort for the production of a new protein-tag from a variant of the hAGT, introducing the SNAP-tag<sup>®</sup> technology in the “biotech” scenario (Keppler et al. 2003, 2004; Gronemeyer et al. 2006; Gautier et al. 2008; Mollwitz et al. 2012): upon the irreversible reaction, hAGT keeps the benzyl moiety of the substrate covalently linked to its catalytic cysteine. This makes possible the labelling of this protein (and a relative protein of interest, if expressed as fusions with it) when any chemical group conjugated to the benzyl moiety of *O*<sup>6</sup>-BG (Keppler et al. 2003, 2004; Gautier et al. 2008) (Fig. 1b). Furthermore, the commercially available fluorescent *O*<sup>6</sup>-BG derivatives (as the SNAP-Vista Green<sup>®</sup> and the SNAP-Cell<sup>®</sup> TMR Star; New England Biolabs) allowed the development of new applications in the fluorescent microscopy field, and recently of a new fluorescent assay for all the BG-sensitive AGTs, overcoming the traditional, long and unsafe assays

for this class of proteins based on HPLC procedures and radioactive substrates (Hishiguro et al. 2008; Perugino et al. 2012). The presence of methylated-DNA in this assay, led to the determination of the activity of AGTs on their natural substrates, by applying the classical enzyme inhibition approaches (Fig. 1c) (Perugino et al. 2012, 2015; Miggiano et al. 2013, 2017; Vettone et al. 2016; Morrone et al. 2017).

All the above-mentioned protein-tags have the disadvantage to be employed in mild reaction conditions and in mesophilic organisms. This was successfully overcome by the introduction of a thermostable OGT from the hyperthermophilic archaea *Saccharolobus solfataricus* (formerly *Sulfolobus solfataricus*). Starting from the studies of the wild-type enzyme (*Ss*OGT), which not only displays the same behaviours of the hAGT on BG derivatives, but is also characterized by an exceptional stability at extremes of temperature, pH, ionic strength and the presence of denaturing agents (Perugino et al. 2012), we developed a DNA-bindingless variant (called H<sup>5</sup>), which resulted in a strong “thermostable SNAP-tag<sup>®</sup>”. H<sup>5</sup> was successfully employed in *in vivo* heterologous expression in thermophilic organisms as *Thermus thermophilus* HB27 and *Sulfolobus islandicus* E233S1 (Vettone et al. 2016; Visone et al. 2017), in fusion with a thermostable β-glycosidase and a *S. solfataricus* reverse gyrase (Vettone et al. 2016; Visone et al. 2017), as well as under extreme reaction conditions as gene reporter in *in vitro* transcription/translation systems based on *Sulfolobus* cell lysates (Lo Gullo et al. 2019). Furthermore, H<sup>5</sup> was successfully fused to the N-terminal domain of the ice nucleation protein (INPN) from the Gram-negative bacterium *Pseudomonas syringae*, a transmembrane protein useful for the one-step heterologous expression and the *in vivo* immobilization of proteins of interest on the outer membrane of *E. coli* (Samuelson et al. 2002). This led to the development of the new Anchoring-and-self-labelling-protein-tag (ASL<sup>tag</sup>) (Merlo et al. 2019), which simultaneously allows the immobilization (by INPN) and the quantitative determination of the yield of an immobilized protein (by the fluorescent assay using H<sup>5</sup>). Surprisingly, the presence of the thermostable H<sup>5</sup> between the INPN and a protein of interest resulted in an enhancement of the stability of the latter, as it was the case for the *Sulfurihydrogenibium yellowstonense* α-carbonic anhydrase (Del Prete et al. 2019).

The growing demand to use recent technologies at very extreme conditions in thermophilic bacteria and archaea, such as the *in vivo* CRISPR-Cas immune systems, leads us to search for new protein-tags with very high activity and thermostability. In this regard, the present work is focussed on the characterization of two new OGTs, from the hyperthermophilic organisms *Thermotoga neapolitana* and *Pyrococcus furiosus*. The purified proteins were compared to *Ss*OGT, showing respect to its exceptional characteristics in terms of enzymatic activity and thermostability.

These results open promising perspectives in the development of new protein-tags to employ in these extreme model organisms.

## Materials and methods

### Reagents

Fine chemicals were from Sigma-Aldrich, SNAP-Vista Green<sup>®</sup> fluorescent substrate (hereinafter BG-FL) was from New England Biolabs (Ipswich, MA). SYPRO Orange 5000× (Invitrogen). Synthetic oligonucleotides listed in Table 1 were from Eurofins (Milan, Italy); *Pfu* DNA polymerase was from NZYTech (Portugal). The Bio-Rad protein assay kit (Bio-Rad Pacific) was used for the determination of the protein concentration, using purified BSA as standard.

### DNA constructs

The cloning procedures for the construction of *E. coli* expression plasmids were the same for both the proteins: the ORF CTN1690/PF1878, encoding a putative OGT, was amplified from genomic DNA from *Thermotoga neapolitana* DSMZ 4359<sup>T</sup>/*Pyrococcus furiosus* JFW02 strain genomic DNA, using Lig5:Lig3 *Tn*OGT/*Pfu*OGT oligonucleotides pairs (listed in Table 1) and directly cloned into the expression vector pHTP1 (NZYTech, Portugal), following the instructions described in the NZYEasy Cloning and Expression kit I (NZYTech, Portugal) manual. The ligation mixture was entirely used to transform commercial *E. coli* DH5α cells (NZY5α Competent Cells-NZYTech, Portugal) and positive clones were confirmed by PCR. Subsequently, a DNA fragment from the resulted pHTP1-*Tn*OGT/pHTP1-*Pfu*OGT was removed by digestion with Nco I restriction endonuclease, and was replaced by a double-stranded oligonucleotides (NZY-His Fwd2 and NZY-His Rev2; Table 1), whose DNA sequence expresses a shorter His<sub>6</sub>-tag (MAHHHHHTG-), similar to that at the N-terminal of the *Ss*OGT protein (Perugino

et al. 2012). Positive clones after transformation of the ligation mixture in *E. coli* KRX competent cells were confirmed by DNA sequencing.

### Protein purification

*Tn*OGT and *Pfu*OGT were expressed in the *E. coli* BL21 (DE3) cells grown overnight at 37 °C in Luria–Bertani (LB) selective medium supplemented with 50 mg/L kanamycin and 30 mg/L chloramphenicol, and the protein expression was induced with 1.0 mM isopropyl-thio-β-D-galactopyranoside (IPTG), when an absorbance value of 0.5–0.6 A<sub>600 nm</sub> was reached. Harvested cells were resuspended 1:3 (w/v) in Buffer A (50 mM phosphate, 300 mM NaCl; pH 8.0) supplemented with 1% Triton X-100 and stored overnight at –20 °C. After this first step of lysis, the biomass was treated with lysozyme and DNase for 60 min in ice, followed by a sonication step. Finally, the lysate was centrifuged for 30 min at 60,000×g in and the cell extract recovered. To remove *E. coli* contaminants, all cell extracts were incubated 20 min at 70.0 °C and 20 min at 65 °C, respectively, followed by a centrifugation at 13,000×g at 4.0 °C; the supernatant was diluted 1:2 (v/v) in purification Buffer A and applied to a Protino Ni–NTA Column 1.0 mL (Macherey–Nagel) for His<sub>6</sub>-tag affinity chromatography. After two washing steps of 10 column volumes of Buffer A and 10 column volumes of Buffer A supplemented with 25.0 mM imidazole, the elution was performed in 20 column volumes of buffer A, by applying a linear gradient of 25.0–250.0 mM imidazole. The fractions containing the protein were collected and analysed by SDS-PAGE. *Tn*OGT was dialysed against PBS 1 × buffer (phosphate buffer 20 mM, NaCl 150 mM; pH 7.3), whereas the fractions of the *Pfu*OGT protein were pooled, concentrated and subjected to a further gel-filtration chromatography, using a Superdex 75 10/300 GL column (GE Healthcare Life Sciences). Finally, both the proteins were concentrated and loaded on 15% SDS-PAGE gel to confirm its purity and stored at –20.0 °C.

**Table 1** List of oligonucleotides used in this study

Name	Sequence	Notes
Lig5- <i>Pfu</i> OGT	5'-TCAGCAAGGGCTGAGGCCATGGTATTGGAAGTTAGG-3'	Nco I site underlined
Lig3- <i>Pfu</i> OGT	5'-CCTCAGCGGAAGCTGAGGTTAGCTTGCCATCCTTCC-3'	
Lig5- <i>Tn</i> OGT	5'-TCAGCAAGGGCTCAGGCCATGGGAGATCGA-3'	Nco I site underlined
Lig3- <i>Tn</i> OGT	5'-CCTCAGCGGAAGCTGAGGTTATCGACTACCTCGC-3'	
NZY-His fwd2	5'-catgGCACACCATCACCATCACCATACGGG-3'	Inserting an His <sub>6</sub> -tag (MAHHHHHTG-; underlined) upstream the <i>Pfu</i> OGT and <i>Tn</i> OGT sequence
NZY-His rev2	5'-catgCCCGTATGGTGATGGTGATGGTGTC-3'	
Fwd <sup>m4</sup>	5'-ggcMgtaggcctagcatgacaatctgcattgtgatcacgg-3'	From Perugino et al. (2015); M = O <sup>6</sup> -methyl-guanine
Rev4	5'-ccgtgatcaccaatgcagattgtcatgctaggcctaccgc-3'	From Perugino et al. (2015)

## In vitro alkyl-transferase assay

The fluorescent substrate BG-FL was used for the determination of the catalytic activity of all thermostable enzymes analysed, as previously described (Perugino et al. 2012, 2015; Miggiano et al. 2013; Vettone et al. 2016; Visone et al. 2017; Merlo et al. 2019; Del Prete et al. 2019). Briefly, 5.0  $\mu\text{M}$  of protein (ca. 0.1 mg/mL) was incubated with 10.0  $\mu\text{M}$  of BG-FL in Fluo Buffer 1 $\times$  (50.0 mM phosphate, 100.0 mM NaCl, 1.0 mM DTT; pH 6.5) at different temperatures and times, as indicated; each reaction was stopped by adding a Laemmli buffer 1 $\times$  (formamide 95%; EDTA 20.0 mM; bromophenol 0.05%), followed by denaturation at 100.0  $^{\circ}\text{C}$  and the direct loading of the sample on SDS-PAGE. The gel was first analysed by fluorescence imaging on a VersaDoc 4000<sup>TM</sup> system (Bio-Rad) by applying as excitation/emission parameters a blue LED/530 bandpass filter, and then was stained by Coomassie. Assuming the irreversible mechanism with 1:1 BG-FL/OGT ratio, fluorescence intensity data were corrected for the amount of loaded protein, and fitted by applying the second-order rate equation, to determine the relative amount of covalently modified protein in time-course experiments. (Gautier et al. 2008; Miggiano et al. 2013; Perugino et al. 2012, 2015).

## Competitive assay and IC<sub>50</sub> calculation

Competitive assay using the fluorescent substrate in the presence of double strands (ds) oligonucleotides pairs (Fwd<sup>m4</sup>: Rev4; Table 1), containing a single O<sup>6</sup>-methyl-guanine, was performed as described (Perugino et al. 2015) to determine the half maximal inhibitory concentration (IC<sub>50</sub>), that is the concentration of methylated DNA needed to reduce the fluorescence intensity of the OGT band by 50.0%. Reactions incubated at fixed temperatures with increasing concentrations (0.0–10.0  $\mu\text{M}$ ) of ds-Fwd<sup>m4</sup> and keeping constant the BG-FL concentration (5.0  $\mu\text{M}$ ) were performed for 10 min at 50.0  $^{\circ}\text{C}$ . Corrected data of fluorescence intensity were fitted with the IC<sub>50</sub> equation (Perugino et al. 2015; Morrone et al. 2017).

## Protein stability analysis

The stability at several conditions of the thermostable OGTs was analysed by the differential scan fluorimetry method (DSF), adapted by a protocol previously described for the SsOGT and relative mutants (Niesen et al. 2007; Perugino et al. 2015; Vettone et al. 2016; Morrone et al. 2017). Triplicates of each condition containing 25.0  $\mu\text{M}$  of enzyme (ca. 0.5 mg/mL) in PBS 1 $\times$  buffer and SYPRO Orange dye 1 $\times$  were subjected to a scan of 70 cycles at temperatures from 25.0 to 94.0  $^{\circ}\text{C}$  for 10 min/ $^{\circ}\text{C}$   $\times$  cycle, and analysed in a Real-Time Light Cycler<sup>TM</sup> (Bio-Rad). Relative fluorescence

data were then normalized to the maximum fluorescence value within each scan. Obtained plots of fluorescence intensity vs temperature displayed sigmoidal curves (typical of a two-state transition), which allowed the determination of the inflection points ( $T_m$  values) by fitting the Boltzmann equation (Niesen et al. 2007; Perugino et al. 2015; Vettone et al. 2016; Morrone et al. 2017).

## Data analysis and softwares

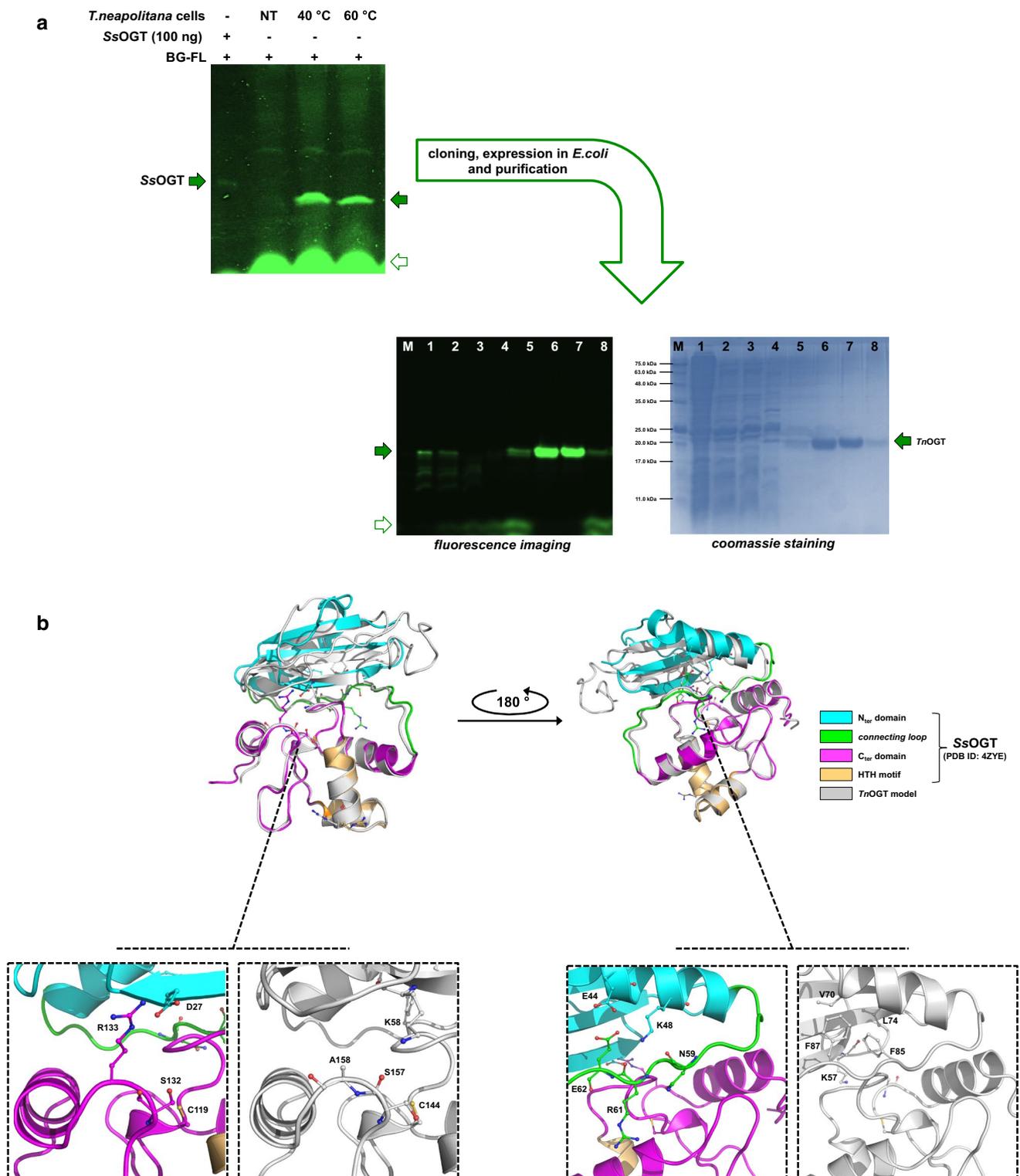
Prism Software Package (GraphPad Software) and GraFit 5.0 Data Analysis Software (Erithacus Software) were used for the data analysis from activity, competitive inhibition and stability assays.

## Results and discussion

### The OGT from *Thermotoga neapolitana*

*Thermotoga neapolitana* is a hyperthermophilic Gram-negative bacterium of the order *Thermotogales* (Belkin et al. 1986; Jannasch et al. 1988) that include good candidates for genetic engineering and biotechnological applications (Connors et al. 2006; Zhang et al. 2015; Fink et al. 2016; Donaldson et al. 2017; Han and Xu 2017). Several *Thermotogales* have been studied for the fermentative production of hydrogen (H<sub>2</sub>) by [FeFe] hydrogenases with yield close to the theoretical Thauer limit of four moles of H<sub>2</sub> per mole of consumed sugar (d'Ippolito et al. 2010; Pradhan et al. 2015, 2016). *T. neapolitana* also shows a novel, anaerobic process named capnophilic lactic fermentation (CLF) that leads to the synthesis of lactic acid (LA) without affecting H<sub>2</sub> production under CO<sub>2</sub> trigger (Di Pasquale et al. 2014; d'Ippolito et al. 2014; Pradhan et al. 2017).

The ORF CTN1690 of *T. neapolitana* encodes a putative 174-aa polypeptide, with a calculated molecular weight of 19.9 kDa, and a clear homology to the O<sup>6</sup>-alkylguanine-DNA-alkyl-transferases. Furthermore, lyophilised *T. neapolitana* cells incubated with the BG-FL substrate showed a strong fluorescent signal close to that of SsOGT by gel imaging after SDS-PAGE (Fig. 2a). The observed molecular weight and, mostly the sensitivity to the benzyl-guanine derivative BG-FL substrate, led to the cloning and the heterologous expression in *E. coli* of the His-tagged version of this protein. After purification by affinity chromatography, the protein was fully active on BG-FL (Fig. 2a) and was further subjected to a biochemical characterization. The inhibition assay in the presence of methylated-dsDNA (Table 2) and the fluorescent substrate BG-FL allowed the determination of an IC<sub>50</sub> value similar to that obtained with the *S. solfataricus* enzyme, thus confirming a role in DNA repair of this thermophilic



protein, hereinafter properly named *TnOGT* (Table 2; Perugino et al. 2015). Surprisingly, the enzyme from *T. neapolitana* displayed a very high activity at low temperatures (Table 3), similar to that shown by the *SsOGT*-H<sup>5</sup> mutant (Perugino et al. 2012; Vettone et al. 2016). This

feature hampered the determination of second order constants at temperatures above 50.0 °C, since its reaction rate went beyond the technical limits of our assay. Such a high activity at moderate temperatures was obtained with the H<sup>5</sup> mutant by replacing the conserved S132 with

**Fig. 2** The OGT from *Thermotoga neapolitana*. **a** Lyophilized *T. neapolitana* cells grown in the presence of CO<sub>2</sub>, were resuspended in PBS 1× buffer and BG-FL 5.0 μM and incubated 120 min at the indicated temperatures; NT, resuspended cells immediately loaded on SDS-PAGE. The *Tn*OGT gene was expressed in *E. coli* and protein was purified by His-tag affinity chromatography, as described in the “Materials and methods”. Lane M: molecular weight marker; lane 1: cell-free extract; lane 2: flowthrough; lanes 3 and 4: column washing; lanes 5–8: eluted protein by imidazole gradient. Filled and empty green arrows indicate labelled proteins and the BG-FL substrate, respectively. **b** Superimposition between the *Ss*OGT 3D structure (PDB ID: 4ZYE; coloured as described in the legend) and a model of the *Tn*OGT (in gray). Insets represent a zoom-view of local ionic interactions in *Ss*OGT compared with the same positions in *Tn*OGT. Atoms are coloured according the CPK convention (carbon, in the corresponding colour of each 3D structure; nitrogen, in blue; oxygen, in red; sulphur, in yellow)

a glutamic acid residue. This replacement was also performed on the SNAP-tag®: the substitution led to a strong enhancement of the activity of both these engineered OGTs towards the O<sup>6</sup>-BG derivative substrates (Juillerat et al. 2003; Perugino et al. 2012). However, the superimposition analysis between a *Tn*OGT model (constructed by the i-TASSER freeware; <https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) and the 3D structure of the free form of *Ss*OGT (PDB ID: 4ZYE) revealed the presence of the conserved serine in *Tn*OGT (Fig. 2b): evidently, other residues in the active site contribute to the exceptional catalytic activity of this enzyme.

On the other hand, *Tn*OGT was unusually prone to degradation during the storage at –20.0 °C, showing a series of bands below the full length, both in fluorescence imaging and in coomassie staining (Fig. S1a). This lability hampered measurements of thermostability from the DSF analysis. The reason for this degradation is unexpected: the recombinant enzyme is in its full-length form in *E. coli* cells, as well as after heat treatment of the cell free extract, and finally after purification by affinity chromatography (see “Materials and methods”; Fig. 2a). From coomassie staining analysis, three main bands of approx. 17.5, 16.1 and 14.7 kDa were detected (Fig. S1). Since they were also fluorescent (and, therefore, catalytically active), it is probable that these cuts are at the expense of the N-terminal domain, still keeping the polypeptides joined together, and making the enzyme catalytically active in solution. Again, from the superimposition, in *Tn*OGT some residues that play an important role in the stabilization of *Ss*OGT are missing. In particular, the ionic interactions that have a crucial role in the stability of the *Saccharolobus* enzyme at high temperatures, as the R133-D27 pair (Perugino et al. 2015) and the K-48 network (Morrone et al. 2017), are mainly replaced by hydrophobic residues in the *Thermotoga* homolog.

## The *Pyrococcus furiosus* O<sup>6</sup>-alkylguanine-DNA-alkyl-transferase

*Pyrococcus furiosus* is one of the best-studied representatives among microorganisms able to thrive above the boiling point of water (Kengen 2017). It was originally isolated anaerobically from geothermally heated marine sediments of Porto Levante (Vulcano Island, Italy), and described in 1986 by Karl Stetter and Gerhard Fiala (Fiala and Stetter 1986). This microorganism in the last decades was a source of thermostable enzymes, with potential applications in various industrial processes: the most famous example is the DNA polymerase I described in 1991 (Lundberg et al. 1991), possessing an associated 3′–5′ exonuclease activity (Kengen 2017). Recently, the discovery of the CRISPR-Cas systems in *P. furiosus* provides fundamental knowledge for new biomedical and biotechnological applications (Hael et al. 2009; Terns and Terns 2013).

In 1998, Margison and co-workers demonstrated the presence of an OGT activity in *P. furiosus*, identifying a 22.0 kDa size band by an SDS-PAGE fluorography assay

**Table 2** DNA repair activity of OGTs by competitive inhibition studies in the presence of BG-FL (used as substrate) and a ds oligonucleotide containing an O<sup>6</sup>-methylated-guanine (as inhibitor; see Table 1)

	IC <sub>50</sub> (μM)	Notes
<i>Ss</i> OGT	1.01 ± 0.08	From Perugino et al. (2015)
<i>Tn</i> OGT	0.53 ± 0.13	This study
<i>Pfu</i> OGT	0.88 ± 0.10 <sup>a</sup>	This study

<sup>a</sup>Performed at 65 °C

**Table 3** Catalytic activities as a function of temperature of thermostable OGTs, expressed as second-order rate constant values in the presence of the sole BG-FL substrate

	T (°C)	K (s <sup>-1</sup> M <sup>-1</sup> )	Notes
<i>Ss</i> OGT	25	2.80 × 10 <sup>3</sup>	From Perugino et al. (2012)
	50	1.50 × 10 <sup>4</sup>	This study
	70	5.33 × 10 <sup>4</sup>	From Perugino et al. (2012)
	80	ND	This study
<i>Tn</i> OGT	25	4.65 × 10 <sup>4</sup>	This study
	50	2.19 × 10 <sup>4</sup>	This study
	70	ND	This study
	80	ND	This study
<i>Pfu</i> OGT	25	ND	This study
	50	1.80 × 10 <sup>3</sup>	This study
	70	2.30 × 10 <sup>3</sup>	This study
	80	1.20 × 10 <sup>4</sup>	This study
	90	1.50 × 10 <sup>5</sup>	This study

ND not determined

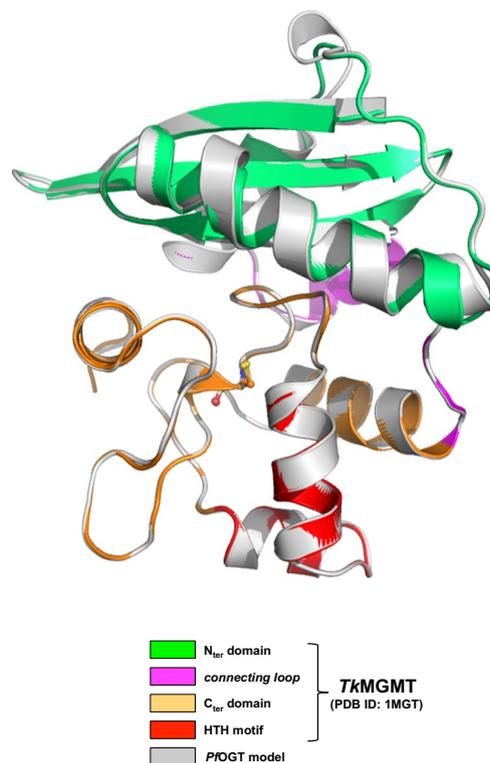
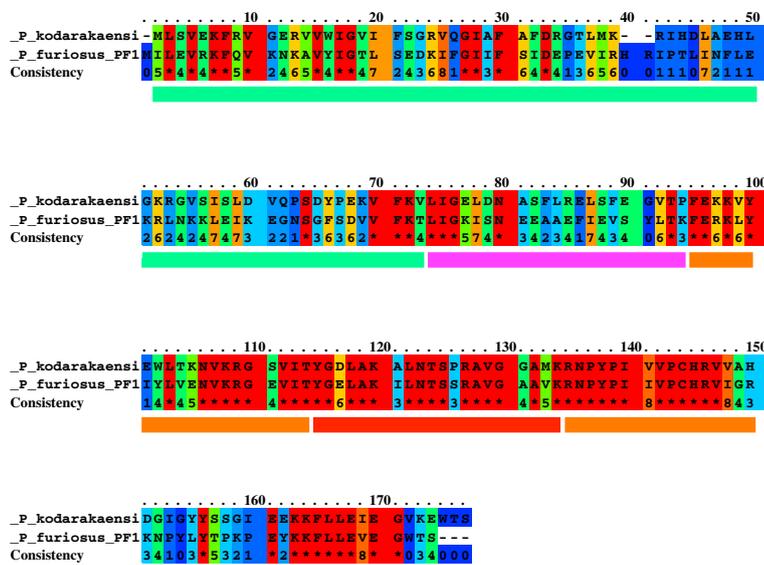
(Skorvaga et al. 1998). Furthermore, this activity was completely abolished by the treatment with the  $O^6$ -benzyl-guanine ( $O^6$ -BG) inhibitor. This information opened the possibility of employing a hyper-thermostable enzyme as SNAP-tag<sup>®</sup> in in vivo CRISPR-Cas system-based applications. The ORF PF1878 is relative to a 174-aa polypeptide, with an expected m.w. of 20.1 kDa: the primary structure is closed to that from the MGMT from *Thermococcus kodakarensis* (*Tk*-MGMT), a well-known enzyme, in terms of structure and thermal stability (Fig. 3) (Leclere et al. 1998; Hashimoto et al. 1999; Nishikori et al. 2005). *Tk*-MGMT is a very thermostable enzyme, and from its solved 3D structure (PDB ID: 1MGT) emerges that a lot of intra-helix ion-pairs contribute to reinforce stability of  $\alpha$ -helices, whereas the presence of inter-helix ion-pairs stabilize internal packing of tertiary structure (Hashimoto et al. 1999).

The cloning of the ORF PF1878 and the subsequent expression and purification of the relative protein allowed to a complete characterization of this new enzyme. Likewise *Tn*OGT, the enzyme from *P. furiosus* is fully active on BG-FL substrate and displayed a clear ability to repair methylated DNA, as shown in the IC<sub>50</sub> experiment listed in Table 2. On the other hand, *Pfu*OGT is a strong thermophilic enzyme, displaying a measurable catalytic activity only at very high temperatures (Table 4), whereas at moderate

temperatures the rate of the reaction is slow, making difficult to perform the fluorescent assay. For this reason, the competitive inhibition in the presence of methylated DNA was performed at 65.0 °C instead of the standard procedure at 50.0 °C (Perugino et al. 2015; Morrone et al. 2017): however, at this temperature, the activity of *Pfu*OGT on single- and/or double-methylated DNA cannot be excluded.

The stability of this enzyme was compared with that of *Ss*OGT by the Differential Scan Fluorimetry analysis: the latter was previously treated using a scan rate of 5 min/°C × cycle (Perugino et al. 2015; Vettone et al. 2016) instead of the classical 1 min/°C × cycle (Niesen et al. 2007). Due to its very high thermal stability, a further increase of the time (10 min/°C × cycle) was necessary for *Pfu*OGT, to obtain the sigmoidal curve to fit with the Boltzmann equation (Niesen et al. 2007). In these new conditions, the  $T_m$  value of *Ss*OGT drops by approx. 13.0 °C ( $67.9 \pm 1.1$ ; Table 4), whereas *Pfu*OGT displayed a  $T_m$  value over 80.0 °C. The stability of *Pfu*OGT was also tested in the presence of perturbants, as high ionic strength or detergents. In the first case, we tested the importance of ionic interactions involved in the maintenance of the structure: as expected, the presence of 4.0 M NaCl strongly affected the stability of *Pfu*OGT, whereas *Ss*OGT has even shown an increase of the  $T_m$  value. Nevertheless, *Ss*OGT is more sensitive to the SDS,

Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved



**Fig. 3** Primary structure alignment and superimposition between the *Pfu*OGT and the OGT from *T. kodakarensis* (*Tk*-MGMT; PDB ID: 1MGT). Conservation of residues and protein domains are coloured on the basis of the respective legends

**Table 4** Protein thermal stability by DSF method.

	Conditions	$T_m$ (°C)	Rate (min/°C×cycle)	Notes
SsOGT	PBS 1×	80.0±0.4	5	From Perugini et al. (2015)
	PBS 1×	67.9±1.1	10	This study
	PBS 1×; NaCl 1.0 M	79.6±0.3	10	This study
	PBS 1×; NaCl 4.0 M	82.9±0.4	10	This study
PfuOGT	PBS 1×; SDS 0.01%	46.5±1.7	10	This study
	PBS 1×	78.8±0.4	10	This study
	PBS 1×; NaCl 1.0 M	83.7±0.3	10	This study
	PBS 1×; NaCl 4.0 M	50.4±2.1	10	This study
	PBS 1×; SDS 0.01%	71.8±0.3	10	This study

$T_m$  values were obtained by plotting the relative fluorescence intensity as a function of temperature. Data were achieved from three independent experiments

showing a collapsed  $T_m$  value up to 22.0 °C, while *Pfu*OGT drops by about 10.0 °C. Taken together, the data obtained clearly indicate a difference in the strategies of protein stabilization adopted by these two enzymes (Table 4).

## Conclusion and perspectives

The modification of AGTs to produce new protein-tags for use in the “SNAP-tag® technology” offers a lot of advantages for the specific labelling of a protein of interest with an innumerable number of chemical groups conjugated to a classical inhibitor of this class of enzymes, the *O*<sup>6</sup>-BG (Fig. 1b) (Hinner and Johnsson 2010). A further step forward was to adapt to this new biotechnology to a thermostable OGT from a microorganism that lives at high temperatures. After the production of *Ss*OGT-H<sup>5</sup> mutant, here we propose two new thermostable OGTs, which will be modified in the future to abolish their ability to bind DNA, without, however, decreasing their activity and stability to heat and to denaturing agents. The new protein-tags can be used in *T. neapolitana* and *P. furiosus*, to analyze in vivo the functions of proteins and enzymes of interest in these model systems.

**Acknowledgements** This work was supported by Fondazione CARI-PLO (Ricerca biomedical condotta da giovani ricercatori, project 2016-0604).

## References

- Aliye N, Fabbretti A, Lupidi G, Tsekoa T, Spurio R (2015) Engineering color variants of green fluorescent protein (GFP) for thermostability, pH-sensitivity, and improved folding kinetics. *Appl Microbiol Biotechnol* 99:1205–1216
- Aoki K, Natsume A (2019) Overview of DNA methylation in adult diffuse gliomas. *Brain Tumor Pathol* 36:84–91
- Ashby MC, Ibaraki K, Henley JM (2004) It's green outside: tracking cell surface proteins with pH-sensitive GFP. *Trends Neurosci* 27:257–261
- Belkin S, Wirsén CO, Jannasch HW (1986) A new sulfur-reducing, extremely thermophilic eubacterium from a submarine thermal vent. *Appl Environ Microbiol* 51:1180–1185
- Campbell TN, Choy FYM (2000) The effect of pH on green fluorescent protein: a brief review. *Mol Biol Today* 2:1–4
- Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC (1994) Green fluorescent protein as a marker for gene expression. *Science* 263:802–805
- Connors SB, Mongodin EF, Johnson MR, Montero-Ci Nelson KE, Kelly RM (2006) Microbial biochemistry, physiology, and biotechnology of hyperthermophilic *Thermotoga* species. *FEMS Microb Rev* 30:872–905
- Coulter R, Blandino M, Tomlinson JM, Pauly GT, Krajewska M, Moschel RC, Peterson LA, Pegg AE, Spratt TE (2007) Differences in the rate of repair of *O*<sup>6</sup>-alkylguanines in different sequence contexts by *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase. *Chem Res Toxicol* 20:1966–1971
- d'Ippolito G, Dipasquale L, Vella FM, Romano I, Gambacorta A, Cutignano A, Fontana A (2010) Hydrogen metabolism in the extreme thermophile *Thermotoga neapolitana*. *Int J Hydrog En* 35:2290–2295
- d'Ippolito G, Dipasquale L, Fontana A (2014) Recycling of carbon dioxide and acetate as lactic acid by the hydrogen-producing bacterium *Thermotoga neapolitana*. *Chemoschem* 7:2678–2683
- Daniels DS, Mol CD, Arvai AS, Kanugula S, Pegg AE, Tainer JA (2000) Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. *EMBO J* 19:1719–1730
- Daniels DS, Woo TT, Luu KX, Noll DM, Clarke ND, Pegg AE, Tainer JA (2004) DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat Struct Mol Biol* 11:714–720
- Del Prete S, Merlo R, Valenti A, Mattossovich R, Rossi M, Carginale M, Supuran CT, Perugini G, Capasso C (2019) Thermostability enhancement of the  $\alpha$ -carbonic anhydrase from *Sulfolobus solfataricus* by using the anchoring-and-self-labelling-protein-tag system (ASL<sup>tag</sup>). *J Enz Inhib Med Chem* 34:946–954
- Dipasquale L, d'Ippolito G, Fontana A (2014) Capnophilic lactic fermentation and hydrogen synthesis by *Thermotoga neapolitana*: an unexpected deviation from the dark fermentation model. *Int J Hydrog En* 39:4857–4862

- Donaldson T, Iozzino L, Deacon LJ, Billones H, Ausili A, D'Auria S, Dattelbaum JD (2017) Engineering a switch-based biosensor for arginine using a *Thermotoga maritima* periplasmic binding protein. *An Biochem* 525:60–66
- Fang Q, Kanugula S, Pegg AE (2005) Function of domains of human *O*<sup>6</sup>-alkyl-guanine-DNA alkyltransferase. *Biochemistry* 44:15396–15405
- Fiala G, Stetter KO (1986) *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaeobacteria growing optimally at 100 °C. *Arch Microb* 145:56–61
- Fink M, Trunk S, Hall M, Schwab H, Steiner K (2016) Engineering of TM1459 from *Thermotoga maritima* for increased oxidative alkene cleavage activity. *Front Microb*. <https://doi.org/10.3389/fmicb.2016.01511>
- Gautier A, Juillerat A, Heinis C, Corrêa IR Jr, Kindermann M, Beaufile F, Johnsson K (2008) An engineered protein-tag for multi-protein labeling in living cells. *Chem Biol* 15:128–136
- Gronemeyer T, Chidley C, Juillerat A, Heinis C, Johnsson K (2006) Directed evolution of *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Prot Eng Des Sel* 19:309–316
- Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, Terns RM, Terns MP (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 139:945–956
- Han D, Xu Z (2017) Development of a pyrE-based selective system for *Thermotoga* sp. strain RQ7. *Extremophiles* 21:297–306
- Hashimoto H, Inouel T, Nishioka M, Fujiwara S, Takagi M, Imanaka T, Kai Y (1999) Hyperthermostable protein structure maintained by intra and inter-helix ion-pairs in archaeal *O*<sup>6</sup>-methylguanine-DNA methyltransferase. *J Mol Biol* 292:707–716
- Hinner MJ, Johnsson K (2010) How to obtain labeled proteins and what to do with them. *Curr Opin Biotechnol* 21:766–776
- Ishiguro K, Shyam K, Penketh PG, Sartorelli AC (2008) Development of an *O*<sup>6</sup>-alkyl-guanine-DNA alkyltransferase assay based on covalent transfer of the benzyl moiety from [benzene-3H]*O*<sup>6</sup>-benzylguanine to the protein. *Anal Biochem* 383:44–51
- Jannasch HW, Huber R, Belkin S, Stetter KO (1988) *Thermotoga neapolitana* sp. nov. of the extremely thermophilic, eubacterial genus *Thermotoga*. *Arch Microbiol* 150:103–104
- Juillerat A, Gronemeyer T, Keppler A, Gendreizig S, Pick H, Vogel H, Johnsson K (2003) Directed evolution of *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. *Chem Biol* 10(4):313–317
- Kengen SWM (2017) *Pyrococcus furiosus*, 30 years on. *Microb Biotechnol* 10:1441–1444
- Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel Johnsson K (2003) A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat Biotechnol* 21:86–89
- Keppler A, Pick H, Arrivoli C, Vogel H, Johnsson K (2004) Labeling of fusion proteins with synthetic fluorophores in live cells. *Proc Natl Acad Sci U S A* 10:9955–9959
- Leclere MM, Nishioka M, Yuasa T, Fujiwara S, Takagi M, Imanaka T (1998) The *O*<sup>6</sup>-methylguanine-DNA methyltransferase from the hyperthermophilic archaeon *Pyrococcus* sp. KOD1: a thermostable repair enzyme. *Mol Gen Genet* 258:69–77
- Lo-Gullo G, Mattosovich R, Perugino G, La Teana A, Londei P, Benelli D (2019) Optimization of an in vitro transcription/translation system based on *Sulfolobus solfataricus* cell lysate. *Archaea* 1:1–10
- Lundberg KS, Shoemaker DD, Adams MWW, Short JM, Sorge JA, Mathur EJ (1991) High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* 108:1–6
- Luu KX, Kanugula S, Pegg AE, Pauly GT, Moschel RC (2002) Repair of oligodeoxyribonucleotides by *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase. *Biochemistry* 41:8689–8697
- Merlo R, Del Prete S, Valenti A, Mattosovich R, Carginale V, Supuran CT, Capasso C, Perugino G (2019) An AGT-based protein-tag system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane. *J Enz Inhib Med Chem* 34:490–499
- Miggiano R, Casazza V, Garavaglia S, Ciaramella M, Perugino G, Rizzi M, Rossi F (2013) Biochemical and structural studies of the *Mycobacterium tuberculosis* *O*<sup>6</sup>-methylguanine methyltransferase and mutated variants. *J Bacteriol* 195:2728–2736
- Miggiano R, Valenti A, Rossi F, Rizzi M, Perugino G, Ciaramella M (2017) Every OGT is illuminated ... by fluorescent and synchrotron lights. *Int J Mol Sci* 18:2613–2631
- Mishina Y, Duguid EM, He C (2006) Direct reversal of DNA alkylation damage. *Chem Rev* 106:215–232
- Mollwitz B, Brunk E, Schmitt S, Pojer F, Bannwarth M, Schiltz M, Rothlisberger U, Johnsson K (2012) Directed evolution of the suicide protein *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase for increased reactivity results in an alkylated protein with exceptional stability. *Biochemistry* 51:986–994
- Morrone C, Miggiano R, Serpe M, Massarotti A, Valenti A, del Monaco G, Rossi M, Rossi F, Rizzi M, Perugino G, Ciaramella M (2017) Interdomain interactions rearrangements control the reaction steps of a thermostable DNA alkyltransferase. *Biochem Biophys Acta* 1861:86–96
- Niesen FH, Berglund H, Vedadi M (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* 2:2212–2221
- Nishikori S, Shiraki K, Fujiwara S, Imanaka T, Takagi M (2005) Unfolding mechanism of a hyperthermophilic protein *O*<sup>6</sup>-methylguanine-DNA methyltransferase. *Biophys Chem* 116:97–104
- Pegg AE (2011) Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA damage, resistance to chemotherapy, and research tools. *Chem Res Toxicol* 24:618–639
- Pegg AE, Goodtzova K, Loktionova NA, Kanugula S, Pauly GT, Moschel RC (2001) Inactivation of human *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase by modified oligodeoxyribonucleotides containing *O*<sup>6</sup>-benzylguanine. *J Pharmacol Exp Ther* 296:958–965
- Perugino G, Vettone V, Illiano G, Valenti A, Ferrara MC, Rossi M, Ciaramella M (2012) Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage. *J Biol Chem* 287:4222–4231
- Perugino G, Miggiano R, Serpe M, Vettone A, Valenti A, Lahiri S, Rossi F, Rossi M, Rizzi M, Ciaramella M (2015) Structure-function relationships governing activity and stability of a DNA alkylation damage repair thermostable protein. *Nucleic Acids Res* 43:8801–8816
- Pradhan N, Dipasquale L, d'Ippolito G, Panico A, Lens PNL, Esposito G, Fontana A (2015) Hydrogen production by the thermophilic bacterium *Thermotoga neapolitana*. *Int J Mol Sci* 16:12578–12600
- Pradhan N, Dipasquale L, d'Ippolito G, Fontana A, Panico A, Lens PNL, Pirozzi F, Esposito G (2016) Kinetic modeling of fermentative hydrogen production by *Thermotoga neapolitana*. *Int J Hydrog Energy* 41:4931–4940
- Pradhan N, Dipasquale L, d'Ippolito G, Panico A, Lens PNL, Esposito G, Fontana A (4359T) Hydrogen and lactic acid synthesis by the wild-type and a laboratory strain of the hyperthermophilic bacterium *Thermotoga neapolitana* DSMZ 4359T under capnophilic lactic fermentation conditions. *Int J Hydrog Energy* 42:16023–16030
- Samuelson P, Gunneriusson E, Nygren PA, Ståhl S (2002) Display of proteins on bacteria. *J Biotechnol* 96:129–154
- Serpe M, Forenza C, Adamo A, Russo N, Perugino G, Ciaramella M, Valenti A (2019) The DNA Alkylguanine DNA

- Alkyltransferase-2 (AGT-2) of *Caenorhabditis Elegans* is involved in meiosis and early development under physiological conditions. *Sci Rep* 9:6889
- Skorvaga M, Raven NDH, Margison GP (1998) Thermostable archaeal *O*<sup>6</sup>-alkylguanine-DNA alkyltransferases. *Proc Natl Acad Sci USA* 95:6711–6715
- Sun G, Zhao L, Zhong R, Peng Y (2018) The specific role of *O*<sup>6</sup>-methylguanine-DNA methyltransferase inhibitors in cancer chemotherapy. *Future Med Chem* 10:1971–2199
- Terns RM, Terns MP (2013) The RNA- and DNA-targeting CRISPR-Cas immune systems of *Pyrococcus furiosus*. *Biochem Soc Trans* 41:1416–1421
- Tsien RY (1998) The green fluorescent protein. *Annu Rev Biochem* 67:509–544
- Tubbs JL, Pegg AE, Tainer JA (2007) DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair* 6:1100–1115
- Vettone A, Serpe M, Hidalgo A, Berenguer J, del Monaco G, Valenti A, Rossi M, Ciaramella M, Perugino G (2016) A novel thermostable protein-tag: optimization of the *Sulfolobus solfataricus* DNA-alkyl-transferase by protein engineering. *Extremophiles* 20:1–13
- Visone V, Han W, Perugino G, del Monaco G, She Q, Rossi M, Valenti A, Ciaramella M (2017) In vivo and in vitro protein imaging in thermophilic archaea by exploiting a novel protein tag. *PLoS ONE* 12:e0185791
- Yang CG, Garcia K, He C (2009) Damage detection and base flipping in direct DNA alkylation repair. *ChemBioChem* 10:417–423
- Zhang J, Shi H, Xu L, Zhu X, Li X (2015) Site-directed mutagenesis of a hyperthermophilic endoglucanase Cel12B from *Thermotoga maritima* based on rational design. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0133824>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Review

# $O^6$ -alkylguanine-DNA Alkyltransferases in Microbes Living on the Edge: From Stability to Applicability

Rosanna Mattosovich <sup>1,†</sup>, Rosa Merlo <sup>1,†</sup>, Riccardo Miggiano <sup>2</sup>, Anna Valenti <sup>1,\*</sup> and Giuseppe Perugino <sup>1,\*</sup>

<sup>1</sup> Institute of Bioscience and BioResources, National Research Council of Italy, Via Pietro Castellino 111, 80131 Naples, Italy; rosanna.mattosovich@ibbr.cnr.it (R.M.), rosa.merlo@ibbr.cnr.it (R.M.)

<sup>2</sup> Department of Pharmaceutical Sciences, University of Piemonte Orientale, Via Bovio 6, 28100 Novara, Italy; riccardo.miggiano@uniupo.it

\* Correspondence: anna.valenti@ibbr.cnr.it (A.V.); giuseppe.perugino@ibbr.cnr.it (G.P.); Tel.: +39-081-6132-247 (A.V.); Tel.: +39-081-6132-496 (G.P.); Fax: +39-081-6132-646 (A.V. & G.P.)

† these authors contribute equally to this work

Received: 26 March 2020; Accepted: 16 April 2020; Published: 20 April 2020

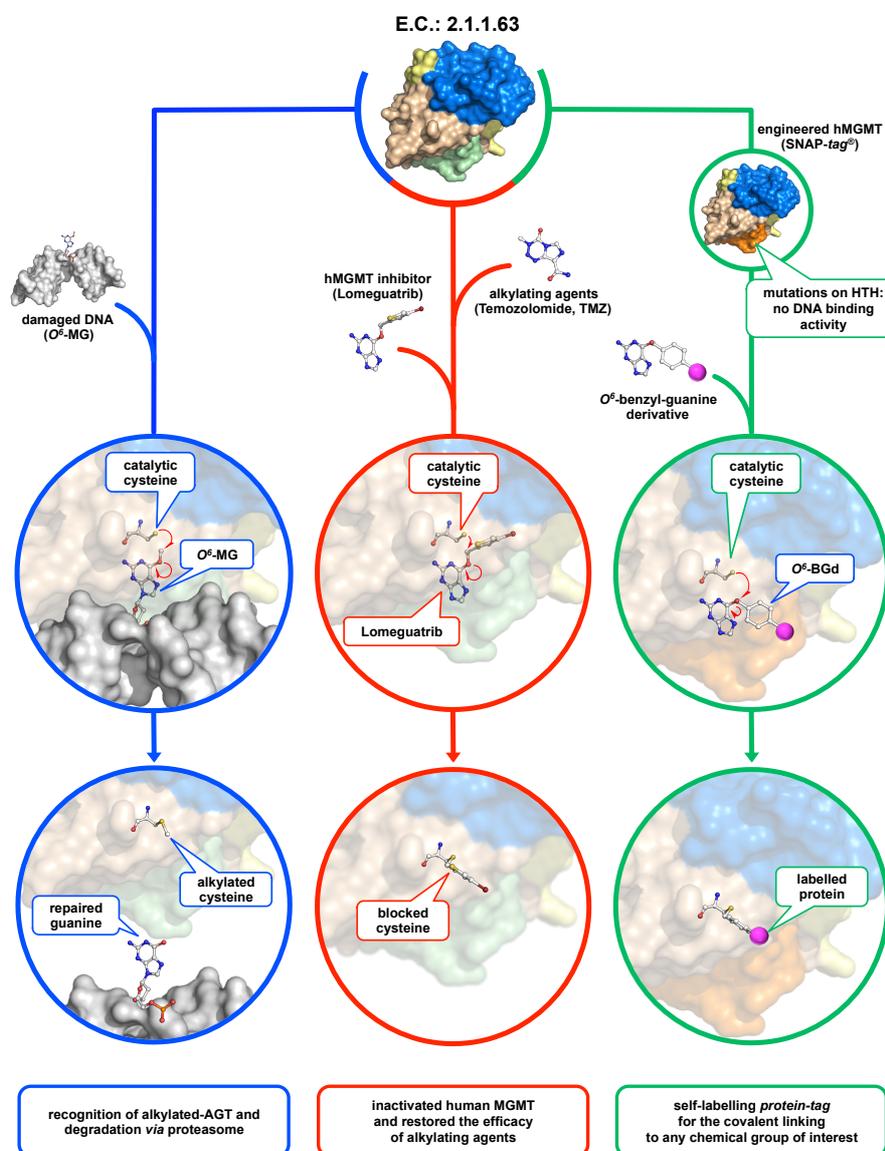
**Abstract:** The genome of living cells is continuously exposed to endogenous and exogenous attacks, and this is particularly amplified at high temperatures. Alkylating agents cause DNA damage, leading to mutations and cell death; for this reason, they also play a central role in chemotherapy treatments. A class of enzymes known as AGTs (alkylguanine-DNA-alkyltransferases) protects the DNA from mutations caused by alkylating agents, in particular in the recognition and repair of alkylated guanines in  $O^6$ -position. The peculiar irreversible self-alkylation reaction of these enzymes triggered numerous studies, especially on the human homologue, in order to identify effective inhibitors in the fight against cancer. In modern biotechnology, engineered variants of AGTs are developed to be used as *protein tags* for the attachment of chemical ligands. In the last decade, research on AGTs from (hyper)thermophilic sources proved useful as a model system to clarify numerous phenomena, also common for mesophilic enzymes. This review traces recent progress in this class of *thermozymes*, emphasizing their usefulness in basic research and their consequent advantages for in vivo and in vitro biotechnological applications.

**Keywords:** thermophilic sources; DNA repair; biotechnological tools, alkylation damage; AGT

## 1. Introduction

Monofunctional alkylating agents, a class of mutagenic and carcinogenic agents present in the environment, induce DNA alkylation in several positions including guanine at  $O^6$  ( $O^6$ -MG; 6% of adducts formed), the  $N^7$  of guanine ( $N^7$ -MG; 70%), and the  $N^3$  of adenine ( $N^3$ -MA; 9%) [1]. Alkylation of guanine ( $O^6$ -AG) is a cytotoxic lesion, although the specific mechanism of this cytotoxicity is not yet fully understood. It was proposed that the toxic effect occurs after DNA replication, because the  $O^6$ -AG incorrectly base-pairs with thymine generating a transition from G:C to A:T [2]. The mutations caused by  $O^6$ -MG that occur at the time of replication are recognized by the post-replication mismatch repair system with potential harmful implications for cell viability. Apart from conventional DNA repair pathways as Mismatch Excision Repair (MMR), Nucleotide Excision Repair (NER), Base Excision Repair (BER), alkylated-DNA protein alkyl-transferases (called  $O^6$ -alkyl-guanine-DNA-alkyl-transferase (AGT or OGT) or  $O^6$ -methyl-guanine-DNA-alkyl-transferase (MGMT); EC: 2.1.1.63) perform the direct repair of alkylation damage in DNA [3,4]. They represent the major factor in counteracting the effects of alkylating agents that form such adducts [4]. These are small enzymes (17–22 kDa) that are widely present in organisms of the three kingdoms (bacteria, archaea, eukaryotes) but apparently absent from plants,

*Schizosaccharomyces pombe*, *Thermus thermophilus*, and *Deinococcus radiodurans*. The reaction mechanism of AGTs is based on the recognition of the damaged nucleobase on DNA [5], followed by a one-step  $SN_2$ -like mechanism, in which the alkyl group of the damaged guanine is irreversibly transferred to a cysteine residue in its active site [5–8] (Figure 1, blue path).



**Figure 1.** The O<sup>6</sup>-alkyl-guanine-DNA-alkyl-transferase (AGT)s' world. AGTs are small enzymes composed by a N-terminal domain (in sky blue), a C-terminal domain (in light brown) connected by a loop (in yellow). In the C-terminal domain, a helix-turn-helix motif (in light green, or in orange in the SNAP-tag) is responsible of the DNA-binding activity. The peculiar irreversible reaction mechanism of these enzymes plays a pivotal role in the physiological DNA repair (blue path), and it has important repercussions in cancer cell treatment (red path) and biotechnological applications (green path). Atoms are coloured by the CPK colour convention.

For these reasons, they are also called *suicide* or *kamikaze* proteins, showing a 1:1 stoichiometry of their reaction with the natural substrate. The disadvantage of this elegant catalysis is that, upon alkylation, the protein is self-inactivated and destabilized, triggering its recognition by cellular systems to be degraded by the proteasome [8,9].

### 1.1. AGTs as Targets in Cancer Therapy

Alkylation damage to DNA occurs in various living conditions, and for this reason the widespread presence of AGT protects cells from killing by alkylating agents. However, human AGT (hMGMT) is a *double-edged sword*—on the one hand, it protects healthy cells from these genotoxic and carcinogenic effects, but also counteracts alkylating agents-based chemotherapy by protecting cancer cells from the killing effect of these drugs [10,11]. Consequently, hMGMT has emerged as a crucial factor in anticancer therapies [12]—an inverse relationship has been discovered between the presence of hMGMT and the sensitivity of cells to the cytotoxic effects of alkylating agents, such as temozolomide (TMZ), in different types of cancer cells, including prostate, breast, colon, and lung cancer cells [13].

The resistance to chemotherapy may be reduced by inhibition of these enzymes; as described before, after removing the lesion, the alkylated form of the protein is inactivated and enters the intracellular degradation pathways. Hence, in order to counteract the action of hMGMT in chemotherapy regimens, a large number of studies aimed to the develop hMGMT inhibitors to be used in combination with alkylating agents. In view of this therapeutic relevance, much success has been obtained through the design of hMGMT pseudo-substrates, namely, the *O*<sup>6</sup>-benzylguanine (*O*<sup>6</sup>-BG) and the strong inactivator *O*<sup>6</sup>-[4-bromophenyl]-guanine (*O*<sup>6</sup>-BTG, Lomeguatrib) [13,14]. These compounds mimic damaged guanine on DNA and react with the protein by the covalent transfer of the alkyl adduct to the active site cysteine residue, thus irreversibly inactivating the enzyme (Figure 1, red path). Therapeutically, *O*<sup>6</sup>-BG is not toxic on its own, but renders cancer cells 2 to 14 times more sensitive to alkylating agents' effects. The oligonucleotides containing several *O*<sup>6</sup>-BG are potent inhibitors and represent a valid alternative to the use of free modified guanines, thereby improving the activity of the alkylating chemotherapy drug in the treatment of some tumours [15–17].

## 1.2. AGTs and Biotechnology

The specific labelling of proteins with synthetic probes is an important advance for the study of protein function. To achieve this, the protein of interest is expressed in a fusion with additional genetically encoded polypeptides, called *tags*, which mediate the labelling. The first example of an *autofluorescent tag* was the *Aequorea victoria* green fluorescent protein (GFP) allowing the *in vivo* localization of fusion proteins in cellular and molecular biology [18,19]. Among *affinity tags*, of particular importance are the poly(His)-*tag*, the chitin-binding protein, the maltose-binding protein [20], the Strep-*tag* [21], and the glutathione-S-transferase (GST-*tag*) [22], which allow fast and specific purification of proteins of interest from their crude biological source using affinity techniques. *Solubilization tags* are especially useful to assist the proper folding of recombinant proteins expressed in chaperone-deficient species such as *Escherichia coli*, avoiding protein precipitation and the use of alternative expression protocols [23,24]—these include thioredoxin [25] and poly(NANP).

However, all the *tags* listed above are limited by the fact that each of them can be used for one or only a few applications. The need therefore emerged to develop a *universal tag* that could widely cover several applications.

In 2003, the group headed by Kai Johnsson pioneered the use of an engineered hMGMT variant as a fusion protein for *in vitro* and *in vivo* biotechnology applications, which led then to its commercialization, namely, the SNAP-*tag* (New England Biolabs) [26–29]. They started from the knowledge that hMGMT tolerates the presence of groups conjugated to the pseudo-substrate *O*<sup>6</sup>-BG (*O*<sup>6</sup>-BG derivatives)—the unusual covalent bond with the benzyl moiety can therefore be exploited for “biotech” purposes (Figure 1, green path). Thanks to its small size, the engineered hMGMT (SNAP-*tag*) can be fused with other proteins of interest. The expression of the fusion protein inside the cells followed by incubation with opportune fluorescent derivatives leads to *in vivo* labelling of fusion proteins with the probe, which can be used for localization studies [26]. The same principle has also been used for the immobilization of tagged fusion proteins *in vitro* [30]. This offers a delicate condition for fixing and disposing in a better orientation of a wide range of proteins/enzymes on a surface. The SNAP-*tag* technology was successfully applied to surface

plasmon resonance (SPR) for the covalent immobilization of proteins of interest [31]. Another interesting application of this protein-tag is the possibility to produce new antibody fragments (scFv-SNAP) to be employed in the SPR analysis [32].

Despite the need to use a specific substrate, SNAP-tag offers endless applications—the possibility to covalently link a desired chemical group (conjugated to the O<sup>6</sup>-BG) to a protein of interest (genetically fused to it) makes it decidedly advantageous, if compared to traditional *protein tags* currently in use. Table 1 shows a brief comparison between some examples of *protein tags* and the SNAP-tag in several application fields.

**Table 1.** The use of *protein tags* in some applicative examples.

Applications	FPs	Affinity Tag	SNAP-Tag	Notes
In vivo imaging	+ <sup>a</sup>		+	
Substrate utilization	+	–	–	FPs do not need of any substrate for their fluorescence
Emission spectra	±	–	+	FPs are in a limited number with respect to chemical probes
Time-resolved fluorescence	±	–	+	
Multi-colour fluorescence	±	–	+	For FPs, multi-cloning and expression is necessary
In vitro applications	±	±	+	
Variety of chemical group labelling	–	–	+	
Pulse-chase analysis	–	–	+	Fresh synthesized FPs cannot be efficiently quenched
Anaerobic conditions	–	+	+	FPs' fluorophore formation requires oxygen
Protein purification	+	+	+	Utilization of the GFP-trap matrix
Protein immobilization	+	+	+	Utilization of the GFP-trap matrix
Pull-down experiments	+	+	+	Utilization of the GFP-trap matrix

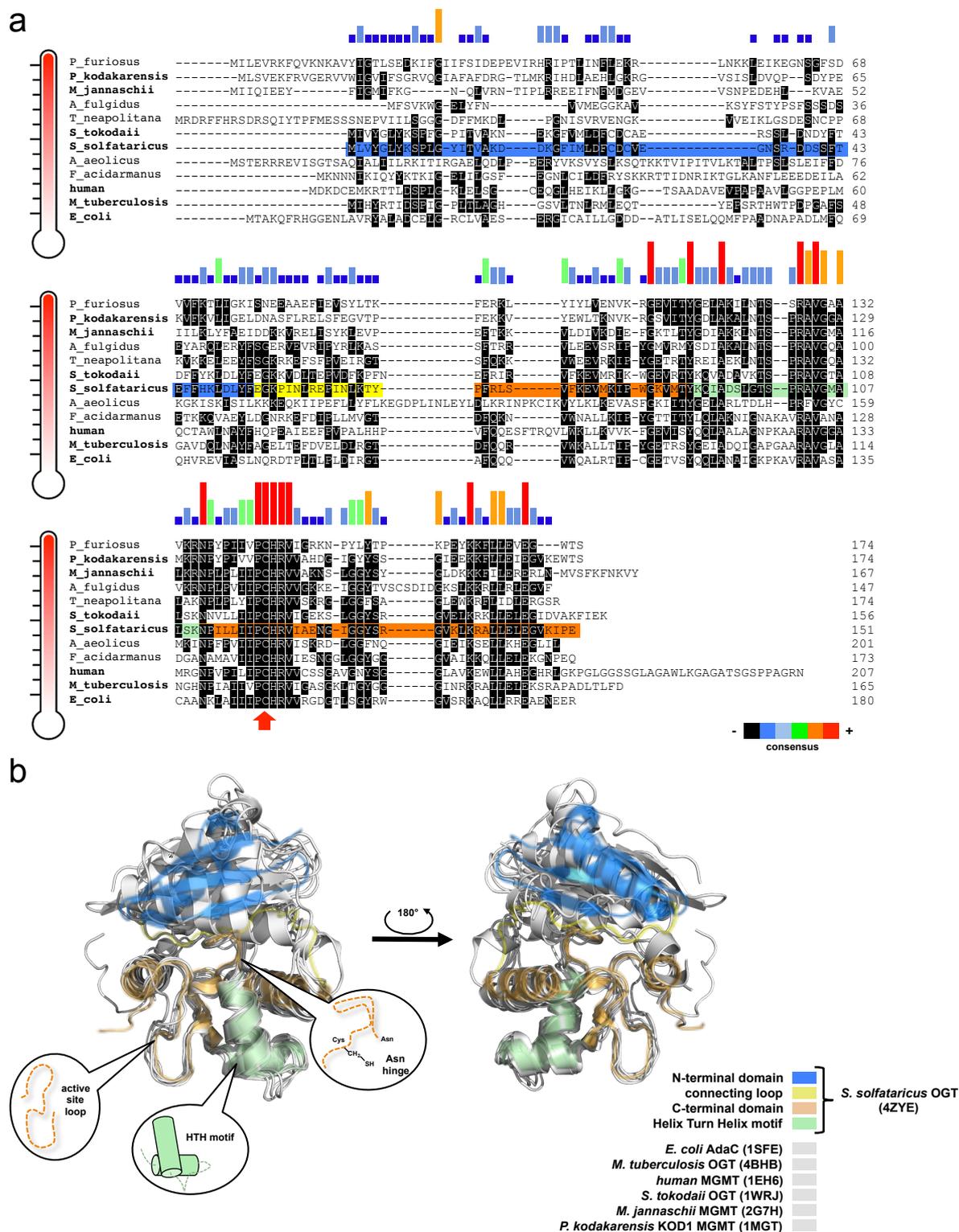
<sup>a</sup>+, fully applicable or advantageous; ±, limited applicability; –, not applicable or disadvantageous; FPs = fluorescent proteins.

## 2. Thermophilic and Thermostable AGTs

As for organisms living under mesophilic conditions, environmental and endogenous alkylating agents also attack the genome of thermophilic and hyper-thermophilic organisms. Additionally, high temperatures accelerate the process of alkylation, leading to DNA breaks [33], because alkylating agents are chemically unstable at the physiological conditions of these organisms, and their collateral decomposition may worsen the formation of DNA alkylation products [34]. Thus, the presence of AGTs and methylpurine glycosylases in hyperthermophilic organisms implies that they are naturally exposed to endogenous methylating agents [34], thus supporting the crucial role of AGTs [35,36].

Apart from some studies on Archaea using cell-free extracts, a few examples of biochemical studies of AGTs from thermophilic sources include the enzymes from *Pyrococcus* sp. KOD1 [35] conducted by Imanaka and co-workers from *Aquifex aeolicus* and *Archaeoglobus fulgidus* performed by the group of Prof. Pegg in 2003 (Figure 2) [34]. Intriguingly, *A. aeolicus* AGT, whose organism

was identified as the most primitive bacterium, is closer to the mammalian AGTs than other bacterial homologues in terms of  $O^6$ -BG sensitivity [34].



**Figure 2.** (a) Alignment of biochemically and structurally (in bold) characterized AGTs. DNA sequences are listed in decreasing order of temperature. The histograms in different colours show the sequence consensus, and the red arrow indicates the highly conserved catalytic cysteine. (b) Superimposition of all known AGT structures in their free form (in grey). All common domains and elements are coloured only for the  $O^6$ -alkyl-guanine-DNA-alkyl-transferase protein from the archaeon *Saccharolobus solfataricus* (hereinafter SsOGT) enzyme. Coloured bars behind the SsOGT sequence in (a) recall the enzyme domains highlighted in the structure and in the legend in (b).

### 2.1. The Common Themes in AGTs' Tertiary Structure and the Intrinsic Factors of Stability

Despite the different primary structures (Figure 2a), thermophilic enzymes show a typical AGT protein architecture, consisting of two domains [37]: a highly conserved C-terminal domain (CTD), surprisingly superimposable for all available AGT structures (Figure 2b), and a N-terminal domain (NTD), which is very different among AGTs and whose function is not well understood (likely involved in regulation, cooperative binding, and stability [6,38,39]). The CTD contains the DNA binding *helix-turn-helix* motif (HTH); the *Asn hinge*, which precedes the -V/IPCHRVV/I- amino acid sequence containing the conserved catalytic cysteine (except the *Caenorhabditis elegans* AGT-2 that has the -PCHP- sequence [40,41]); and the *active site loop*, responsible for the substrate specificity.

A comparative structural analysis performed on AGT proteins whose structures are in the Protein Data Bank revealed significant differences of the intrinsic structural features that have been considered to be relevant for thermostability, such as helix capping, intramolecular contacts (hydrogen bonds, ion-pairs), and solvent-accessible surface areas. Helix capping plays a central role in the stability of  $\alpha$ -helices, due to lack of intra-helical hydrogen bonds in the first and last turn [42,43], and its effect results in an overall structural stabilization of protein folding [44]. By inspecting the crystal structure of SsOGT (PDB ID: 4ZYE), considered here as the thermophilic reference AGT protein, we verified that the five  $\alpha$ -helices of the composing the protein tertiary structure are characterized by the presence of helix capping, this possibly increasing the thermal stability. In particular, the helix *H1* at the NTD is stabilised by a peculiar double serine sequence (S40–S41) and a glutamic acid (E54) at its CTD, the latter is strictly conserved in all AGTs from thermophilic organisms (see Figure 2a). The HTH motif, built on helices *H3* and *H4*, is stabilized at the level of *H3* by a highly conserved threonine residue (T89) as N-cap and a serine (S96), distinctive of SsOGT, as C-cap. Furthermore, helix *H4* contains two serine-based capping among which the one placed at NTD (S100) is strictly conserved in all thermophilic AGTs and is followed by a proline (P101) that fits well in the first turn of the helix thanks to its own backbone conformation. Finally, the helix *H5* is protected by glutamic acid capping that is present in all the AGTs from different species. Another feature contributing to thermal stability is the solvent-accessible surface area (SASA). Indeed, the decrease of SASA and the increase of hydrophobic residues that are buried from the solvent have stabilizing principles for thermostable protein [45]. As described in Table 2, SsOGT shows the smaller total SASA value, in line with its exceptional stability. On the contrary, OGT from *Mycobacterium tuberculosis* [38,46] has a higher value due to the peculiar conformation of both the active site loop and the C-terminal tail that are exposed to the bulk solvent and are less heat stable (Table 2).

Finally, by comparing hyperthermophilic AGTs with the orthologs from mesophilic organisms, in terms of atomic contacts between charged residues as well as intramolecular hydrogen bonds (Table 2), significant differences emerged in the number of charged residues contacts. As expected for thermostable proteins [47], SsOGT, as well as the proteins from *Sulfurisphaera tokodaii* and *Pyrococcus kodakaraensis*, shows a larger number of electrostatic contacts, characterized by higher bond-dissociation energy, with respect to hydrogen bonds for which we did not detect significant differences among the analysed structures, apart from MGMT of *P. kodakaraensis* (*Pk*-MGMT) [48].

Although the number of H-bonds is approximately similar across the AGTs from different organisms, there should be differences in the position-related role of such bonds, supporting overall stability of thermophilic variants. With reference to *Pk*-MGMT, Hashimoto and co-workers detected the same number of ion-pairs between the extremophilic protein and *E. coli* Ada-C [49]; however, more intra- and inter-helix ion pairs were found in *Pk*-MGMT. Although the absence of a correlation between ion pairs' position and stabilization in Ada-C exists, the intra-helix ion pairs act in the secondary structure of *Pk*-MGMT, stabilizing helices, and the inter-helix ion pairs consolidate the inter-domain interactions, enhancing the stability of the tertiary structure packing.

**Table 2.** Comparison of solvent-accessible surface area and intramolecule contacts.

T <sub>opt</sub> (°C)	Enzyme (PDB ID)	Total SASA (Å <sup>2</sup> )	Charged Residues Contacts	Intramolecule H-bonds <sup>a</sup>	References
37	<i>Escherichia coli</i> Ada-C (1SFE)	8421.8	74	141	[49]
37	<i>Mycobacterium tuberculosis</i> OGT (4BHB)	9535.2	56	143	[38]
37	<i>Homo sapiens</i> MGMT (1EH6)	8764.3	71	127	[6]
80	<i>Saccharolobus solfataricus</i> OGT (4ZYE)	8054.1	94	137	[39]
80	<i>Sulfurisphaera tokodaii</i> OGT (1WRJ)	8049.5	124	134	PDB <sup>b</sup>
80	<i>Methanocaldococcus jannashii</i> MGMT (2G7H)	17,770.8 <sup>c</sup>	N.D.	N.D.	[50]
85	<i>Pyrococcus kodakaraensis</i> MGMT (1MGT)	8302.8	111	157	[48]

<sup>a</sup> Excluding intra-residues H-bonds. <sup>b</sup> <https://www.rcsb.org/structure/1wrj>. <sup>c</sup> The structure has been solved by means of NMR explaining the high SASA value.

### 3. The O<sup>6</sup>-Alkylguanine-DNA-Alkyltransferase from *Saccharolobus solfataricus*

In the last decade, SsOGT has been characterized through detailed physiological, biochemical, and structural analysis. Due to its intrinsic stability, the SsOGT protein has proven to be an outstanding model for clarifying the relationships between function and structural characteristics.

*Saccharolobus solfataricus* (previously known as *Sulfolobus solfataricus*) is a microorganism first isolated and discovered in 1980 in the Solfatara volcano (Pisciarelli-Naples, Italy) [51], which thrives in volcanic hot springs at 80 °C and a pH 2.0–4.0 range. In order to protect its genome in these harsh conditions, *S. solfataricus* evolved several efficient protection and repair systems [33,52]. *S. solfataricus* is highly sensitive to the alkylating agent methyl methane sulfonate (MMS), showing a transient growth arrest when treated with MMS concentrations in the range of > 0.25 mM to 0.7 mM [33,52]. Interestingly, although the *ogt* RNA level increases after MMS treatment, the relative enzyme concentration decreases, suggesting its degradation in cells in response to the alkylating agent and, in general, to a cellular stress [52]. Under these treatment conditions, however, the protein level rises after few hours, and, in parallel, the growth of *Saccharolobus* starts again [52], indicating a role of SsOGT in efficient DNA repair by alkylation damage.

#### 3.1. Innovative OGT Assays

Various assays to measure AGT activity are reported in the literature. The first methods were based on the use of oligonucleotides carrying radioactive (<sup>3</sup>H or <sup>14</sup>C) O<sup>6</sup>-alkylguanine groups. Proteinase K digestion was then carried out to measure the levels of marked S-methyl-cysteine in the lysate in an automatic amino acid analyser [53]. A very similar, but simpler and faster radioactive assay was used in another procedure with a <sup>32</sup>P-terminal labelled oligonucleotide containing a modified guanine in a methylation-sensitive restriction enzyme sequence (as *Mbo* I). The AGT DNA repair activity thereby allowed the restriction enzyme to cut [54]. This procedure was also used by Ciaramella's group to identify for the first time the activity of SsOGT [52]. This

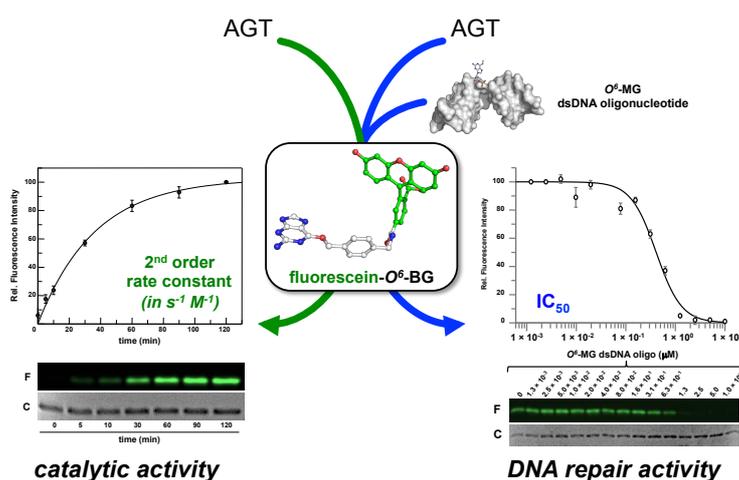
test has the advantage of analysing the digested fragment directly by electrophoresis on a polyacrylamide gel [55].

It was therefore improved in terms of precision by the subsequent separation of the digested oligonucleotides by HPLC. The chromatographic separation allowed the calculation of the concentration of active AGT after measuring the radioactivity of the peak corresponding to the digested fragment [55]. Similarly, Moschel's group developed the analysis of hMGMT reaction products based on HPLC separation in 2002. This test investigated the degree of inhibition of oligonucleotides with  $O^6$ -MG or  $O^6$ -BG in different positions that varied from the 3' to the 5' end and whether they could be used as chemotherapy agents.  $IC_{50}$  values were obtained by quantifying the remaining active protein after the radioactive DNA reaction [56].

Although the assay measures the protein activity, the use of radioactive materials and chromatographic separations made these assays long, tedious, and unsafe.

An alternative approach was proposed in 2010 by the group of Carme Fàbrega, who set up an assay based on the thrombin DNA aptamer (TBA), a single-stranded 15 mer DNA oligonucleotide identified *via* Systematic Evolution of Ligands by EXponential enrichment (SELEX), which in its quadruplex form binds thrombin protease with high specificity and affinity [57]. In this assay, they put a fluorophore and a quencher to the TBA—the quadruplex structure of this oligonucleotide is compromised if a central  $O^6$ -MG is present, preventing the two probes to stay closer. An AGT's repair activity on the oligonucleotide allows the folding of the quadruplex structure and the Förster Resonance Energy Transfer (FRET) energy transfer takes place, resulting in a decrease of the fluorescence intensity [58].

Recently, the introduction of fluorescent derivatives of the  $O^6$ -BG (as SNAP Vista Green, New England Biolabs) made possible the development of a novel DNA alkyl-transferase assay. Because AGT covalently binds a benzyl-fluorescein moiety of its substrate after reaction, it is possible to immediately load the protein product on a SDS-PAGE—the *gel-imaging* analysis of the fluorescence intensity gives a direct measure of the protein activity because of the 1:1 stoichiometry of protein/substrate (Figure 3). Signals of fluorescent protein (corrected by the amount of loaded protein by Coomassie staining analysis) obtained at different times are plotted, and a second order reaction rate is determined [38,39,46,52,59,60]. This method can be applied to all AGTs that bind  $O^6$ -BG, with the exception of the *E. coli* Ada-C [61,62].



**Figure 3.** Innovative fluorescent AGT assay. The substrate could be used alone for the determination of the AGT catalytic activity, or in combination with a competitive non-fluorescent substrate (alkylated-DNA). In the latter case, an indirect measure of the DNA repair activity on natural substrates is determined (adapted from [63]).

Furthermore, an alkylated double strand DNA (dsDNA) oligonucleotide can be included in a competition assay with the fluorescein substrate. This non-fluorescent substrate lowers the final fluorescent signal on *gel imaging* analysis, depending on its concentration. In this way, it is possible

to measure the activity of AGTs for their natural substrate, giving an indirect measure of methylation repair efficiency (Figure 3) [38,39,46,52,59,60]. By using this methodology, it was even possible to discriminate the SsOGT activity regarding the position of the O<sup>6</sup>-MG on DNA (see below; [39]), in line with previous data on hMGMT [64].

### 3.2. Biochemical Properties of *S. solfataricus* OGT

The recombinant SsOGT protein, heterologously expressed in *E. coli*, has been fully characterized using the fluorescent assay described and summarized in Section 3.1, and some results are compiled in Table 3. In agreement with its origin, the protein showed optimal catalytic activity at 80 °C, although retaining a residual activity at lower temperatures (Table 3), and in a pH range between 5.0 and 8.0. As for the most part of many thermophilic enzymes, SsOGT is resistant over a wide range of reaction conditions, such as ionic strength, organic solvents, common denaturing agents, and proteases [52,59]. Interestingly, chelating agents do not affect the activity of this enzyme. Crystallographic data clarified this observation, as the archaeal enzyme lacks a zinc ion in the structure [39], whereas this ion is important for correct folding of hMGMT [6].

### 3.3. Crystal Structure of SsOGT

All catalytic steps of the AGTs' activity (alkylated DNA recognition, DNA repair, irreversible trans-alkylation of the catalytic cysteine, recognition, and degradation of the alkylated protein) have been structurally characterized. Most information comes from the classic studies on hMGMT, as well as the Ada-C and OGT from *Escherichia coli* [5–8,49]. Other AGTs' structures are also available in the Protein Data Bank site (Figure 2a) (<http://www.rcsb.org/pdb/results/results.do?tabtoshow=Current&qrid=D3B02F3B>).

As shown in Figure 1, all AGTs are inactivated after the reaction and degraded via proteasome, whereas in higher organisms, the degradation is preceded by protein ubiquitination [9]. It is a common view that the recognition of alkylated-AGTs is due by a conformational change; however, data on structure and properties of alkylated AGTs are limited because alkylation greatly destabilizes their folding [39]. The methylated-hMGMT and benzylated-hMGMT 3D structures were only obtained by flash-frozen crystals, showing that alkylation of the catalytic cysteine (C145) induces subtle conformational changes [6,7,65]. Consequently, these structures might not reflect the physiological conformation of the alkylated hMGMT [39].

**Table 3.** Biochemical properties comparison among SNAP-tag, SsOGT and the relative H<sup>5</sup> mutant.

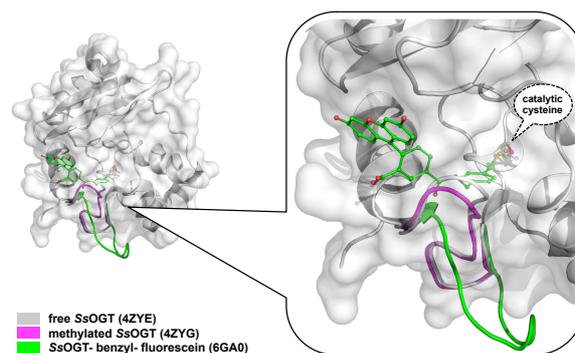
Features	SNAP-tag <sup>®</sup>	SsOGT	SsOGT-H <sup>5</sup>
Molecular weight (kDa) <sup>a</sup>	23.0	17.0	17.0
T <sub>opt</sub> (°C)	37.0	80.0	75.0
Relative activity	at 25.0°C	80%	25%
	at 37.0°C	100%	45%
	at 80°C	-	100%
Catalytic activity at 37 °C (M <sup>-1</sup> s <sup>-1</sup> )	2.8 × 10 <sup>4</sup>	2.8 × 10 <sup>3</sup>	<b>1.6 × 10<sup>4</sup></b>
pH <sub>opt</sub>	6.0	7.5	6.0
Thermal stability T <sub>1/2</sub> (°C)	6 h (37)	3 h (70)	3 h (70)
Thermal stability T <sub>1/2</sub> at 37 °C (h)	6	> 24	> 24
Additives	NaCl	< 0.3 M	> 1.0 M
	EDTA	no	yes
	sarcosyl	no	> 0.5%
	DDT	yes	no

<sup>a</sup> Data from [50,57,64]. <sup>b</sup> Enhancement with respect to the SsOGT (in bold).

Concerning the interactions with the DNA, SsOGT binds methylated oligonucleotides. However, the repair activity depends on the position of the alkyl-group [39]. To efficiently repair the alkylated base on DNA double helix, the protein requires at least three bases from either the 5'

or the 3' end. This is due to the necessary interactions formed with the double helix. Structural analysis confirmed these data [39].

To overcome the serious limitation to obtain structural data from mesophilic AGTs after reaction, studies have moved to thermostable homologues, based also on the knowledge that all AGTs share a common CTD domain structure (Figure 2b). In contrast to the human counterpart, alkylated SsOGT was soluble and relatively stable, thus allowing *in-deep* analysis of the protein in its post-reaction form [39]. Structural and biochemical analysis of the archaeal OGT, as well as after the reaction with a bulkier adduct in the active site (benzyl-fluorescein; [66]), suggested a possible mechanism of alkylation-induced SsOGT unfolding and degradation (Figure 4).



**Figure 4.** Conformational changes of the SsOGT *active site loop* after reaction with an  $O^6$ -MG dsDNA oligonucleotide (in magenta; [39]), or with SNAP Vista Green substrate (in green; [66]).

On the basis of their data, Perugino and co-workers suggested a general model for the mechanism of post-reaction AGT destabilization—the so called *active-site loop* moves towards the bulk solvent as a result of the covalent binding of alkyl adduct on the catalytic cysteine and the extent of the loop movement and dynamic correlates with the steric hindrance of the adduct [39,66] (Figure 4). The destabilization of this protein region triggers then the recognition of the alkylated protein by degradation pathway.

### 3.4. Biotechnological Applications of an Engineered SsOGT—the $H^5$ Mutant

As described in Section 1.2, the introduction of the SNAP-tag technology enabled a wide in vivo and in vitro labelling variety for biological studies by fusing any protein of interest (POI) to this *protein tag* [67]. However, being originated from hMGMT, the extension to extremophilic organisms and/or harsh reaction conditions is seriously limited.

By following the same approach used for the hMGMT as Kai Johnsson [26–30], an engineered version of SsOGT was produced [52,59]. This protein, called SsOGT- $H^5$ , contains five mutations in the helix-turn-helix domain, abolishing any DNA-binding activity [52]. In addition, a sixth mutation was made—in the *active site loop*, where serine residue was replaced by a glutamic acid at position 132 (S132E). This modification increased the catalytic activity of SsOGT [52,59], as it was observed in the engineered version of the hMGMT during the SNAP-tag development [26]. SsOGT- $H^5$  shows slightly lower heat stability in respect to the wild-type protein (Table 3), whereas the resistance to other denaturing agents is maintained. Moreover, SsOGT- $H^5$  is characterised by a surprisingly high catalytic activity at lower temperatures, keeping the rate of reaction to the physiological ones (Table 3) [52,59]. These characteristics make this mutant a potential alternative to SNAP-tag for in vivo and in vitro biotechnological applications. The stability against thermal denaturation allowed Miggiano and co-workers to obtain the structure of the protein after the reaction with the fluorescent substrate SNAP-Vista Green, revealing the peculiar destabilization of the *active site loop* after the alkylation of the active cysteine [66].

### 3.4.1 In vitro Thermostable H<sup>5</sup>-Based Chimeras

The *Saccharolobus* OGT mutant has been firstly tested as *protein tag* fused to two thermostable *S. solfataricus* proteins heterologously expressed in *E. coli*. The chimeric proteins were correctly folded, and the *tag* did not interfere with the enzymatic activity of the tetrameric *S. solfataricus*  $\beta$ -glycosidase (*Ss* $\beta$ gly) [59], nor with the hyperthermophile-specific DNA topoisomerase reverse gyrase [68–72]. Furthermore, the stability of H<sup>5</sup> made possible a heat treatment of the cell-free extract to remove most of the *E. coli* proteins and performing the  $\beta$ -glycosidase assay at high temperatures without the need of removing the *tag* [60].

### 3.4.2 Expression in Thermophilic Organisms Models

As the applicability of the thermostable *tag* under in vivo conditions is very important, the *Ss*OGT-H<sup>5</sup> was also expressed in thermophilic organisms. The fluorescent AGT assay allows for the detection of the presence of *Ss*OGT-H<sup>5</sup> both in living cells as well as in vitro in cell-free extracts [59,72]. To assay the activity to *Ss*OGT-H<sup>5</sup>, it was necessary to choose models in which the endogenous AGT activity is suppressed. *Thermus thermophilus* is an *ogt*-species, showing only one *ogt* homologue (TTHA1564), whose annotation corresponds to an alkyltransferase-like protein (ATL) [73]. ATLs are a class of proteins present in prokaryotes and lower eukaryotes [74], presenting aminoacidic motifs similar to those of AGTs' CTD, in which a tryptophan residue replaces the cysteine in the active site [75]. Like AGTs, ATLs use a helix-turn-helix motif to bind the minor groove of the DNA, but they do not repair it as they only recruit and interact with proteins involved in the nucleotide excision repair system [76,77].

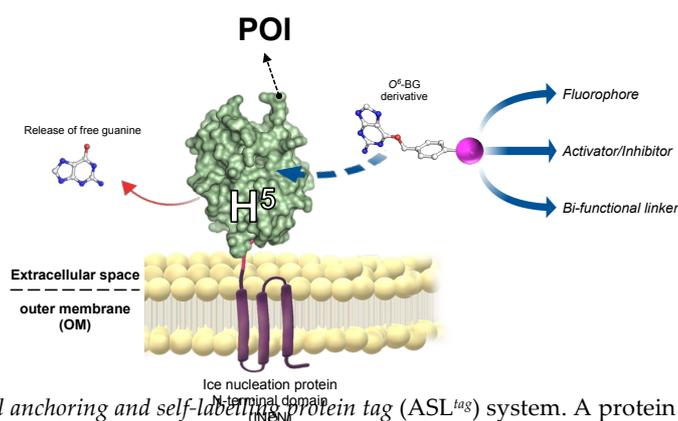
Although *T. thermophilus* is a natural *ogt* knockout organism, *Sulfolobus islandicus* possesses an *ogt* gene very similar to that of *S. solfataricus*, which was silenced by a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-based technique and then used as a host organism [72].

The fluorescent signal obtained by SDS-PAGE *gel imaging* revealed that *Ss*OGT-H<sup>5</sup> not only is efficiently expressed in these thermophilic microorganisms, but it also showed that this *tag* was correctly folded and active, demonstrating the fact that *Ss*OGT-H<sup>5</sup> might be used as an in vivo *protein tag* at high temperatures [59,72]. As is the case with SNAP-*tag* in human cells, the utilization of *Ss*OGT-H<sup>5</sup> with different fluorescent substrates gives the opportunity to perform a multi-colour fluorescence study (see Table 1), by following a POI inside living “*thermo* cells” at different stages and localization.

### 3.4.3 The ASL<sup>tag</sup> System

As most biotechnological processes require harsh operational conditions, the immobilization of very robust enzymes on solid supports is often essential [78]. By definition, an immobilized enzyme is a “physically confined biocatalyst, which retains its catalytic activity and can be used repeatedly” [79]. Protein immobilisation offers several advantages, such as the catalysts' recovery and reuse, as well as the physical separation of the enzymes from the reaction mixture. Currently, different immobilisation strategies are available, from physical adsorption to covalent coupling [80–83]. However, all these procedures require purified biocatalysts and suffer from problems related to steric hindrance between the catalyst, the substrate, and the solid support, with increasing of costs and time for the production processes.

The introduction of “cell-based” immobilisation systems resulted in a significant improvement and reduces both time and costs of the process. One of the most widely used display strategies is the simultaneous heterologous expression of enzymes and their in vivo immobilisation on the external surface of Gram-negative bacteria cells, by the utilisation of the ice nucleation protein (INP) from *Pseudomonas syringae* [84,85]. Most recently, the N-terminal domain of INP (INPN) was used to produce a *novel anchoring and self-labelling protein tag* (hereinafter ASL<sup>tag</sup>). The ASL<sup>tag</sup> consists of two moieties, the INPN and the engineered and *Ss*OGT-H<sup>5</sup> mutant (Figure 5) [86].



**Figure 5.** The novel anchoring and self-labeling protein tag (ASL<sup>tag</sup>) system. A protein of interest (POI) is genetically encoded with the tag, which in turn makes it anchored in the outer membrane and accessible for the covalent linkage to a desired chemical group (magenta sphere) by the activity of SsOGT-H<sup>5</sup> (adapted from [87]).

The INPN allows an *in vivo* immobilisation on *E. coli* outer membrane of enzymes of interest and their exposition to the solvent. The significant reduction of the costs related to the purification and immobilization is added to the overcoming of problems related to the recovery of enzymes by simple filtration or centrifugation methods [88]. SsOGT-H<sup>5</sup>, in turn, gives the unique opportunity to label immobilized enzymes with any desired chemical groups (opportunistically conjugated to the benzyl-guanine; in magenta in Figure 5) [27,59], dramatically expanding biotechnological applications of this new tool. Depending of the chemical group of choice, modulating the activity of enzymes fused with the ASL<sup>tag</sup> can be possible by introducing activator or inhibitor molecules (Figure 5). The ASL<sup>tag</sup> system was successfully employed for the expression and immobilization of monomeric biocatalysts, such as the thermostable carbonic anhydrase from *Saccharolobus solfataricus* (SspCA), as well as the tetrameric Ssβgly, without affecting their folding and catalytic activity [86]. Moreover, SspCA fused to the ASL<sup>tag</sup> showed an increase in residual activity of up to 30 % for a period of 10 days at 70 °C [87], representing a huge advantage in pushing beyond reactions in bioreactors and in the reutilization of biocatalysts.

#### 4. *Pyrococcus furiosus* and *Thermotoga neapolitana* OGT

To extend the SNAP-tag technology to hyperthermophilic microorganisms for *in vivo* studies, an O<sup>6</sup>-alkylguanine-DNA alkyltransferase has been recently characterized from the archaeon *Pyrococcus furiosus* [89]. This extremophilic microorganism was originally isolated from hot marine sediments in Vulcano Island (Italy) [90], with an optimum growth temperature around 100 °C, thus thriving under extremely harsh conditions. Like those of other thermophilic Archaea, its enzymes are extremely thermostable and can be used in various biotechnological applications. For example, DNA polymerase I, also known as PfuDNA polymerase, is one of the most famous and frequently used enzymes from *P. furiosus* because of its high activity, thermostability, and strong 3'-5' proof-reading activity [91]. The first demonstration of an OGT activity in *P. furiosus* was in 1998, when Margison and co-workers identified a protein of 22 kDa, whose catalytic activity was abolished by the O<sup>6</sup>-BG pseudo-substrate. The PF1878 ORF is relative to a protein of 20.1 kDa. From its primary structure, the relative polypeptide seems to be closely related to the MGMT from *Pyrococcus kodakarensis* KOD1 (*Pk*-MGMT) [48,89,92]. The extreme thermostability was confirmed by *in vitro* biochemical studies on the heterologous expressed and characterized OGT protein from *P. furiosus* PfuOGT. This enzyme was active on BG-fluorescent substrates, thus allowing the competitive assay with methylated dsDNA. However, the experiments were performed at 65 °C instead of the standard procedure at 50 °C, as described for SsOGT [39,59,60], due to the strong thermophilicity of this enzyme. This behaviour was effectively confirmed by differential scan

fluorimetry analysis where the temperature melting ( $T_m$ ) of *Pfu*OGT was found to be 80 °C, much higher than that of *Ss*OGT (68 °C) [89]. It is worth noting that, in order to obtain a sigmoidal melting curve for *Pfu*OGT, a slower heating rate (10 min/°C × cycle) was set up, whereas the  $T_m$  value measurement is usually performed at 1 min/°C × cycle [93].

*Thermotoga neapolitana* is a hyperthermophilic Gram-negative bacterium of the order of Thermotogales [94–96], which are excellent models for genetic engineering and biotechnological applications [97–100]. The CTN1690 ORF shows a clear homology of the  $O^6$ -alkylguanine-DNA-alkyl-transferase. SDS-PAGE *gel imaging* analysis on lyophilized *T. neapolitana* cells incubated with the AGT fluorescent substrate showed a strong fluorescent signal with a molecular weight close to that of *Ss*OGT. The observed molecular weight and, above all, the sensitivity to the  $O^6$ -BG derivative, led to the cloning and heterologous expression of the *Thermotoga neapolitana* OGT protein (*Tn*OGT) in *E. coli* [89]. This protein, like most AGTs, has a role in DNA repair, as confirmed by competitive fluorescent assay in the presence of methylated dsDNA. As shown in Figure 3, the  $IC_{50}$  value was similar to that obtained for *Ss*OGT. Surprisingly, the enzyme from *T. neapolitana* exhibited a very high activity at low temperatures [89], similar to that possessed by the mutant *Ss*OGT-H<sup>5</sup> (Table 3) [52,59]. Superimposition analysis between a *Tn*OGT 3D model and the free form of *Ss*OGT (ID PDB: 4ZYE) revealed in both structures the presence of a serine residue in the *active site loop* (S132 in *Ss*OGT, see Figure 2a), which was replaced in *Ss*OGT-H<sup>5</sup> by a glutamic acid to improve its activity at lower temperatures. Interestingly, some residues are missing in *Tn*OGT that play an important role in stabilizing *Ss*OGT. In particular, the ionic interactions that play a crucial role in the stability of the *Saccharolobus* enzyme at high temperatures, such as the pair R133-D27 [39] and the *K-48 network* [60], are largely replaced by hydrophobic residues in the *Thermotoga* homolog. Evidently, different residues and mechanisms of stabilization may contribute to its exceptional catalytic activity at moderate temperatures and the high thermal stability.

## 5. Future Perspectives

The interest shown from the important insights of this class of small proteins led to novel biotechnological applications [101]. Studies on thermophilic AGTs represent a unique opportunity for structural analysis and, in the case of the *S. solfataricus* protein, for the identification of conformational changes after the trans-alkylation reaction, which are detectable with mesophilic AGTs, as the alkylated form are rapidly destabilized [6]. These results could have a wide impact, especially in medical fields for the design of novel hMGMT inhibitors to be used in cancer therapy [102]. Furthermore, given their small size, thermophilic enzymes are very useful for studying general stabilization mechanisms at high temperatures (as for *Pk*-MGMT and *Ss*OGT), which can then be applied to mesophilic enzymes. Searching for alternative *Ss*OGT homologues was clearly useful, leading to the identification of AGTs that are more resistant to thermal denaturation (*Pfu*OGT) or to enzymes with a higher reaction rate at all tested temperatures (*Tn*OGT).

Concerning biotechnology, the use of a modified hMGMT as *protein tag* opened the possibility to generalise this method—a targeted mutagenesis on a thermostable OGT by following a *rational approach* led to the characterization of *Ss*OGT-H<sup>5</sup>, applicable to in vitro harsh reaction conditions and to in vivo (hyper)thermophilic model organisms. On the other hand, by an *irrational approach* (random mutagenesis) it is also possible to enhance their catalytic activity [103], or modify the substrate specificity of these enzymes, making them active on benzyl-cytosine ( $O^2$ -BC) derivatives, such as that which happened for the production of the CLIP-tag [104].

This knowledge could be the starting point of developing a new engineered *thermo*-SNAP-tag to be employed in particular biotechnological fields, from in vivo studies in (hyper)thermophilic microorganisms (such as the in vivo CRISPR-Cas immune system in *P. furiosus* [105,106]) to industrial processes that require high temperatures or, in general, harsh reaction conditions.

**Author Contributions:** R.Ma. and R.Me. equally contributed to the present review article by writing the manuscript and joined as First Author. R.Mi. analysed the 3D structures, prepared Table 2, and wrote the relative text. A.V. contributed to the editing of the manuscript and to the collecting part of the references. G.P.

contributed to the editing of the texts and the drawing of figures. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Ministero dell'Istruzione Università e Ricerca (MIUR) National Operational Program (PON) Research and Innovation 2014-2020 (CCI 2014IT16M2OP005), European Social Fund, Action I.1 "Innovative Doctorates with Industrial characterization".

**Acknowledgments:** G.P. would like to thank all the authors, Miss Elena and Elisa Perugino for their efforts in writing this work and in their technical assistance, but mainly for their human support during the difficult and delicate period of staying at home following the COVID-19 outbreak. All authors strongly thank the reviewers for their important corrections and changes, increasing the scientific level of this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

AGT	<i>O</i> <sup>6</sup> -alkyl-guanine-DNA-alkyl-transferase
CLIP-tag	engineered version of SNAP-tag active on <i>O</i> <sup>2</sup> -BC
FP	auto-fluorescent protein
hMGMT	human <i>O</i> <sup>6</sup> -methyl-guanine-DNA-alkyl-transferase
MGMT	<i>O</i> <sup>6</sup> -methyl-guanine-DNA-alkyl-transferase
<i>O</i> <sup>2</sup> -BC	<i>O</i> <sup>2</sup> -benzyl-cytosine
<i>O</i> <sup>6</sup> -AG	<i>O</i> <sup>6</sup> -alkyl-guanine
<i>O</i> <sup>6</sup> -BG	<i>O</i> <sup>6</sup> -benzyl-guanine
<i>O</i> <sup>6</sup> -MG	<i>O</i> <sup>6</sup> -methyl-guanine
OGT	<i>O</i> <sup>6</sup> -alkyl-guanine-DNA-alkyl-transferase
SNAP-tag	engineered version of hMGMT for biotech purposes

## References

- Liu, L.; Gerson, S.L. Targeted modulation of MGMT: Clinical implications. *Clin. Cancer Res.* **2006**, *12*, 328–331.
- Leonard, G.A.; Thomson, J.; Watson, W.P.; Brown, T. High-resolution structure of a mutagenic lesion in DNA. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9573–9576.
- Drabløs, F.; Feyzi, E.; Aas, P.A.; Vaagbø, C.B.; Kavli, B.; Bratlie, M.S.; Peña-Díaz, J.; Otterlei, M.; Slupphaug, G.; Krokan, H.E. Alkylation damage in DNA and RNA-repair mechanisms and medical significance. *DNA Repair* **2004**, *11*, 1389–1407.
- Pegg, A.E. Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA damage, resistance to chemotherapy, and research tools. *Chem. Res. Toxicol.* **2011**, *24*, 618–639.
- Duguid, E.M.; Rice, P.A.; He, C. The structure of the human AGT protein bound to DNA and its implications for damage detection. *J. Mol. Biol.* **2005**, *350*, 657–666.
- Daniels, D.S.; Mol, C.D.; Arvai, A.S.; Kanugula, S.; Pegg, A.E.; Tainer, J.A. Active and alkylated human AGT structures: A novel zinc site, inhibitor and extrahelical base binding. *EMBO J.* **2000**, *19*, 1719–1730.
- Daniels, D.S.; Woo, T.T.; Luu, K.X.; Noll, D.M.; Clarke, N.D.; Pegg, A.E.; Tainer, J.A. DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat. Struct. Mol. Biol.* **2004**, *11*, 714–720.
- Tubbs, J.L.; Pegg, A.E.; Tainer, J.A. DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by *O*<sup>6</sup>-alkylguanine-DNA-alkyltransferase and its implications for cancer chemotherapy. *DNA Repair* **2007**, *6*, 1100–1115.
- Xu-Welliver, M.; Pegg, A.E. Degradation of the alkylated form of the DNA repair protein, *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase. *Carcinogenesis* **2002**, *23*, 823–830.
- Gerson, S.L. MGMT: Its role in cancer aetiology and cancer therapeutics. *Nat. Rev. Cancer.* **2004**, *4*, 296–307.
- Sabharwal, A.; Middleton, M.R. Exploiting the role of *O*<sup>6</sup>-methylguanine-DNA-methyltransferase (MGMT) in cancer therapy. *Curr. Opin. Pharmacol.* **2006**, *6*, 355–363.
- Zhong, Y.; Huang, Y.; Huang, Y.; Zhang, T.; Ma, C.; Zhang, S.; Fan, W.; Chen, H.; Qian, J.; Lu, D. Effects of *O*<sup>6</sup>-methylguanine-DNA methyltransferase (MGMT) polymorphisms on cancer: A meta-analysis. *Mutagenesis* **2010**, *25*, 83–95.

13. Kaina, B.; Christmann, M. DNA repair in personalized brain cancer therapy with temozolomide and nitrosoureas. *DNA Repair* **2019**, *78*, 128–141.
14. Khan, O.; Middleton, M.R. The therapeutic potential of O<sup>6</sup>-alkylguanine DNA alkyltransferase inhibitors. *Expert Opin. Investig. Drugs* **2007**, *10*, 1573–1584.
15. Paranjpe, A.; Zhang, R.; Ali-Osman, F.; Bobustuc, G.C.; Srivenugopal, K.S. Disulfiram is a direct and potent inhibitor of human O<sup>6</sup>-methylguanine-DNA methyltransferase (MGMT) in brain tumor cells and mouse brain and markedly increases the alkylating DNA damage. *Carcinogenesis* **2014**, *35*, 692–702.
16. Rabik, C.A.; Dolan, M.E. Molecular mechanisms of resistance and toxicity associated with platinating agents. *Cancer Treat. Rev.* **2007**, *33*, 9–23.
17. Kaina, B.; Margison, G.P.; Christmann, M. Targeting O<sup>6</sup>-methylguanine-DNA methyltransferase with specific inhibitors as a strategy in cancer therapy. *Cell Mol. Life Sci.* **2010**, *67*, 3663–3681.
18. Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W.W.; Prasher, D.C. Green fluorescent protein as a marker for gene expression. *Science* **1994**, *263*, 802–805.
19. Tsien, R.Y. The green fluorescent protein. *Annu. Rev. Biochem.* **1998**, *67*, 509–554.
20. di Guan, C.; Li, P.; Riggs, P.D.; Inouye, H. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* **1988**, *67*, 21–30.
21. Schmidt, T.G.M.; Koepke, J.; Frank, R.; Skerra, A. Molecular Interaction Between the Strep-tag Affinity Peptide and its Cognate Target, Streptavidin. *J. Mol. Biol.* **1996**, *255*, 753–766.
22. Ren, L.; Chang, E.; Makky, K.; Haas, A.L.; Kaboord, B.; Walid Qoronfleh, M. Glutathione S-transferase pull-down assays using dehydrated immobilized glutathione resin. *Anal. Biochem.* **2003**, *322*, 164–169.
23. Donini, S.; Ferraris, D.M.; Miggiano, R.; Massarotti, A.; Rizzi, M. Structural investigations on orotate phosphoribosyltransferase from *Mycobacterium tuberculosis*, a key enzyme of the *de novo* pyrimidine biosynthesis. *Sci. Rep.* **2017**, *7*, 1180.
24. Donini, S.; Garavaglia, S.; Ferraris, D.M.; Miggiano, R.; Mori, S.; Shibayama, K.; Rizzi, M. Biochemical and structural investigations on phosphoribosylpyrophosphate synthetase from *Mycobacterium smegmatis*. *PLoS ONE* **2017**, *12*, e0175815.
25. LaVallie, E.R.; DiBlasio, E.A.; Kovacic, S.; Grant, K.L.; Schendel, P.F.; McCoy, J.M. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Bio Technol.* **1993**, *11*, 187–193.
26. Juillerat, A.; Gronemeyer, T.; Keppler, A.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K. Directed evolution of O<sup>6</sup>-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules *in vivo*. *Chem. Biol.* **2003**, *10*, 313–317.
27. Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. A general method for the covalent labeling of fusion proteins with small molecules *in vivo*. *Nat. Biotechnol.* **2003**, *21*, 86–89.
28. Kindermann, M.; George, N.; Johnsson, N.; Johnsson, K. Covalent and selective immobilization of fusion proteins. *J. Am. Chem. Soc.* **2003**, *125*, 7810–7811.
29. Gronemeyer, T.; Chidley, C.; Juillerat, A.; Heinis, C.; Johnsson, K. Directed evolution of O<sup>6</sup>-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Eng. Des. Sel.* **2006**, *19*, 309–316.
30. Hinner, M.J.; Johnsson, K. How to obtain labeled proteins and what to do with them. *Curr. Opin. Biotechnol.* **2010**, *21*, 766–776.
31. Huber, W.; Perspicace, S.; Kohler, J.; Müller, F.; Schlatter, D. SPR-based interaction studies with small molecular weight ligands using hAGT fusion proteins. *Anal. Biochem.* **2004**, *333*, 280–288.
32. Niesen, J.; Sack, M.; Seidel, M.; Fendel, R.; Barth, S.; Fischer, R.; Stein, C. SNAP-tag technology: A useful tool to determine affinity constants and other functional parameters of novel anti-body fragments. *Bioconjug. Chem.* **2016**, *27*, 1931–1941.
33. Valenti, A.; Napoli, A.; Ferrara, M.C.; Nadal, M.; Rossi, M.; Ciaramella, M. Selective degradation of reverse gyrase and DNA fragmentation induced by alkylating agent in the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res.* **2006**, *34*, 2098–2108.
34. Kanugula, S.; Pegg, A.E. Alkylation damage repair protein O<sup>6</sup>-alkyl-guanine-DNA alkyltransferase from the hyperthermophiles *Aquifex aeolicus* and *Archaeoglobus fulgidus*. *Biochem. J.* **2003**, *375*, 449–455.
35. Leclere, M.M.; Nishioka, M.; Yuasa, T.; Fujiwara, S.; Takagi, M.; Imanaka, T. The O<sup>6</sup>-methylguanine-DNA methyltransferase from the hyperthermophilic archaeon *Pyrococcus* sp. KOD1: A thermostable repair enzyme. *Mol Gen Genet*, **1998**, *258*, 69–77.

36. Skorvaga, M.; Raven, N.D.; Margison, G.P. Thermostable archaeal O<sup>6</sup>-alkylguanine-DNA alkyltransferases. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6711–6715.
37. Fang, Q.; Kanugula, S.; Pegg, A.E. Function of domains of human O<sup>6</sup>-alkyl-guanine-DNA alkyltransferase. *Biochemistry* **2005**, *44*, 15396–15405.
38. Miggiano, R.; Casazza, V.; Garavaglia, S.; Ciaramella, M.; Perugino, G.; Rizzi, M.; Rossi, F. Biochemical and structural studies of the *Mycobacterium tuberculosis* O<sup>6</sup>-methylguanine methyltransferase and mutated variants. *J. Bacteriol.* **2013**, *195*, 2728–2736.
39. Perugino, G.; Miggiano, R.; Serpe, M.; Vettone, A.; Valenti, A.; Lahiri, S.; Rossi, F.; Rossi, M.; Rizzi, M.; Ciaramella, M. Structure-function relationships governing activity and stability of a DNA alkylation damage repair thermostable protein. *Nucleic Acids Res.* **2015**, *43*, 8801–8816.
40. Kanugula, S.; Pegg, A.E. Novel DNA Repair Alkyltransferase from *Caenorhabditis elegans*. *Env. Mol. Mut.* **2001**, *38*, 235–243.
41. Serpe, M.; Forenza, C.; Adamo, A.; Russo, N.; Perugino, G.; Ciaramella, M.; Valenti, A. The DNA alkylguanine DNA alkyltransferase-2 (AGT-2) of *Caenorhabditis elegans* is involved in meiosis and early development under physiological conditions. *Sci. Rep.* **2019**, *9*, 6889.
42. Presta, L.G.; Rose, G.D. Helix signals in proteins. *Science* **1988**, *240*, 1632–1641.
43. Richardson, J.S.; Richardson, D.C. Amino acid preferences for specific locations at the ends of  $\alpha$ -helices. *Science* **1988**, *240*, 1648–1652.
44. Koscielska-Kasprzak, K.; Cierpicki, T.; Otlewski, J. Importance of alpha-helix N-capping motif in stabilization of betabetaalpha fold. *Protein Sci.* **2003**, *12*, 1283–1289.
45. Chan, M.K.; Mukund, S.; Kletzin, A.; Adams, M.W.W.; Rees, D.C. Structure of a hyperthermo-philic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science* **1995**, *267*, 1463–1469.
46. Miggiano, R.; Perugino, G.; Ciaramella, M.; Serpe, M.; Rejman, D.; Páv, O.; Pohl, R.; Garavaglia, S.; Lahiri, S.; Rizzi, M. Crystal structure of *Mycobacterium tuberculosis* O<sup>6</sup>-methylguanine-DNA methyltransferase protein clusters assembled on to damaged DNA. *Biochem. J.* **2016**, *473*, 123–133.
47. Rice, D.W.; Yip, K.S.; Stillman, T.J.; Britton, K.L.; Fuentes, A.; Connerton, J.; Pasquo, A.; Scandura, R.; Engel, P.C. Insight into the molecular basis of thermal stability from the structure determination of *Pyrococcus furiosus* glutamate dehydrogenase. *FEMS Microbiol. Rev.* **1996**, *18*, 105–117.
48. Hashimoto, H.; Inoue, T.; Nishioka, M.; Fujiwara, S.; Takagi, M.; Imanaka, T.; Kai, Y. Hyperthermostable protein structure maintained by intra and inter-helix ion-pairs in archaeal O<sup>6</sup>-methylguanine-DNA methyltransferase. *J. Mol. Biol.* **1999**, *292*, 707–716.
49. Moore, M.H.; Gulbis, J.M.; Dodson, E.J.; Demple, B.; Moody, P.C. Crystal structure of a suicidal DNA repair protein: The Ada O<sup>6</sup>-methylguanineDNA methyltransferase from *E. coli*. *EMBO J.* **1994**, *13*, 1495–1501.
50. Roberts, A.; Pelton, J.G.; Wemmer, D.E. Structural studies of MJ1529, an O<sup>6</sup>-methylguanine-DNA methyltransferase. *Magn. Reson. Chem.* **2006**, *44*, S71–S82.
51. Zillig, W.; Stetter, K.O.; Wunderl, S.; Schulz, W.; Priess, H.; Scholz, I. The Sulfolobus-“Caldariella” group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases. *Arch. Microbiol.* **1980**, *125*, 259–269.
52. Perugino, G.; Vettone, A.; Illiano, G.; Valenti, A.; Ferrara, M.C.; Rossi, M.; Ciaramella, M. Activity and regulation of archaeal DNA alkyltransferase: Conserved protein involved in repair of DNA alkylation damage. *J. Biol. Chem.* **2012**, *287*, 4222–4231.
53. Olsson, M.; Lindahl, T. Repair of alkylated DNA in *Escherichia coli*. Methyl group transfer from O<sup>6</sup>-methylguanine to a protein cysteine residue. *J. Biol. Chem.* **1980**, *255*, 10569–10571.
54. Wu, R.S.; Hurst-Calderone, S.; Kohn, K.W. Measurement of O<sup>6</sup>-alkylguanine-DNA-alkyltransferase activity in human cells and tumor tissues by restriction endonuclease inhibition. *Cancer Res.* **1987**, *47*, 6229–6235.
55. Klein, S.; Oesch, F. Assay for O<sup>6</sup>-alkylguanine-DNA-alkyltransferase using oligonucleotides containing O<sup>6</sup>-methylguanine in a BamHI recognition site as substrate. *Anal. Biochem.* **1992**, *205*, 294–299.
56. Luu, K.X.; Kanugula, S.; Pegg, A.E.; Pauly, G.T.; Moschel, R.C. Repair of oligodeoxyribonucleotides by O(6)-alkylguanine-DNA alkyltransferase. *Biochemistry* **2002**, *41*, 8689–8697.
57. Bock, L.C.; Griffin, L.C.; Latham, J.A.; Vermaas, E.H.; Toole, J.J. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* **1992**, *355*, 564–566.

58. Tintoré, M.; Aviñó, A.; Ruiz, F.M.; Eritja, R.; Fábrega, C. Development of a novel fluorescence assay based on the use of the thrombin binding aptamer for the detection of O<sup>6</sup>-alkyl-guanine-DNA alkyltransferase activity. *J. Nuc. Ac.* **2010**, *2010*, 632041.
59. Vettone, A.; Serpe, M.; Hidalgo, A.; Berenguer, J.; del Monaco, G.; Valenti, A.; Ciaramella, M.; Perugino, G. A novel thermostable protein-tag: Optimization of the *Sulfolobus solfataricus* DNA-alkyl-transferase by protein engineering. *Extremophiles* **2016**, *20*, 1–13.
60. Morrone, C.; Miggiano, R.; Serpe, M.; Massarotti, A.; Valenti, A.; Del Monaco, G.; Rossi, M.; Rossi, F.; Rizzi, M.; Perugino, G. Interdomain interactions rearrangements control the reaction steps of a thermostable DNA alkyltransferase. *Biochim. Biophys. Acta* **2017**, *1861*, 86–96.
61. Elder, R.H.; Margison, G.P.; Rafferty, J.A. Differential inactivation of mammalian and *Escherichia coli* O<sup>6</sup>-alkylguanine-DNA alkyltransferases by O<sup>6</sup>-benzylguanine. *Biochem. J.* **1994**, *298*, 231–235.
62. Goodtzova, K.; Kanugula, S.; Edara, S.; Pauly, G.T.; Moschel, R.C.; Pegg, A.E. Repair of O<sup>6</sup>-benzylguanine by the *Escherichia coli* Ada and Ogt and the human O<sup>6</sup>-alkylguanine-DNA alkyltransferase. *J. Biol. Chem.* **1997**, *272*, 8332–8339.
63. Miggiano, R.; Valenti, A.; Rossi, F.; Rizzi, M.; Perugino, G.; Ciaramella, M. Every OGT Is Illuminated ... by Fluorescent and Synchrotron Lights. *Int. J. Mol. Sci.* **2017**, *18*, 2613–2630.
64. Melikishvili, M.; Rasimas, J.J.; Pegg, A.E.; Fried, M.G. Interactions of human O(6)-alkylguanine-DNA alkyltransferase (AGT) with short double-stranded DNAs. *Biochemistry* **2008**, *47*, 13754–13763.
65. Brunk, E.; Mollwitz, B.; Rothlisberger, U. Mechanism to Trigger Unfolding in O<sup>6</sup>-Alkylguanine-DNA Alkyltransferase. *ChemBioChem* **2013**, *14*, 703–710.
66. Rossi, F.; Morrone, C.; Massarotti, A.; Ferraris, D.M.; Valenti, A.; Perugino, G.; Miggiano, R. Crystal structure of a thermophilic O<sup>6</sup>-alkylguanine-DNA alkyltransferase-derived self-labeling protein-tag in covalent complex with a fluorescent probe. *Biochem. Biophys. Res. Comm.* **2018**, *500*, 698–703.
67. Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K. Labeling of fusion proteins with synthetic fluorophores in live cells. *Proc. Natl. Acad. Sci. USA* **2004**, *10*, 9955–9959.
68. Valenti, A.; Perugino, G.; D’Amaro, A.; Cacace, A.; Napoli, A.; Rossi, M.; Ciaramella, M. Dissection of reverse gyrase activities: Insight into the evolution of a thermostable molecular machine. *Nucleic Acids Res.* **2008**, *36*, 4587–4597.
69. Valenti, A.; Perugino, G.; Nohmi, T.; Rossi, M.; Ciaramella, M. Inhibition of translesion DNA polymerase by archaeal reverse gyrase. *Nucleic Acids Res.* **2009**, *37*, 4287–4295.
70. Perugino, G.; Valenti, A.; D’Amaro, A.; Rossi, M.; Ciaramella, M. Reverse gyrase and genome stability in hyperthermophilic organisms. *Biochem. Soc. Trans.* **2009**, *37*, 69–73.
71. Valenti, A.; Perugino, G.; Rossi, M.; Ciaramella, M. Positive supercoiling in thermophiles and mesophiles: Of the good and evil. *Biochem. Soc. Trans.* **2011**, *39*, 58–63.
72. Visone, V.; Han, W.; Perugino, G.; Del Monaco, G.; She, Q.; Rossi, M.; Valenti, A.; Ciaramella, M. *In vivo* and *in vitro* protein imaging in thermophilic archaea by exploiting a novel protein tag. *PLoS ONE* **2017**, *10*, e0185791.
73. Morita, R.; Nakagawa, N.; Kuramitsu, S.; Masui, R. An O<sup>6</sup>-methylguanine-DNA methyltransferase-like protein from *Thermus thermophilus* interacts with a nucleotide excision repair protein. *J. Biochem.* **2008**, *144*, 267–277.
74. Schärer, O.D. Alkyltransferase-like Proteins: Brokers Dealing with Alkylated DNA Bases. *Mol. Cell* **2012**, *47*, 3–4.
75. Tubbs, J.L.; Latypov, V.; Kanugula, S.; Butt, A.; Melikishvili, M.; Kraehenbuehl, R.; Fleck, O.; Marriott, A.; Watson, A.J.; Verbeek, B.; et al. Flipping of alkylated DNA damage bridges base and nucleotide excision repair. *Nature* **2009**, *459*, 808–813.
76. Latypov, V.F.; Tubbs, J.L.; Watson, A.J.; Marriott, A.S.; McGown, G.; Thorncroft, M.; Wilkinson, O.J.; Senthong, P.; Butt, A.; Arvai, A.S.; et al. AtI1 regulates choice between global genome and transcription-coupled repair of O(6)-alkylguanines. *Mol Cell.* **2012**, *47*, 50–60.
77. Lahiri, S.; Rizzi, M.; Rossi, F.; Miggiano, R.; *Mycobacterium tuberculosis* UvrB forms dimers in solution and interacts with UvrA in the absence of ligands. *Proteins* **2018**, *86*, 98–109.
78. Zhou, Z.; Hartmann, M. Progress in enzyme immobilization in ordered mesoporous materials and related applications. *Chem Soc Rev* **2013**, *42*, 3894–3912.
79. Tosa, T.; Mori, T.; Fuse, N.; Chibata, I. Studies on continuous enzyme reactions. I. Screening of carriers for preparation of water-insoluble aminoacylase. *Enzymologia* **1966**, *31*, 214–224.

80. Mohamad, N.R.; Che Marzuki, N.H.; Buang, N.A.; Huyop, F.; Wahab, R.A. An overview of technologies for immobilization of enzymes and surface analysis techniques for immobilized enzymes. *Biotech. Biotech. Equip.* **2015**, *29*, 205–220.
81. Nguyen, H.H.; Kima, M. An Overview of Techniques in Enzyme Immobilization. *App. Sci. Conv. Tech.* **2017**, *26*, 157–163.
82. Jakub Zdarta, J.; Meyer, A.S.; Jesionowski, T.; Pinelo, M. A General Overview of Support Materials for Enzyme Immobilization: Characteristics, Properties, Practical Utility. *Catalysts* **2018**, *8*, 92.
83. Sirisha, V.L.; Jain, A.; Jain, A. Chapter Nine—Enzyme Immobilization: An Overview on Methods, Support Material, and Applications of Immobilized Enzymes. *Adv. Food Nut. Res.* **2016**, *79*, 179–211.
84. Gurian-Sherman, D.; Lindow, S.E. Bacterial ice nucleation: Significance and molecular basis. *FASEB J.* **1993**, *7*, 1338–1343.
85. Cochet, N.; Widehem, P. Ice crystallization by *Pseudomonas syringae*. *Appl. Microbiol. Biotechnol.* **2000**, *54*, 153–161.
86. Merlo, R.; Del Prete, S.; Valenti, A.; Mattosovich, R.; Carginale, V.; Supuran, C.T.; Capasso, C.; Perugino, P. An AGT-based protein-tag system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane. *J. Enz. Inhib. Med. Chem.* **2019**, *34*, 490–499.
87. Del Prete, S.; Merlo, R.; Valenti, A.; Mattosovich, R.; Rossi, M.; Carginale, M.; Supuran, C.T.; Perugino, G.; Capasso, C. Thermostability enhancement of the  $\alpha$ -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* by using the anchoring-and-selflabelling-protein-tag system (ASL<sup>tag</sup>). *J. Enz. Inhib. Med. Chem.* **2019**, *34*, 946–954.
88. Del Prete, S.; Perfetto, R.; Rossi, M.; Alasmay, F.A.S.; Osman, S.M.; AlOthman, Z.; Supuran, C.T.; Capasso, C. A one-step procedure for immobilising the thermostable carbonic anhydrase (SspCA) on the surface membrane of *Escherichia coli*. *J. Enz. Inhib. Med. Chem.* **2017**, *32*, 1120–1128.
89. Mattosovich, R.; Merlo, R.; Fontana, A.; d’Ippolito, G.; Terns, M.P.; Watts, E.A.; Valenti, A.; Perugino, P. A journey down to hell: New thermostable protein-tags for biotechnology at high temperatures. *Extremophiles* **2020**, *24*, 81–91.
90. Fiala, G.; Stetter, K.O. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaeobacteria growing optimally at 100 °C. *Arch. Microb.* **1986**, *145*, 56–61.
91. Lundberg, K.S.; Shoemaker, D.D.; Adams, M.W.W.; Short, J.M.; Sorge, J.A.; Mathur, E.J. High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* **1991**, *108*, 1–6.
92. Nishikori, S.; Shiraki, K.; Yokota, K.; Izumikawa, N.; Fujiwara, S.; Hashimoto, H.; Imanaka, T.; Takagi, M. Mutational effects on O<sup>6</sup>-methylguanine-DNA methyltransferase from hyperthermophile: Contribution of ion-pair network to protein thermostability. *J. Biochem.* **2004**, *135*, 525–532.
93. Niesen, F.H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc.* **2007**, *2*, 2212–2221.
94. Belkin, S.; Wirsén, C.O.; Jannasch, H.W. A new sulfur-reducing, extremely thermophilic eubacterium from a submarine thermal vent. *App. Environ. Microbiol.* **1986**, *51*, 1180–1185.
95. Jannasch, H.W.; Huber, R.; Belkin, S.; Stetter, K.O. *Thermotoga neapolitana* sp. nov. of the extremely thermophilic, eubacterial genus *Thermotoga*. *Arch. Microbiol.* **1988**, *150*, 103–104.
96. Conners, S.B.; Mongodin, E.F.; Johnson, M.R.; Montero, C.I.; Nelson, K.E.; Kelly, R.M. Microbial biochemistry, physiology, and biotechnology of hyperthermophilic *Thermotoga* species. *FEMS Microb. Rev.* **2006**, *30*, 872–905.
97. Zhang, J.; Shi, H.; Xu, L.; Zhu, X.; Li, X. Site-directed mutagenesis of a hyperthermophilic endoglucanase Cel12B from *Thermotoga maritima* based on rational design. *PLoS ONE* **2015**, doi:10.1371/journal.pone.0133824.
98. Fink, M.; Trunk, S.; Hall, M.; Schwab, H.; Steiner, K. Engineering of TM1459 from *Thermotoga maritima* for increased oxidative alkene cleavage activity. *Front. Microb.* **2016**, *7*, 1–9.
99. Donaldson, T.; Iozzino, L.; Deacon, L.J.; Billones, H.; Ausili, A.; D’Auria, S.; Dattelbaum, J.D. Engineering a switch-based biosensor for arginine using a *Thermotoga maritima* periplasmic binding protein. *An. Biochem.* **2017**, *525*, 60–66.
100. Han, D.; Xu, Z. Development of a pyrE-based selective system for *Thermotoga* sp. strain RQ7. *Extremophiles* **2017**, *21*, 297–306.
101. Pegg, A.E. Repair of O<sup>6</sup>-alkylguanine by alkyltransferases. *Mutat. Res.* **2000**, *462*, 83–100.

102. Margison, G.P.; Povey, A.C.; Kaina, B.; Santibanez Koref, M.F. Variability and regulation of *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase. *Carcinogenesis* **2003**, *24*, 625–635.
103. Mollwitz, B.; Brunk, E.; Schmitt, S.; Pojer, F.; Bannwarth, M.; Schiltz, M.; Rothlisberger, U.; Johnsson, K. Directed evolution of the suicide protein *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase for increased reactivity results in an alkylated protein with exceptional stability. *Biochemistry* **2012**, *51*, 986–994.
104. Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I.R., Jr.; Kindermann, M.; Beaufils, F.; Johnsson, K. An engineered protein-tag for multi-protein labeling in living cells. *Chem. Biol.* **2008**, *15*, 128–136.
105. Hale, C.R.; Zhao, P.; Olson, S.; Duff, M.O.; Graveley, B.R.; Wells, L.; Terns, R.M.; Terns, M.P. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **2009**, *139*, 945–956.
106. Terns, R.M.; Terns, M.P. The RNA- and DNA-targeting CRISPR-Cas immune systems of *Pyrococcus furiosus*. *Biochem. Soc. Trans.* **2013**, *41*, 1416–1421.



© 2020 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



## The SNAP-tag technology revised: an effective *chemo-enzymatic approach* by using a universal azide-based substrate

Rosa Merlo , Diego Caprioglio , Michele Cillo , Anna Valenti , Rosanna Mattosovich , Castrese Morrone , Alberto Massarotti , Franca Rossi , Riccardo Miggiano , Antonio Leonardi , Alberto Minassi & Giuseppe Perugino

To cite this article: Rosa Merlo , Diego Caprioglio , Michele Cillo , Anna Valenti , Rosanna Mattosovich , Castrese Morrone , Alberto Massarotti , Franca Rossi , Riccardo Miggiano , Antonio Leonardi , Alberto Minassi & Giuseppe Perugino (2021) The SNAP-tag technology revised: an effective *chemo-enzymatic approach* by using a universal azide-based substrate, Journal of Enzyme Inhibition and Medicinal Chemistry, 36:1, 85-97, DOI: [10.1080/14756366.2020.1841182](https://doi.org/10.1080/14756366.2020.1841182)

To link to this article: <https://doi.org/10.1080/14756366.2020.1841182>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 29 Oct 2020.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)

SHORT COMMUNICATION



## The SNAP-tag technology revised: an effective *chemo-enzymatic approach* by using a universal azide-based substrate

Rosa Merlo<sup>a\*</sup>, Diego Caprioglio<sup>b\*</sup>, Michele Cillo<sup>c\*</sup>, Anna Valenti<sup>a</sup>, Rosanna Mattosovich<sup>a</sup>, Castrese Morrone<sup>b</sup>, Alberto Massarotti<sup>b,d</sup>, Franca Rossi<sup>b</sup>, Riccardo Miggiano<sup>b,d</sup>, Antonio Leonardi<sup>c</sup>, Alberto Minassi<sup>b</sup> and Giuseppe Perugino<sup>a</sup>

<sup>a</sup>Institute of Biosciences and BioResources, National Research Council of Italy, Naples, Italy; <sup>b</sup>Department of Pharmaceutical Sciences, University of Piemonte Orientale, Novara, Italy; <sup>c</sup>Department of Molecular Medicine and Medical Biotechnology, University of Naples "Federico II", Naples, Italy; <sup>d</sup>IXTAL srl, Novara, Italy

### ABSTRACT

SNAP-tag<sup>®</sup> is a powerful technology for the labelling of protein/enzymes by using benzyl-guanine (BG) derivatives as substrates. Although commercially available or ad hoc produced, their synthesis and purification are necessary, increasing time and costs. To address this limitation, here we suggest a revision of this methodology, by performing a *chemo-enzymatic approach*, by using a BG-substrate containing an azide group appropriately distanced by a spacer from the benzyl ring. The SNAP-tag<sup>®</sup> and its relative thermostable version (*SsOGT-H<sup>5</sup>*) proved to be very active on this substrate. The stability of these tags upon enzymatic reaction makes possible the exposition to the solvent of the azide-moiety linked to the catalytic cysteine, compatible for the subsequent conjugation with DBCO-derivatives by azide-alkyne Huisgen cycloaddition. Our studies propose a strengthening and an improvement in terms of biotechnological applications for this self-labelling *protein-tag*.

### ARTICLE HISTORY

Received 16 September 2020  
Revised 5 October 2020  
Accepted 14 October 2020

### KEYWORDS

*Protein-tag*; protein labelling; enzymatic reaction; click chemistry; biotechnology

## 1. Introduction

The advent of the self-labelling *protein-tags* (SLPs) has led to a huge push in modern biotechnology, especially in the field of cell biology, where auto-fluorescent proteins (AFPs) for a long time dominated for their versatility in the localisation experiments of proteins, organelles, and membranes<sup>1</sup>. But the use of SLPs clearly goes beyond: they catalyse the covalent, highly specific and irreversible attachment of a part of their synthetic ligands upon reaction. This offers the opportunity to label them by conjugation to those ligands of an infinite number of chemical groups, such as fluorescent dyes, affinity molecules, or solid surfaces, expanding the application fields<sup>2</sup>. Among SLPs, of particular note are the *Halotag*<sup>®</sup>, the *SpyTag*<sup>3</sup> the SNAP- and the CLIP-tag<sup>®</sup>. The Promega *Halotag*<sup>®</sup> is a halo-alkane dehalogenase with a genetically modified active site, which reacts irreversibly with primary alkyl-halides<sup>4,5</sup>.

SNAP-tag<sup>®</sup> from New England Biolabs (NEB) is the engineered variant of the natural suicide human O<sup>6</sup>-methylguanine DNA-methyltransferase protein (hMGMT). Alkylated DNA-alkyltransferases (AGTs, MGMTs or OGTs, E.C. 2.1.1.63) are ubiquitous and conserved proteins involved in the repair of the DNA alkylation damage, in particular, they remove alkyl adducts at the level of O<sup>6</sup>-position on guanine base<sup>6,7</sup>. The peculiar single-step mechanism are called "*suicide enzymes*," in which the alkylated base is directly repaired by the irreversible transfer of the alkyl group

from the damaged guanine to the catalytic cysteine in the protein active site<sup>8</sup>. The protein is permanently inactivated upon the trans-alkylation reaction and susceptible to *in vivo* degradation via the proteasome.

In 2003, the group of Kai Johnsson developed a new strategy to exploit the hMGMT suicidal reaction in biotechnology, adopting a directed-evolution approach to engineer a variant to be used as an innovative *protein-tag*, that is, the SNAP-tag<sup>®</sup>. The rationale behind the SNAP-tag technology is the low substrate specificity of some AGT proteins, being able to efficiently recognise also the O<sup>6</sup>-benzyl-guanine (BG) nucleobase<sup>9</sup>. Likely, the reaction of these enzymes with BG-derivatives could happen: upon the irreversible transfer to the catalytic cysteine, they indeed demonstrated the specific labelling of the hMGMT with molecules, as fluorophores, previously conjugated to the 4-position of the BG benzyl ring. Because of the small dimension of this protein, it was mutagenized to abolish any DNA binding activity and utilised as *protein-tag* for the indirect labelling of proteins of interest genetically fused to it (Figure 1)<sup>9–13</sup>. Later, the same group further engineered the SNAP-tag<sup>®</sup> to obtain the CLIP-tag<sup>®</sup>, which specifically reacts with O<sup>2</sup>-benzyl-cytosine derivatives, expanding that technology for *in vivo* and *in vitro* multi-protein labelling<sup>14</sup>.

Apart from cell biology and fluorescence imaging, hundreds of papers are present in the literature showing many applications of SNAP-tag<sup>®</sup> in several fields, among which RNA-editing<sup>15</sup>, the

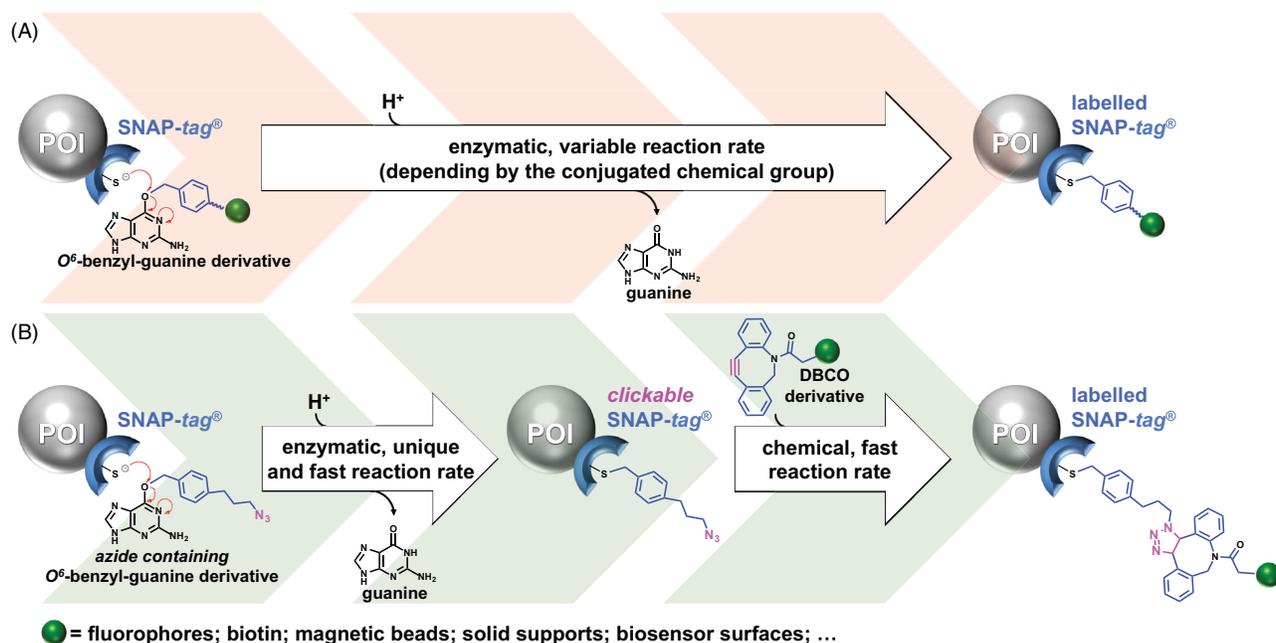
**CONTACT** Giuseppe Perugino ✉ [giuseppe.perugino@ibbr.cnr.it](mailto:giuseppe.perugino@ibbr.cnr.it) 📧 Institute of Biosciences and BioResources, National Research Council of Italy, Via Pietro Castellino 111, Naples 80131, Italy; Alberto Minassi ✉ [alberto.minassi@uniupo.it](mailto:alberto.minassi@uniupo.it) 📧 Department of Pharmaceutical Sciences, University of Piemonte Orientale, Largo Donegani 2/3, Novara 28100, Italy

\*These Authors contributed equally and joined as First Author.

📄 Supplemental data for this article can be accessed [here](#).

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** Single-step reaction vs chemo-enzymatic approach. (A) The SNAP-tag<sup>®</sup> technology is based on BG-derivatives singularly synthesised and purified, and not excluding that the conjugated chemical group (green sphere) could affect the enzymatic reaction rate. (B) The SNAP-tag<sup>®</sup> technology revised uses a unique and universal azide BG-derivative, converting SNAP-tag<sup>®</sup> in a clickable form, prone to perform a fast and efficient cycloaddition with DBCO-based chemical groups. POI, protein of interest genetically fused to the SNAP-tag<sup>®</sup>.

development of SNAP-based sensors for small molecules<sup>16–18</sup> and ions<sup>19,20</sup>, and protein-DNA complexes in “DNA Origami” structures<sup>21</sup>.

Following the same approach, Perugino and co-workers expanded this technology to extremophilic organisms and to all the applications which require harsh reaction conditions, not fully suitable for the employing of the mesophilic SNAP-tag<sup>®</sup>. To this aim, they developed a “thermo-SNAP-tag” by the production of a variant of the OGT from *Saccharolobus solfataricus* (previously *Sulfolobus solfataricus*, *SsOGT-H<sup>5</sup>*, hereinafter *H<sup>5</sup>*), an enzyme which revealed extremely resistant to high temperature, high ionic strength, proteases attack, and, in general, to common physical and chemical denaturants<sup>22,23</sup>. The intrinsic stability of *H<sup>5</sup>* made it compatible with expression and utilisation *in vivo* as protein-tag in thermophilic organisms, as *Thermus thermophilus*<sup>24</sup> and *Sulfolobus islandicus*<sup>25</sup> as well as in an *in vitro* expression system using *Sulfolobus* lysates<sup>26</sup>. Recently, *H<sup>5</sup>* became a part of the new ASL<sup>tag</sup> system<sup>27</sup>, which was particularly useful for the *in vivo* immobilisation and contemporary labelling of proteins and enzymes of interest, stabilising them without any purification procedures needed<sup>28</sup>.

SNAP-tag<sup>®</sup> technology is essentially based on BG-substrates: although many of them are commercially available, the possibility of conjugation of infinite desired molecules to the 4-position on BG leads to the synthesis of *ad hoc* substrates. This is generally possible through the crosslinking reaction of the so-called “BG-building block” (such as the amine-reactive BG-NH<sub>2</sub>) with NHS-ester derivative compounds. The main disadvantage is the need to purify the final compounds before the reaction with the enzyme, increasing the times and costs of the experiments (Figure 1(A)). Furthermore, the presence of chemical groups conjugated to the benzyl moiety of the BG could affect the reaction efficiency of the SNAP-tag<sup>®</sup><sup>29–33</sup>, sometimes making this enzyme not fully applicable to particular requests.

In this work, we analysed and confirmed the catalytic dependence of SNAP-tag<sup>®</sup> and *H<sup>5</sup>* by several substrates having different

chemical groups conjugated to the O<sup>6</sup>-position of the guanine. To overcome these limitations, in the current study we suggest a further improvement of this technology with the application of a chemo-enzymatic approach, by using a unique and universal azide decorated BG-derivative, to obtain the specific labelling of the tag (clickable-SNAP), that can be easily coupled with a potentially infinite number of commercially available di-benzo-cyclo-octyl (DBCO)-based molecules, through the copper-free azide-alkyne Huisgen cycloaddition (Figure 1(B)). This approach could mainly offer the advantage to take into account of a unique reaction rate for the enzyme (with the azide-based BG), saving costs and times for the linking to the tag of an infinite number of commercially available DBCO-molecules. Here, we successfully proved the labelling of the SNAP-tag<sup>®</sup> with several DBCO-based fluorophores and the covalent immobilisation of this protein on alkyne-coated surface sensors.

## 2. Materials and methods

### 2.1. Reagents

**BG** was from Activate Scientific GmbH (UK), whereas **MGPA** was a gift of Prof D. Prosperi (University of Bicocca, Milan, Italy). SNAP-Vista<sup>®</sup> Green (**SVG**), SNAP Cell<sup>®</sup> Block (**SCB**), SNAP Cell<sup>®</sup> 430 (**SC430**), BG-PEG-NH<sub>2</sub> (**BGPA**), pSNAP-tag(m) plasmid, DNA restriction endonucleases and DNA modification enzymes were purchased from New England Biolabs (USA). Molecular biology kits for plasmid preparations were from Macherey-Nagel GmbH (Germany). Oligonucleotides synthesis and DNA sequencing service were performed by Eurofins Genomics (Germany). **BDP FL alkyne**, **BDP FL DBCO**, **Cy5 DBCO** were purchased from Lumiprobe GmbH (Germany). **DBCO-PEG<sub>4</sub>-Fluor 545**, Tris(2-carboxyethyl)phosphine (TCEP), Tris [(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) were from Sigma-Aldrich (St. Louis, MO). Pierce<sup>™</sup> Premium Grade Sulfo-NHS (N-hydroxy-sulfo-succinimide) and Pierce<sup>™</sup> Premium Grade 1-ethyl-3-(3-dimethyl-amino-propyl)-

carbodiimide hydrochloride (EDC) were from Thermo Fisher Scientific (Carlsbad, CA).

## 2.2. Compounds synthesis: general procedures

$^1\text{H}$  (400 MHz) and  $^{13}\text{C}$  (100 MHz) NMR spectra were measured on Bruker Advance Neo 400 MHz spectrometer. Chemical shifts were referenced to the residual solvent signal ( $\text{CDCl}_3$ :  $\delta_{\text{H}} = 7.26$ ,  $\delta_{\text{C}} = 77.0$ ; DMSO:  $\delta_{\text{H}} = 2.50$ ,  $\delta_{\text{C}} = 39.5$ ). Low-resolution ESI-MS were obtained on an LTQ OrbitrapXL (Thermo Scientific) mass spectrometer. IR spectra were registered on Shimadzu DR 8001 spectrophotometer. Silica gel 60 (70–230 mesh) used for gravity column chromatography (CC) was purchased from Macherey–Nagel. Reactions were monitored by TLC on Merck 60 F254 (0.25 mm) plates, visualised by staining with 5%  $\text{H}_2\text{SO}_4$  in ethanol or  $\text{KMnO}_4$  and heating. Organic phases were dried with  $\text{Na}_2\text{SO}_4$  before evaporation. Chemical reagents and solvents were from Aldrich, Alfa Aesar, and TCI and were used without any further purification unless stated otherwise.

## 2.3. Synthesis of BGN3

**BGN3** was synthesised according to the method of Zhang et al.<sup>34</sup>, whose experimental spectra were comparable. White solid.  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta$  7.84 (s, 1H), 7.53 (d,  $J = 7.8$  Hz, 2H), 7.39 (d,  $J = 7.8$  Hz, 2H), 6.28 (bs, 2H), 5.49 (s, 2H), 4.45 (s, 2H) (Figure S1(A)).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta$  159.69, 136.80, 135.43, 128.82, 128.58, 66.44, 53.37. IR (KBr)  $\text{cm}^{-1}$ : 3638, 3462, 3322, 2799, 2132, 1424, 1257, 1163, 912, 790, 656, 514. ESI/MS:  $m/z$  [ $\text{M} + \text{H}^+$ ] 297 (Figure S1(B)).

## 2.4. Synthesis of BGSN3

**BGSN3** was synthesised by following the scheme in Figure S2.

### 2.4.1 Synthesis of 4-azido-N-(4-(hydroxymethyl) benzyl) butanamide (compound 3)

A stirred solution of compound **1** (see Figure S2; 1.176 g, 9.115 mmol, 1 eq/mol) was prepared according to the method by Huang et al.<sup>35</sup> in DCM (30 ml), compound **2** (1.500 g, 10.939 mmol, 1.2 eq/mol; prepared according to the method by Leng et al.<sup>36</sup> and TEA (5.08 ml, 36.460 mmol, 4 eq/mol) were added. The mixture was stirred for 10 min at room temperature, then T3P (50% solution in EtOAc, 10.85 ml, 18.230 mmol, 2 eq/mol) was slowly added dropwise, and the stirred reaction was left overnight at room temperature until the complete conversion of the starting material (TLC: PE-EtOAc 4:6; Rf 1 = 0.47; Rf 3 = 0.16). The reaction was quenched by the addition of BRINE and extraction with DCM. After drying ( $\text{Na}_2\text{SO}_4$ ) and evaporation, the residue was purified by gravity column chromatography on silica gel (gradient PE-EtOAc from 6:4 to 3:7) to afford compound **3** as a white solid (660 mg, 30%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.28 (d,  $J = 8.0$  Hz, 2H), 7.20 (d,  $J = 7.9$  Hz, 2H), 6.36 (t,  $J = 5.6$  Hz, 1H), 4.63 (s, 2H), 4.35 (d,  $J = 5.7$  Hz, 2H), 3.31 (t,  $J = 6.6$  Hz, 2H), 2.27 (t,  $J = 7.3$  Hz, 2H), 1.89 (p,  $J = 6.9$  Hz, 2H) (Figure S3(A)).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  171.93, 140.30, 137.28, 127.69, 127.17, 64.53, 50.65, 43.19, 32.96, 24.69. IR (KBr)  $\text{cm}^{-1}$ : 3276, 3055, 2921, 2880, 2103, 1635, 1540, 1418, 1257, 1015, 827, 747, 677, 553. ESI/MS:  $m/z$  [ $\text{M} + \text{H}^+$ ] 249 (Figure S3(B)).

### 2.4.2. Synthesis of (N-(4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)-4-azidobutanamide) (BGSN3)

To a cooled solution (0 °C) of compound **3** (400 mg, 1.611 mmol, 1 eq/mol) in dry DMF (10 ml) in a dry flask under  $\text{N}_2$  atmosphere, NaH (60% dispersion in mineral oil, 202 mg, 5.059 mmol, 3.14 eq/mol) was slowly added. The mixture was stirred at 0 °C for 10 min, then DMAP (16 mg, 0.129 mmol, 0.08 eq/mol) and compound **4** (451 mg, 1.772 mmol, 1.1 eq/mol; prepared according to the method by Kindermann et al.<sup>37</sup> were sequentially added. The reaction was then heated at room temperature and stirred for 4 h until the complete conversion of the starting material (TLC: DCM-MeOH 9:1; Rf 4 = 0.70; Rf BGSN = 0.55), then quenched by slow addition of BRINE and extraction with EtOAc. After drying ( $\text{Na}_2\text{SO}_4$ ) and evaporation, the residue was purified by gravity column chromatography on silica gel (gradient DCM-MeOH from pure DCM to 20:1) to afford **BGSN3** as a white solid (413 mg, 67%).  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta$  12.48 (bs, NH purine, 1H), 8.43 (t,  $J = 5.9$  Hz, 1H), 7.85 (s, 1H), 7.49 (d,  $J = 7.7$  Hz, 2H), 7.30 (d,  $J = 7.8$  Hz, 2H), 6.31 (bs, NH2 purine 2H), 5.50 (s, 2H), 4.30 (d,  $J = 5.9$  Hz, 2H), 3.36 (t,  $J = 6.8$  Hz, 2H), 2.26 (t,  $J = 7.4$  Hz, 2H), 1.81 (p,  $J = 7.1$  Hz, 2H) (Figure S4(A)).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta$  171.38, 159.91, 159.69, 155.23, 139.48, 137.90, 135.31, 128.60, 127.34, 113.57, 66.59, 50.36, 41.95, 32.23, 24.58. IR (KBr)  $\text{cm}^{-1}$ : 3647, 3484, 3379, 3282, 2794, 2100, 1580, 1403, 1282, 1163, 938, 835, 789, 635, 553. ESI/MS:  $m/z$  [ $\text{M} + \text{H}^+$ ] 382 (Figure S4(B)).

## 2.5 Plasmids and protein purification

The cloning procedures in the pQE31 expression vector (Qiagen, Germany) were similar for both proteins. In particular, the pSNAP-tag(m) Vector was used as a template to amplify the DNA fragment relative to the SNAP-tag<sup>®</sup> gene, by using QE\_SNAP-Fwd/QE\_SNAP-Rev oligonucleotides pairs (5'-ATGGCAGGATCCAA TGGACAAAGACTGCGAAATG-3'/5'-CTATCAAAGCTTAACCCAGCCCAG GCTTGCCCA G-3'; BamH I and Hind III sites, respectively, are underlined). Afterwards, the resulting fragment and the pQE31 vector were digested with BamH I and Hind III restriction enzymes and ligated, leading to the final pQE-SNAP plasmid. The final SNAP-tag<sup>®</sup> protein was expressed with an extra N-terminal amino-acidic sequence, comprising a His<sub>6</sub>-tag (MRGSHHHHHHTDP-). The ligation mixture was used to transform *E. coli* KRX competent cells and positive colonies were confirmed by PCR and DNA sequence analyses.

$H^5$  was cloned as previously described<sup>22</sup>. SNAP-tag<sup>®</sup> and  $H^5$  proteins were expressed in *E. coli* ABLE C cells, grown at 37 °C in Luria–Bertani (LB) medium supplemented with 50 mg/l kanamycin and 100 mg/l ampicillin. The protein expression was induced with 1 mM isopropyl-thio- $\beta$ -D-galactoside (IPTG) at an absorbance value of 0.5–0.6  $A_{600}$  nm. After overnight growth, cells were collected and resuspended 1:3 (w/v) in purification buffer (50 mM phosphate, 300 mM NaCl; pH 8.0) supplemented with 1% Triton X-100 and stored overnight at –20 °C. Subsequently, the biomass was treated in ice with lysozyme and DNase for 60 min and then sonicated as described (Perugino et al., 2012). After centrifugation of 30 min at 60,000  $\times g$ , the cell extract was recovered and applied to a Protino Ni-NTA Column 1 ml (Macherey–Nagel) for His<sub>6</sub>-tag affinity chromatography. The eluted fractions containing the protein were collected and dialysed against phosphate-buffered saline (PBS 1 $\times$ , 20 mM phosphate buffer, NaCl 150 mM, pH 7.3). Pooled fractions were concentrated and protein purification was confirmed by SDS-PAGE analysis. Aliquots were finally stored at –20 °C.

**Table 1.** Substrate specificity of SNAP-tag<sup>®</sup> and H<sup>5</sup> by competitive inhibition method (IC<sub>50</sub>) by using SVG as substrate, and second order rate constant of the enzymatic reaction of these protein-tags only on the BGSN3 substrate.

Structure	Name	SNAP-tag <sup>®</sup>		SsOGT-H <sup>5</sup>		Note
		IC <sub>50</sub> (μM)	k <sup>a</sup> (s <sup>-1</sup> M <sup>-1</sup> )	IC <sub>50</sub> (μM)	k (s <sup>-1</sup> M <sup>-1</sup> )	
	SVG	–	2.8 × 10 <sup>4</sup> <sup>b</sup>	–	1.6 × 10 <sup>4</sup>	[14,24]
	BG	36.8 ± 5.6	–	10.1 ± 1.0	–	This work
	SCB	2.1 ± 0.5	–	4.4 ± 0.8	–	This work
	BGN3	15.6 ± 0.3	–	23.5 ± 1.0	–	This work
	BG430	ND <sup>c</sup>	–	ND	–	This work
	BGPA	86.0 ± 6.7	–	14.3 ± 1.9	–	This work
	MGPA <sup>d</sup>	–	–	268.9 ± 19.1 <sup>e</sup>	–	This work
	BGSN3	17.8 ± 1.1	4.64 ± 1.04 × 10 <sup>5</sup>	10.0 ± 0.7	1.40 ± 0.47 × 10 <sup>4</sup>	This work

For each compound, the guanine moiety is drawn *in black* and the chemical group conjugated to the benzyl ring *in blue*. The fluorescein moiety of the SVG is *in green*, whereas SCB differs from the other derivatives by the presence of a benzylic ring (*in red*). Azide group is conventionally coloured *in magenta*.

<sup>a</sup>Reaction rates at 25 °C; <sup>b</sup>this value was obtained by using a BG-fluorescein substrate (BG-FL) very similar to SVG; <sup>c</sup>not determined; <sup>d</sup>this molecule is a O<sup>6</sup>-methyl-guanine derivative; <sup>e</sup>competitive assay for H<sup>5</sup> was performed at 65 °C.

## 2.6 AGTs' substrate assay by competitive inhibition method

Competitive inhibition assay was performed as described<sup>23,38</sup>. Briefly, by using a fixed concentration of the fluorescent SVG (5 μM) and enzymes (5 μM), an increasing amount of guanine-derivatives (0–2 mM) was added to the mixtures. The reactions were incubated for 30 min at 25 °C and 50 °C for SNAP-tag<sup>®</sup> and H<sup>5</sup> respectively, and loaded on SDS-PAGE. Subsequently, fluorescent bands were measured by *gel-imaging* on a VersaDoc 4000<sup>™</sup> system (Bio-Rad), by applying a blue LED/530 bandpass filter. Obtained data were finally plotted by Equation (1),

$$RF = \frac{100\%}{1 + \left(\frac{[I]}{IC_{50}}\right)[S]} \quad (1)$$

where RF is the obtained Relative Fluorescence, [I] and [S] are the concentration of the inhibitor and the substrate, respectively, and

finally the IC<sub>50</sub> is the concentration needed to reduce by 50% the fluorescence intensity of the protein band.

We evaluated the activity of SNAP-tag<sup>®</sup> and H<sup>5</sup> enzymes on BGN3 and BGSN3 by the afore-mentioned IC<sub>50</sub> method (Figure S5(A,B)) and other O<sup>6</sup>-guanine-derivatives (Table 1).

## 2.7. In vitro Huisgen Cu(I)-catalysed cycloaddition reaction

The Huisgen chemical reaction was evaluated on SNAP-tag<sup>®</sup> and H<sup>5</sup> previously incubated with BGN3 and BGSN3. An opportune amount of purified proteins was incubated within an equimolar ratio of these substrates for 60–120 min at 25 °C and 37 °C respectively, to ensure the complete enzymatic labelling reaction. Later, we performed the subsequent cycloaddition using 5 μM of an alkylne-derivative of the fluorescein (BDP FL alkylne), in the

presence of copper (1 mM), TCEP (1 mM), TBTA (0.1 mM) and, where indicated, of SDS (0.05%). Finally, mixtures were loaded on SDS-PAGE and analysed as described in Section 4 (Figure S5(C,D)).

## 2.8. Molecular modelling

All molecular modelling studies were performed on a Tesla workstation equipped with two Intel Xeon X5650 2.67 GHz processors and Ubuntu 14.04 (<http://www.ubuntu.com>). The protein structures and 3D chemical structures were generated in PyMOL (The PyMOL Molecular Graphics System, version 2.2.3, Schrödinger LLC, 2019).

## 2.9. Molecular dynamics (MD) simulation

The MD simulations were carried out using the Desmond simulation package of Schrödinger LLC (Schrödinger Release 2019–1: Desmond Molecular Dynamics System; D. E. Shaw Research: New York, NY, 2019; Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2019). The X-ray structure of the  $H^5$  covalently bound to **SVG** was used in this study, entry code 6GAO<sup>39</sup>, water molecules were removed, and all hydrogen atoms and charges were added. The NPT ensemble with the temperature of 300 K and a pressure 1 bar was applied in all runs. The simulation length was 100 ns with relaxation time 1 ps. The OPLS\_2005 force field parameters were used in all simulations<sup>40</sup>. The long-range electrostatic interactions were calculated using the particle mesh Ewald method<sup>41</sup>. The cut-off radius in Coulomb interactions was 9.0 Å. The water molecules were explicitly described using the simple point charge model<sup>42</sup>. The Martyna–Tuckerman–Klein chain coupling scheme<sup>43</sup> with a coupling constant of 2.0 ps was used for the pressure control and the Nosé–Hoover chain coupling scheme<sup>44</sup> for the temperature control. Non-bonded forces were calculated using an r-RESPA integrator where the short-range forces were updated every step and the long-range forces were updated every three steps. The trajectory sampling was done at an interval of 1.0 ps. The behaviour and interactions between the ligands and protein were analysed using the Simulation Interaction Diagram tool implemented in the Desmond MD package. The stability of MD simulations was monitored by looking at the RMSD of the ligand and protein atom positions in time.

## 2.10. Determination of the rate constants of the chemo-enzymatic labelling approach

Rate constants of the enzymatic reactions with the only **BGSN3** were determined by the method of Gautier et al.<sup>14</sup>. In this case, purified proteins (5  $\mu$ M) were incubated with the substrate (5  $\mu$ M) in PBS 1 $\times$  buffer at 25 °C. Aliquots were taken at different times, the reactions were immediately stopped in Leamli Buffer 1 $\times$  in addition with 10  $\mu$ M of **Cy5 DBCO** fluorophore and placing tubes on ice.

Rate constants for the chemical reaction needed for the preliminary achievement of the *clickable*-SNAP and *clickable*- $H^5$  with **BGSN3**, which was obtained by the afore-described protocol, in order to get the complete labelling. Then, to each aliquot of 5  $\mu$ M of *clickable* proteins, 20  $\mu$ M of **DBCO-PEG<sub>4</sub>-Fluor 545** fluorophore was added. At different times, an excess of sodium azide (NaN<sub>3</sub>, 300 mM) was immediately added to each aliquot and then placing tubes on ice, in order to stop the click reaction between the azide group on the **BGSN3** and the **DBCO-PEG<sub>4</sub>-Fluor 545** molecule.

Finally, for both the experiments, all aliquots were boiled in an SDS buffer for 5 min, and immediately loaded on a SDS-PAGE, for the *gel-imaging* and *coomassie staining* analyses, as previously described. Data were fitted to a pseudo-first-order reaction model using the GraFit 5.0 software package (Erithacus Software Ltd.). Second-order rate constants  $k$  (in s<sup>-1</sup> M<sup>-1</sup>) were then obtained by dividing the pseudo-first-order constant by the concentration of the substrate (Figure 2 and Table 1). Values given are an average of at least three independent measurements.

## 2.11. In vitro Huisgen copper-free cycloaddition reaction with different DBCO-fluorophores

For the copper-free click reaction, aliquots of 5  $\mu$ M of each *clickable*-protein were incubated for 60 min at room temperature in the dark with 5  $\mu$ M of fluorescent DBCO-derivative substrates (**BDP FL DBCO**, **Cy5 DBCO**, and **DBCO-PEG<sub>4</sub>-Fluor 545**) in a total volume of 10  $\mu$ l of PBS 1 $\times$  buffer (Figure 4 and Figure S6). The reactions were finally stopped in Leamli Buffer 3 $\times$ , loaded on SDS-PAGE, and analysed as described in Section 4, by applying a blue LED/530 bandpass filter, red LED/695 bandpass filter and green LED/605 bandpass filter as excitation/emission parameters for each DBCO-fluorophores, respectively. The click reaction was also performed on 5  $\mu$ M of both the enzymes, but in the presence of an *EcCFE* diluted in PBS 1 $\times$  buffer.

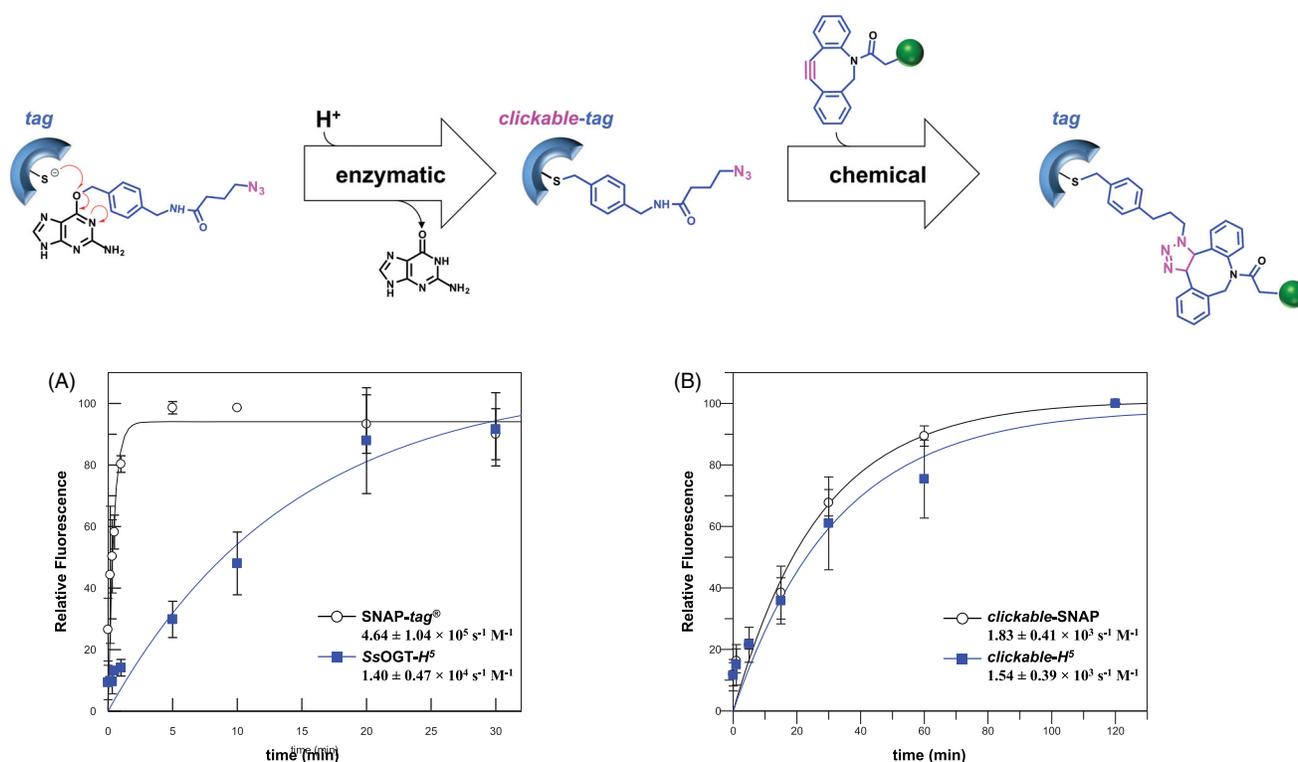
## 2.12. Procedure for protein immobilisation on bio layer interferometry (BLI), by following the chemo-enzymatic approach

OctetRED96<sup>TM</sup> (ForteBio, Fremont, CA) was used to immobilise specifically SNAP-*tag*<sup>®</sup> and  $H^5$  with the *chemo-enzymatic approach* (Figure 5(A,B)). Samples and reaction buffers were located in black 96-well plates (OptiPlate-96 Black, Black Opaque 96-well Microplate, PerkinElmer, Billerica, MA) in a maximum reaction volume of 300  $\mu$ l per well with 800 rpm shaking for each step. For the immobilisation procedure, AR2G sensors were first wetted in 200  $\mu$ l of pure water for at least 15 min, followed by an equilibration step (3 min) in acetate buffer 0.1 M, pH 5.0. Afterwards, they were activated with 20 mM 1-ethyl-3-(3-dimethyl-amino-propyl) carbodiimide hydrochloride (EDC)/20 mM N-hydroxy-sulfo-succinimide (sulfo-NHS) mixture in acetate buffer (60 min) and covered with 2 mM **propargyl-PEG<sub>3</sub>-amine** bifunctional linker (BroadPharm, San Diego, CA) in Loading step (20 min). To avoid the presence of any free amine groups on the biosensors, a Blocking step with Ethanolamine 1 M (30 min) was performed. Subsequently, a Washing step (15 min) with water and an Equilibration step in click-reaction buffer (15 min) are followed.

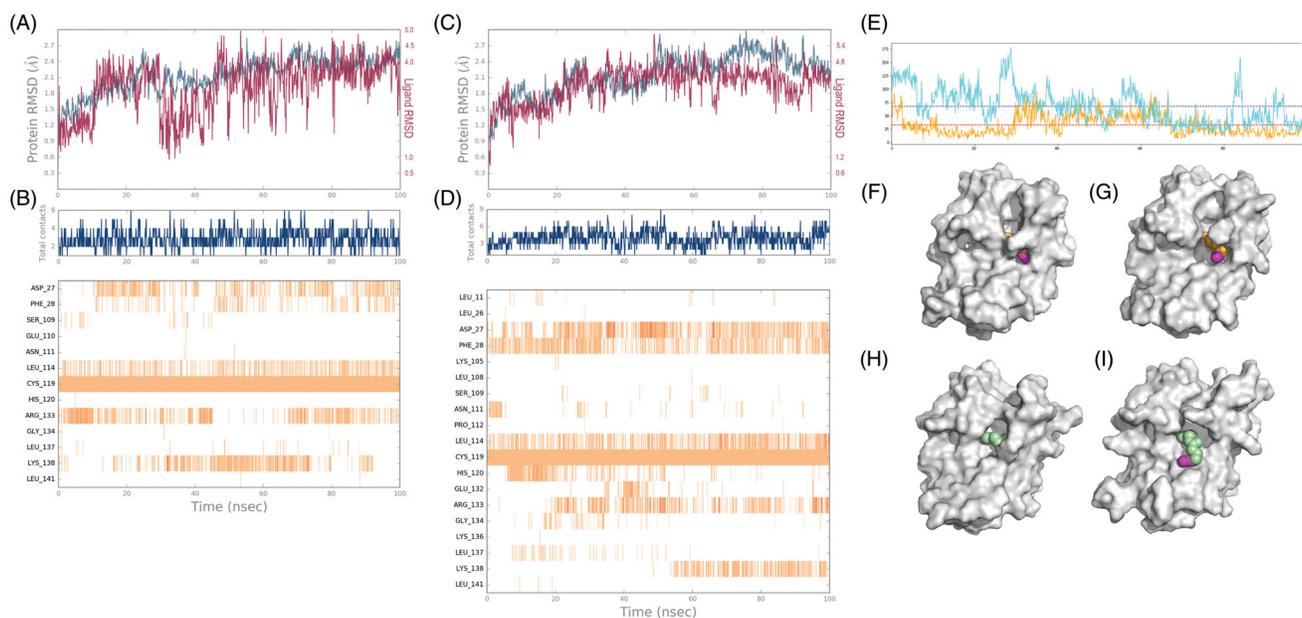
During the afore-described procedure, proteins were labelled with **BGSN3**. Finally, the immobilisation step for each sample via Huisgen reaction was carried out at 30 °C for 80 min, followed by a Washing step (20 min), in order to remove all the unbound molecules. This procedure was the same in the presence of the *EcCFE*. All measurements were performed in triplicates.

## 2.13. Permeability of eukaryotic and prokaryotic cells to BGSN3

HEK293T cells were maintained at 37 °C with 5% CO<sub>2</sub> in Dulbecco's Modified Essential Medium (Invitrogen, Carlsbad, CA) supplemented with 10% Foetal Bovine Serum (FBS) (Invitrogen) and 100 U/ml Penicillin/Streptomycin (Roche, Switzerland). HEK293T cells were transfected with SNAP-*tag*<sup>®</sup> plasmid by using



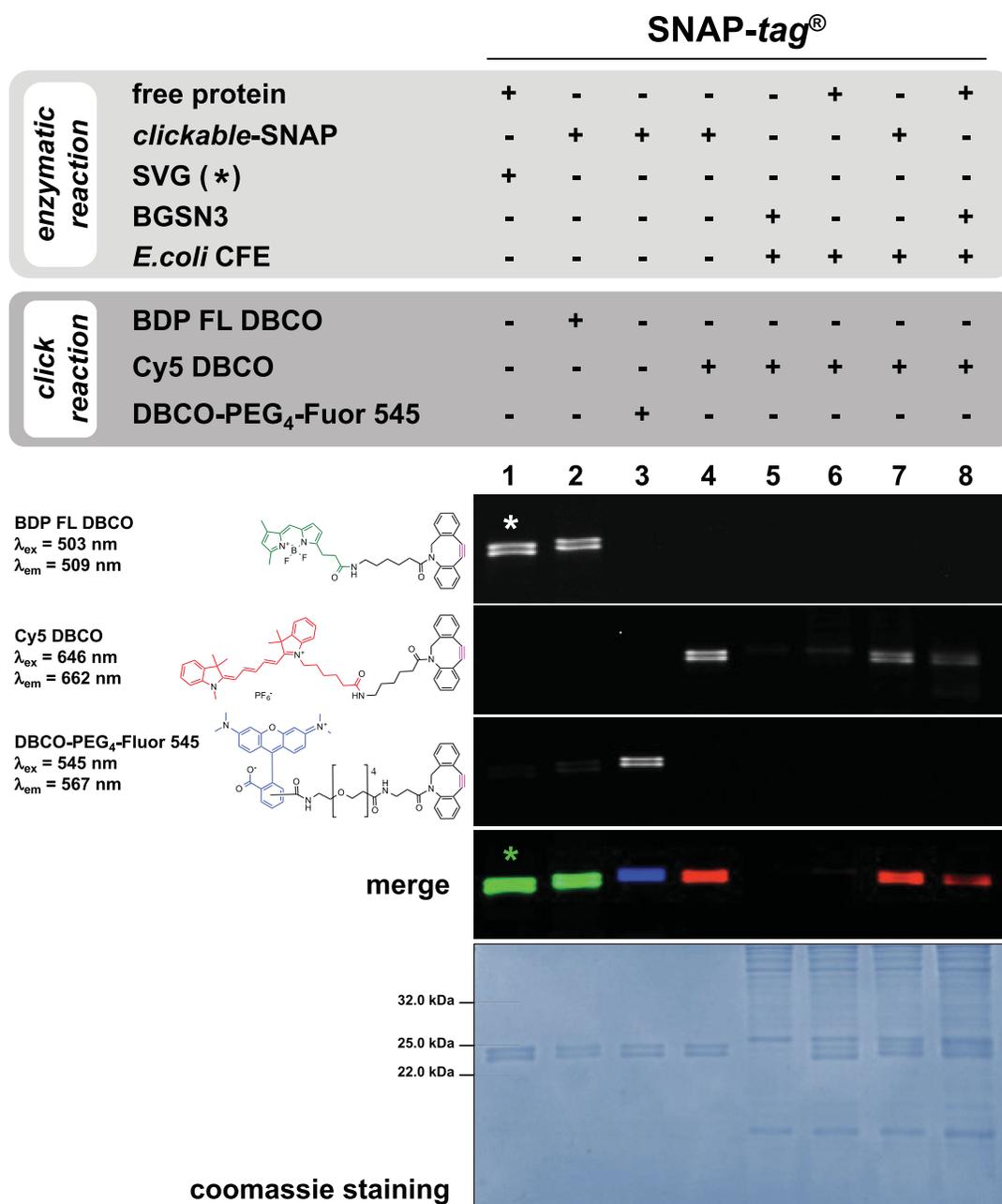
**Figure 2.** Reaction rates of the chemo-enzymatic approach. Pseudo-first-order reaction of *protein-tags* for (A) the enzymatic reaction with BGSN3 (see  $k$  values also in Table 1), and of *clickable-tags* for (B) Huisgen reaction with DBCO-PEG<sub>4</sub>-Fluor 545 (see values in the main text). Values given are an average of three independent measurements. The reaction scheme was an exemplification of Figure 1(B) in the main text. Data are represented as mean  $\pm$  SEM.



**Figure 3.** Molecular modelling on H<sup>5</sup> with BG-azides. (A) RMSD of the atomic positions for the compound BGN3 (Lig fit Prot, in red) and the protein H<sup>5</sup> (C $\alpha$  positions, in blue) of the 100 ns molecular dynamics simulations using Desmond package. (B) A timeline representation of the interactions and contacts (H-bonds, Hydrophobic, Ionic, Water bridges). (C) RMSD of the atomic positions for the compound BGSN3 (Lig fit Prot, in red) and the protein H<sup>5</sup> (C $\alpha$  positions, in blue) of the 100 ns molecular dynamics simulations using Desmond package. (D) A timeline representation of the interactions and contacts (H-bonds, Hydrophobic, Ionic, Water bridges). (E) Solvent Accessible Surface Area (SASA) of BGN3/H<sup>5</sup> (in orange) and BGSN3/H<sup>5</sup> (in cyan) complexes over the MD simulation time (mean values are depicted as dot lines). Frames of H<sup>5</sup>-probe complexes with lower (F, H) and higher (G, I) SASA value for BGN3 (F, G) and BGSN3 (H, I), respectively.

Lipofectamine 2000 (Invitrogen) following manufacturer's protocol. The treatment with BGSN3 were performed, at the concentration and time indicated for each experiment. Twenty-four hours after transfection, we treated cells with BGSN3 for 2 h at different

concentrations ranging (from 1 to 25  $\mu$ M), directly dissolving the compound in complete culture medium. Then cells were harvested, washed with PBS 1 $\times$  buffer and lysed with 50mM Tris-HCl pH 7.4, 150mM NaCl, 0.5mM EDTA, 0.1% Triton X-100



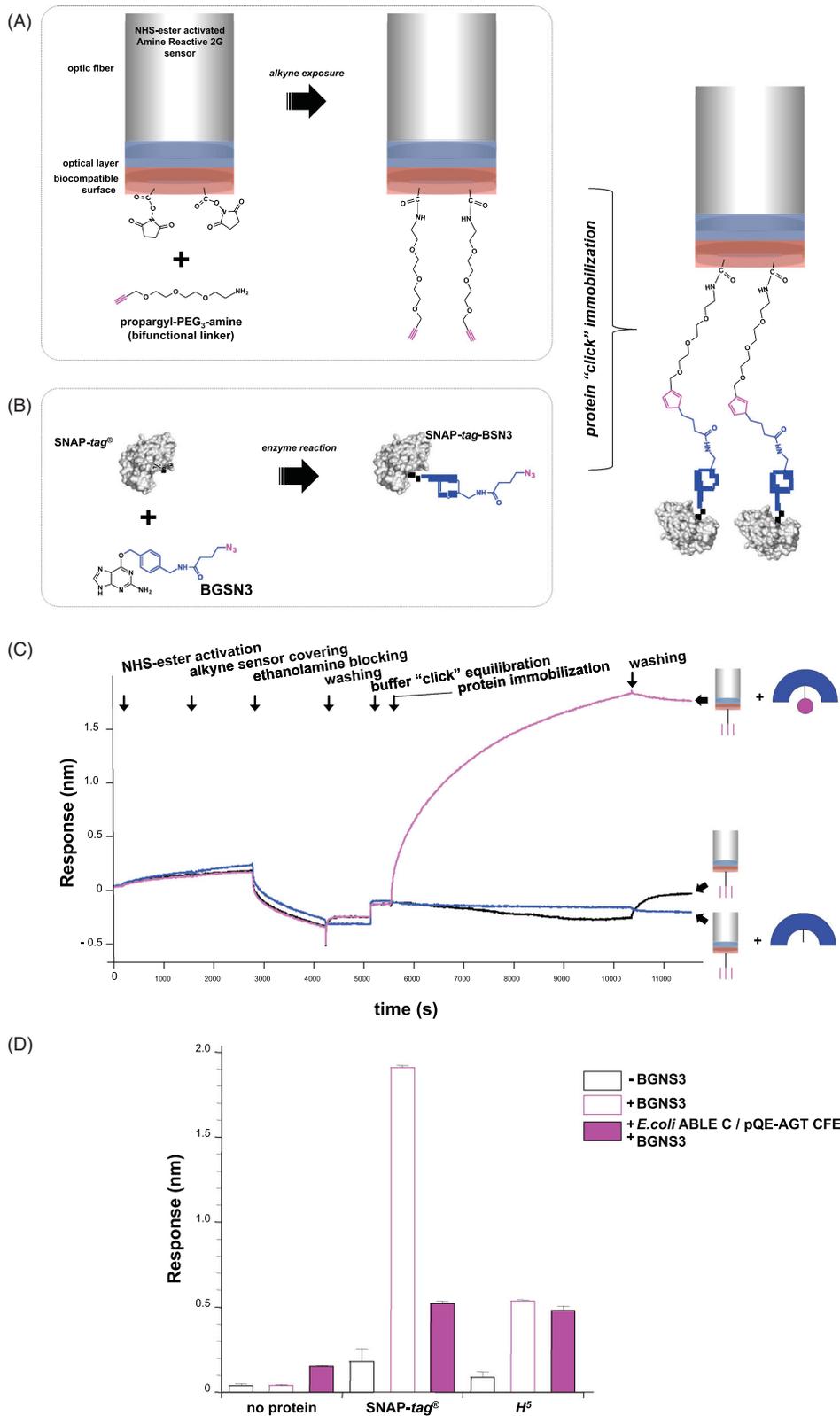
**Figure 4.** Specificity of the Huisgen reaction. *Gel-imaging* analysis of SNAP-tag<sup>®</sup> labelling by a chemo-enzymatic approach with BGSN3 and three different DBCO-derivative fluorophores. Protein (5 μM) was incubated with 5 μM of the azide-based BG for 60 min at 25 °C; then, an equimolar amount of DBCO-based substrate was added for the chemical click reaction, keeping the same time and temperature conditions. As control, SNAP-tag<sup>®</sup> was incubated only with SVG (lane 1, signal marked with an asterisk).

supplemented with complete protease (Roche, Switzerland) and phosphatase (SERVA Electrophoresis, Germany) inhibitors. Afterwards, transfected cells were treated with a fixed concentration of BGSN3 (10 μM) at different time points (from 30 to 120 min). Again, HEK293T cells were washed and lysed as described before. To confirm the reaction with BGSN3, the same amount of protein extract (0.91 μg/μL for each sample) was incubated for 30 min at 25 °C with SVG. Subsequently, proteins were loaded on SDS-PAGE and analysed by *gel-imaging* on a VersaDoc 4000<sup>™</sup> system (Bio-Rad), by applying a blue LED/530 bandpass filter (Figure 6).

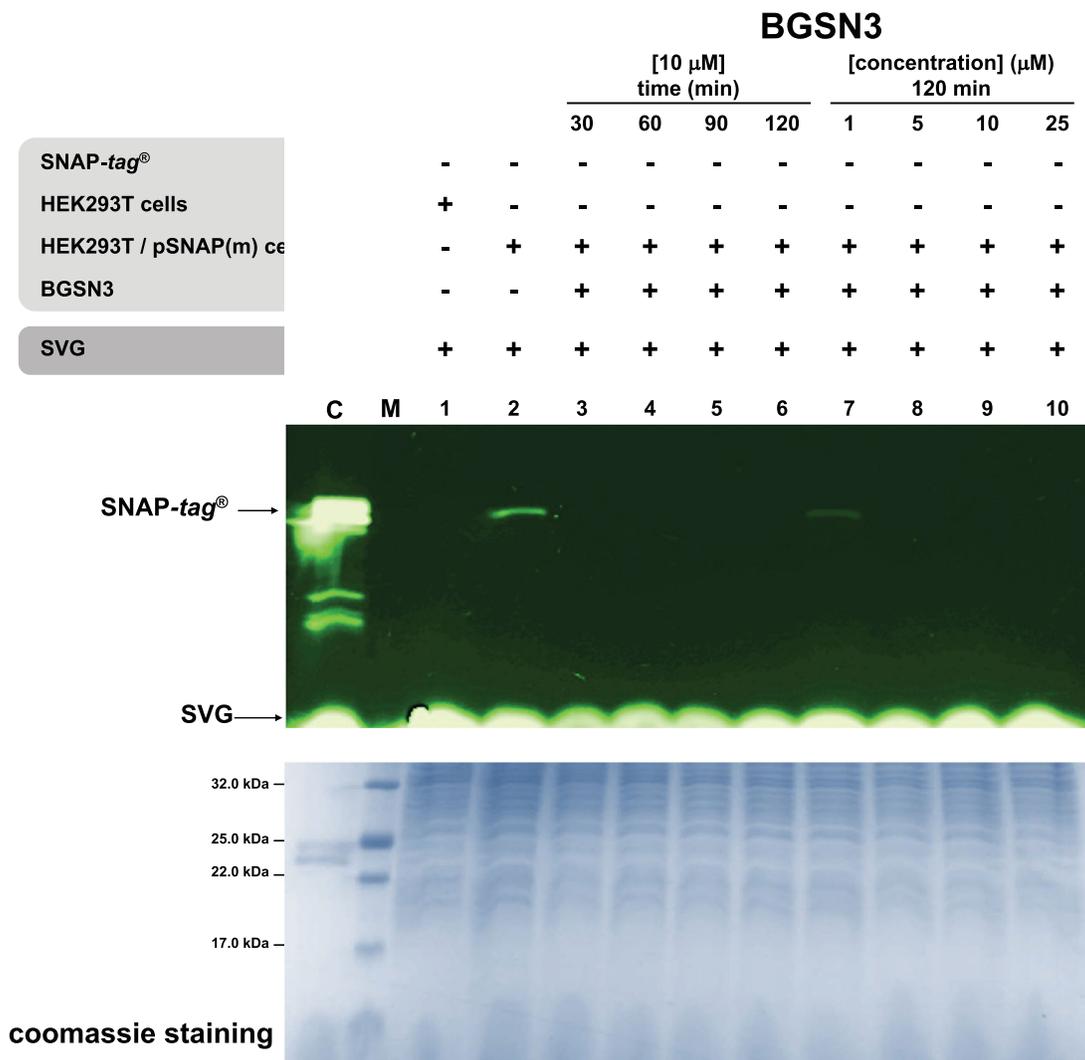
For flow cytometry analysis, HeLa cells were seeded in 24-well plates and transfected with SNAP-tag<sup>®</sup> plasmid by using

Lipofectamine 2000 (Invitrogen, USA) following manufacturer's protocol. Twenty-four hours after the transfection, cells were treated with 25 μM BGSN3 for 1 h, and the excess of the substrate was washed out by 2 × 15 min, followed by 1 × 30 min washes. Cells were then treated with 2.5 μM BDP FL DBCO for 30 min and unbound fluorophore was removed by following the same procedure performed for the BGSN3. All treatments and washes were performed at 37 °C in a complete culture medium. Lastly, cells were harvested by trypsinization, and fluorescence was measured using FACS CANTO II instrument. The analysis was performed on live singlet cells using FlowJo software (Figure S7(A)).

*E. coli* ABLE C strain was transformed with SNAP-tag<sup>®</sup> plasmid and protein expressed as previously described. After overnight



**Figure 5.** Covalent immobilisation of clickable-tags on the BLI sensor. (A) Covering of the BLI sensor with a bi-functional linker, exposing alkyne groups for the Huisgen cycloaddition reaction; (B) reaction of the SNAP-tag<sup>®</sup> with BGSN3; (C) chemo-enzymatic SNAP-tag<sup>®</sup> immobilisation on BLI. The alkyne-covered sensor (*silver cylinder*) was immersed in wells containing the buffer (*in black*), the free SNAP-tag<sup>®</sup> (*in blue*) and the *clickable*-SNAP (*in magenta*); (D) column chart relative to the BLI immobilisation of purified *protein-tags* alone (*black-bordered bars*) or in the presence of BGSN3 (*magenta-bordered bars*). Filled magenta bars represent the BLI immobilisation using the *EcCFE* upon heterologous expression of *protein-tags*. Standard deviations were obtained from three independent experiments. Data are represented as mean  $\pm$  SEM.



**Figure 6.** Eukaryotic permeability to BGSN3. SDS-PAGE analysis by *gel-imaging* and *coomassie staining* of HEK293T cell lysates. After BGSN3 in medium treatment, lysates were incubated with SVG.

growth, samples of 2 ml were treated with 100  $\mu$ M of **BGSN3** for 2 h at 25 °C and then collected by centrifugation at 2000  $\times g$ . Cell pellets of 0.05 g were resuspended 1:3 (w/v) in PBS 1 $\times$  supplemented with 1% Triton X-100 and subjected to cell lysis, by applying 5 cycles of freeze-thawing. After a centrifugation at 13,000  $\times g$ , the supernatants containing the protein extract were incubated 30 min at 25 °C with **SVG**, and proteins were loaded on SDS-PAGE. Finally, fluorescent bands were analysed by *gel-imaging* techniques (Figure S7(B)).

### 3. Result and discussion

#### 3.1. Substrate specificity of AGTs on BG-based substrates

Following the irreversible reaction shown in Figure 1, we evaluated the activity of two enzymes in our possession on several *O*<sup>6</sup>-guanine-derivatives (Table 1). Because most of them are non-fluorescent compounds, we performed an AGTs' competitive inhibition assay by using the fluorescein-derivative SNAP-Vista<sup>®</sup> Green as substrate (**SVG**), as previously described<sup>22–24,45</sup>. Briefly, the reaction of an AGT with **SVG** led to a fluoresceinated protein, which can be visualised as a fluorescent band in *gel-imaging* analysis after SDS-PAGE. The presence of increasing amounts of a

non-fluorescent competitor in the reaction causes a decrease of the fluorescent signals, which can be measured and plotted for the IC<sub>50</sub> values determination<sup>23,46</sup>. As shown in Table 1, SNAP-*tag*<sup>®</sup> and *H*<sup>5</sup> displayed different behaviours versus these competitors, without any rationale for the dimension and/or polarity of the conjugated chemical groups. While SNAP-Cell<sup>®</sup> 430 (**SC430**) completely lost the competition with **SVG**, both the enzymes are extremely active on the SNAP Cell<sup>®</sup> Block (**SCB**), displaying the lowest IC<sub>50</sub> value measured. This result was expected, because **SCB** has a structure very similar to the Lomeguatrib, one of the most efficient inhibitors of the hMGMT protein, employed in the cancer treatment in combination with alkylating agents-based chemotherapeutics<sup>47</sup>.

In general, all commercially available products used (**SVG**, **SCB**, **BG430**, and **BG-PEG-NH2**, **BGPA**) are good substrates for the SNAP-*tag*<sup>®</sup> and *H*<sup>5</sup> enzymes, completing their labelling reaction in few hours (data not shown). However, based on our results, the choice of the chemical group to be conjugated to the *O*<sup>6</sup>-guanine for customized substrates may present risks, with consequent decreases in the reaction rate for these *protein-tags*. This was the case of methyl-guanine-PEG-NH2 (**MGPA**), which is an *O*<sup>6</sup>-methyl-guanine derivative, used for the immobilisation of SNAP-*tag*<sup>®</sup> on nanoparticles<sup>48</sup>. The latter is not a preferred substrate, probably

because of the absence of the benzyl ring, which leads to complete labelling of the SNAP-tag<sup>®</sup> and H<sup>5</sup> after over-night incubation at 4 °C<sup>48</sup> and 65 °C (data not shown), respectively.

### 3.2. In vitro enzymatic reaction of engineered AGTs with BG-azide substrates

Recent studies were focussed on the synthesis of alternative “BG-building blocks,” which offer the opportunity to produce SNAP-substrates by following easier and faster protocols: an alkyne substituted O<sup>6</sup>-BG was employed in the synthesis of compounds by the Huisgen cycloaddition with azide-based fluorescent probes<sup>49</sup> or, inversely, by using the O<sup>6</sup>-BG-N<sub>3</sub> (**BGN3**, Figure S1) for the conjugation with alkyne-based chemical groups<sup>34</sup>. We evaluated the enzymatic reaction of the H<sup>5</sup> and the SNAP-tag<sup>®</sup> directly on **BGN3** and a synthesised BG-derivative containing a benzyl ring opportunely spaced from the azide group (**BGSN3**, Figure S4): after the reaction, no fluorescent signal was obtained on SDS-PAGE *gel-imaging* upon the addition of **SVG** (Figure S5(A,B)). This indicates that the catalytic cysteine was completely blocked by the benzyl-azide moiety, impeding the access of the fluorescent substrate to the active site. Compared to the classical BG-derivatives, these *protein-tags* showed a reasonable activity on both these BG-azides, as resulted by the calculated IC<sub>50</sub> (Table 1 and Figure S5(A,B)).

After the enzymatic reaction of H<sup>5</sup> with **BGN3** and **BGSN3**, we performed the subsequent cycloaddition using an alkyne-derivative of the fluorescein (**BDP FL alkyne**): however, the chemical reaction was less efficient using the former substrate (Figure S5(C), lane 2). In this case, the complete fluorescein labelling of the protein was achieved only in the presence of a small amount of SDS during the cycloaddition step (lane 3), suggesting that the protein is still folded after the enzymatic reaction and the azide is hidden in the active site core. The addition of the denaturant could have slightly opened the protein structure, favouring a better exposure of the azide group to the solvent, and allowing the click reaction to occur.

On the contrary, using **BGSN3** as substrate, the labelling of both the enzymes was comparable to the classical reaction with **SVG** without any denaturing agent, likely the longer spacer of **BGSN3** could sufficiently move away from the azide group from the protein surface for the Huisgen reaction (Figure S5(D), lanes 2 and 4). From now on, experiments were only performed by using the longer BG-azide. We first calculated the rate of the enzymatic reaction, demonstrating that both *protein-tags* show a high catalytic activity comparable to the commercial BG-derivatives currently used (Figure 2(A) and Table 1), also indicating that the complete protein labelling in less of an hour can be performed<sup>13,14,24</sup>.

### 3.3. Molecular modelling on the H<sup>5</sup> with BG-azides

**BGN3** and **BGSN3** differ in length since the chemical spacer between the benzyl ring and the active azide makes the latter potentially more prone to the labelling reaction. It could be assumed that this aspect alone influences the availability of the azide moiety to react. However, proteins are not a static system, the amino acids side-chain movements could mask the azide and prevent the “click” chemistry reaction. The covalent complexes of these compounds with H<sup>5</sup> were analysed with Molecular Dynamics (MD) simulations using the Desmond package (see Experimental Section). The complexes were simulated for 100 ns at 300 K using a standard protocol. The protein structure has been stabilised, as shown in the RMSDs for both the IDO1 C $\alpha$  and the ligand (Figure

3(A,C)). The MD results were analysed in terms of Solvent Accessible Surface Area (SASA) of the compounds: more time the compounds are exposed to the solvent, the higher is the possibility to react<sup>50</sup>. In Figure 3 is reported the fluctuation of the SASA values over the simulation time together with the structure model of the H<sup>5</sup> protein in complex with **BGN3** and **BGSN3**, respectively. The former is less exposed to the solvent with a SASA value of  $32.967 \pm 18.573 \text{ \AA}^2$  compared to **BGSN3**, which shows a higher SASA value  $68.302 \pm 32.455 \text{ \AA}^2$ . This simulation confirmed our biochemical data, proposing the BG-derivative with the spacer as a better substrate for our *chemo-enzymatic approach*.

### 3.4. Specificity and versatility of the chemo-enzymatic reaction

The O<sup>6</sup>-BG-based **BGSN3** is a good substrate for the two *protein-tags* used (Table 1 and Figure 2(A)) and offering the advantage to sufficiently expose the azide group for the Huisgen reaction. This was the starting point to examine: (i) the labelling efficiency of the *clickable*-SNAP and *clickable*-H<sup>5</sup> by using different DBCO-based fluorophores; (ii) the specificity of the “click” reaction.

Upon the reaction with **BGSN3**, all cycloaddition reactions with three different DBCO-based fluorophores were complete in ca. 30–45 min in PBS 1 $\times$  buffer (Figure 4, lanes 2–4), with a protein-labelling as efficient as the enzymatic reaction using the sole **SVG** (lane 1). We quantitatively evaluated the rate (*k*) of the click reaction by using the **DBCO-PEG4-Fluor 545** fluorophore: as expected, both the *clickable-tags* were labelled with the same efficiency ( $1.83 \pm 0.41 \times 10^3 \text{ s}^{-1} \text{ M}^{-1}$  for SNAP-tag<sup>®</sup>;  $1.54 \pm 0.39 \times 10^3 \text{ s}^{-1} \text{ M}^{-1}$  for H<sup>5</sup>), demonstrating that the chemical reaction is sufficiently fast and independent from the *tags* (Figure 2(B)).

Concerning the specificity, we added a crude protein extract from *Escherichia coli* ABLE C (ECCFE), without any AGT activity at the *gel-imaging* analysis (Figure 4, lane 5). In this context, the only presence of the free *protein-tag* and the DBCO-fluorophore also did not result in any fluorescent signal (lanes 6), whereas the previously purified *clickable*-SNAP (lane 7), as well as its free form in the presence of **BGSN3** (lane 8), was specifically able to complete the chemo-enzymatic reaction, giving an evident fluorescent signal. The high specificity of our approach was also confirmed by using the H<sup>5</sup> enzyme, which displays a better labelling reaction than the mesophilic SNAP-tag<sup>®</sup> (Figure S6). Probably, something in the extract might impede SNAP-tag<sup>®</sup> activity. These results clearly demonstrated the high efficiency of our *chemo-enzymatic approach* for the labelling of both the *protein-tags* used.

### 3.5. Application to the bio layer interferometry

The possibility to apply the SNAP-tag<sup>®</sup> technology to the Surface Plasmon Resonance (SPR) for the covalent immobilisation of a protein of interest was first explored by the group of Kai Johnsson<sup>37</sup>, followed by other groups with the same substrate<sup>51</sup> or a biotin BG-derivative<sup>52</sup>. Their approaches, again, required preliminarily the synthesis and the purification of a compatible substrate to cover the sensor chip surface. We used, instead, the **BGSN3** substrate for the immobilisation of the SNAP-tag<sup>®</sup> directly on an alkyne-derived sensor chip of the bio layer interferometry (BLI) equipment, as shown in Figure 5. This technique is more advantageous with respect to the SPR because: (i) it needs a smaller amount of sample, making it more compatible to higher throughput (the capacity of running up to 96 samples in a parallel); (ii) the possibility to reuse samples, and (iii) of the total independency from any microfluidic issues.

Given the lack of any available BLI alkyne-derived sensors, we first activated the AR2G type by a bi-functional linker (**propargyl-PEG 3-amine**) in order to expose an alkyne group on the surface (Figure 5(A)). This modified protocol provides first the coating of the sensor tips with alkyne groups (approx. 80 min), during that the reaction between the *protein-tag* and **BGSN3** inside the 96-wells rack takes place (Figure 5(B)). Only the contemporary presence of the *clickable*-SNAP and the alkyne-coated sensor led to a measurable response (Figure 5(C)). After washing procedures, the signal did not significantly drop-down, given the covalent reaction between the protein and the sensor. We successfully achieved results with both the enzymes, although temperature and times of the enzymatic reaction on BLI (30 °C) favoured the SNAP-*tag*<sup>®</sup> respect to the thermophilic *H*<sup>513,24</sup>. Furthermore, in *EcCFEs* where both the enzymes were expressed, a specific and efficient immobilisation on BLI sensor tips occurred (Figure 5(D)), although the SNAP-*tag*<sup>®</sup> displayed a lower labelling efficiency in the *EcCFE*, as expected (compare lane 8 in Figure 4 and Figure S6). As for other techniques, this specific surface immobilisation of SNAP-*tag*<sup>®</sup> gives the opportunity to perform a directly *on-chip purification* of a tagged-POI from a crude lysate. without any purification step, in an indirect manner, which favours a better orientation of the POI for its biological activities.

### 3.6. Permeability of eukaryotic and prokaryotic cells to BG-azides

One of the major applications of the SNAP-*tag*<sup>®</sup> technology concerns the field of cell biology, where detecting fluorescent-tagged-POIs in living cells represents an important tool to study protein functions and locations<sup>53</sup>. To test our *chemo-enzymatic approach*, we first investigated the permeability of **BGSN3**. Lysates of HEK293T cells pre-treated with **BGSN3** were then incubated with the **SVG** substrate: the absence of any fluorescent signal by *gel-imaging* only in BG-azide treated lysates demonstrated that the internalisation of **BGSN3** was fast (ca. 30 min; Figure 6, lane 3) and at concentrations comparable with commercial cell biology BG-substrates (in the range of <5 μM; Figure 6, lane 8). Preliminary experiments by FACS analysis confirmed that the *in vivo* cycloaddition between **BGSN3** and the **BDP-FL DBCO** occurred (Figure S7(A)). This was also confirmed for *E. coli* bacterial cells (Figure S7(B), lane 2).

## 4. Conclusions and perspectives

We developed an innovative modification of the SNAP-*tag*<sup>®</sup> technology, in order to overcome times and costs relative to the production and the utilisation of commercial or purified customised BG-derivatives. Although they are compatible in terms of catalytic activity as for the SNAP-*tag*<sup>®</sup>, as well as for the others AGTs<sup>22–24,37,46,54</sup> the risk of lowering the catalytic activity of these *tags* with customised BG-derivatives should not be underestimated (Table 1). We started by the knowledge that: (i) self-labelling *protein-tags* are still folded and enough stability in their benzylated form after the enzymatic reaction<sup>13,24</sup>; (ii) the Huisgen cycloaddition is extremely versatile, fast and specific. Recently, the latter was used for the entrapment of catalytic activities by azide-based pseudo-substrates in a well-known powerful method, the *in vivo* activity-based protein profiling (ABPP)<sup>55</sup>. For these reasons, a *chemo-enzymatic approach* (Figure 1(B)) with an opportunely selected azide-based BG-substrate (**BGSN3**) was set up: the efficient exposition of the azide outside the protein surface allows the reaction with a huge number of commercially DBCO-based

molecules, more than those BG-derivatives, keeping high the specificity in the presence of *in vitro* “perturbing” proteins (like in cell lysates) and the *in vivo* labelling of expressed SNAP-*tag*<sup>®</sup> in eukaryotic cells. Finally, **BGSN3** proved to be a good substrate for the direct immobilisation of these *tags* on solid surfaces. We demonstrated that splitting the SNAP-*tag*<sup>®</sup> reaction into two fast steps, as experimentally measured (Figure 2(A,B)), does not affect the overall rate and efficiency of the protein labelling<sup>13,24</sup>, thus opening new perspectives and widening the applications of this powerful biotechnology.

### Authors' contributions

Conceptualisation, G.P. and A.Mi.; Methodology, R.Me., G.P.; Investigation, R.Me., D.C. and M.C.; Formal Analysis, C.M., R.Mi. and A.Ma.; Validation, R.Me., R.Ma., A.Mi. and G.P.; Writing – Original Draft, R.Me. and G.P.; Writing – Review & Editing, C.M., F.R. and A.V. and G.P.; Funding Acquisition, A.V. and G.P.; Resources, R.Me., R.Ma. and D.C.; Supervision, A.L., A.Mi. and G.P.

### Acknowledgement

We are grateful to prof. Davide Prosperi (University of Milan Bicocca) for the kind gift of the **MGPA** substrate.

### Disclosure statement

No potential conflict of interest was reported by the author(s).

### Funding

This research was funded by Ministero dell'Istruzione Università e Ricerca (MIUR) National Operational Program (PON) Research and Innovation 2014–2020 [CCI 2014IT16M2OP005], European Social Fund, Action 10.1 “Innovative Doctorates with Industrial characterization.” This research was also funded by Università del Piemonte Orientale (project: RIVmiggianorFAR\_2019).

### References

1. Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA. Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol Rev* 2010;90:1103–63.
2. Liss V, Barlag B, Nietschke M, Hensel M. Self-labelling enzymes as universal tags for fluorescence microscopy, super-resolution microscopy and electron microscopy. *Sci Rep* 2015;5:17740–52.
3. Hatlem D, Trunk T, Linke D, Leo JC. Catching a SPY: using the SpyCatcher-SpyTag and related systems for labeling and localizing bacterial proteins. *Int J Mol Sci* 2019;20:2129–39.
4. Los GV, Encell LP, McDougall MG, et al. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* 2008;3:373–82.
5. England CG, Luo H, Cai W. HaloTag technology: a versatile platform for biomedical applications. *Bioconjug Chem* 2015; 26:975–86.
6. Pegg AE. Repair of O(6)-alkylguanine by alkyltransferases. *Mutat Res* 2000;462:83–100.
7. Daniels DS, Mol CD, Arvai AS, et al. Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. *Embo J* 2000;19:1719–30.

8. Daniels DS, Woo TT, Luu KX, et al. DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat Struct Mol Biol* 2004;11:714–20.
9. Keppler A, Gendreizig S, Gronemeyer T, et al. A general method for the covalent labeling of fusion proteins with small molecules *in vivo*. *Nat Biotechnol* 2003;21:86–9.
10. Keppler A, Pick H, Arrivoli C, et al. Labeling of fusion proteins with synthetic fluorophores in live cells. *Proc Natl Acad Sci USA* 2004;10:9955–9.
11. Juillerat A, Gronemeyer T, Keppler A, et al. Directed evolution of O6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules *in vivo*. *Chem Biol* 2003;10:313–7.
12. Gronemeyer T, Chidley C, Juillerat A, et al. Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Eng Des Sel* 2006;19:309–16.
13. Mollwitz B, Brunk E, Schmitt S, et al. Directed evolution of the suicide protein O6-alkylguanine-DNA alkyltransferase for increased reactivity results in an alkylated protein with exceptional stability. *Biochemistry* 2012;51:986–94.
14. Gautier A, Juillerat A, Heinis C, et al. An engineered protein tag for multiprotein labeling in living cells. *Chem Biol* 2008;15:128–36.
15. Vogel P, Moschref Merkle MM, Li Q, et al. Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. *Nat Methods* 2018;15:535–8.
16. Srikun D, Albers AE, Nam CI, et al. Organelle-targetable fluorescent probes for imaging hydrogen peroxide in living cells via SNAP-Tag protein labeling. *J Am Chem Soc* 2010;132:4455–65.
17. Abo M, Minakami R, Miyano K, et al. Visualization of phagosomal hydrogen peroxide production by a novel fluorescent probe that is localized via SNAP-tag labeling. *Anal Chem* 2014;86:5983–90.
18. Griss R, Schena A, Reymond L, et al. Bioluminescent sensor proteins for point-of-care therapeutic drug monitoring. *Nat Chem Biol* 2014;10:598–603.
19. Tomat E, Nolan EM, Jaworski J, Lippard SJ. Organelle-specific zinc detection using zinpyr-labeled fusion proteins in live cells. *J Am Chem Soc* 2008;130:15776–7.
20. Kamiya M, Johnsson K. Localizable and highly sensitive calcium indicator based on a BODIPY fluorophore. *Anal Chem* 2010;82:6472–9.
21. Saccà B, Meyer R, Erkelenz M, et al. Orthogonal protein decoration of DNA origami. *Angew Chem Int Ed Engl* 2010;49:9378–83.
22. Perugino G, Vettone A, Illiano G, et al. Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage. *J Biol Chem* 2012;287:4222–31.
23. Perugino G, Miggiano R, Serpe M, et al. Structure-function relationships governing activity and stability of a DNA alkylation damage repair thermostable protein. *Nucleic Acids Res* 2015;43:8801–16.
24. Vettone A, Serpe M, Hidalgo A, et al. A novel thermostable protein-tag: optimization of the *Sulfolobus solfataricus* DNA-alkyl-transferase by protein engineering. *Extremophiles* 2016;20:1–13.
25. Visone V, Han W, Perugino G, et al. *In vivo* and *in vitro* protein imaging in thermophilic archaea by exploiting a novel protein tag. *PLoS One* 2017;12:e0185791.
26. Lo Gullo G, Mattosovich R, Perugino G, et al. Optimization of an *in vitro* transcription/translation system based on *Sulfolobus solfataricus* cell lysate. *Archaea* 2019;2019:9848253.
27. Merlo R, Del Prete S, Valenti A, et al. An AGT-based protein-tag system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane. *J Enzyme Inhib Med Chem* 2019;34:490–9.
28. Del Prete S, Merlo R, Valenti A, et al. Thermostability enhancement of the  $\alpha$ -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* by using the anchoring-and-self-labelling-protein-tag system (ASL<sup>tag</sup>). *J Enzyme Inhib Med Chem* 2019;34:946–54.
29. Moschel RC, McDougall MG, Dolan ME, et al. Structural features of substituted purine derivatives compatible with depletion of human O6-alkylguanine-DNA alkyltransferase. *J Med Chem* 1992;35:4486–91.
30. Chae MY, McDougall MG, Dolan ME, et al. Substituted O6-benzylguanine derivatives and their inactivation of human O6-alkylguanine-DNA alkyltransferase. *J Med Chem* 1994;37:342–7.
31. Terashima I, Kawate H, Sakumi K, et al. Substrate specificity of human O6-methylguanine-DNA methyltransferase for O6-benzylguanine derivatives in oligodeoxynucleotides. *Chem Res Toxicol* 1997;10:1234–9.
32. McElhinney RS, Donnelly DJ, McCormick JE, et al. Inactivation of O6-alkylguanine-DNA alkyltransferase. 1. Novel O6-(hetaryl)methylguanines having basic rings in the side chain. *J Med Chem* 1998;41:5265–71.
33. Sun G, Fan T, Zhang N, et al. Identification of the structural features of guanine derivatives as MGMT inhibitors using 3D-QSAR modeling combined with molecular docking. *Molecules* 2016;21:823–44.
34. Zhang CJ, Li L, Chen GYJ, et al. One- and two-photon live cell imaging using a mutant SNAP-Tag protein and its FRET substrate pairs. *Org Lett* 2011;13:4160–3.
35. Huang M, Deng Z, Tian J, Liu T. Synthesis and biological evaluation of salinomycin triazole analogues as anticancer agents. *Eur J Med Chem* 2017;127:900–8.
36. Leng S, Qiao Q, Miao L, et al. A wash-free SNAP-tag fluorogenic probe based on the additive effects of quencher release and environmental sensitivity. *Chem Commun* 2017;53:6448–51.
37. Kindermann M, George N, Johnsson N, Johnsson K. Covalent and selective immobilization of fusion proteins. *J Am Chem Soc* 2003;125:7810–1.
38. Mattosovich R, Merlo R, Fontana A, et al. A journey down to hell: new thermostable protein-tags for biotechnology at high temperatures. *Extremophiles* 2020;24:81–91.
39. Rossi F, Morrone C, Massarotti A, et al. Crystal structure of a thermophilic O6-alkylguanine-DNA alkyltransferase-derived self-labeling protein-tag in covalent complex with a fluorescent probe. *Biochem Biophys Res Commun* 2018;500:698–703.
40. Banks JL, Beard HS, Cao Y, et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J Comput Chem* 2005;26:1752–80.
41. Toukmaji AY, Board JA. Jr., Ewald summation techniques in perspective: a survey. *Comput Phys Commun* 1996;95:73–92.
42. Zielkiewicz J. Structural properties of water: comparison of the SPC, SPCE, TIP4P, and TIP5P models of water. *J Chem Phys* 2005;123:104501.
43. Martyna GJ, Klein ML, Tuckerman M. Reversible multiple time scale molecular dynamics. *J Chem Phys* 1992;97:2635–43.

44. Evans DJ, Holian BL. The Nose–Hoover thermostat. *J Chem Phys* 1985;3:4069.
45. Miggiano R, Perugino G, Ciaramella M, et al. Crystal structure of *Mycobacterium tuberculosis* O6-methylguanine-DNA methyltransferase protein clusters assembled on to damaged DNA. *Biochem J* 2016;473:123–33.
46. Morrone C, Miggiano R, Serpe M, et al. Interdomain interactions rearrangements control the reaction steps of a thermostable DNA alkyltransferase. *Biochim Biophys Acta Gen Subj* 2017;1861:86–96.
47. Ranson M, Middleton MR, Bridgewater J, et al. Lomeguatrib, a potent inhibitor of O6-alkylguanine-DNA-alkyltransferase: phase I safety, pharmacodynamic, and pharmacokinetic trial and evaluation in combination with temozolomide in patients with advanced solid tumors. *Clin Cancer Res* 2006;12:1577–84.
48. Colombo M, Mazzucchelli S, Montenegro JM, et al. Protein oriented ligation on nanoparticles exploiting O6-alkylguanine-DNA transferase (SNAP) genetically encoded fusion. *Small* 2012;8:1492–7.
49. Song X, Wang C, Han Z, et al. Terminal alkyne substituted O6-benzylguanine for versatile and effective syntheses of fluorescent labels to genetically encoded SNAP-tags. *RSC Adv* 2015;5:23646–9.
50. Ferraris DM, Spallek R, Oehlmann W, et al. Structure of *Thermococcus litoralis*  $\Delta^1$ -pyrroline-2-carboxylate reductase in complex with NADH and L-proline. *Proteins* 2014;82:2268–74.
51. Huber W, Perspicace S, Kohler J, et al. SPR-based interaction studies with small molecular weight ligands using hAGT fusion proteins. *Anal Biochem* 2004;333:280–8.
52. Niesen J, Sack M, Seidel M, et al. SNAP-Tag technology: a useful tool to determine affinity constants and other functional parameters of novel antibody fragments. *Bioconjug Chem* 2016;27:1931–41.
53. Testa I, Wurm CA, Medda R, et al. Multicolor fluorescence nanoscopy in fixed and living cells by exciting conventional fluorophores with a single wavelength. *Biophys J* 2010;99:2686–94.
54. Miggiano R, Valenti A, Rossi F, et al. Every OGT is illuminated ... by fluorescent and synchrotron lights. *Int J Mol Sci* 2017;18:2613–31.
55. Zweerink S, Kallnik V, Ninck S, et al. Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea. *Nat Commun* 2017;8:15352.

**DNA Nanotechnology**

# Folding-upon-Repair DNA Nanoswitches for Monitoring the Activity of DNA Repair Enzymes

Nada Farag<sup>†</sup>, Rosanna Mattossovich<sup>†</sup>, Rosa Merlo, Łukasz Nierzwicki, Giulia Palermo, Alessandro Porchetta,\* Giuseppe Perugino,\* and Francesco Ricci\*

**Abstract:** We present a new class of DNA-based nanoswitches that, upon enzymatic repair, could undergo a conformational change mechanism leading to a change in fluorescent signal. Such folding-upon-repair DNA nanoswitches are synthetic DNA sequences containing *O*<sup>6</sup>-methyl-guanine (*O*<sup>6</sup>-MeG) nucleobases and labelled with a fluorophore/quencher optical pair. The nanoswitches are rationally designed so that only upon enzymatic demethylation of the *O*<sup>6</sup>-MeG nucleobases they can form stable intramolecular Hoogsteen interactions and fold into an optically active triplex DNA structure. We have first characterized the folding mechanism induced by the enzymatic repair activity through fluorescent experiments and Molecular Dynamics simulations. We then demonstrated that the folding-upon-repair DNA nanoswitches are suitable and specific substrates for different methyltransferase enzymes including the human homologue (hMGMT) and they allow the screening of novel potential methyltransferase inhibitors.

## Introduction

The genetic information inside cells is protected against DNA damage by multi-enzymatic DNA repair mechanisms, such as base excision, nucleotide excision repair,<sup>[1]</sup> or a direct damage reversal by *O*<sup>6</sup>-methylguanine-DNA-methyltransferases (abbreviated here as AGTs).<sup>[2–4]</sup> The latter is a class of evolutionarily conserved biocatalysts, able to directly and irreversibly remove alkyl groups at the *O*<sup>6</sup>-position of guanines on DNA in a single S<sub>N</sub>2-like reaction mechanism.

These enzymes, indeed, represent the major factor in contrasting the effects of alkylating agents that form such adducts.<sup>[2]</sup> On the other hand, human methyltransferase (hMGMT) activity impacts the alkylating agent-based chemotherapy in cancer cells.<sup>[2,5]</sup> For this reason, different hMGMT inactivators/inhibitors are usually employed in combination with this kind of chemotherapy.<sup>[2,6]</sup>

Because of the clinical relevance of AGTs, assays able to measure their activity in a reliable and rapid way are needed.<sup>[7]</sup> Most of the assays developed so far are based on the use of oligonucleotides carrying radioactive *O*<sup>6</sup>-methyl-guanine (*O*<sup>6</sup>-MeG) groups and chromatographic separations.<sup>[8–11]</sup> These are, however, time-consuming, tedious, and unsafe. Alternative fluorescence-based assays that allow simple and fast detection of hMGMT activity have been recently developed. For example, Kool and co-workers designed a chemosensor that couples fluorescence change to the bond-breaking step occurring during repair activity.<sup>[12,13]</sup> Other approaches employing optically-labelled DNA aptamers,<sup>[11]</sup> DNA-based electrochemical sensors<sup>[14]</sup> or *O*<sup>6</sup>-MeG-containing double strands DNA (dsDNA) oligonucleotides in competition with fluorescent substrates<sup>[4,15–18]</sup> were also proposed. While these latter systems provide several advantages including ease of use and high sensitivity, there is still an urgent need for finding new strategies to achieve efficient activity-based monitoring of DNA repair enzymes that can be versatile enough to be suitable for a wide range of repair activities.

Recently, DNA nanoswitches have emerged as a new class of programmable probes that allow the sensitive and rapid detection of a wide range of molecular targets.<sup>[19]</sup> DNA nanoswitches usually undergo a binding-induced conformational change that can provide a measurable output in the presence of a specific input. A number of strategies employing DNA nanoswitches have been developed for the detection of different targets including pH,<sup>[20,21]</sup> metal ions,<sup>[22]</sup> small molecules,<sup>[23,24]</sup> proteins,<sup>[25,26]</sup> or specific antibodies.<sup>[27,28]</sup> Critically, because their signaling is linked to a change in the physics of the DNA probe induced by the recognition with a target, such conformational switching sensors are generally highly specific and selective.<sup>[29]</sup>

Motivated by the above considerations, we demonstrate here the rational design of a new class of DNA nanoswitches, here named folding-upon-repair DNA nanoswitches, that can be conveniently applied for monitoring DNA repair activity. As an applicative example, we initially focused on methyltransferase enzymes and designed a fluorescent-labelled DNA nanoswitch containing *O*<sup>6</sup>-MeG that, upon enzymatic

[\*] N. Farag,<sup>[†]</sup> Dr. A. Porchetta, Prof. F. Ricci  
Department of Chemistry, University of Rome Tor Vergata  
Via della Ricerca Scientifica, 00133 Rome (Italy)  
E-mail: alessandro.porchetta@uniroma2.it  
francesco.ricci@uniroma2.it

R. Mattossovich,<sup>[†]</sup> R. Merlo, Dr. G. Perugino  
Institute of Biosciences and BioResources  
National Research Council of Italy  
Via Pietro Castellino 111, 80131 Naples (Italy)  
E-mail: giuseppe.perugino@ibbr.cnr.it

Dr. Ł. Nierzwicki, Prof. G. Palermo  
Department of Bioengineering, University of California Riverside  
900 University Avenue, Riverside, CA 92512 (USA)

Prof. G. Palermo  
Department of Chemistry, University of California Riverside  
900 University Avenue, Riverside, CA 92512 (USA)

[†] These authors contributed equally to this work.

 Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:  
 <https://doi.org/10.1002/anie.202016223>.

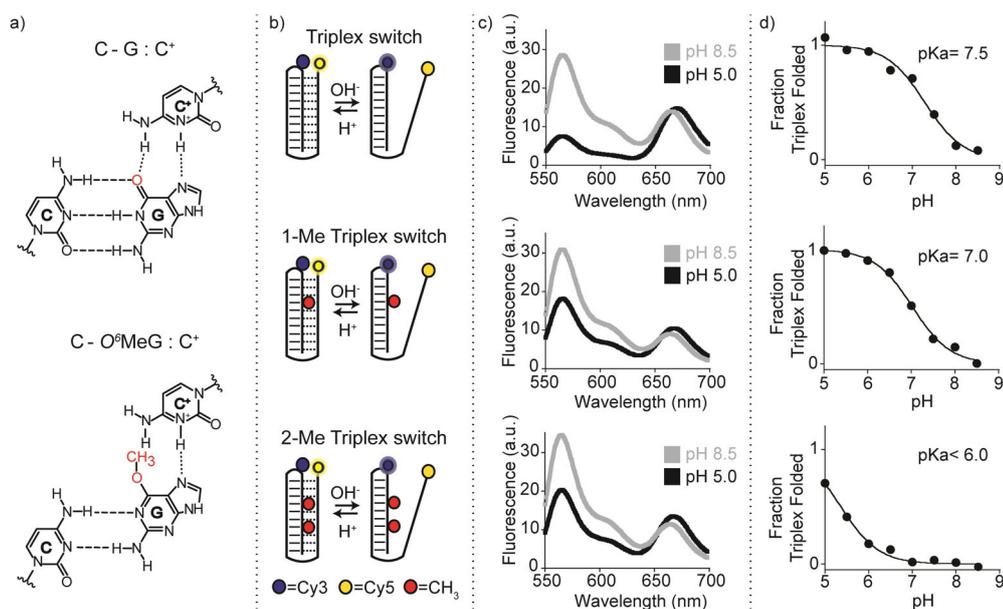
repair, could undergo a conformational change associated with a change in the output signal. This approach allows the direct measurement of the activity of, in principle, any methyltransferase enzyme thus opening the possibility to develop an easy methodology for the high-throughput determination of methyltransferase enzyme inhibitors.

## Results and Discussion

Our strategy to design programmable nucleic acid nanoswitches for the detection of methyltransferase activity is based on the use of a single-stranded DNA capable of forming an intramolecular triplex structure through hydrogen bonds (Hoogsteen interactions) between a hairpin duplex domain and a single-strand triplex-forming portion.<sup>[30]</sup> More specifically, we have designed three nanoswitches displaying the same triplex-forming domains (i.e., 4 cytosines + 6 thymines) but differing in the content of *O*<sup>6</sup>-MeGs in the hairpin duplex domain (0, 1, and 2). The idea underlying the molecular design of the triplex nanoswitches is that the methyl group at the guanine *O*<sup>6</sup> position should affect the formation of the Hoogsteen hydrogen bond with the cytosine on the single-strand DNA (Figure 1 a), thus influencing the pH-dependent triplex folding/unfolding behaviour of the nanoswitch. The nanoswitches are also labelled with a pH-insensitive FRET

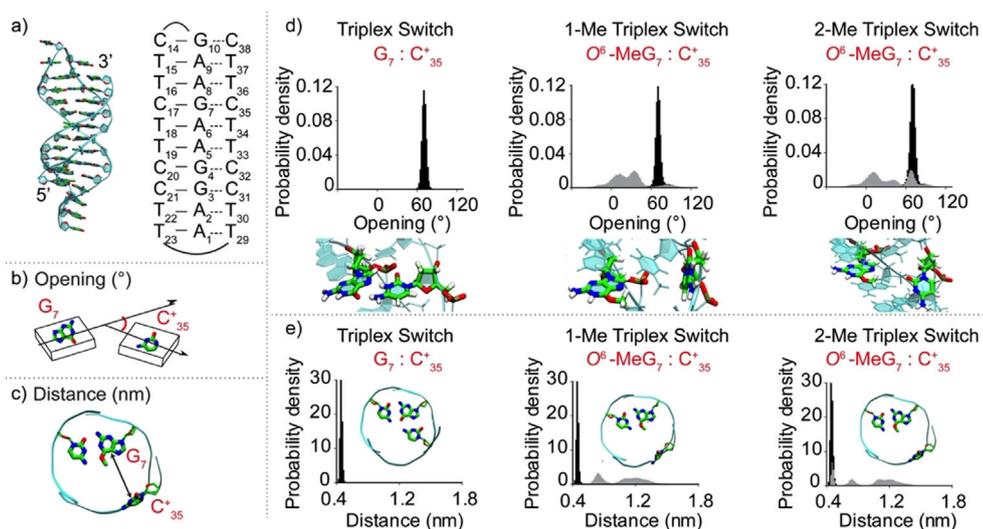
pair to follow pH-dependent folding/unfolding (Figure 1 b). To demonstrate the effect of *O*<sup>6</sup>-MeG on the triplex formation, we have first tested the pH-dependent folding/unfolding behaviour of our three nanoswitches by measuring the fluorescence signal of our FRET pair at a fixed concentration of the nanoswitch and different pH values. As expected, our control triplex nanoswitch (Triplex switch, Figure 1 b, top) that lacks any *O*<sup>6</sup>-MeG in its sequence, shows signals that are consistent with the formation of a folded triplex at acidic pH (pH 5.0) and suggest the unfolding of the triplex structure at basic pH values (pH 8.5). The pH of semi-protonation (here defined as  $pK_a$ , the average  $pK_a$  due to several interacting protonation sites) for this triplex nanoswitch is 7.5 (Figure 1 c,d top, and S1). The presence of one *O*<sup>6</sup>-MeG (1-Me Triplex switch, Figure 1 b, center) in the hairpin duplex (at position 7) strongly destabilizes triplex formation; as a result, we observe at pH 5.0 a FRET signal that is consistent with a partially unfolded triplex structure and we obtain a  $pK_a$  of 7.0 (Figures 1 c,d center, and S1). Finally, the nanoswitch containing two *O*<sup>6</sup>-MeG (2-Me Triplex switch, at position 4 and 7, Figure 1 b, bottom) is even more destabilized and the pH titration curve with this switch does not reach a plateau at lower pH values suggesting a  $pK_a$  lower than 6.0 (Figures 1 c,d bottom, and S1). The formation of a folded triplex structure of the Triplex switch was confirmed by melting and urea denaturation experiments performed at different pHs (Figures S2, S3). Also, CD spectra of the three nanoswitches further support the formation of pH-dependent Hoogsteen interactions (Figure S4).

To better understand the effect that guanine methylation has on the stability of our triplex nanoswitches we performed overall molecular dynamics (MD) simulations. We first considered the non-methylated nanoswitch (Triplex switch) where all cytosines involved in Hoogsteen base pairing are protonated (Figure 2 a). Then, we introduced the *O*<sup>6</sup>-MeG at position 7 (creating the 1-Me Triplex switch), and at both positions 4 and 7 (2-Me Triplex switch). The obtained trajectories encompassed > 4.5  $\mu$ s of sampling, including multiple replicas and providing solid statistics for the analysis of the DNA triplex structural changes. We focused on the



**Figure 1.** Folding-upon-repair DNA nanoswitches. a) DNA parallel triplets formed between two cytosines and one guanine and involving Watson–Crick (three hydrogen bonds) and Hoogsteen interactions (two hydrogen bonds) require the protonation of the N<sup>3</sup> of cytosine in the third strand (top, left) and thus are only stable at acidic pH values (average  $pK_a$  of protonated cytosines in C-G:C triplet is  $\approx$ 6.5). Because it affects Hoogsteen base pairing efficiency, methylation of the guanine in the *O*<sup>6</sup>-position destabilizes the triplex conformation. b) Programmable DNA-based triplex nanoswitches designed to form an intramolecular triplex structure with 0 (Triplex), 1 (1-Me Triplex) or 2 (2-Me Triplex) *O*<sup>6</sup>-MeG in the sequence. c) Fluorescence spectra were obtained for each nanoswitch at pH 5.0 and pH 8.5. d) pH-titration curves of the triplex nanoswitches. Triplex-to-duplex transition is monitored through a pH-insensitive FRET pair at the 3'-end (Cy5) and internally located (Cy3). The pH titration experiments were performed at 25 °C, [nanoswitch] = 50 nM by measuring the fluorescence signal at different pH values in 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 150 mM NaCl buffer. Spectra were obtained by excitation at 530 ( $\pm$  5) nm and acquisition between 545 and 700 nm ( $\pm$  10) nm.





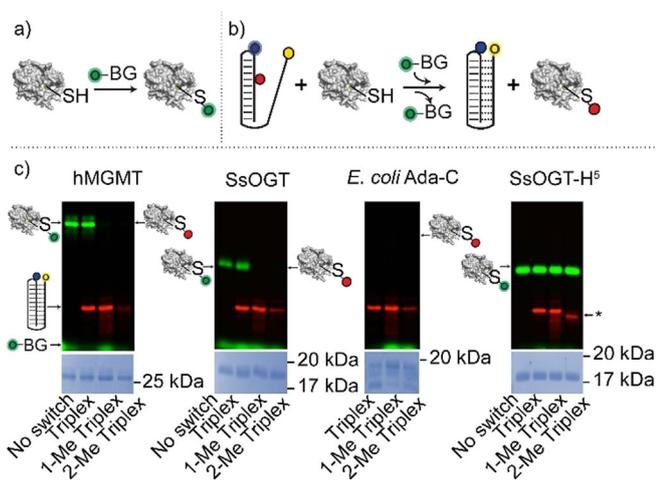
**Figure 2.** Molecular dynamics simulations of triplex nanoswitches. a) Structure (left) and schematic representation (right) of the Triplex switch with numeration of DNA nucleotides. Graphical representation of b) opening angle and c) distance parameters for the  $G_7:C_{35}^+$  pair. d) Opening angle distribution for the Triplex switch (left), 1-Me Triplex switch (center), and 2-Me Triplex switch (right). Below each graph, side-view snapshots of the analysed base pairs are reported. The opening distribution for 1Me- and 2Me- Triplex switches are shown in grey. The distribution of the non-methylated Triplex switch (black) is also reported in each graph as a comparison. e) Distance parameter for the  $G_7:C_{35}^+$  Hoogsteen interaction for the Triplex switch, 1-Me Triplex switch, and 2-Me Triplex switch. In each graph, a top-view snapshot of the triplet most probable configuration is shown. The distance distribution for 1Me- and 2Me- triplex switches is shown in grey. The distribution of the non-methylated Triplex switch (black) is also shown in each graph as a comparison.

$C_{17}-G_7:C_{35}^+$  triplet, as this triplet share the methylated guanine in both the 1-Me and 2-Me Triplex switches. Our analysis took into consideration two parameters: 1) the base-pair opening angle, which describes the in-plane opening between  $G_7:C_{35}^+$  paired Hoogsteen-forming bases, as defined by Lu and Olson<sup>[31]</sup> (Figure 2b; see Tables S1–S4 for all DNA base-pair parameters); 2) the distance between the centers of mass of the atoms of the  $G_7:C_{35}^+$  Hoogsteen-forming bases at the interface with each other (Figure 2c).

The narrow distributions of both the opening and distance parameters for the  $C_{17}-G_7:C_{35}^+$  triplet in the Triplex switch (with an average value of  $73^\circ$  and 0.45 nm, respectively) correspond to a well-defined triplex structure (Figure 2d,e, left and Table S1). Conversely, in the case of the 1-Me and 2-Me Triplex switches, distributions of both parameters are broader and have multiple maxima, showing that the same triplet adopts multiple conformational states (Figure 2d,e, center, right). Remarkably, in case of 1-Me and 2-Me Triplex switches the distance distribution shifts towards higher values. This clearly shows that the  $O^6$ -methylation of guanine hampers the efficient interaction with the Hoogsteen pairing cytosine, leading to a local unwinding of the triplex (see representative structures, Figure 2d, and highlights on Figure 2e, center, right and Figures S5–8). Similar shifts and broadening of opening and distance distributions were also observed in the case of  $C_{20}-G_4:C_{32}^+$  triplet in 2Me-Triplex switch, whereas in the case of 1Me-Triplex switch, these distributions were virtually the same as in case of non-methylated Triplex switch (Figures S9–11). Although at a lower extent, we also observed the broadening of the opening and distance distributions for the  $C_{17}-G_7:C_{35}^+$  triplet during addi-

tional  $\approx 1.5 \mu\text{s}$  long simulations of the Triplex switch in the deprotonated form (Figure S12), further supporting the notion that methylation strongly destabilizes the formation of the triplet. Finally, to achieve more meaningful structural information on the triplex destabilization induced by the methylation, we obtained free energy profiles that can be associated with the increase of the distance between the two fluorophores on each switch. As expected, under protonated conditions the control Triplex is the most stable, followed by the 1-Me Triplex and then by the 2-Me Triplex (Figure S13). These results provide an atomic-level understanding of the triplex stability, indicating that the efficient Hoogsteen base pairing is critical for the triplex formation and both cytosine deprotonation at high pH and guanine methylation should reduce the propensity of DNA hairpin to adopt triplex conformation.

Because the methyl group at the guanine  $O^6$  position prevents the efficient formation of Hoogsteen interactions with the respective cytosine in the triplex-forming domain, the  $O^6$ -MeG-containing triplex nanoswitches characterized above could be used as suitable probes for the monitoring of methyltransferase activity. The enzymatic removal of the methyl group of  $O^6$ -MeG in the nanoswitches would restore the optimal conditions for Hoogsteen interactions and triplex formation by the nanoswitch. To achieve this goal, we initially performed an indirect assay using a fluorescein derivative of  $O^6$ -benzyl-guanine ( $O^6$ -BG)<sup>[4,15,32,33]</sup> (Figure 3a) to demonstrate that  $O^6$ -MeG-containing triplex nanoswitches are suitable substrates for methyltransferase enzymes (Figure 3b). We first demonstrated that no interaction occurs between the nanoswitch and the fluorescent-labelled substrate in the absence of the protein. As expected, visible separated bands for both the  $O^6$ -BG fluorescent derivative and the three DNA nanoswitches tested can be observed (Figure S14). It is worth noting that the nanoswitch containing two  $O^6$ -MeGs (2-Me Triplex) shows a slightly, but significant, higher mobility (marked by an asterisk in Figure 3c) compared to the single-methylated triplex (1-Me Triplex) and the non-methylated triplex (Triplex) (Figure S14). Although speculative, a possible explanation of the different migration of the 2-Me Triplex could be that this oligonucleotide adopts a partial folding in denaturing conditions. We have pre-incubated the methyltransferase enzymes with our nanoswitches and then added an equimolar amount of the  $O^6$ -BG

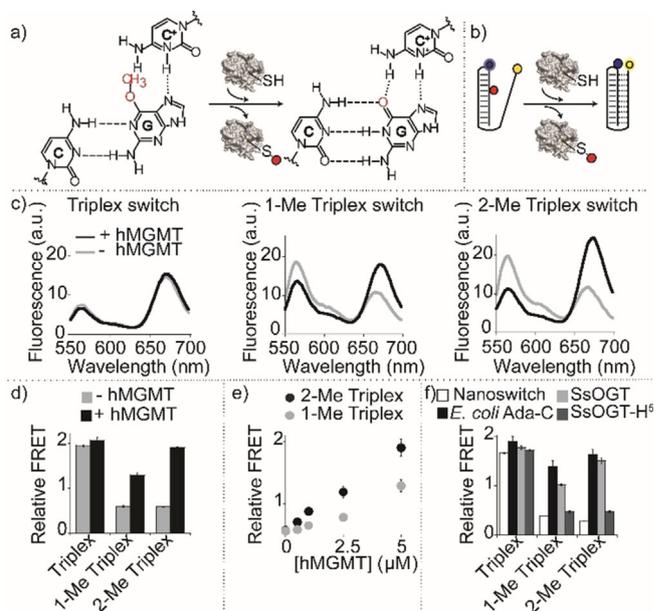


**Figure 3.** Characterization of triplex nanoswitches as suitable substrates of methyltransferase enzymes. a) Suicide reaction of an AGT enzyme with a fluorescein derivative of  $O^6$ -BG. b) The alkyl transfer suicide reaction between the enzyme and a DNA methylated triplex hampers the covalent bond between the enzyme and the fluorescein group. c) SDS-PAGE images of reactions in which different enzymes (i.e. hMGMT, SsOGT, *E. coli* Ada-C, SsOGT-H<sup>5</sup>) after pre-incubation with the DNA nanoswitches and then reacted with the  $O^6$ -BG fluorescent derivative. The enzymes that have reacted with the fluorescein derivative appear as green bands on the gel. The absence of any fluorescent bands indicates that the enzyme has not reacted with the fluorescent substrate. Fluorescence-labelled triplex nanoswitches appear as red bands. The asterisk shows the higher mobility of the 2-Me Triplex nanoswitch compared to the other Triplex switches (see also Figure S14). The experiments were performed by pre-incubating for 60 min the different AGT enzymes with the relevant triplex nanoswitches and then adding the  $O^6$ -BG fluorescent derivative. Samples were then loaded on 15% acrylamide SDS-PAGE. Gels underwent Coomassie staining for the determination and correction of the protein amount loaded.

fluorescent derivative that interacts with the free enzymes in the solution (Figure 3 b). In the case of hMGMT, a fluorescent protein band was observed in the presence of the non-methylated Triplex, suggesting that no repair reaction on this DNA occurred (Figure 3c). On the contrary, 1-Me Triplex nanoswitch and 2-Me Triplex nanoswitch, showed a complete absence of the fluorescent band (Figure 3c). This is likely due to the suicidal nature of methyltransferase enzymes,<sup>[3]</sup> as the methyl group is irreversibly transferred to the enzyme, which is no longer available for the  $O^6$ -BG fluorescent derivative, thus demonstrating that methylated nanoswitches are effective substrates for hMGMT. The same results were further confirmed using the thermostable *Saccharolobus solfataricus* homologue (SsOGT), an enzyme responsible for the direct repair of  $O^6$ -alkylguanine in double-stranded DNA at high temperatures<sup>[4,16,32,33]</sup> and the *Escherichia coli* homologue (Ada-C), which is reported to be insensitive to  $O^6$ -BG derivatives.<sup>[17,18]</sup> In the latter case, the absence of the enzyme fluorescent bands despite being correctly loaded as confirmed by the Coomassie staining analysis (Figure 3c, *E. coli* Ada-c), and the similar migration of all the triplexes, confirm that the DNA-repair occurred. The methylated DNA triplex nanoswitches are also highly specific as they showed no enzymatic activity when incubated with a mutant of SsOGT (SsOGT-H<sup>5</sup>)

that was previously reported to be catalytically active on  $O^6$ -BG derivatives, but unable to bind and react with ds-DNA.<sup>[4,32–35]</sup>

Prompted by the results described above, we tested whether the methyltransferase repair activity on our nano-switches could result in a conformational switch, and a consequent measurable FRET signal change, providing a means of direct detection of enzymatic activity (Figure 4a,b). Initially, we focused on the detection of the activity of hMGMT; as expected, no significant difference in the FRET signal of the control nanoswitch (non-methylated Triplex) before ( $1.9 \pm 0.1$ ) and after ( $2.1 \pm 0.1$ ) hMGMT incubation can be observed (Figure 4c, left). Under the experimental conditions used, both the fluorescence spectra suggest that the non-methylated triplex switch is completely folded (at pH 5.0). The same experiment carried out using the methylated triplex nanoswitches shows, instead, a strong difference in FRET signals before and after hMGMT incubation (Figure 4c, center, right). For both nanoswitches, the FRET signals observed before hMGMT incubation confirm a partially unfolded configuration (1-Me Triplex,  $0.6 \pm 0.1$ , 2-Me



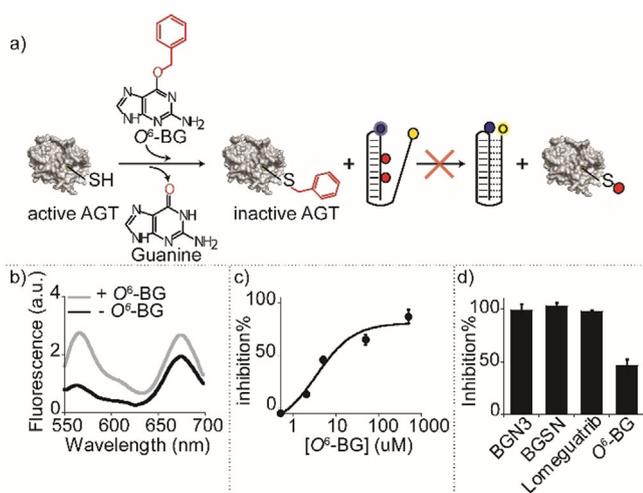
**Figure 4.** Detection of methyltransferase activity with triplex nanoswitches. a) Methyltransferase enzymatic activity on DNA methylated nanoswitches leads to folding-upon-repair of the triplex DNA structure. b) Enzymatic activity detection can be achieved by monitoring folding/unfolding of the Triplex nanoswitch by fluorescence FRET signaling. c) Spectra and d) relative FRET signals obtained with the Triplex nanoswitches before and after incubation with 5  $\mu$ M ( $0.1 \mu\text{g} \mu\text{L}^{-1}$ ) hMGMT. e) Relative FRET signals observed at different hMGMT concentrations for 1-Me and 2-Me Triplex nanoswitches. f) Relative FRET signals with different methyltransferase enzymes. Spectra were obtained by excitation at 530 ( $\pm 5$ ) nm and acquisition between 545 and 700 nm ( $\pm 10$ ) nm. Relative FRET signals were obtained by first incubating the DNA nanoswitches (0.5  $\mu$ M) in the absence or presence of AGTs (0.1  $\mu\text{g} \mu\text{L}^{-1}$ ) in 10  $\mu$ L solution of 50 mM  $\text{Na}_2\text{HPO}_4$  buffer, 150 mM NaCl, pH 7.5 at 30°C for 60 minutes. The reaction mixtures were then diluted to 100  $\mu$ L using 50 mM  $\text{Na}_2\text{HPO}_4$  buffer, 250 mM NaCl at pH 5.0, and heat-inactivated for 2 minutes at 70°C before performing the fluorescence spectra at 25°C.



Triplex,  $0.6 \pm 0.1$ , Figure 4d). Upon hMGMT incubation, the enzymatic demethylation of  $O^6$ -MeG in the triplex nanoswitches restores their ability to form a triplex structure. This is particularly clear for 2-Me triplex nanoswitch that shows a FRET signal ( $1.9 \pm 0.1$ ) after hMGMT incubation that is within the error of the control non-methylated nanoswitch ( $2.1 \pm 0.1$ ) (Figure 4d). Of note, the folding dynamics of these switches are extremely rapid ( $K_{\text{folding}}$  and  $K_{\text{unfolding}} = 10 \text{ s}^{-1}$  and  $2 \text{ s}^{-1}$ , respectively)<sup>[36]</sup> and thus the rate-determining step in these measurements is given by the enzyme-catalyzed reaction. We found out that with hMGMT a saturation of signal is observed after 15 minutes of enzymatic reaction (Figure S15). Both methylated triplex nanoswitches show a change in the relative FRET values that are linearly dependent on the concentration of hMGMT in the range between 0.5 and 5  $\mu\text{M}$  ( $R^2_{\text{2-Me Triplex}} = 0.99$ ,  $R^2_{\text{1-Me Triplex}} = 0.95$ ) (Figure 4e). Similar experiments were also performed with other methyltransferase enzymes, both methylated triplex nanoswitches upon incubation with SsOGT and *E. coli* Ada-C enzymes gave FRET signal changes consistent with triplex folding suggesting efficient enzymatic activity (Figure 4f). The same experiment performed with SsOGT-H<sup>2</sup> (a DNA repair defective mutant, but catalytically active) produced no effect on the FRET signal of both nanoswitches (Figure 4f) once again confirming the specificity of the folding-upon-repair mechanism. To demonstrate the possibility of using this platform in

more complex media, we have also performed the measurement of the enzymatic activity of hMGMT in 10% serum and observed FRET signal changes well distinguishable from the control experiment in absence of enzyme (Figure S16).

To demonstrate the utility of our platform for the study and characterization of new hMGMT inhibitors as possible drug candidates,<sup>[37]</sup> we used our nanoswitches to measure the activity of the human enzyme in the presence of inhibitors, inactivators, or pseudo-substrates. The addition of an hMGMT inhibitor in the reaction mixture should prevent the enzyme from repairing the methylated DNA, and no significant change of the nanoswitch FRET signal should be observed (Figure 5a). We initially tested  $O^6$ -BG, a widely characterized methyltransferase inactivator.<sup>[38,39]</sup> The relative FRET signal of the nanoswitch incubated with hMGMT and  $O^6$ -BG (Figure 5b, grey line) is, at  $0.8 \pm 0.1$ , similar to the values observed in the absence of enzyme (Figure 5b, black line) an effect consistent with the inactivation of hMGMT activity by  $O^6$ -BG. The measured inactivation, as expected, follows a concentration-dependent behaviour with a measured  $\text{IC}_{50} = 3.5 \pm 2 \mu\text{M}$  (Figure 5c). Moreover, our platform is able to measure the inactivation efficiency of different molecules including 4-azido-*N*-(4-(hydroxymethyl) benzyl) butanamide (BGN3), (*N*-(4-(((2-amino-9*H*-purin-6-yl)oxy)methyl)benzyl)-4-azidobutanamide (BGSN), Lomeguatrib,<sup>[40,41]</sup>  $O^6$ -Benzylguanine ( $O^6$ -BG),<sup>[42]</sup> demonstrating once again its versatility (Figure 5d).



**Figure 5.** Detection of methyltransferase inhibitors. a) Pre-incubation of the enzyme with an inhibitor (here  $O^6$ -BG) leads to irreversible inhibition of the enzyme that is thus not able to demethylate the DNA nanoswitch. b) Spectra obtained with the 2-Me Triplex nanoswitch in the presence of hMGMT (5  $\mu\text{M}$ ) before and after incubation with the inhibitor  $O^6$ -BG (5  $\mu\text{M}$ ). c) % Inhibition plot obtained at different concentrations of  $O^6$ -BG inhibitor. d) % Inhibition obtained with the 2-Me Triplex nanoswitch with an equimolar concentration of different enzymatic inhibitors and hMGMT (5  $\mu\text{M}$ ). Spectra were obtained by excitation at  $530(\pm 5)$  nm and acquisition between 545 and 700 nm ( $\pm 10$ ) nm. Relative FRET signals were obtained by enzymatic incubation of 5  $\mu\text{M}$  hMGMT and inhibitor at  $30^\circ\text{C}$  for 60 minutes in 10  $\mu\text{L}$  solution of 50 mM  $\text{Na}_2\text{HPO}_4$  buffer, 150 mM NaCl at pH 7.5. 2-Me Triplex nanoswitch (0.5  $\mu\text{M}$ ) was then added to the reaction mixtures and incubated for another 60 min at  $30^\circ\text{C}$ . The reaction mixtures were diluted to 100  $\mu\text{L}$  using 50 mM  $\text{Na}_2\text{HPO}_4$  buffer, 250 mM NaCl at pH 5.0, and heat-inactivated for 2 minutes at  $70^\circ\text{C}$  before performing the fluorescence spectra at  $25^\circ\text{C}$ .

## Conclusion

Here, we have designed folding-upon-repair DNA nanoswitches containing  $O^6$ -MeG nucleobases that undergo a conformational switch from a duplex to a triplex conformation upon enzymatic demethylation. We have first demonstrated that the presence of a methyl group at the guanine  $O^6$  position in the duplex portion of the nanoswitch strongly affects the triplex formation. Molecular simulations were used to determine how the  $O^6$ -methylation of guanine bases prevents the efficient formation of Hoogsteen interactions with the cytosine in the triplex-forming domain, resulting in a local unwinding of the triplex structure. Such methylated triplex-based nanoswitches are versatile tools for the direct measurement of methyltransferase activity as they form the natural enzymatic substrate duplex conformation when methylated and, upon enzymatic demethylation, they can fold into a measurable triplex structure. We showed that the methylated nanoswitches were efficiently recognized by these enzymes leading to measurable FRET signal changes upon repair activity. Finally, we have also measured hMGMT activity in the presence of several enzyme inactivators, demonstrating the possibility of using in a high-throughput system our nanoswitches for the screening of novel potential inactivators/pseudo-substrates/inhibitors of hMGMT.<sup>[37]</sup>

Our folding-upon-repair DNA nanoswitches are reagent-less, highly specific, and versatile. These features make our proposed approach convenient and easy to perform. Although the current experimental protocol requires a change in the pH of the solution after the enzymatic reaction making



it less amenable to automatization, we believe that the principle we have described together with the ease with which these probes are designed make the approach suitable to develop a suite of activity-based DNA nanoswitches for other DNA repair enzymes.

### Acknowledgements

This work was supported by the European Research Council, ERC (project n.336493) (FR), by Associazione Italiana per la Ricerca sul Cancro, AIRC (project n. 14420) (FR), by the Italian Ministry of Health (project n. GR-2010-2317212) and by the Italian Ministry of University and Research (Project of National Interest, PRIN, 2017YER72K). Computational work has been supported by the National Science Foundation under Grant No. CHE-1905374 and by the National Institute of Health under Grant No. R01-EY027440 (GP). Computer time for MD simulations has been awarded by XSEDE via the Grant No. TG-MCB160059 (GP). We would like to thank Mateusz Kogut for valuable suggestions regarding DNA triplex folding simulations. F.R., G.P., and A.P. would like to thank all the authors for their efforts in writing this work, technical assistance, but mainly for human support during the difficult and delicate period of stay-at-home following the COVID-19 outbreak.

### Conflict of interest

The authors declare no conflict of interest.

**Keywords:** conformational change mechanism · DNA nanoswitches · DNA nanotechnology · DNA repair enzymes · triplex DNA

- [1] E. C. E. T. Madders, J. L. Parsons in *Advances in Experimental Medicine and Biology*, Vol. 1241 (Eds.: W. E. Crusio, J. D. Lambris, H. H. Radeke, N. Rezaei), Springer, Cham, **2020**, pp. 59–75.
- [2] S. L. Gerson, *Nat. Rev. Cancer* **2004**, *4*, 296–307.
- [3] A. E. Pegg, *Chem. Res. Toxicol.* **2011**, *24*, 618–639.
- [4] R. Mattosovich, R. Merlo, R. Miggiano, A. Valenti, G. Perugino, *Int. J. Mol. Sci.* **2020**, *21*, 2878.
- [5] D. Fu, J. A. Calvo, L. D. Samson, *Nat. Rev. Cancer* **2012**, *12*, 104–120.
- [6] S. L. Gerson, *J. Clin. Oncol.* **2002**, *20*, 2388–2399.
- [7] Z. D. Nagel, A. A. Beharry, P. Mazzucato, G. J. Kitange, J. N. Sarkaria, E. T. Kool, L. D. Samson, *PLoS One* **2019**, *14*, e0208341.
- [8] S. Hurst-Calderone, K. W. Kohn, *Cancer Res.* **1987**, *47*, 6229–6235.
- [9] S. Klein, F. Oesch, *Anal. Biochem.* **1992**, *205*, 294–299.
- [10] K. X. Luu, S. Kanugula, A. E. Pegg, G. T. Pauly, R. C. Moschel, *Biochemistry* **2002**, *41*, 8689–8697.
- [11] M. Tintoré, A. Aviñó, F. M. Ruiz, R. Eritja, C. Fàbrega, *J. Nucleic Acids* **2010**, *2010*, 632041.
- [12] A. A. Beharry, Z. D. Nagel, L. D. Samson, E. T. Kool, *PLoS One* **2016**, *11*, e0152684.
- [13] D. L. Wilson, A. A. Beharry, A. Srivastava, T. R. O'Connor, E. T. Kool, *Angew. Chem. Int. Ed.* **2018**, *57*, 12896–12900; *Angew. Chem.* **2018**, *130*, 13078–13082.
- [14] S. Chen, J. Su, Z. Zhao, Y. Shao, Y. Dou, F. Li, W. Deng, J. Shi, Q. Li, X. Zuo, S. Song, C. Fan, *Nano Lett.* **2020**, *20*, 7028–7035.
- [15] A. Vettone, M. Serpe, A. Hidalgo, J. Berenguer, G. del Monaco, A. Valenti, M. Rossi, M. Ciaramella, G. Perugino, *Extremophiles* **2016**, *20*, 1–13.
- [16] C. Morrone, R. Miggiano, M. Serpe, A. Massarotti, A. Valenti, G. del Monaco, M. Rossi, F. Rossi, M. Rizzi, G. Perugino, M. Ciaramella, *Biochim. Biophys. Acta* **2017**, *1861*, 86–96.
- [17] R. H. Elder, G. P. Margison, J. A. Rafferty, *Biochem. J.* **1994**, *298*, 231–235.
- [18] K. Goodtzova, S. Kanugula, S. Edara, G. T. Pauly, R. C. Moschel, A. E. Pegg, *J. Biol. Chem.* **1997**, *272*, 8332–8339.
- [19] S. G. Harroun, C. Prévost-Tremblay, D. Lauzon, A. Desrosiers, X. Wang, L. Pedro, A. Vallée-Bélisle, *Nanoscale* **2018**, *10*, 4607–4641.
- [20] S. Modi, M. G. Swetha, D. Goswami, G. D. Gupta, S. Mayor, Y. Krishnan, *Nat. Nanotechnol.* **2009**, *4*, 325–330.
- [21] T. Patino, A. Porchetta, A. Jannasch, A. Lladó, T. Stumpp, E. Schäffer, F. Ricci, S. Sánchez, *Nano Lett.* **2019**, *19*, 3440–3447.
- [22] A. Porchetta, A. Vallée-Bélisle, K. W. Plaxco, F. Ricci, *J. Am. Chem. Soc.* **2013**, *135*, 13238–13241.
- [23] M. Hansen-Bruhn, L. D. F. Nielsen, K. V. Gothelf, *ACS Sens.* **2018**, *3*, 1706–1711.
- [24] M. Rossetti, R. Ippodrino, B. Marini, G. Palleschi, A. Porchetta, *Anal. Chem.* **2018**, *90*, 8196–8201.
- [25] L. Liu, H.-C. Wu, *Angew. Chem. Int. Ed.* **2016**, *55*, 15216–15222; *Angew. Chem.* **2016**, *128*, 15440–15446.
- [26] N. A. W. Bell, U. F. Keyser, *J. Am. Chem. Soc.* **2015**, *137*, 2035–2041.
- [27] S. Ranallo, M. Rossetti, K. W. Plaxco, A. Vallée-Bélisle, F. Ricci, *Angew. Chem. Int. Ed.* **2015**, *54*, 13214–13218; *Angew. Chem.* **2015**, *127*, 13412–13416.
- [28] R. Arts, S. K. J. Ludwig, B. C. B. Van Gerven, E. M. Estirado, L. G. Milroy, M. Merckx, *ACS Sens.* **2017**, *2*, 1730–1736.
- [29] A. Vallée-Bélisle, F. Ricci, K. W. Plaxco, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13802–13807.
- [30] A. Idili, A. Vallée-Bélisle, F. Ricci, *J. Am. Chem. Soc.* **2014**, *136*, 5836–5839.
- [31] X. J. Lu, W. K. Olson, *Nucleic Acids Res.* **2003**, *31*, 5108–5121.
- [32] G. Perugino, A. Vettone, G. Illiano, A. Valenti, M. C. Ferrara, M. Rossi, M. Ciaramella, *J. Biol. Chem.* **2012**, *287*, 4222–4231.
- [33] G. Perugino, R. Miggiano, M. Serpe, A. Vettone, A. Valenti, S. Lahiri, F. Rossi, M. Rossi, M. Rizzi, M. Ciaramella, *Nucleic Acids Res.* **2015**, *43*, 8801–8816.
- [34] F. Rossi, C. Morrone, A. Massarotti, D. M. Ferraris, A. Valenti, G. Perugino, R. Miggiano, *Biochem. Biophys. Res. Commun.* **2018**, *500*, 698–703.
- [35] R. Merlo, S. Del Prete, A. Valenti, R. Mattosovich, V. Carginale, C. T. Supuran, C. Capasso, G. Perugino, *J. Enzyme Inhib. Med. Chem.* **2019**, *34*, 490–499.
- [36] D. Mariottini, A. Idili, M. A. D. Nijenhuis, G. Ercolani, F. Ricci, *J. Am. Chem. Soc.* **2019**, *141*, 11367–11371.
- [37] B. Kaina, M. Christmann, *DNA Repair* **2019**, *78*, 128–141.
- [38] J. A. Quinn, S. X. Jiang, D. A. Reardon, A. Desjardins, J. J. Vredenburgh, J. N. Rich, S. Gururangan, A. H. Friedman, D. D. Signer, J. H. Sampson, R. E. McLendon, J. E. Herndon, A. Walker, H. S. Friedman, *J. Clin. Oncol.* **2009**, *27*, 1262–1267.
- [39] J. Reinhard, U. Eichhorn, M. Wiessler, B. Kaina, *Int. J. Cancer* **2001**, *93*, 373–379.
- [40] M. R. Middleton, J. Kelly, N. Thatcher, D. J. Donnelly, R. S. McElhinney, T. B. H. McMurry, J. E. McCormick, G. P. Margison, *Int. J. Cancer* **2000**, *85*, 248–252.
- [41] M. Ranson, M. R. Middleton, J. Bridgewater, S. M. Lee, M. Dawson, D. Jowle, G. Halbert, S. Waller, H. McGrath, L.



Gumbrell, R. S. McElhinney, D. Donnelly, T. B. H. McMurry, G. P. Margison, *Clin. Cancer Res.* **2006**, *12*, 1577–1584.  
[42] R. Merlo, D. Caprioglio, M. Cillo, A. Valenti, R. Mattosovich, C. Morrone, A. Massarotti, F. Rossi, R. Miggiano, A. Leonardi,

A. Minassi, G. Perugino, *J. Enzyme Inhib. Med. Chem.* **2021**, *36*, 85–97.

Manuscript received: December 7, 2020  
Accepted manuscript online: January 8, 2020  
Version of record online: ■ ■ ■ ■, ■ ■ ■ ■

---



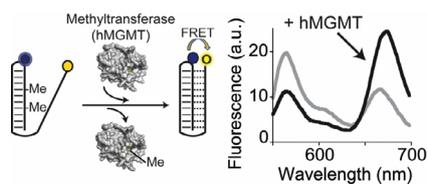
## Research Articles



## DNA Nanotechnology

N. Farag, R. Mattosovich, R. Merlo,  
Ł. Nierzwicki, G. Palermo, A. Porchetta,\*  
G. Perugino,\* F. Ricci\* — ■■■■—■■■■

Folding-upon-Repair DNA Nanoswitches  
for Monitoring the Activity of DNA Repair  
Enzymes



We present a new class of DNA-based nanoswitches, named folding-upon-repair DNA nanoswitches, that upon enzymatic repair could undergo a conformational change mechanism leading to a change in fluorescent signal. The nanoswitches are suitable substrates for different methyltransferases and allow the screening of novel potential methyltransferase inhibitors.

# Book Publisher International



F. No. SDI/BP/8226D/4267

Dated 30-December-2020

To  
**Rosa Merlo, Rosanna Mattosovich, Anna Valenti, and Giuseppe Perugino.**

*Institute of Biosciences and BioResources, National Research Council of Italy, Via Pietro Castellino 111, 80131 Naples, Italy.*

**Subject: Acceptance letter for manuscript (2020/BP/8226D) as a book chapter of Recent Research Advances in Biology**

Dear Dr. Giuseppe Perugino,

We are pleased to inform that your manuscript (Ref. no. **2020/BP/8226D**) entitled "**Recent advances in the use of the SNAP-tag® in the modern biotechnology.**" is ACCEPTED for publication as a book chapter in the following book: [Recent Research Advances in Biology](#)

Thank you for submitting your manuscript in [Recent Research Advances in Biology](#)

Thanking you.

A handwritten signature in blue ink, appearing to read "M. Basu".

**Dr. M. Basu**  
**Chief Managing Editor**  
[Book Publisher International](#)

---

Reg. Offices:

India: Book Publisher International, Guest House Road, Street no - 1/6, Hooghly, West Bengal, India, Corp. Firm Registration Number: L77527, Tele: +91 8617752708

UK: Book Publisher International, Third Floor, [207 Regent Street, London, W1B 3HH, UK](#)

Fax: +44 20-3058-1429

# Recent advances in the use of the SNAP-tag<sup>®</sup> in the modern biotechnology.

Rosa Merlo,<sup>‡</sup> Rosanna Mattossovich,<sup>‡</sup> Anna Valenti,<sup>\*</sup> and Giuseppe Perugino.<sup>\*</sup>

*Institute of Biosciences and BioResources, National Research Council of Italy, Via Pietro Castellino 111, 80131 Naples, Italy.*

<sup>\*</sup> corresponding author: [anna.valenti@ibbr.cnr.it](mailto:anna.valenti@ibbr.cnr.it); Lead corresponding author: [giuseppe.perugino@ibbr.cnr.it](mailto:giuseppe.perugino@ibbr.cnr.it).

<sup>‡</sup> these authors contributed equally and joined as First Author.

## ABSTRACT

Alkylguanine-DNA-alkyltransferases (AGTs) have a natural role in the protection of DNA from mutations caused by alkylating agents. Their peculiar irreversible self-alkylation reaction led to the development as new tools in the modern biotechnology. SNAP-tag<sup>®</sup> is a powerful enzyme for the specific labelling of protein/enzymes by using benzyl-guanine (BG) derivatives as substrates. This technology has some limitations, as the mesophilic nature of the tag (an engineered variant of the human enzyme) and the needs of purify each substrate. Recently, the SNAP-tag<sup>®</sup> technology was successfully implemented by the employment of thermostable “SNAP-tag-like” variants from (hyper)thermophilic sources, and by the utilization of BG-substrates containing an azide group to be combined with DBCO-derivatives by azide-alkyne Huisgen cycloaddition. The introduction of these new actors on the scene made possible the expansion of the methodology to *in vivo* and *in vitro* harsh reaction conditions, as well as the utilization of more chemical groups in the overall reaction enzyme labelling.

## KEYWORDS

*Protein-tag; protein labelling; enzymatic reaction; click-chemistry; biotechnology.*

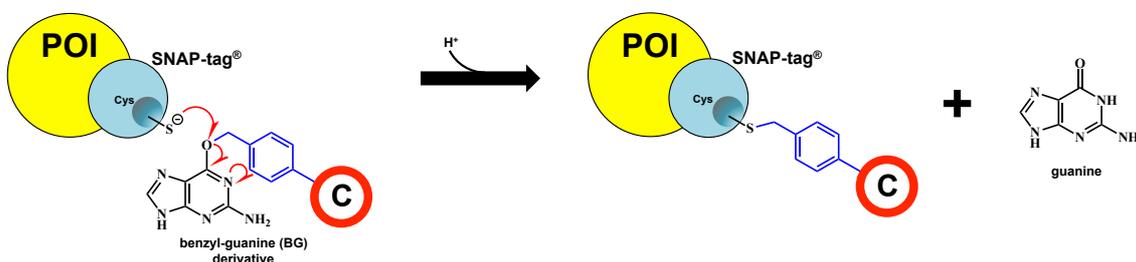
## 1. THE SNAP-tag<sup>®</sup> TECHNOLOGY

Self-labelling *protein-tags* (SLP), as the *Halotag*<sup>®1,2</sup>, the *SpyTag/SpyCatcher*<sup>3</sup>, the SNAP- and the CLIP-tag<sup>®</sup> are gradually replacing auto-fluorescent proteins<sup>1</sup> (AFPs) in modern biotechnology, especially in the field of cell biology for the *in vivo* localization of proteins, organelles and membranes. Despite AFPs, SLPs catalyse an auto-labelling reaction, by a covalent, highly specific and irreversible attachment of a moiety of their synthetic ligands. The conjugation to those ligands to a huge numbers of available chemical groups (fluorescent dyes, affinity molecules, or solid surfaces), led to the enormous success of these tags in various application fields<sup>4</sup>. The Promega *Halotag*<sup>®</sup> is a halo-alkane dehalogenase with a genetically modified active site, which reacts irreversibly with primary alkyl-halides. The *SpyTag/SpyCatcher* system is a method to irreversibly conjugate two different recombinant proteins: the 13-mer *SpyTag* peptide spontaneously reacts with the small protein *SpyCatcher* (12.3 kDa) to form an intermolecular isopeptide bond between the pair<sup>5</sup>. By genetically introducing each component into two distinct protein of interest

51 (POI), the obtained fusion proteins can be covalently linked each other through the  
52 *SpyTag/SpyCatcher* system, when mixed in a reaction.

53 *SNAP-tag*<sup>®</sup> from New England Biolabs (NEB) originates from a laboratory engineering of  
54 the human *O*<sup>6</sup>-methyl-guanine-DNA-methyltransferase (hMGMT). It belongs to the  
55 alkylated DNA-alkyl-transferases family (AGTs, MGMTs or OGTs, E.C. 2.1.1.63),  
56 evolutionary conserved proteins present in most organisms (except in plants and some  
57 microorganisms, as *Thermus thermophilus* and *Saccharomices pombe*<sup>6</sup>), and involved in  
58 the direct repair of the DNA alkylation damage. The activity of these enzymes is the  
59 removal of alkyl adducts from the *O*<sup>6</sup>-position on guanine nucleobases<sup>7,8</sup>. The peculiar  
60 irreversible single-step mechanism led to their definition as “*suicide enzymes*”, since the  
61 alkylated group on the guanine is directly transferred to the catalytic cysteine in the protein  
62 active site<sup>9</sup>, repairing the DNA. Upon reaction, the alkylated enzyme is not longer active  
63 because permanently inactivated and susceptible to *in vivo* degradation *via* proteasome.

64 Some AGTs, as the hMGMT, display low substrate specificity, being able to efficiently  
65 recognise also bulky adducts on the *O*<sup>6</sup>-position of the guanine nucleobase<sup>10</sup>.  
66 Furthermore, it was known that *O*<sup>6</sup>-benzyl-guanine (BG) is a potent inactivator of hMGMT<sup>11</sup>.  
67 From this knowledge, prof. Kai Johnsson and co. engineered a DNA binding-less variant  
68 of hMGMT, by adopting a directed-evolution approach: its suicidal reaction was exploited  
69 for the introduction of a new protagonist in biotechnology, the *SNAP-tag*<sup>®</sup>. Given the  
70 irreversibility transfer of the reaction, they successfully obtained a covalent labelling of the  
71 enzyme with molecules, as fluorophores, previously conjugated to the 4-position of the BG  
72 benzyl ring. The small dimension of *SNAP-tag*<sup>®</sup> opened the possibility to introduce an  
73 innovative methodology for the indirect labelling of POIs genetically fused to it (**Figure**  
74 **1**)<sup>10,12-15</sup>.



84 **Figure 1.** The *SNAP-tag*<sup>®</sup> irreversible enzymatic reaction. This tag is able to covalently link a desired  
85 chemical group (indicated as C) to a protein of interest (POI) avoiding its direct labelling, sometimes risky for  
86 the maintenance of the biological activity.

87  
88  
89 Although most of applications are in cell biology and fluorescence-imaging fields, to date  
90 there are in literature several examples of utilization of this tag, as RNA-editing<sup>16</sup>, the  
91 development of SNAP-based biosensors for small molecules<sup>17-19</sup> and ions<sup>20,21</sup>, and  
92 fascinating “DNA Origami” protein-DNA structures<sup>22</sup>. In general, the possibility to  
93 synthesise BG-substrates through a crosslinking reaction between “BG-building blocks”  
94 (as amine reactive BG-NH<sub>2</sub>) with NHS-ester derivative compounds, make this technique  
95 more desirable than the conventional protein-tags.

96 Later, *SNAP-tag*<sup>®</sup> was further engineered to obtain its orthogonal enzyme, the *CLIP-tag*<sup>®</sup>,  
97 which specifically reacts with *O*<sup>2</sup>-benzyl-cytosine derivatives, thus expanding that  
98 technology for *in vivo* and *in vitro* multi-protein labelling<sup>23</sup>.

103 **2. IMPROVING THE TECHNOLOGY**

104

105 SNAP-tag<sup>®</sup> is a powerful tool widely applicable to every need; however, its technology  
106 has some limitations, in terms of the enzyme and of the relative BG-derivative substrates.  
107 In the former case, this engineered version of hMGMT is suitable only for mesophilic  
108 model organisms and mild reaction conditions. On the other side, apart from commercially  
109 available BG-derivatives, the synthesis of *ad hoc* substrates has the disadvantage of  
110 tedious and complex purifications, increasing times and costs. Furthermore, the  
111 conjugation to the benzyl moiety of the BG of chemical groups is not without any risk for  
112 the efficiency of the SNAP-tag<sup>®</sup> reaction<sup>24-28</sup>. Recently, the group of Dr. Perugino focused  
113 its attention on these two aspects, aim to improving this technology.

114

115

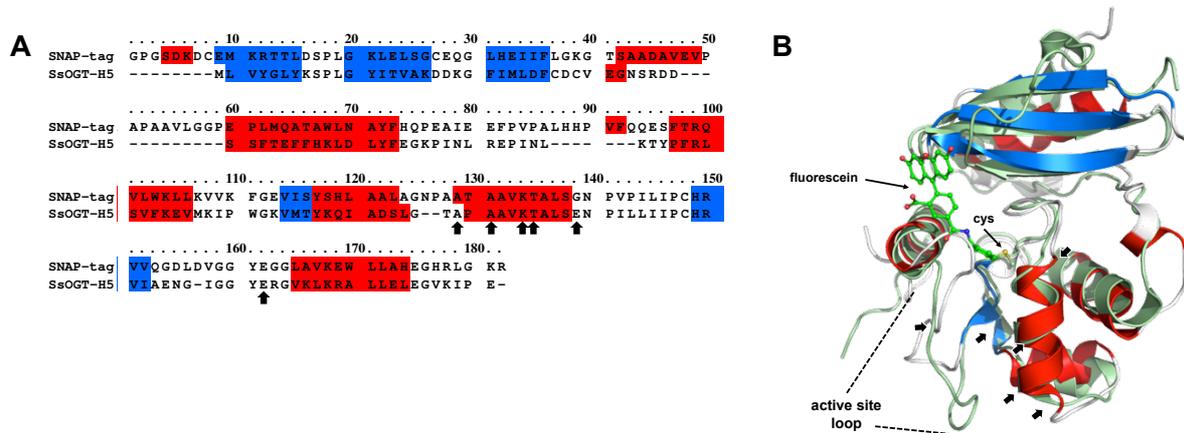
116 **2.1 Acting on the enzyme**

117 In recent decades, the study of proteins and enzymes from thermophilic and  
118 hyperthermophilic sources has had a notable development, especially for all those  
119 (industrial) applications requiring non-moderate reaction conditions. The main reason for  
120 their success lies not only in their intrinsic stability and resistance to heat, but also to other  
121 physical and chemical denaturants. Therefore, to overcome the limitations of the  
122 mesophilic SNAP-tag<sup>®</sup>, it was necessary to identify an AGT from (hyper)thermophilic  
123 sources.

124 In 1999, Imanaka and co. shed light on the thermostability of the AGT from the  
125 hyperthermophilic Euryarchaeota *Pyrococcus kodakaraensis* (Pk-MGMT)<sup>29</sup>: by the 3D  
126 structure and by the comparison with the *E. coli* homologous (AdaC) they observed that  
127 intra-helix ion-pairs mainly contribute to reinforce stability. However, from this interesting  
128 “thermozyme” no application was made. Later, Dr. Kanugula and Prof. Pegg cloned and  
129 purified AGTs from two extremely thermophilic organisms, the bacterium *Aquifex aeolicus*  
130 and the archaeon *Archaeoglobus fulgidus*<sup>5</sup>. These enzymes, although resulted active by  
131 using a radioactive assay, were insoluble and recovered in the inclusion bodies: this  
132 hampered the successive utilization as protein-tag for “biotech” applications.

133 In the last decade, a putative AGT (ORF SSO2487) from the thermo-acidophilic  
134 archaeon *Saccharolobus solfataricus* (previously *Sulfolobus solfataricus*) (SsOGT) was  
135 cloned and widely characterised by the group of Perugino<sup>30-32</sup>. By using a new fluorescent  
136 assay with a fluorescein BG-derivative (SNAP-Vista<sup>®</sup> Green from NEB), this enzyme  
137 displayed a similar catalytic efficiency of hMGMT at its own temperature optimum (80 °C).  
138 Furthermore, it displays a marked stability over a wide range of temperature, pH, ionic  
139 strength and to common denaturing agents<sup>30</sup> (**Table 1**). By following the same approach of  
140 Prof. Johnsson, SsOGT was used as starting material for the development of a “thermo-  
141 SNAP-tag”: SsOGT-*H*<sup>5</sup> (hereinafter *H*<sup>5</sup>) is an *in silico* engineered variant, by hitting five  
142 residues in the helix-turn-helix motif (S100A, R102A, G105K, M106T, K110E) making the  
143 enzyme completely unable to bind DNA (**Figure 2**). A sixth mutation falls on the *active site*  
144 *loop*, a region involved in the substrate recognition and specificity: the replace of a  
145 conserved serine with a glutammic acid was demonstrated in the SNAP-tag<sup>®</sup> to enhance  
146 the catalytic activity towards the BG-substrates<sup>15</sup>. *H*<sup>5</sup> maintained the same behaviours of  
147 the *wt* counterpart in terms of stability, but unexpectedly showed a very high catalytic  
148 activity also at moderate temperatures<sup>30,33</sup> (**Table 1**). The intrinsic stability of thermostable  
149 AGTs allowed their purification after the irreversible reaction, and the subsequent  
150 crystallization in the alkylated form<sup>31</sup>. A 3D structure of *H*<sup>5</sup> covalently linked to a bulky  
151 molecule was achieved, showing for the first time a strong destabilization of the active site  
152 loop<sup>34</sup> (**Figure 2**). *H*<sup>5</sup> is rightly compatible for the biotechnology in thermophilic organisms,  
153 as demonstrated for its utilization as protein-tag in *in vivo* experiments in *Thermus*

154 *thermophilus*<sup>33</sup> and *Sulfolobus islandicus*<sup>35</sup>. Recently, it was a convenient enzyme reporter  
 155 in *in vitro* expression system using *Saccharolobus lysates*<sup>36</sup>.



169 **Figure 2.** SNAP-tag<sup>®</sup> vs H<sup>5</sup>. (A) Primary structure alignment and (B) 3D structure superimposition of the  
 170 human engineered MGMT and the thermostable variant of *Saccharolobus* homologue. 3D structure of H<sup>5</sup>  
 171 after reaction with SNAP-Vista<sup>®</sup> Green (PDB ID: 6GA0) is in light green, whereas the structure of SNAP-tag<sup>®</sup>  
 172 (PDB ID: 3KZY) is in white with helices and strands shown in red and blue, respectively. Black arrows  
 173 indicate the position of mutation in H<sup>5</sup>.

176 The will to push the application of the SNAP-tag<sup>®</sup> technology on the boundaries of  
 177 reaction conditions (as at the boiling point of water) led to the characterisation of  
 178 alternative AGTs: the *Pyrococcus furiosus*<sup>37</sup> was originally isolated from hot marine  
 179 sediments in Vulcano Island (Italy)<sup>38</sup>. This microorganism began famous mainly for the  
 180 DNA polymerase I, also known as *Pfu*DNA polymerase, for its high activity, thermostability,  
 181 and strong 3'-5' proof-reading activity<sup>39</sup>. An OGT activity in *P. furiosus* cell extracts was  
 182 first measured by Margison, identifying a protein of ca. 22.0 kDa<sup>40</sup>. Noteworthy, the  
 183 catalytic activity was completely abolished by the O<sup>6</sup>-BG pseudo-substrate: this behaviour  
 184 makes this enzyme appealed for the SNAP-tag<sup>®</sup>-based applications, which use BG-  
 185 derivatives. The product of the PF1878 ORF was obtained as recombinant protein  
 186 (*Pfu*OGT), confirming its activity on BG-fluorescent substrates, as well as the DNA repair  
 187 activity on natural methylated double strand DNA oligonucleotides<sup>37</sup>. *Pfu*OGT is an  
 188 extreme thermostable enzyme (**Table 1**), although molecular strategies of this resistance  
 189 to denaturation are different from the *Saccharolobus* homologue. The former display a  
 190 worse stability in the presence of high ionic strength conditions, probably due to the  
 191 extensive ionic networks on its surface<sup>37</sup>.

192 The CTN1690 ORF from *Thermotoga neapolitana* shows a clear homology to the AGTs  
 193 family: by fluorescent gel imaging analysis on lyophilized cells, a strong signal with an  
 194 expected molecular weight was revealed<sup>37,41</sup>. The sensitivity to a fluorescent O<sup>6</sup>-BG  
 195 derivative was convincing for the cloning and the heterologous expression in *E. coli*<sup>37</sup>.  
 196 Apart from the activity on artificial substrates, this enzyme (*Tn*OGT) is able to repair  
 197 methylated DNA and, surprisingly, it is very active at low temperatures<sup>37</sup>, in the same order  
 198 of magnitude of H<sup>5</sup> (**Table 1**)<sup>37</sup>. No 3D information on *Tn*OGT are currently available,  
 199 stimulating the research to achieve crystal of this protein to identify thermal resistance and  
 200 activity strategies, which are clearly different from the *SsOGT* enzyme, on the basis of the  
 201 primary structure alignment.

		SNAP-tag®	SsOGT	SsOGT-H <sup>5</sup>	TnOGT	PfuOGT
mol. weight		23.0 kDa	17.0 kDa	17.0 kDa	19.9 kDa	20.1 kDa
T <sub>opt</sub>		37.0 °C	70.0 °C	70.0 °C	- <sup>a</sup>	90.0 °C
relative activity	at 25 °C	80 %	25 %	50 %	100 %	ND <sup>b</sup>
	at 37 °C	100 %	45 %	65 %	- <sup>a</sup>	ND
	at 50 °C	-	-	-	100 %	1 %
	at 80 °C	-	100 %	100 %	- <sup>a</sup>	10 %
	at 90 °C	-	-	-	- <sup>a</sup>	100 %
catalytic activity (s <sup>-1</sup> M <sup>-1</sup> )	at 25 °C	2.8 × 10 <sup>4</sup>	2.8 × 10 <sup>3</sup>	1.6 × 10 <sup>4</sup>	4.7 × 10 <sup>4</sup>	ND
	at T <sub>opt</sub>		5.3 × 10 <sup>4</sup>	3.8 × 10 <sup>4</sup>	- <sup>a</sup>	1.5 × 10 <sup>5</sup>
pH <sub>opt</sub>		7.6	6.0	6.0	6.5	6.5
thermal stability (hrs)		6.0 (37 °C)	2.8 (70 °C)	3.0 (70 °C)	-	-
thermal stability at DSF (T <sub>m</sub> )		67.0 °C (1 min /°C × cycle)	80.0 °C (5 min /°C × cycle)	75.0 °C (5 min /°C × cycle)	- <sup>a</sup>	79.0 °C (10 min /°C × cycle)
thermal stability at 37 °C (hrs)		6.0	> 24	> 24	-	-
additives	NaCl	< 0.25 M	> 4.0 M	> 4.0 M	- <sup>a</sup>	> 1.0 M
	EDTA	inhibitor	> 10.0 mM	> 10.0 mM	-	-
	sarcosyl	no	> 0.5 %	> 0.5 %	-	-
	DTT	required	not required	not required	-	-
	SDS	no	≤ 0.01 %	≤ 0.01 %	-	> 0.01 %

**Table 1.** Biochemical properties of SNAP-tag® and thermostable OGTs. The values are obtained by the fluorescent-based assay with the SNAP Vista® Green and methylated dsDNA oligonucleotides. <sup>a</sup>Not measured for technical limitation of the end-point assay. <sup>b</sup>Not determined. (from refs 15, 23, 30, 31, 33 and 41).

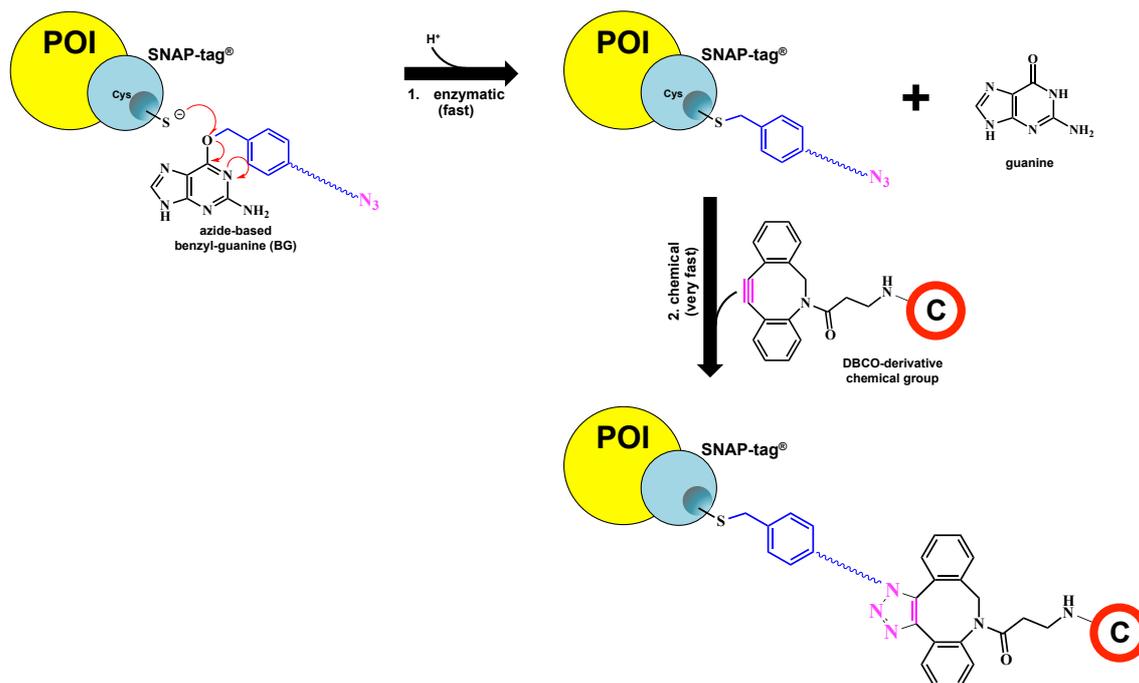
## 2.2 Acting on the substrate

The relatively high time and costs for the production and the utilization of commercial or purified customized BG-derivatives should be overcome by a slight and innovative modification of the SNAP-tag® technology. The starting assumptions are that AGTs keep the correct folding in their benzylated form after the enzymatic reaction<sup>15,33</sup>, and the chemical “click” reaction by the copper-free azide-alkyne Huisgen cycloaddition is extremely versatile, fast and specific. Hence, a *chemo-enzymatic approach* (**Figure 3**) with a selected azide-based BG-substrate could be possible: if the azide group is efficiently exposed outside the protein surface, it should allow the reaction with a di-benzo-cyclo-octyl (DBCO)-based molecule. To date, BG-N<sub>3</sub> and BG-DBCO derivatives were already used as building block for the synthesis of BG-based molecules<sup>42,43</sup>, but the group of Perugino for the first time used a BG-azide directly with the SNAP-tag® and the H<sup>5</sup> variant<sup>44</sup>.

The reasons of this choice lie on the fact that DBCO-based molecules are more commercially available and affordable compounds than BG-derivatives. Furthermore, the need of purifying every BG-substrate is completely avoided, because only that of BG-azide substrate would take place. Finally, the presence of bulky chemical groups conjugated to the benzyl moiety sometime affect the catalytic efficiency of these protein-tags<sup>44,45</sup>, whereas a selected azide-substrate could bypass this problem.

Perugino and his group used a BG-derivative in which the azide group is opportunely distanced by the benzyl moiety of the molecule<sup>44</sup>. This spacer pushes away the azide from the active site better than the well-known BG-N<sub>3</sub> derivative<sup>42</sup>: this was also confirmed by molecular simulation, by determining the Solvent Accessible Surface Area (SASA) data

258 with both BG-azide substrates<sup>44</sup>. The SNAP-tag<sup>®</sup> previously incubated with the azide-  
 259 substrate was efficiently prone (“clickable”) to the subsequent labelling with several DBCO-  
 260 based fluorophores and with alkyne-coated surface sensors for its direct immobilization<sup>44</sup>.  
 261 The high specificity of both the reactions was verified in the presence of *in vitro*  
 262 “perturbing” proteins (like in cell lysates) and the *in vivo* labelling of expressed SNAP-tag<sup>®</sup>  
 263 in eukaryotic and prokaryotic cells<sup>44</sup>. In the latter case, both the cell types resulted very  
 264 permeable to the BG-azide substrate used. In conclusion, Perugino’s group experimentally  
 265 demonstrated that splitting the SNAP-tag<sup>®</sup> reaction into two fast steps does not affect the  
 266 overall rate and efficiency of the protein labelling<sup>15,33</sup> (**Figure 3**). This new approach could  
 267 open new perspectives and widening the applications of this powerful biotechnology.



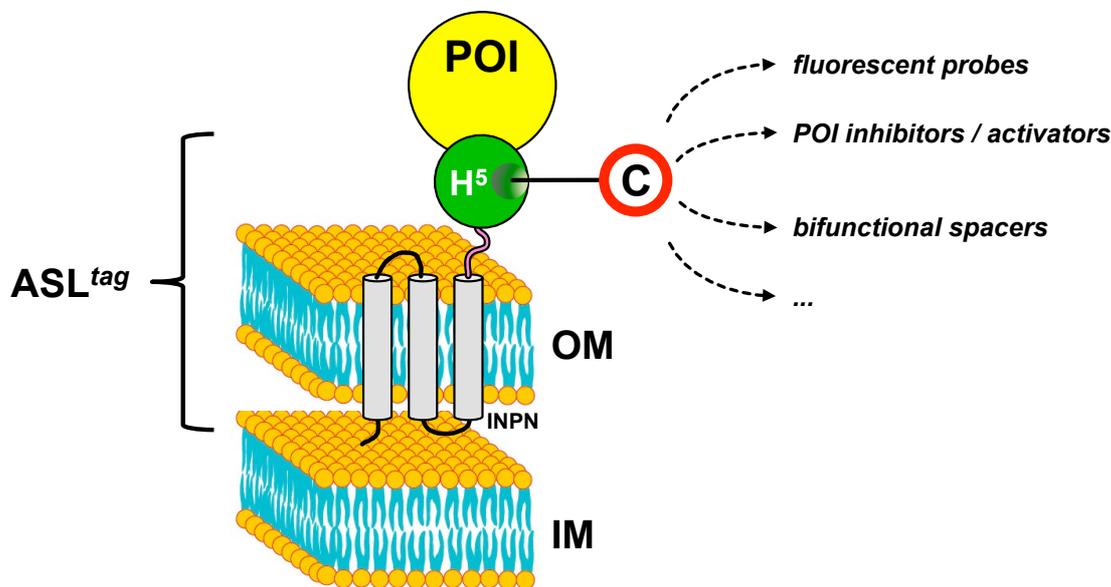
289 **Figure 3.** The chemo-enzymatic approach by using a universal azide-based BG substrate for the SNAP-tag<sup>®</sup>  
 290 labelling.

### 294 3. THE INNOVATIVE ASL<sup>tag</sup> SYSTEM

296 Enzyme immobilization on solid supports sometimes requires harsh operational  
 297 conditions<sup>46</sup>. By definition, an immobilized enzyme is a “physically confined biocatalyst,  
 298 which retains its catalytic activity and can be used repeatedly”<sup>47</sup>. The catalysts’ recovery  
 299 and reuse, as well as the physical separation of the enzymes from the reaction mixture are  
 300 some advantages of protein immobilisation. Actually, different immobilisation approaches  
 301 are available, from physical adsorption to covalent coupling<sup>48-51</sup>. However, most of the  
 302 immobilization protocols need of purified biocatalysts, and the “random” interaction  
 303 between the enzyme and the solid support could suffer of the steric hindrance and the  
 304 correct orientation of the former, increasing of costs and time for the production processes.

305 The introduction of the N-terminal domain of the ice nucleation protein (INPN) from  
 306 *Pseudomonas syringae*<sup>52,53</sup> allowed a “cell-based” *in vivo* immobilisation system,  
 307 significantly improving and reducing time and costs of the process. Recently, this widely  
 308 used display strategy on the external surface of Gram<sup>(-)</sup> bacteria cells was used in  
 309 combination with the H<sup>5</sup> enzyme, to produce a novel anchoring and self-labelling protein

310 tag (ASL<sup>tag</sup>), as shown in **Figure 4**<sup>54</sup>. INPN immobilises a protein or an enzyme of interest  
 311 on *E. coli* outer membrane, allowing its exposition to the solvent and its correct orientation  
 312 for biological activities. This transmembrane domain overcomes the problems related to  
 313 the purification, the immobilization and the recovery of biocatalysts<sup>55</sup>. The H<sup>5</sup> moiety, on  
 314 the other side, because catalyses the covalent labelling with any desired chemical groups  
 315 (opportunistically conjugated to the benzyl-guanine)<sup>54</sup>, offers the possibility to modulate the  
 316 enzyme activity, by introducing activator or inhibitor molecules (**Figure 4**).



334 **Figure 4.** The ASL<sup>tag</sup> system. This protein-tag allows the *in vivo* POIs immobilization on the bacterial Gram<sup>(-)</sup>  
 335 outer membrane (OM) by the INPN domain, and at the same time the possibility to combine their biological  
 336 activities with desired chemical groups (C) by the covalent labelling operated by the H<sup>5</sup> enzyme.

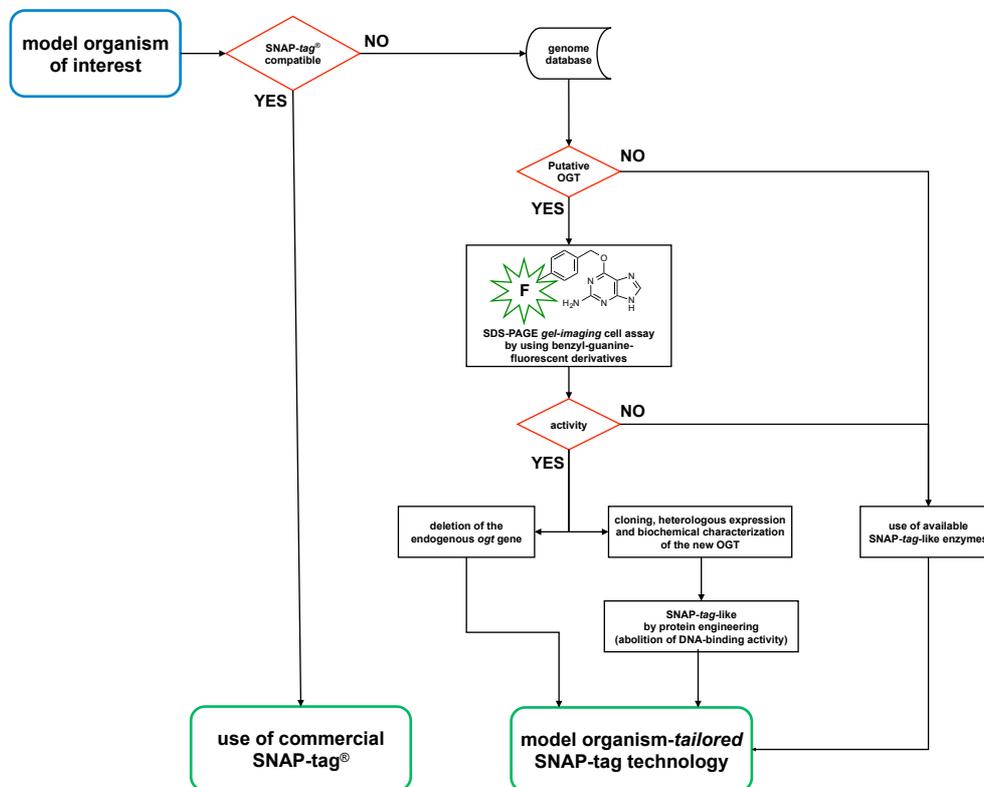
339 The ASL<sup>tag</sup> system was successfully employed for the expression and immobilization of  
 340 two thermostable enzymes, the carbonic anhydrase from *Sulfurihydrogenibium*  
 341 *yellowstonense* (SspCA), and the tetrameric β-glycosyl hydrolase from *Saccharolobus*  
 342 *solfatarius* (Ssβgly). Moreover, the presence of this tag does not interfere with their  
 343 folding and catalytic activity<sup>54</sup>, but makes SspCA more stable to thermal denaturation,  
 344 respect to the same enzyme fused only to the INPN<sup>55,56</sup>. This aspect represents a clear  
 345 advantage in pushing beyond reactions in bioreactors and in the reutilization of  
 346 biocatalysts.

347 This innovative protein-tag is particularly useful for the *in vivo* immobilization and the  
 348 contemporary labelling of proteins and/or enzymes of interest, stabilizing them without any  
 349 purification procedures needed.

#### 353 4. CONCLUSION AND PERSPECTIVES

355 The knowledge acquired in recent years on AGTs and their use in modern  
 356 biotechnologies allows the development of a “workflow” aim to optimize this protein-tag for  
 357 different needs (**Figure 5**). In the case of model organisms thriving in extreme conditions  
 358 of temperature, salinity, pressure, etc., in which the commercial SNAP-tag<sup>®</sup> is not  
 359 particularly suitable, it would be possible to imagine an identification of a relative AGT and,  
 360 through mutagenesis, obtain a “SNAP-tag-like” tool (**Figure 2**). This strategy has to  
 361 necessarily take into account the fact that this identified AGT should be sensitive to O<sup>6</sup>-BG

362 derivatives: some members of this family, such as *E. coli* AdaC, are resistant to the action  
 363 of O6-BG<sup>57,58</sup>. In parallel, it is possible to perform an *ogt* gene deletion in the model  
 364 organism, in order to eliminate any endogenous activity. This workflow was successfully  
 365 realised in thermophilic microorganisms, as the archaeon *S. islandicus*<sup>35</sup> and in the  
 366 bacterium *T. thermophilus*. In the latter case, it was not necessary to eliminate the *ogt*  
 367 gene, because this organism lacks of it<sup>33</sup>.



391  
 392 **Figure 5.** The SNAP-tag<sup>®</sup> workflow.  
 393

## 394 REFERENCES

- 395
- 396 1. Los G.V.; Encell L.P.; McDougall M.G.; Hartzell D.D.; Karassina N.; Zimprich C.; Wood M.G.;  
 397 Learish R.; Ohana R.F.; Urh M.; Simpson D. **HaloTag: a novel protein labeling technology for**  
 398 **cell imaging and protein analysis.** *ACS Chem. Biol.* 2008; **3**: 373-382. DOI: 10.1021/cb800025k.
  - 399 2. England C.G.; Luo H.; Cai W. **HaloTag technology: a versatile platform for biomedical**  
 400 **applications.** *Bioconjug. Chem.* 2015; **26**: 975-86. DOI: 10.1021/acs.bioconjugchem.5b00191.
  - 401 3. Hatlem D.; Trunk T.; Linke D.; Leo J.C. **Catching a SPY: Using the SpyCatcher-SpyTag and**  
 402 **Related Systems for Labeling and Localizing Bacterial Proteins.** *Int. J. Mol. Sci.* 2019; **20**: 2129-  
 403 2139. DOI: doi: 10.3390/ijms20092129.
  - 404 4. Liss V.; Barlag B.; Nietschke M.; Hensel M. **Self-labelling enzymes as universal tags for**  
 405 **fluorescence microscopy, super-resolution microscopy and electron microscopy.** *Sci. Rep.*  
 406 2015; **5**: 17740-17752. DOI: 10.1038/srep17740.
  - 407 5. Zakeri B., Fierer J.O.; Celik E.; Chittock E.C.; Schwarz-Linek U.; Moy V.T.; Howarth M. **Peptide tag**  
 408 **forming a rapid covalent bond to a protein, through engineering a bacterial adhesin.** *Proc.*  
 409 *Natl. Acad. Sci. USA* 2012, **109**: E690-7. doi:10.1073/pnas.1115485109. PMC 3311370.
- 410  
 411  
 412  
 413  
 414  
 415  
 416

- 417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478
6. Kanugula S.; Pegg A.E. **Alkylation damage repair protein O<sup>6</sup>-alkylguanine-DNA alkyltransferase from the hyperthermophiles *Aquifex aeolicus* and *Archaeoglobus fulgidus*.** *Biochem. J.* 2003, **375**: 449-455.
  7. Pegg A.E. **Repair of O(6)-alkylguanine by alkyltransferases.** *Mutat. Res.* 2000; **462**: 83-100. DOI: 10.1016/s1383-5742(00)00017-x.
  8. Daniels D.S.; Mol C.D.; Arvai A.S.; Kanugula S.; Pegg A.E.; Tainer J.A. **Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding.** *EMBO J.* 2000; **19**: 1719-1730. DOI: 10.1093/emboj/19.7.1719.
  9. Daniels D.S.; Woo T.T.; Luu K.X.; Noll D.M.; Clarke N.D.; Pegg A.E.; Tainer J.A. **DNA binding and nucleotide flipping by the human DNA repair protein AGT.** *Nat. Struct. Mol. Biol.* 2004; **11**: 714-720. DOI: 10.1038/nsmb791.
  10. Keppler A.; Gendreizig S.; Gronemeyer T.; Pick H.; Vogel H.; Johnsson K.A. **A general method for the covalent labeling of fusion proteins with small molecules *in vivo*.** *Nat. Biotechnol.* 2003; **21**: 86-89. DOI: 10.1038/nbt765.
  11. Kaina B.; Christmann M. **DNA repair in personalized brain cancer therapy with temozolomide and nitrosoureas.** *DNA Repair* 2019, **78**: 128-141.
  12. Keppler A.; Pick H.; Arrivoli C.; Vogel H.; Johnsson K. **Labeling of fusion proteins with synthetic fluorophores in live cells.** *Proc. Natl. Acad. Sci. USA.* 2004; **10**: 9955-9959. DOI: 10.1073/pnas.0401923101.
  13. Juillerat A.; Gronemeyer T.; Keppler A.; Gendreizig S.; Pick H.; Vogel H.; Johnsson K. **Directed evolution of O<sup>6</sup>-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules *in vivo*.** *Chem. Biol.* 2003; **10**: 313-317. DOI: 10.1016/S1074-5521(03)00068-1.
  14. Gronemeyer T.; Chidley C.; Juillerat A.; Heinis C.; Johnsson K. **Directed evolution of O<sup>6</sup>-alkylguanine-DNA alkyltransferase for applications in protein labeling.** *Prot. Eng. Des. Sel.* 2006; **19**: 309-316. DOI: 10.1093/protein/gz1014.
  15. Mollwitz B.; Brunk E.; Schmitt S.; Pojer F.; Bannwarth M.; Schiltz M.; Rothlisberger U.; Johnsson K. **Directed evolution of the suicide protein O<sup>6</sup>-alkylguanine-DNA alkyltransferase for increased reactivity results in an alkylated protein with exceptional stability.** *Biochemistry.* 2012; **51**: 986-994. DOI: 10.1021/bi2016537.
  16. Vogel P.; Moschref Merkle M.M.; Li Q.; Selvasarayanan K.D.; Li J.B.; Stafforst T. **Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs.** *Nat. Meth.* 2018; **15**: 535-538. DOI: 10.1038/s41592-018-0017-z.
  17. Srikun D.; Albers A.E.; Nam C.I.; Iavarone A.T.; Chang C.J. **Organelle-targetable fluorescent probes for imaging hydrogen peroxide in living cells via SNAP-Tag protein labelling.** *J. Am. Chem. Soc.* 2010; **132**: 4455-4465. DOI: 10.1021/ja100117u.
  18. Abo M.; Minakami R.; Miyano K.; Kamiya M.; Nagano T.; Urano Y.; Sumimoto H. **Visualization of phagosomal hydrogen peroxide production by a novel fluorescent probe that is localized via SNAP-tag labeling.** *Anal. Chem.* 2014; **86**: 5983-5990. DOI: 10.1021/ac501041w.
  19. Griss R.; Schena A.; Reymond L.; Patiny L.; Werner D.; Tinberg C.E.; Baker D.; Johnsson K. **Bioluminescent sensor proteins for point-of-care therapeutic drug monitoring.** *Nat. Chem. Biol.* 2014; **10**: 598-603. DOI: 10.1038/nchembio.1554.
  20. Tomat E.; Nolan E.M.; Jaworski J.; Lippard S.J. **Organelle-specific zinc detection using zinpyr-labeled fusion proteins in live cells.** *J. Am. Chem. Soc.* 2008; **130**: 15776-15777. DOI: 10.1021/ja806634e.
  21. Kamiya M.; Johnsson K. **Localizable and highly sensitive calcium indicator based on a BODIPY fluorophore.** *Anal. Chem.* 2010; **82**: 6472-6479. DOI: 10.1021/ac100741t.

479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540

22. Saccà B.; Meyer R.; Erkelenz M.; Kiko K.; Arndt A.; Schroeder H.; Rabe K.S.; Niemeyer C.M. **Orthogonal protein decoration of DNA origami.** *Angew. Chem. Int. Ed.* 2010; **49**: 9378 -9383. DOI: 10.1002/anie.201005931.
23. Gautier A.; Juillerat A.; Heinis C.; Corrêa I.R. Jr.; Kindermann M.; Beaufils F.; Johnsson K. **An engineered protein-tag for multi- protein labeling in living cells.** *Chem. Biol.* 2008; **15**: 128-136. DOI: 10.1016/j.chembiol.2008.01.007.
24. Moschel R.C.; McDougall M.G.; Dolan M.E.; Stine L.; Pegg A.E. **Structural features of substituted purine derivatives compatible with depletion of human O<sup>6</sup>-alkylguanine-DNA alkyltransferase.** *J. Med. Chem.* 1992; **35**: 4486-4491. DOI: 10.1021/jm00101a028.
25. Chae M.Y.; McDougall M.G.; Dolan M.E.; Swenn K.; Pegg A.E.; Moschel R.C. **Substituted O<sup>6</sup>-benzylguanine derivatives and their inactivation of human O<sup>6</sup>-alkylguanine-DNA alkyltransferase.** *J. Med. Chem.* 1994; **37**: 342-347. DOI: 10.1021/jm00029a005.
26. Terashima I.; Kawate H.; Sakumi K.; Sekiguchi M.; Kohda K. **Substrate specificity of human O<sub>6</sub>-methylguanine-DNA methyltransferase for O<sub>6</sub>-benzylguanine derivatives in oligodeoxynucleotides.** *Chem. Res. Toxicol.* 1997; **10**: 1234-1239. DOI: 10.1021/tx9700580.
27. McElhinney R.S.; Donnelly D.J.; McCormick J.E.; Kelly J.; Watson A.J.; Rafferty J.A.; Elder R.H.; Middleton M.R.; Willington M.A.; McMurry T.B.H.; Margison GP. **Inactivation of O<sup>6</sup>-alkylguanine-DNA alkyltransferase. 1. Novel O<sup>6</sup>-(hetarylmethyl)-guanines having basic rings in the side chain.** *J. Med. Chem.* 1998; **41**: 5265-5271. DOI: 10.1021/jm9708644.
28. Sun G.; Fan T.; Zhang N.; Ren T.; Zhao L.; Zhong R. **Identification of the Structural Features of Guanine Derivatives as MGMT Inhibitors Using 3D-QSAR Modeling Combined with Molecular Docking.** *Molecules.* 2016; **21**: 823-844. DOI: 10.3390/molecules21070823.
29. Hashimoto, H.; Inoue, T.; Nishioka, M.; Fujiwara, S.; Takagi, M.; Imanaka, T.; Kai, Y. **Hyperthermostable Protein Structure Maintained by Intra and Inter-helix Ion-pairs in Archaeal O<sub>6</sub>-Methyl-guanine-DNA Methyltransferase.** *J. Mol. Biol.* 1999, **292**: 707-716.
30. Perugino G.; Vettone A.; Illiano G.; Valenti A.; Ferrara M.C.; Rossi M.; Ciaramella M. **Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage.** *J. Biol. Chem.* 2012; **287**: 4222-4231. DOI: 10.1074/jbc.M111.308320.
31. Perugino G.; Miggiano R.; Serpe M.; Vettone A.; Valenti A.; Lahiri S.; Rossi F.; Rossi M.; Rizzi M.; Ciaramella M. **Structure- function relationships governing activity and stability of a DNA alkylation damage repair thermostable protein.** *Nuc. Ac. Res.* 2015; **43**: 8801-8816. DOI: 10.1093/nar/gkv774
32. Morrone C.; Miggiano R.; Serpe M.; Massarotti A.; Valenti A.; Del Monaco G.; Rossi M.; Rossi F.; Rizzi M.; Perugino G.; Ciaramella M. **Interdomain interactions rearrangements control the reaction steps of a thermostable DNA alkyltransferase.** *Biochim. Biophys. Acta.* 2017; **1861**: 86-96. DOI: 10.1016/j.bbagen.2016.10.020.
33. Vettone A.; Serpe M.; Hidalgo A.; Berenguer J.; Del Monaco G.; Valenti A.; Rossi M.; Ciaramella M.; Perugino G. **A novel thermostable protein-tag: optimization of the *Sulfolobus solfataricus* DNA-alkyl-transferase by protein engineering.** *Extremophiles.* 2016; **20**: 1-13. DOI: 10.1007/s00792-015-0791-9.
34. Rossi F.; Morrone C.; Massarotti A.; Ferraris D.M.; Valenti A.; Perugino G.; Miggiano R. **Crystal structure of a thermophilic O<sup>6</sup>-alkylguanine-DNA alkyltransferase-derived self-labeling protein-tag in covalent complex with a fluorescent probe.** *Biochem. Biophys. Res. Comm.* 2018; **500**: 698-703. DOI: 10.1016/j.bbrc.2018.04.139.
35. Visone V.; Han W.; Perugino G.; Del Monaco G.; She Q.; Rossi M.; Valenti A.; Ciaramella M. **In vivo and in vitro protein imaging in thermophilic archaea by exploiting a novel protein tag.** *PLoS ONE.* 2017; **12**: 1-19. DOI: 10.1371/journal.pone.0185791.

- 541 36. Lo Gullo G.; Mattossovich R.; Perugino G.; La Teana A.; Londei P.; Benelli D. **Optimization of an *in***  
542 ***vitro* transcription/ translation system based on *Sulfolobus solfataricus* cell lysate.** *Archaea*,  
543 2019; Article ID 9848253. DOI: 10.1155/2019/9848253.  
544
- 545 37. Mattossovich R.; Merlo R.; Fontana A.; d'Ippolito G.; Terns M.P.; Watts E.A.; Valenti A.; Perugino G.  
546 **A journey down to hell: new thermostable protein-tags for biotechnology at high**  
547 **temperatures.** *Extremophiles*. 2020; **24**: 81-91. DOI: 10.1007/s00792-019-01134-3.  
548
- 549 38. Fiala G.; Stetter K.O. ***Pyrococcus furiosus* sp. nov. represents a novel genus of marine**  
550 **heterotrophic archaeobacteria growing optimally at 100 °C.** *Arch. Microb.* 1986, **145**: 56-61.  
551
- 552 39. Lundberg K.S.; Shoemaker D.D.; Adams M.W.W.; Short J.M.; Sorge J.A.; Mathur E.J. **High-fidelity**  
553 **amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*.** *Gene*  
554 1991, **108**: 1-6.  
555
- 556 40. Skorvaga M.; Raven N.D.H.; Margison G.P. **Thermostable archaeal O<sup>6</sup>-alkylguanine-DNA**  
557 **alkyltransferases.** *Proc. Natl. Acad. Sci. USA* 1998, **95**: 6711-6715.  
558
- 559 41. Mattossovich R.; Merlo R.; Miggiano R.; Valenti A.; Perugino, G. **O<sup>6</sup>-alkylguanine-DNA**  
560 **Alkyltransferases in Microbes Living on the Edge: From Stability to Applicability.** *Int. J. Mol.*  
561 *Sci.* 2020, **21**: 2878. doi:10.3390/ijms21082878  
562
- 563 42. Zhang C.J.; Li L.; Chen G.Y.J.; Xu Q-H.; Yao S.Q. **One- and two-photon live cell imaging using a**  
564 **mutant SNAP-Tag protein and its FRET substrate pairs.** *Org. Lett.* 2011; **13**: 4160-4163. DOI:  
565 10.1021/ol201430x.  
566
- 567 43. Song X.; Wang C.; Han Z.; Xu Y.; Xiao Y. **Terminal alkyne substituted O<sup>6</sup>-benzylguanine for**  
568 **versatile and effective syntheses of fluorescent labels to genetically encoded SNAP-tags.**  
569 *RSC Adv.* 2011; **5**: 23646-23649. DOI: 10.1039/C4RA17072E.  
570
- 571 44. Merlo R.; Caprioglio D.; Cillo M.; Valenti A.; Mattossovich R.; Morrone C.; Massarotti A.; Rossi F.;  
572 Miggiano R.; Leonardi A.; Minassi A.; Perugino G. **The SNAP-tag technology revised: an effective**  
573 **chemo-enzymatic approach by using a universal azide-based substrate.** *J. Enz. Inhib. Med.*  
574 *Chem.* 2020, **36**: 85-97. doi: 10.1080/14756366.2020.1841182  
575
- 576 45. Colombo M.; Mazzucchelli S.; Montenegro J.M.; Galbiati E.; Corsi F.; Parak W.J.; Prospero D.  
577 **Protein oriented ligation on nanoparticles exploiting O<sup>6</sup>-alkylguanine-DNA transferase (SNAP)**  
578 **genetically encoded fusion.** *Small*. 2012; **8**: 1492-1497. DOI: 10.1002/smll.201102284.  
579
- 580 46. Zhou Z.; Hartmann M. **Progress in enzyme immobilization in ordered mesoporous materials**  
581 **and related applications.** *Chem. Soc. Rev.* 2013, **42**: 3894-3912.  
582
- 583 47. Tosa T.; Mori T.; Fuse N.; Chibata I. **Studies on continuous enzyme reactions. I. Screening of**  
584 **carriers for preparation of water-insoluble aminoacylase.** *Enzymologia* 1966, **31**: 214-224.  
585
- 586 48. Mohamad N.R.; Che Marzuki N.H.; Buang N.A.; Huyop F.; Wahab R.A. **An overview of**  
587 **technologies for immobilization of enzymes and surface analysis techniques for immobilized**  
588 **enzymes.** *Biotech. Biotech. Equip.* 2015, **29**: 205-220.  
589
- 590 49. Nguyen H.H.; Kima M. **An Overview of Techniques in Enzyme Immobilization.** *App. Sci. Conv.*  
591 *Tech.* 2017, **26**: 157-163.  
592
- 593 50. Jakub Zdarta J.; Meyer A.S.; Jesionowski T.; Pinelo M. **A General Overview of Support Materials**  
594 **for Enzyme Immobilization: Characteristics, Properties, Practical Utility.** *Catalysts* 2018, **8**: 92.  
595
- 596 51. Sirisha V.L.; Jain A. **Chapter Nine - Enzyme Immobilization: An Overview on Methods, Support**  
597 **Material, and Applications of Immobilized Enzymes.** *Adv. Food Nut. Res.* 2016, **79**: 179-211.  
598
- 599 52. Gurian-Sherman D.; Lindow S.E. **Bacterial ice nucleation: Significance and molecular basis.**  
600 *FASEB J.* 1993, **7**: 1338-1343.  
601

- 602 53. Cochet N.; Widehem P. **Ice crystallization by *Pseudomonas syringae***. *Appl. Microbiol. Biotechnol.* 2000, **54**: 153-161.
- 603
- 604
- 605 54. Merlo R.; Del Prete S.; Valenti A.; Mattossovich R.; Carginale V.; Supuran C.T.; Capasso C.; Perugino P. **An AGT-based protein-tag system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane**. *J. Enz. Inhib. Med. Chem.* 2019, **34**: 490-499.
- 606
- 607
- 608
- 609 55. Del Prete S.; Perfetto R.; Rossi M.; Alasmary F.A.S.; Osman S.M.; Al Othman Z.; Supuran C.T.; Capasso C. **A one-step procedure for immobilising the thermostable carbonic anhydrase (SspCA) on the surface membrane of *Escherichia coli***. *J. Enz. Inhib. Med. Chem.* 2017, **32**: 1120-1128.
- 610
- 611
- 612
- 613
- 614 56. Del Prete S.; Merlo R.; Valenti A.; Mattossovich R.; Rossi M.; Carginale M.; Supuran C.T.; Perugino G.; Capasso C. **Thermostability enhancement of the  $\alpha$ -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* by using the anchoring-and-self labelling-protein-tag system (ASLtag)**. *J. Enz. Inhib. Med. Chem.* 2019, **34**: 946-954.
- 615
- 616
- 617
- 618
- 619 57. Elder R.H.; Margison G.P.; Rafferty J.A. **Differential inactivation of mammalian and *Escherichia coli* O6-alkylguanine-DNA alkyltransferases by O6-benzylguanine**. *Biochem. J.* 1994, **298**: 231-235.
- 620
- 621
- 622
- 623 58. Goodtzova K.; Kanugula S.; Edara S.; Pauly G.T.; Moschel R.C.; Pegg A.E. **Repair of O6-benzylguanine by the *Escherichia coli* Ada and Ogt and the human O6-alkylguanine-DNA alkyltransferase**. *J. Biol. Chem.* 1997, **272**: 8332-8339.
- 624
- 625
- 626
- 627
- 628

## 629 **ACKNOWLEDGMENTS**

630 G.P. would like to thank all the authors, Elena and Elisa Perugino for their efforts in writing  
631 this work and in their technical assistance, but mainly for their human support during the  
632 difficult and delicate period of staying at home following the COVID-19 outbreak.

633

634

Before starting with the Acknowledgments, I would like to keep a promise.

*To Maria Ciaramella*

*That day, I made a promise. I would have preferred to tell you in person, but I hope that these words will reach you in one way or another.*

*Firstly, I would like to thank you for the opportunity you gave me, to stay in this lab and grow in it as a young scientist. Thank you for believing in me, even if only for a minute, and for accepting me as PhD student.*

*It all happened by chance and the chance brought me other opportunities that presented greater security. Since then, you have given me your support, your advice ... but what can I say, I'm so stubborn about the things I love. In your studio, at that moment, I chose to stay, without pretensions, with the only request to be able to spend part of my training in another country, to grow and expand my mind even more.*

*Since then, I promised that I would do everything in my power to make you proud of me, to ensure that your offer was rightly honoured and repaid and that you would not regret the decision. Then, only at the end of this path, I would ask you to talk, in your studio, where it all started, and I would have simply said "Thank you. I deeply hope I made you proud".*

*I made the same promise to you That sad day... I hope to have honoured you, even just in part.*

*Now I'm here, sitting at my desk writing these words to you, with tears running down my cheeks. Hope you can hear them; hope you can hear me.*

*Thank you, Maria. It was an honour.*

## Acknowledgements

“She stood in the storm, and when the wind did not blow her way, she adjusted her sails.”

— Elizabeth Edwards

I think there's no better sentence to describe my journey as a PhD student.

They have been intense and formative years, full of challenges, opportunities, experiences and, fortunately, successes. I would like to sincerely thank my supervisor, Dr Giuseppe Perugino, for constantly testing me and allowing me to overcome my limits. Thanks Pino, for offering me so many possibilities (I hope I was able to take them to the best) and for giving me the best training that a PhD student could wish for: thank you for teaching me to THINK like a scientist, to ask myself questions, to always think further, to consider every detail and above all, thanks for having encouraged my passion.

Thanks to all the people who shared this path with me and all those I met; you probably didn't realize it but yours was an important contribution to my growth.

A special thanks to my family, who continues to encourage me in what I do. There are no words for you, for always giving me everything. To Andrea, for his immeasurable patience, for his support and presence, 42 times thank you. And finally, to all of you who have always been there, in heartbeats, meters or kilometres away.