

A novel thermostable protein-tag: optimization of the *Sulfolobus solfataricus* DNA-alkyl-transferase by protein engineering

Antonella Vettone¹ · Mario Serpe¹ · Aurelio Hidalgo² · José Berenguer² · Giovanni del Monaco¹ · Anna Valenti¹ · Mosé Rossi¹ · Maria Ciaramella¹ · Giuseppe Perugino¹

Received: 4 August 2015 / Accepted: 3 October 2015 / Published online: 23 October 2015
© Springer Japan 2015

Abstract In the last decade, a powerful biotechnological tool for the in vivo and in vitro specific labeling of proteins (SNAP-tagTM technology) was proposed as a valid alternative to classical protein-tags (green fluorescent proteins, GFPs). This was made possible by the discovery of the irreversible reaction of the human alkylguanine-DNA-alkyl-transferase (hAGT) in the presence of benzyl-guanine derivatives. However, the mild reaction conditions and the general instability of the mesophilic SNAP-tagTM make this new approach not fully applicable to (hyper-)thermophilic and, in general, extremophilic organisms. Here, we introduce an engineered variant of the thermostable alkylguanine-DNA-alkyl-transferase from the Archaea *Sulfolobus solfataricus* (SsOGT-H⁵), which displays a catalytic efficiency comparable to the SNAP-tagTM protein, but showing high intrinsic stability typical of proteins from this organism. The successful heterologous expression obtained in a thermophilic model organism makes SsOGT-H⁵ a valid candidate as protein-tag for organisms living in extreme environments.

Keywords Archaea · *Sulfolobus solfataricus* · Thermostable proteins · Protein-tag · Biotechnology

Communicated by S. Albers.

A. Vettone and M. Serpe equally contributed to the present work.

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

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

² Centro De Biología Molecular Severo Ochoa, Nicolás Cabrera 1, Cantoblanco, 28049 Madrid, Spain

Introduction

Labeling of proteins with synthetic probes in living cells is an important approach to study protein function (Grone-meyer et al. 2006). However, radioactive and fluorescent probes need complex chemical modification and protein purification procedures, leading only in few cases to a specific labeling of proteins. The introduction of fusion polypeptides (tags), which allow to simplify and optimize purification methods (affinity tags) and/or detection procedures (by using specific antibodies) has been fundamental. Nevertheless, most tags do not satisfy the need of the in vivo specific detection of proteins of interest. The discovery of autofluorescent proteins as the *Aequorea victoria* green fluorescent protein (GFP) made possible the expansion of the studies on fusion proteins in several applicative fields of cellular and molecular biology (Chalfie et al. 1994; Tsien 1998). Despite their wide use in the last decades, GFP tag and its derivatives have some disadvantages and/or limitations: (1) their relative big dimensions (ca. 230 amino acids) in some cases might affect the target protein's function and localization; (2) the isolation of the natural fluorophore in the active site makes GFPs and variants partially sensitive to environmental changes, as pH (Ashby et al. 2004; Campbell and Choy 2000); (3) because the formation of the internal fluorophore is O₂-dependent, GFPs are not suitable tags in applications requiring anaerobic conditions (anaerobic organisms and tumor cells in particular stages); (4) the general use of most protein-tags is restricted to mesophiles and mild reaction conditions. Recently, thermostable GFPs variants have been developed as tags for thermophilic microorganisms (Aliye et al. 2015; Cava et al. 2008; Pédelacq et al. 2006); however, these tools, again, suffer from most of the same limitations listed above for their mesophilic counterparts.

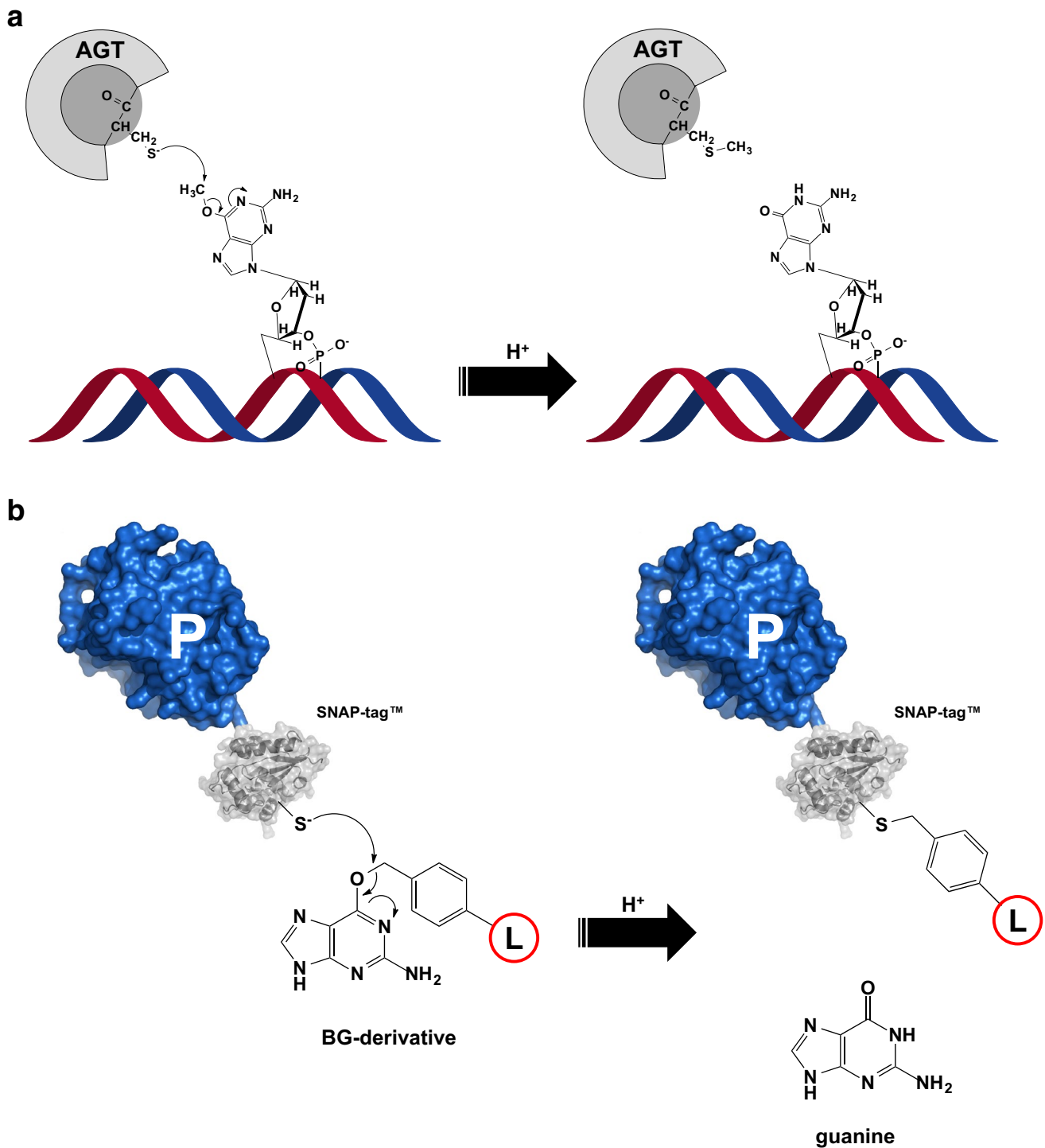


Fig. 1 **a** Direct repair of alkylation DNA damage by AGTs: reaction mechanism of the irreversible transfer of the alkyl group from the damaged guanine to a cysteine in the active site. **b** A variant of

O^6 -alkylguanine-DNA-alkyl-transferases (AGTs or MGMTs, EC: 2.1.1.63) are ubiquitous proteins involved in the direct repair of alkylation damage in DNA. By a one-step SN_2 reaction mechanism, they specifically transfer the alkyl group from the damaged base (O^6 -alkylguanine or

the human AGT (SNAP-tag™) can react in vitro and/or in vivo with BG-derivative substrates, leading to the labeling of a fused protein of interest (*P*) with a desired chemical group (*L*)

O^4 -alkylthymine) to a cysteine residue in their own active site (Daniels et al. 2000, 2004; Fang et al. 2005; Pegg 2011; Tubbs et al. 2007; Yang et al. 2009) (Fig. 1a). AGTs are called suicide or kamikaze proteins because of the 1:1 stoichiometry of their reaction, leading to an irreversible

inactivation and destabilization of the alkylated form of the protein (Fang et al. 2005; Yang et al. 2009). In 2003, the group of Kai Johnsson, studying the human AGT (hAGT), investigated the possibility to employ this protein as protein-tag, using benzyl-guanine-derivative substrates, also accepted by the active site of the protein (Keppler et al. 2003). The covalent and irreversible nature of the alkylated form of hAGT and its small size (ca. 20.0 kDa) allowed new biotechnological approaches, making it an interesting tool for in vivo and in vitro specific labeling of proteins, when fused to a protein of interest (Gautier et al. 2008; Keppler et al. 2003, 2004). A commercial variant of the hAGT (SNAP-tag™, New England Biolabs) is already available to be covalently labelled with a large number of chemical groups conjugated with *O*⁶-benzyl-guanine (BG-derivative) (Fig. 1b).

However, no thermophilic AGTs have been developed as biotechnological tools to date, making the SNAP-tag™ technology suitable only for mesophilic organisms and mild reaction conditions. Hence, a thermostable version of AGT could be useful for the development of protein-tags for organisms growing under harsh reaction conditions. We previously characterized an AGT from the thermoacidophilic archaeon *Sulfolobus solfataricus* (*SsOGT*), which has the same catalytic efficiency of hAGT with fluorescent BG-derivatives, but displays a marked stability over a wide range of temperature, pH, ionic strength and to common denaturing agents (Perugino et al. 2012). Here, we present the use of an engineered variant of the thermostable *SsOGT* protein, which lacks DNA binding activity (Perugino et al. 2012), as a protein-tag for extremophilic organisms.

Materials and methods

Reagents

All chemicals were purchased from Sigma-Aldrich, SNAP-Vista Green™ and CLIP-FI™ fluorescent substrates

(referred to BG-FL and BC-FL throughout, respectively) were from New England Biolabs (Ipswich, MA). Synthetic oligonucleotides were from Primm (Milan, Italy) and listed in Table 1, *Pfu* DNA polymerase were from Stratagene (La Jolla, CA), as well as *Escherichia coli* ABLE C strain. *Thermus thermophilus* HB27^{EC} strain is a derivative of DSM 7039 with enhanced natural competence because of a knockout mutation in the argonaute protein (Swarts et al. 2014). Protein concentration was determined by the BioRad protein assay kit (Bio-Rad Pacific), using BSA as standard.

DNA constructs

The DNA fragment of the beta-glycosidase from *S. solfataricus* gene (*lacS*) was PCR-amplified using the pGEX-K-Gly construct as template (Moracci et al. 1996), and by using the LacS-fwd and LacS-rev oligonucleotides (Table 1), which possess an internal *Pst* I site. This allowed the ligation of this gene downstream and in frame with the *ogtH*⁵ gene in the pQE-*ogtH*⁵ construct (Perugino et al. 2012), and leading to the pQE-*ogtH*⁵-*lacS* plasmid for the heterologous expression of the *SsOGT*-H⁵-*Ssβgl* fusion protein. pMK-*ogtH*⁵ for the heterologous and constitutively expression in *T. thermophilus* of the *SsOGT*-H⁵ mutant (referred to H⁵ throughout) was obtained by multiple rounds of PCR amplification: the PlspA-fwd/OP-rev and PO-fwd/Ogt3' oligonucleotides pairs were first used to amplify the PlspA promoter and the *ogtH*⁵ gene DNA fragments, respectively. In a final round of PCR, the former two DNA fragments were fused to each other by the complementary nature of the PO-fwd and OP-rev oligonucleotides (Table 1), and amplified by using the external PlspA-fwd/Ogt3' oligonucleotides pair. Finally, the obtained DNA fragment was ligated in the multi-cloning site of the pMK184 (Cava et al. 2007) vector by using the *EcoR* I/*Hind* III sites. For all the obtained constructs, regions encoding the cloned genes were verified by DNA sequencing (Primm, Milan, Italy).

Table 1 List of oligonucleotides used in this study

Name	Sequence	Notes
LacS-fwd	5'-CCAGA <u>ACTGCAGATGTACTCATTCCAAATAGC</u> -3'	<i>Pst</i> I site underlined
LacS-rev	5'-CCAGA <u>ACTGCAGTTAGTGCCTTAATGGCTTTACTGG</u> -3'	<i>Pst</i> I site underlined
OP-rev	5'-GGTGATGGTGATGGTGAGATCCTCTC <u>ATatgcctcacacctccttaaggg</u> -3'	PlspA promoter sequence in lowercase; <i>ogt</i> gene sequence in uppercase; <i>Nde</i> I site underlined
PlspA-fwd	5'-aatggagaattcggggatcgcacccggg-3'	<i>EcoR</i> I site underlined
PO-fwd	5'-cccttaaggaggtgtgaggcat <u>ATGAGAGGATCTCACCATCACCATCACC</u> -3'	PlspA promoter sequence in lowercase; <i>ogt</i> gene sequence in uppercase; <i>Nde</i> I site underlined
Ogt-3'	5'-TATTAAGCTTTTATTCTGGTATTTGACTCCCTCTAGT-3'	Perugino et al. (2012); <i>Hind</i> III site underlined

Protein expression and purification

H⁵ and H⁵-Ssβgly were expressed in *E. coli* ABLE C and purified by His₆-tag affinity chromatography, as described (Perugino et al. 2012, 2015), with the exception of a further purification step for the fusion protein: briefly, to remove *E. coli* contaminants, the pool of the eluted fractions from the affinity chromatography was incubated for 20 min at 70 °C, and centrifuged for 30 min at 30,000g. The soluble fraction was dialysed against PBS 1× buffer (phosphate buffer 20 mM, NaCl 150 mM, pH 7.3) and aliquots stored at −20 °C. To assess the purity of the protein samples and determine their concentrations, SDS-PAGE and Bio-Rad protein assay were performed, respectively.

In vitro and in vivo alkyl-transferase assay

The activity of the purified proteins wild-type SsOGT and H⁵, alone or in fusion with Ssβgly, was measured by a fluorescent-based assay using BG-FL and BC-FL substrates in standard conditions, as previously described (Miggiano et al. 2013; Perugino et al. 2012, 2015). Briefly, in a total volume of 10 μL, 5.0 μM of protein (0.1 mg mL⁻¹) was incubated with 20 μM of the substrate in 1× Fluo Reaction Buffer (50 mM phosphate, 0.1 M NaCl, 1.0 mM DTT, pH 6.5) at different temperatures and times, as indicated. Reactions were stopped by denaturing and loading samples on SDS-PAGE, followed by fluorescence imaging analysis on a VersaDoc 4000™ system (Bio-Rad) by applying as excitation/emission parameters a blue LED/530 bandpass filter, respectively. For the determination of the second-order rate constants, assuming the irreversible mechanism with 1:1 substrate/enzyme binding, fluorescence intensity data were corrected for the amount of loaded protein (by coomassie staining analysis), in order to estimate the relative amount of covalently modified protein in time-course experiments. (Gautier et al. 2008; Miggiano et al. 2013; Perugino et al. 2012, 2015). For the in vivo assay, transformed cells of *T. thermophilus* HB27^{EC} strain with the pMK184 and pMK-*ogtH*⁵ plasmids were grown at 60 °C in TB selective medium (tryptone 8 g L⁻¹, yeast extract 4 g L⁻¹, NaCl 3 g L⁻¹, kanamycin 30 mg L⁻¹; in mineral water, pH 7.5) as late as stationary phase (O.D._{600 nm} >1.5) (Cava et al. 2009). Cell pellets from an appropriate volume (typically 1.0 mL) were resuspended in 0.1 mL of TB medium in the presence of 3.0 μM of BG-FL and incubated at several temperatures for different times, as indicated. After the reaction, cells were first washed twice with 1.0 mL of TB medium, then denatured for 10 min at 110 °C in O'Farrell 1× buffer supplemented

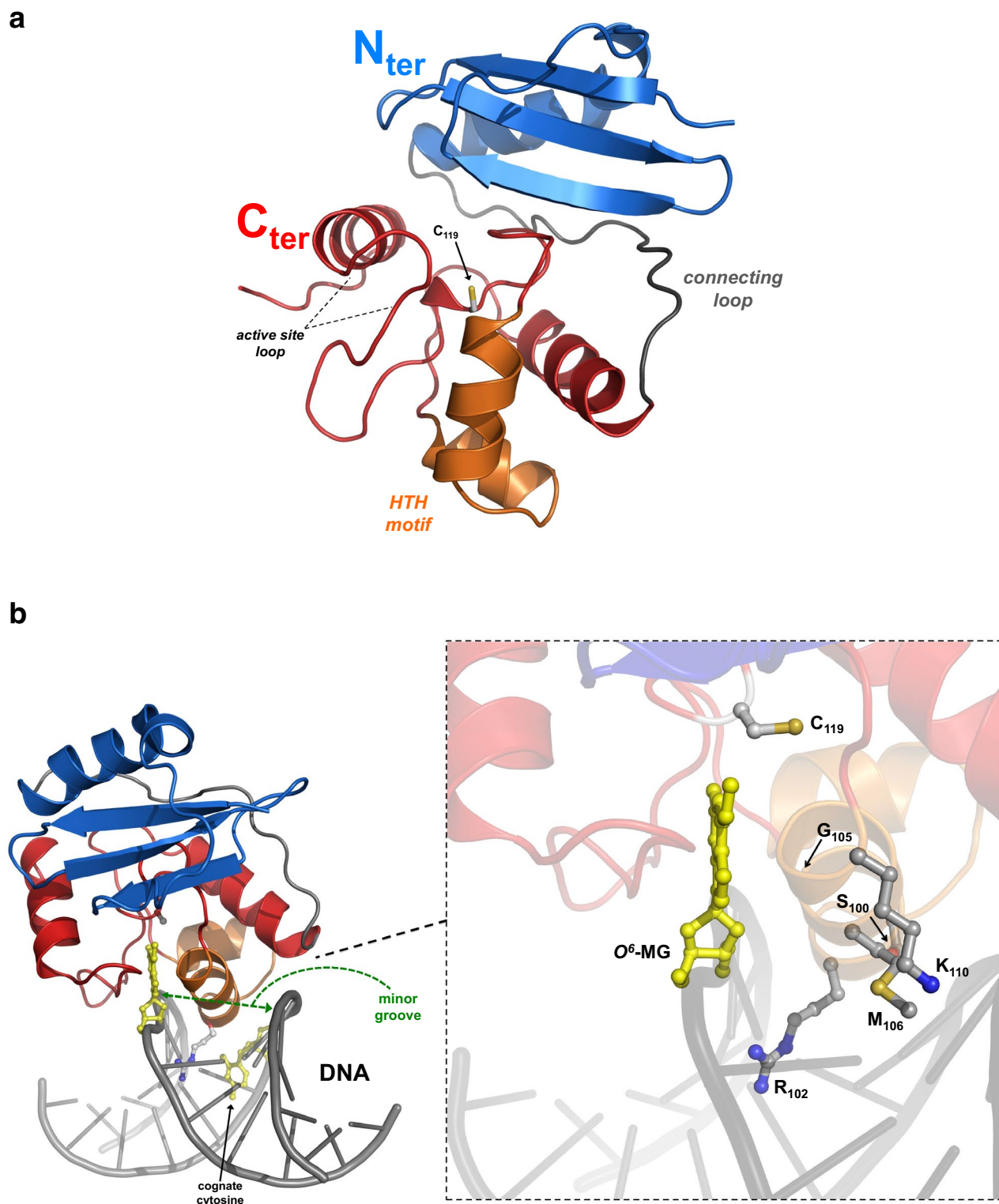
with EDTA 10 mM, and finally loaded on SDS-PAGE for the analysis as described above.

Protein stability analysis

The stability of the wild-type and H⁵ was determined by Differential Scan Fluorimetry (DSF), following an adaptation of a protocol previously described (Niesen et al. 2007; Perugino et al. 2015): samples containing 25 μM of each protein (0.5 mg mL⁻¹) in PBS 1× buffer and SYPRO Orange dye 1× were subjected to a temperature scanning from 20 to 95 °C at 0.2 °C min⁻¹ (5 min per cycle with an increase of 1 °C per cycle) in a Real-Time Light Cycler™ (Bio-Rad). Relative fluorescence data were normalized to the maximum fluorescence value within each scan. Finally, plots of fluorescence intensity as a function of temperature showed increasing fluorescence sigmoidal curves (which described a two-state transition), leading to the determination of the inflection points, indicating T_m values, as described by the Boltzmann equation (Niesen et al. 2007). Stability was also tested by protease attack resistance and by in vivo protein degradation in *T. thermophilus* cells. In the first analysis, in a total volume of 10 μL, 1.0 μM of each protein (0.02 mg mL⁻¹) was incubated for 2 h at 25 °C in the presence of trypsin protease at different substrate/enzyme weight ratio, as indicated. Reactions were then denatured for 5 min at 100 °C in O'Farrell 1× buffer and immediately loaded on SDS-PAGE (0.2 μg per lane). Gels were subjected to Western Blotting analysis by using an anti-SsOGT antibody, as previously described (Perugino et al. 2012). For the preparation of the fluoresceinated form of the wild-type and H⁵, before the trypsin treatment, a reaction with BG-FL for 2.0 h at 25 °C was performed, followed by a gel fluorescence imaging to assess the complete labeling. To determine the stability of alkylated SsOGT in *T. thermophilus* HB27^{EC} strain, we followed the same procedure described above to assay the protein activity; after incubation with BG-FL and washing, cells were incubated at 70 °C for different times, followed by denaturation, SDS-PAGE and fluorescence imaging analysis.

Beta-glycosidase assay

Glycoside hydrolytic activity of H⁵-Ssβgly was detected in vivo on LB Agar selective medium in the presence of the chromogenic substrate X-gal (Sambrook et al. 1989), as well as by calculating the steady-state kinetic constants with 4-nitrophenyl-β-D-galactopyranoside (4Np-gal) and 4-nitrophenyl-β-D-glucopyranoside (4Np-glc), by using a Cary E1 spectrophotometer (Varian) and following the



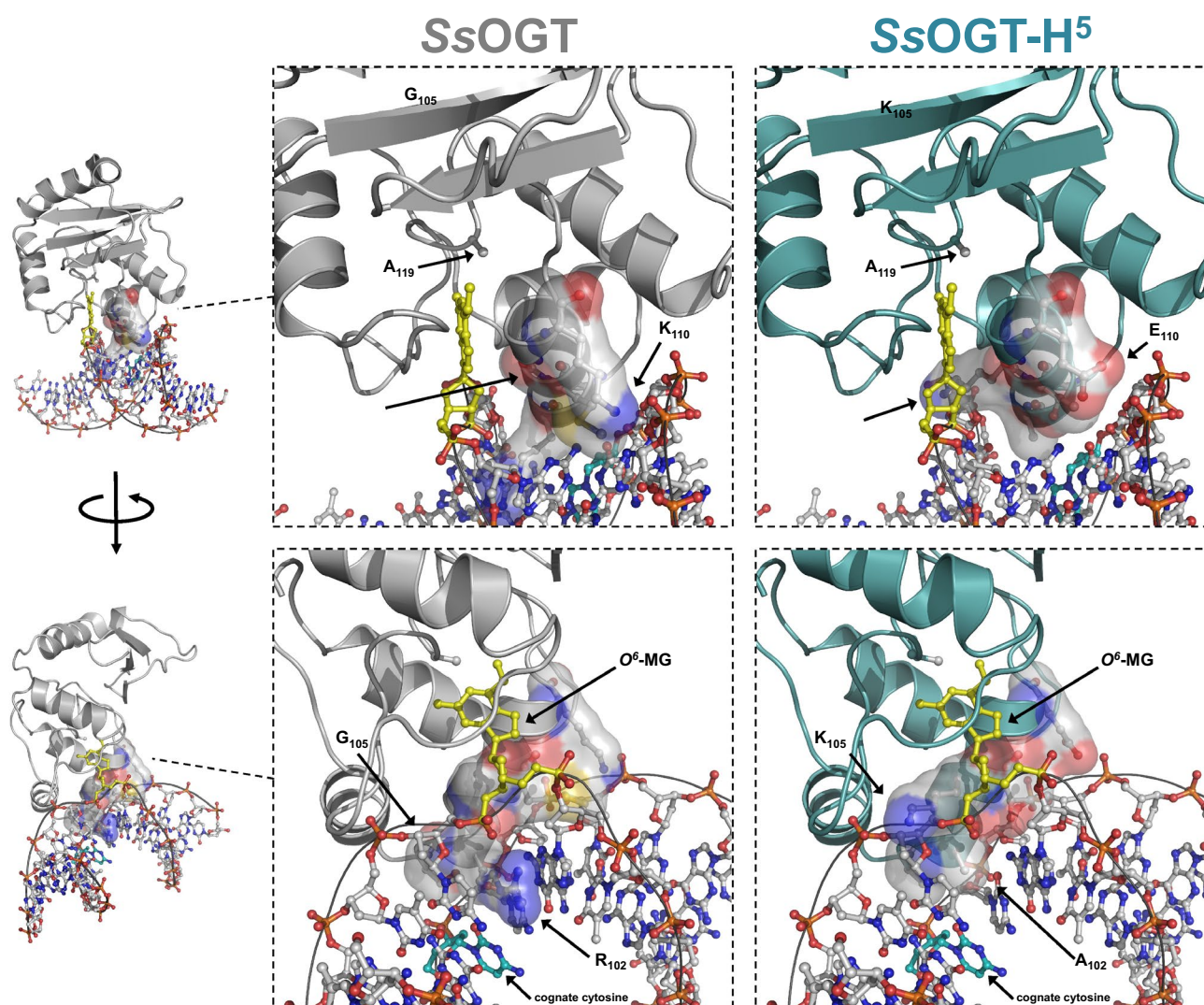


Fig. 3 Comparison of the crystal structure of methylated-DNA::SsOGT-C119A complex of the SsOGT-C119A (gray) with a 3D model of the same DNA molecule in complex with the H⁵ mutant (cyan): DNA, O⁶-MG and residues involved in the mutagenesis are represented as ball and sticks. A spheric representation in transpar-

ency in the SsOGT-C119A and H⁵ for the five residues involved in the mutagenesis is also shown. Atoms are colored by the CPK convention, whereas O⁶-MG and the unpaired cytosine are colored in yellow and cyan, respectively

same procedure used for the Ssβgly enzyme (Moracci et al. 1996, 1998).

Data analysis and softwares

Analysis of the solved 3D structures of the wild-type and the methylated DNA::SsOGT-C119A complex, as well as the modeling of the methylated DNA::H⁵ complex, were performed by using the MacPyMOL freeware (DeLano Scientific FCC). Data from activity and stability assays were fitted to appropriate equations by using the Prism Software Package (GraphPad Software) and GraFit 5.0 Data Analysis Software (Erithacus Software).

Results

Mutations at the HTH motif affect the DNA-binding by SsOGT

The 3D structure of SsOGT, recently solved by X-ray crystallography, is overall very similar to that of all other AGTs and consists of two domains connected by a loop (Perugino et al. 2015) (Fig. 2a). The role of the N_{ter} domain of these proteins is not completely understood and is likely involved in the protein overall stability (Perugino et al. 2015) as well as activity (Miggiano et al. 2013), whereas the highly conserved C_{ter} domain contains all the elements for the

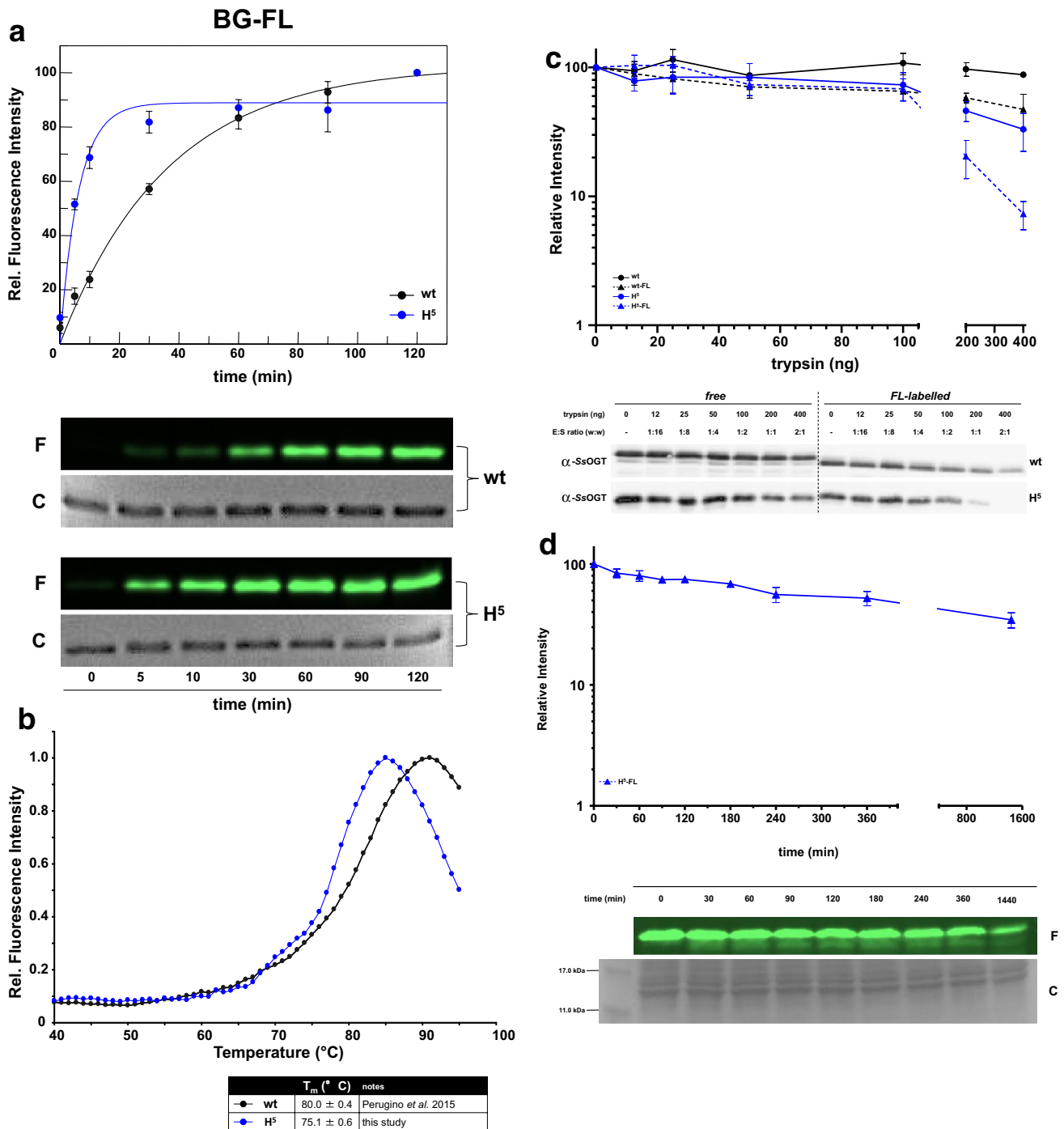


Fig. 4 **a** Time-course of wild-type (Perugino et al. 2012) and H⁵ under standard conditions with 20 μM of SNAP-Vista Green™ (BG-FL) at 25 °C. Analysis and quantification of protein activity, as well as second-order rate constant values determination were performed as previously described (Perugino et al. 2012), and shown in Table 2. **b** Stability of H⁵ in comparison with the wild-type protein by DSF analysis, as described in “Materials and methods”. T_m values were obtained by measuring the relative fluorescence intensity of each protein from three independent experiments, as a function of temperature. **c** Resistance to protease attack of the wild-type SsOGT and H⁵: proteins (ca. 0.2 μg per lane) were incubated for 2 h at 25 °C in the

presence of different amounts of trypsin, as indicated. After SDS-PAGE, gels were analyzed by fluorescence imaging and then subjected to western blot analysis using a polyclonal anti-SsOGT antibody (Perugino et al. 2012). Fluorescein-labelled forms of wild-type and H⁵ are indicated as wt-FL and H⁵-FL, respectively. **d** In vivo stability of fluoresceinated H⁵: after protein labeling procedure with BG-FL (see “Materials and methods”), intact *T. thermophilus* cells were incubated at 70 °C for different times as indicated. Acrylamide gel was analyzed by fluorescence imaging and coomassie staining (to correct fluorescence for the amount of loaded proteins)

Table 2 Second order rate constant values of wild-type and H⁵ under standard conditions on the BG-FL substrate

<i>T</i> (°C)	<i>k</i> (s ⁻¹ M ⁻¹)	References
Wild-type		
70	5.33 × 10 ⁴	Perugino et al. (2012)
25	2.80 × 10 ³	Perugino et al. (2012)
H ⁵		
70	3.78 × 10 ⁴	Perugino et al. (2012)
25	1.60 × 10 ⁴	This study

Incubation temperatures are indicated

recognition and the repair of the damaged guanine. In contrast to most DNA binding proteins, AGTs contact the DNA via the minor groove through their helix turn helix (HTH) motif (Daniels et al. 2004) (Fig. 2b).

We previously prepared a mutant version of SsOGT (H⁵), whose mutations in five conserved residues of the HTH motif (S100A, R102A, G105K, M106T, K110E; zoom in Fig. 2b) impair dramatically the DNA binding capacity of the protein (Perugino et al. 2012). The availability of the crystal structure of the methylated DNA bound SsOGT (Perugino et al. 2015) allowed us to obtain a reliable model of the H⁵-DNA complex 3D structure (Fig. 3), which showed that: i) the substitution of the positively charged K110 with a glutamic acid residue could result in repulsive effects toward DNA phosphate backbone; ii) the replacement of the arginine finger R102 by an alanine residue might hamper the stabilization of the transiently unpaired cytosine opposite the methylated G (Kanugula et al. 1995; Liu et al. 2007; Perugino et al. 2012; Zang et al. 2005); iii) the lysine in place of the G105 residue could produce a dramatic steric hindrance and collide with the phosphate backbone of the DNA strand close to the O⁶-MG. This analysis provided a structural explanation of the complete inability of H⁵ to bind DNA, which is an important pre-requisite for its application as a protein-tag.

Activity and stability of H⁵

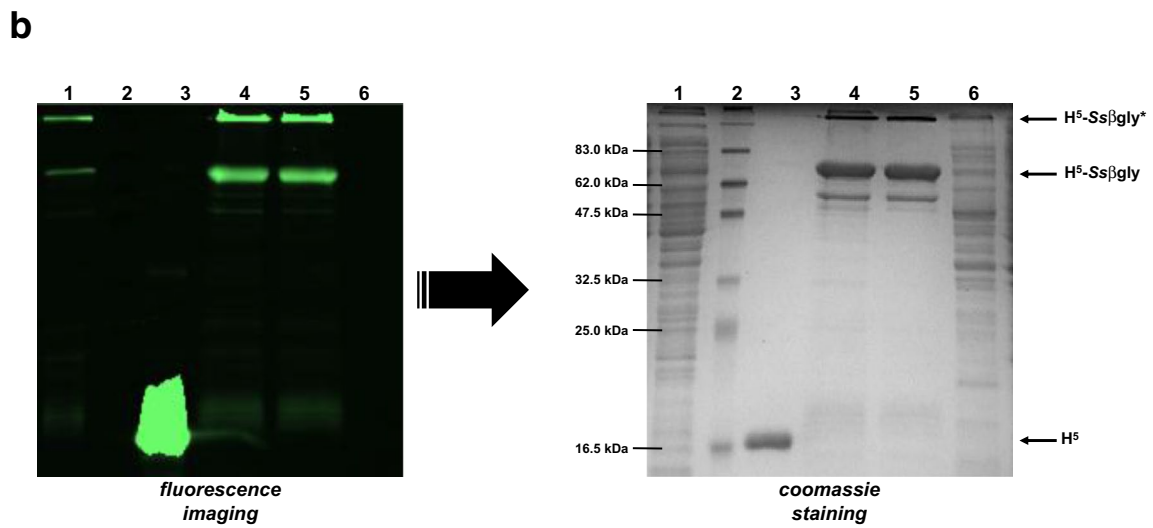
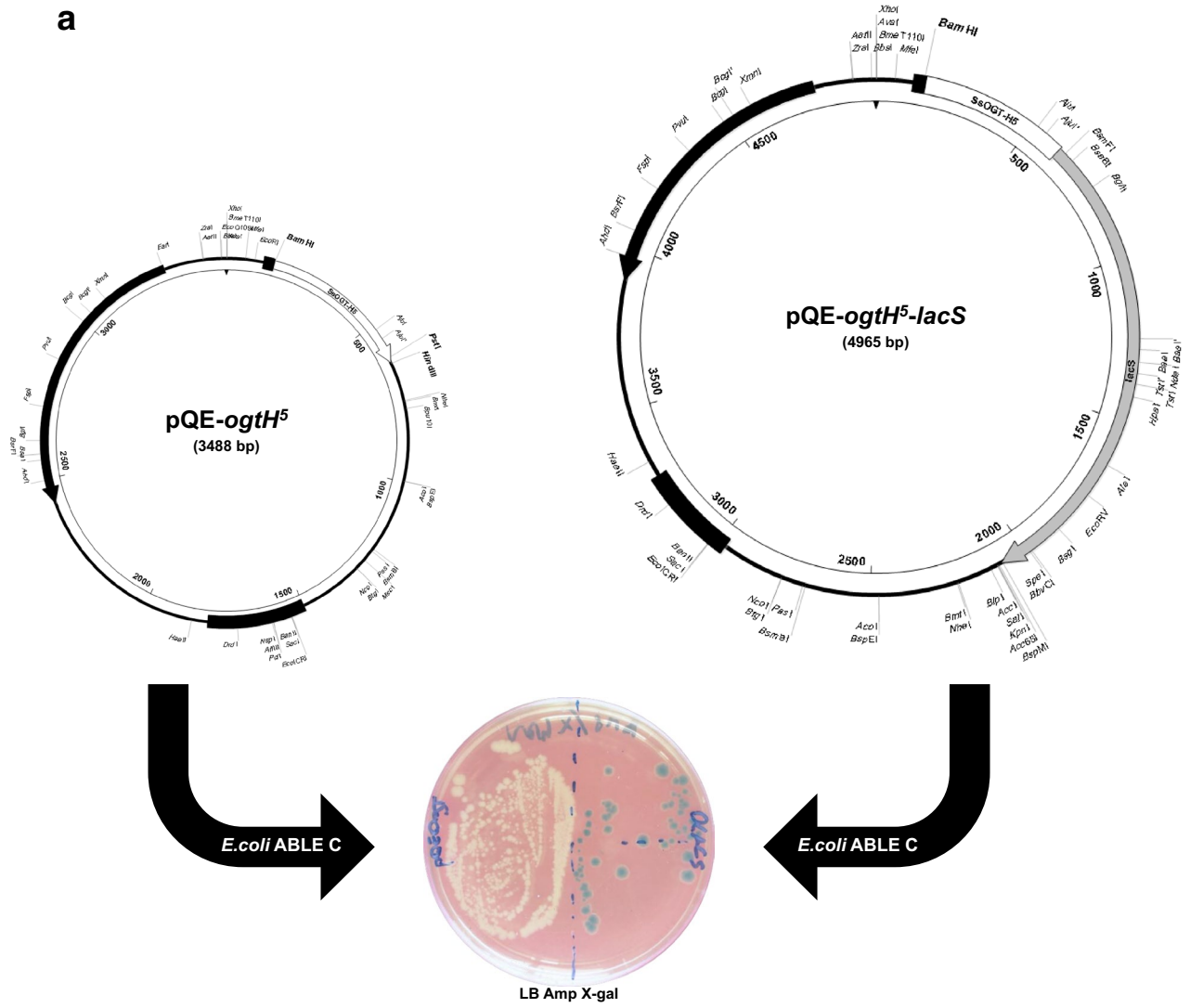
Previous biochemical analysis suggested that H⁵ might be a good candidate thermostable protein-tag because: i) it retained the catalytic activity on the fluorescent BG-FL substrate, and, like SsOGT, showed significant activity over a broad range of reaction conditions (Perugino et al. 2012); ii) the absence of any structural Zn²⁺ ion in the SsOGT structure (Perugino et al. 2015), makes it insensitive to the presence of chelating agents (EDTA up to 10 mM; Perugino et al. 2012). Furthermore, H⁵ displays a significant increase of the activity at any temperature tested, if compared to the wild-type SsOGT (Perugino et al. 2012): the second-order rate constant value at 25 °C on a

Fig. 5 **a** *E. coli* ABLE C strain was transformed by using pQE-*ogtH*⁵ (Perugino et al. 2012) and the pQE-*ogtH*⁵-*lacS* plasmids. Transformed cells were plated on LB agar in the presence of ampicillin and the X-gal chromogenic substrate. **b** SDS-PAGE of the expression and purification by His₆-tag affinity chromatography of the H⁵-Ssβgly fusion protein: 1 cell free extract of *E. coli* ABLE C/pQE-*ogtH*⁵-*lacS*; 2 protein molecular weight marker; 3 2.0 μg of H⁵ protein; 4 pool of fractions eluted by imidazole; 5 soluble fraction of the sample of lane 4, after thermal treatment (see “Materials and methods”); 6 whole *E. coli* ABLE C cells transformed with the pQE31TM empty vector (Qiagen). All samples, except for the protein marker, were incubated with the BG-FL substrate under standard conditions, and loaded on SDS-PAGE. After, the gel was exposed for the fluorescence imaging analysis (left) and then stained by coomassie blue (right), as described in “Materials and Methods”. The high molecular weight band marked by an asterisk correspond to a partially denatured form of the fusion protein, which is particularly resistant to denaturation, as previously shown for the tetrameric Ssβgly (Moracci et al. 1996, 1998)

BG-derivative substrate was one order of magnitude higher than the wild-type (Fig. 4a; Table 2). This result suggests that the replaced residues in the HTH motif mutated in H⁵ play a role in the overall flexibility of the protein during the protein-substrate complex formation.

The substrate specificity of the wild-type and H⁵ proteins was evaluated by reactions in the presence of a benzyl-cytosine-derivative substrate (BC-FL): the failure of complete labeling hampered the determination of second-order rate constants (data not shown), confirming that alkylated cytosines are poor substrates for natural AGTs.

Previous experiments showed that H⁵ was only marginally less stable than the wild-type (Perugino et al. 2012). To compare the stability of the H⁵ protein with that of the SNAP-tagTM, we performed quantitative analysis by DSF. We first followed the protocol used for SNAP-tagTM (ramping of 1 %: 1 min per cycle with an increase of 1 °C per cycle; Mollwitz et al. 2012), leading to a *T_m* value for the SNAP-tagTM protein of ca. 67 °C. However, these conditions were not sufficient to induce the H⁵ protein denaturation, since no fluorescence change was obtained up to 95 °C, as also observed for the wild-type protein (data not shown). Melting temperature values were achieved only increasing the incubation times by heating samples from 20 to 95 °C with a ramping of 0.2 % (5 min per cycle with an increase of 1 °C per cycle). Under these conditions, we calculate a *T_m* value of 75 °C (Fig. 4b; Perugino et al. 2015). Although this value is comparable to that obtained for SNAP-tagTM (67 °C; Mollwitz et al. 2012), the stronger conditions used in the DSF analysis for H⁵ clearly indicate that this protein is more stable than its mesophilic counterpart. Stability was also analyzed by the intrinsic resistance to protease attack of free and labelled form of both wild-type and H⁵. An incubation for 2 h at 25 °C in the presence of different amounts of trypsin protease was performed (Fig. 4c). Due to their intrinsic stability, both proteins were unaffected by the action of the trypsin in most of the conditions tested (from 1:16 to 1:2 *E:S* ratio), notably higher than



those used for hAGT (Kanugula et al. 1998; Mollwitz et al. 2012). To observe a reasonable reduction of the western blot signals, a very high amount of protease was needed (1:1 and 2:1 *E:S* ratio): as expected, alkylated (fluoresceinated) forms of both proteins were less stable than their unlabelled counterparts. Again, a slight difference between the wild-type and H⁵ proteins was found at high *E:S* ratios; overall data confirmed the general stability and resistance to protease attack of these thermostable AGTs.

Application of H⁵ as protein-tag

In order to test the possibility to use H⁵ as a protein-tag, we used as a model the *S. solfataricus lacS* gene coding for a thermophilic β -glycosidase (*Ss* β gly; Moracci et al. 1996); previous studies demonstrated that tags linked at the N-terminus of *Ss* β gly (such as a GST-tag) did not affect the overall folding of the latter (Moracci et al. 1996, 1998). The *lacS* gene was fused downstream to and in frame with the *ogtH⁵* gene, for the expression of the H⁵-*Ss* β gly fusion protein in the *E. coli* ABLE C *lac*⁽⁻⁾ strain. Blue colonies on X-gal LB agar solid medium appeared, indicating the presence of a hexogenous β -galactosidase activity, and confirming that *Ss* β gly is active also as fusion protein even at mesophilic temperatures (Fig. 5a). Purification of the fusion protein by affinity chromatography through a nickel column was performed, exploiting the presence of a His₆-tag at the N-terminus of the H⁵ moiety. Interestingly, the presence of *Ss* β gly did not interfere with the H⁵ activity, making it possible to follow all purification steps by SDS-PAGE and fluorescence imaging of the samples, previously incubated with the BG-FL substrate. Finally, we were able to calculate the kinetic constants of the hydrolytic activity of H⁵-*Ss* β gly at high temperatures, which were comparable to those previously described for the *Ss* β gly enzyme (Table 3) (Moracci et al. 1996). Importantly, the utilization of the thermostable protein-tag

enabled us to proceed with a thermoprecipitation treatment to further purify the pool of affinity chromatography eluted fractions (Fig. 5b, lane 5), and to perform kinetic analysis on a thermophilic enzyme at high temperatures without the need to remove the tag.

Heterologous expression of H⁵ in *T. thermophilus*

Thermus thermophilus HB27^{EC} was chosen as model thermophilic cell host to test our protein-tag because it is a naturally *agt*⁽⁻⁾ strain, thus lacking any endogenous DNA-alkyl-transferase activity. It encodes a single AGT-like protein (ATL), which is inactive in DNA repair, but involved in the recruiting of nucleotide excision repair (NER) proteins (Morita et al. 2008). The *ogtH⁵* gene was cloned in the pMK184 plasmid (Fig. 6a) (Cava et al. 2007), a shuttle vector containing: i) both *E. coli* and *T. thermophilus* HB27 replication origins; ii) the *kat* gene for growth in kanamycin selective medium; iii) a multi-cloning site (MCS) upstream of the α -*lacZ* gene, allowing *E. coli* white/blue colony screening on X-gal containing plates. As in *E. coli* cells (Perugino et al. 2012), it was possible to check the in vivo activity of H⁵, by directly incubating intact cells in the TB medium supplemented with the BG-FL substrate. To test the permeability of *T. thermophilus* (a bacterium with a complex cell envelope, including an outer membrane) to this substrate, we incubated at different temperatures and also in the presence of non-toxic organic solvents, such as DMSO, to improve cell permeabilization. Figure 6b shows fluorescent signals at the same molecular weight of the purified H⁵ protein in whole cell extracts from *Tth*HB27/pMK-*ogtH⁵* transformants, whereas no signals were seen in cells transformed with the empty plasmid (*Tth*HB27/pMK184), confirming the lack of any endogenous alkyl-transferase activity. The maximal activity was obtained at 60 and 70 °C, likely due to the higher permeability of the cells at their physiological temperatures concomitant with an increase of activity of the thermophilic H⁵ protein (Fig. 6, lanes 5, 6 and 9). The presence of organic solvents impaired the protein activity and/or the substrate permeability, as suggested by the absence of the unreacted substrate at the bottom of the gel at the highest concentration of DMSO (lane 8 in Fig. 6). As for most of model organisms used, BG-FL molecule resulted non-toxic for *T. thermophilus* HB27^{EC} cells, as indicated by a TB agar plate (“spot”) assay (data not shown) (Wilkinson et al. 2012).

Moreover, the fluoresceinated form of H⁵ was reasonably stable in vivo in *T. thermophilus* cells, showing a slow fluorescence signal decay: ca. 50 % of the protein was still detectable after 24 h at high temperatures (Fig. 4d), a time span compatible with most of the experiments performed by using the SNAP-tagTM technology in vivo.

Table 3 Steady state kinetic constants of *Ss* β gly and the H⁵-*Ss* β gly fusion protein for the hydrolysis of chromogenic substrates at 65 °C in 50 mM phosphate buffer pH 6.5

Substrate	K_M (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_M (s ⁻¹ mM ⁻¹)	References
<i>Ss</i> β gly				
4Np-gal	1.17 ± 0.06	275.0 ± 4.0	235.0	Moracci et al. (1996)
4Np-glc	0.30 ± 0.04	240.0 ± 7.0	800.0	Moracci et al. (1996)
H ⁵ - <i>Ss</i> β gly				
4Np-gal	2.13 ± 0.50	205.2 ± 14.4	96.5	This study
4Np-glc	0.04 ± 0.02	88.5 ± 7.1	2328.2	This study

if conjugated to a benzyl-guanine, a classical inhibitor of hAGT and a suitable substrate of its commercially available variant (Hinner and Johnsson 2010).

The discovery of a BG-sensitive thermostable OGT from the thermoacidophilic Archaea *S. solfataricus* opened the possibility to widen the SNAP-tag™ technology to organisms which thrive in extreme environments. Because of the intrinsic nature of proteins from hyperthermophiles, SsOGT displayed a strong stability under several harsh conditions, as extremes of pH, ionic strength, temperatures, presence of organic solvents (Perugino et al. 2012), and protease treatments. Mutagenesis at the expense of the HTH motif allowed the construction of a variant of this enzyme (H⁵), which retained most of the advantages of the native protein in terms of stability, and showed even enhanced activity at lower temperatures. We showed that H⁵ is correctly folded, expressed, functionally active and stable in both mesophilic *E. coli* and thermophilic *T. thermophilus* hosts. In order to obtain a smaller version of SsOGT, a preliminary attempt to reduce the wt protein by limited proteolysis was performed, obtaining a truncated polypeptide (ca. 14.0 kDa) which was still active in the presence of BG-FL substrate. Gel purification and N-terminal sequence analysis revealed the loss of the first 36 aminoacids. However, this “mini SsOGT” protein could not be obtained by heterologous expression in *E. coli* (data not shown), suggesting that the first 36 aminoacids are important for the correct folding of the protein during translation. Interestingly, inspection of the 3D structure revealed the presence of a disulfide bond at the N_{ter} (C²⁹-C³¹), which is important for the protein stability (Perugino et al. 2015).

The analysis of the H⁵-Ssbgly fusion protein demonstrated that it is possible to study both mesophilic and thermophilic proteins/enzymes fused to H⁵ under their own physiological conditions, without the need to remove the tag. These properties, together with the complete abolition of binding to DNA, make H⁵ a robust alternative as protein-tag for the application of the SNAP-tag™ technology for in vivo (protein–protein interaction, in situ localization, FRET experiment in combination with thermostable GFP variants, etc.) and in vitro studies (heterologous expression and purification of proteins of interest in extremophilic organisms). Finally, the high substrate specificity toward BG-derivatives makes H⁵ an interesting starting point to be modified by molecular evolution in order to obtain a thermostable variant active on BC-derivative substrates (like the commercial CLIP-tag™): this orthogonal substrate specificity will allow simultaneously and specifically the labeling of different molecular targets in living cells (Gautier et al. 2008).

Acknowledgments We gratefully thank Castrese Morrone for help in some experiments. This work was supported by: (1) Short

Term Mobility Program of the National Research Council of Italy; (2) FIRB-Futuro in Ricerca RBFR12001G_002-Nematic and Merit RBNE08YFN3-Molecular Oncology; (3) BIO2013-44953-R Project.

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
- Ashby MC, Ibaraki K, Henley JM (2004) It's green outside: tracking cell surface proteins with pH-sensitive GFP. *Trends Neurosci* 27:257–261
- Campbell TN, Choy FYM (2000) The effect of pH on green fluorescent protein: a brief review. *Mol Biol Today* 2:1–4
- Cava F, Laptenko O, Borukhov S, Chahlafl Z, Blas-Galindo E, Gómez-Puertas P, Berenguer J (2007) Control of the respiratory metabolism of *Thermus thermophilus* by the nitrate respiration conjugative element NCE. *Mol Microbiol* 64:630–646
- Cava F, de Pedro MA, Blas-Galindo E, Waldo GS, Westblade LF, Berenguer J (2008) Expression and use of superfolder green fluorescent protein at high temperatures in vivo: a tool to study extreme thermophile biology. *Environ Microbiol* 10:605–613
- Cava F, Hidalgo A, Berenguer J (2009) *Thermus thermophilus* as biological model. *Extremophiles* 13:213–231
- Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC (1994) Green fluorescent protein as a marker for gene expression. *Science* 5148:802–805
- 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
- Fang Q, Kanugula S, Pegg AE (2005) Function of domains of human O⁶-alkyl-guanine-DNA alkyltransferase. *Biochemistry* 44:15396–15405
- 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 O⁶-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Prot Eng Des Sel* 19:309–316
- Hinner MJ, Johnsson K (2010) How to obtain labeled proteins and what to do with them. *Curr Opin Biotechnol* 21:766–776
- Kanugula S, Goodtzova K, Edara S, Pegg AE (1995) Alteration of arginine-128 to alanine abolishes the ability of human O⁶-alkyl-guanine-DNA alkyltransferase to repair methylated DNA but has no effect on its reaction with O⁶-benzylguanine. *Biochemistry* 34:7113–7119
- Kanugula S, Goodtzova K, Pegg AE (1998) Probing of conformational changes in human O⁶-alkylguanine-DNA alkyl transferase protein in its alkylated and DNA-bound states by limited proteolysis. *Biochem J* 329:545–550
- Keppeler 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
- Keppeler A, Pick H, Arrivoli C, Vogel H, Johnsson K (2004) Labeling of fusion proteins with synthetic fluorophores in live cells. *Proc Natl Acad Sci USA* 10:9955–9959
- Liu L, Watanabe K, Fang Q, Williams KM, Guengerich FP, Pegg AE (2007) Effect of alterations of key active site residues in O⁶-alkylguanine-DNA alkyltransferase on its ability to modulate

- the genotoxicity of 1,2-dibromoethane. *Chem Res Toxicol* 20:155–163
- Miggiano R, Casazza V, Garavaglia S, Ciaramella M, Perugino G, Rizzi M, Rossi F (2013) Biochemical and structural studies of the *Mycobacterium tuberculosis* O^6 -methylguanine methyltransferase and mutated variants. *J Bacteriol* 195:2728–2736
- Mollwitz B, Brunk E, Schmitt S, Pojer F, Bannwarth M, Schiltz M, Rothlisberger U, Johnsson K (2012) Directed evolution of the suicide protein O^6 -alkylguanine-DNA alkyltransferase for increased reactivity results in an alkylated protein with exceptional stability. *Biochemistry* 51:986–994
- Moracci M, Capalbo L, Ciaramella M, Rossi M (1996) Identification of two glutamic acid residues essential for catalysis in the β -glycosidase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Protein Eng* 12:1191–1195
- Moracci M, Trincone A, Perugino G, Ciaramella M, Rossi M (1998) Restoration of the activity of active-site mutants of the hyperthermophilic beta-glycosidase from *Sulfolobus solfataricus*: dependence of the mechanism on the action of external nucleophiles. *Biochemistry* 37:17262–17270
- Morita R, Nakagawa N, Kuramitsu S, Masui R (2008) An O^6 -methylguanine-DNA methyltransferase-like protein from *Thermus thermophilus* interacts with a nucleotide excision repair protein. *J Biochem* 144:267–277
- 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
- Pédelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotech* 24:79–88
- 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
- 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. *Nucl Ac Res*. doi:10.1093/nar/gkv774
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory Press, New York, pp 1.123–1.125
- Swarts DC, Jore MM, Westra ER, Zhu Y, Janssen JH, Snijders AP, Wang Y, Patel DJ, Berenguer J, Brouns SJJ, van der Oost J (2014) DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* 507:258–261
- 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^6 -alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair* 6:1100–1115
- Wilkinson OJ, Latypov V, Tubbs JL, Millington CL, Morita R, Blackburn H, Marriott A, McGown G, Thorncroft M, Watson AJ, Connolly BA, Grasby JA, Masui R, Hunter CA, Tainer JA, Margison GP, Williams DM (2012) Alkyltransferase-like protein (At11) distinguishes alkylated guanines for DNA repair using cation- π interactions. *Proc Natl Acad Sci* 109:18755–18760
- Yang CG, Garcia K, He C (2009) Damage detection and base flipping in direct DNA alkylation repair. *ChemBioChem* 10:417–423
- Zang H, Fang Q, Pegg AE, Guengerich FP (2005) Kinetic analysis of steps in the repair of damaged DNA by human O^6 -alkylguanine-DNA alkyltransferase. *J Biol Chem* 280:30873–30881