Biochemical and Biophysical Research Communications 500 (2018) 698–703

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Crystal structure of a thermophilic O^6-alkylguanine-DNA alkyltransferase-derived self-labeling protein-tag in covalent complex with a fluorescent probe

Franca Rossi [a], Castrese Morrone [a], Alberto Massarotti [a], Davide Maria Ferraris [a, c], Anna Valenti [b], Giuseppe Perugino [b, **], Riccardo Miggiano [a, c, *]

[a] DSF-Dipartimento di Scienze del Farmaco, University of Piemonte Orientale, Via Bovio 6, 28100 Novara, Italy
[b] Institute of Biosciences and BioResources, National Research Council of Italy, Via Pietro Castellino 111, 80131 Naples, Italy
[c] IXTAL srl, via Bovio 6, 28100, Novara, Italy

ARTICLE INFO

Article history:
Received 12 April 2018
Accepted 17 April 2018
Available online 30 April 2018

Keywords:
O^6-alkylguanine-DNA alkyltransferase
Self-labeling protein tag
Fluorescent probe
Extremophiles
Crystal structure

ABSTRACT

The self-labeling protein tags are robust and versatile tools for studying different molecular aspects of cell biology. In order to be suitable for a wide spectrum of experimental conditions, it is mandatory that these systems are stable after the fluorescent labeling reaction and do not alter the properties of the fusion partner. *Ss*OGT-H^5 is an engineered variant alkylguanine-DNA-alkyl-transferase (OGT) of the hyperthermophilic archaeon *Sulfolobus solfataricus*, and it represents an alternative solution to the SNAP-tag® technology under harsh reaction conditions.

Here we present the crystal structure of *Ss*OGT-H^5 in complex with the fluorescent probe SNAP-Vista Green® (*Ss*OGT-H^5-SVG) that reveals the conformation adopted by the protein upon the trans-alkylation reaction with the substrate, which is observed covalently bound to the catalytic cysteine residue. Moreover, we identify the amino acids that contribute to both the overall protein stability in the post-reaction state and the coordination of the fluorescent moiety stretching-out from the protein active site. We gained new insights in the conformational changes possibly occurring to the OGT proteins upon reaction with modified guanine base bearing bulky adducts; indeed, our structural analysis reveals an unprecedented conformation of the active site loop that is likely to trigger protein destabilization and consequent degradation. Interestingly, the SVG moiety plays a key role in restoring the interaction between the N- and C-terminal domains of the protein that is lost following the new conformation adopted by the active site loop in the *Ss*OGT-H^5-SVG structure. Molecular dynamics simulations provide further information into the dynamics of *Ss*OGT-H^5-SVG structure, highlighting the role of the fluorescent ligand in keeping the protein stable after the trans-alkylation reaction.

1. Introduction

The fluorescent labeling of proteins is a powerful approach to study any aspects of the biology of the cell, including protein function, localization, trafficking and macromolecular interaction networks, both *in vitro* and *in vivo*. In many cases, the labeling is achieved by expressing a chimera, in which the protein of interest is fused to a polypeptide (*i.e.* tag) that behaves as the traceable moiety for protein imaging [1,2]. The self-labeling protein tags specifically and covalently bind different small-molecule fluorophores, a property that finds a huge number of applications [3–7]. A well-characterized example is the SNAP-tag® (New England Biolabs) [8], an engineered form of the hAGT, responsible for alkylated-DNA direct repair [9,10]. AGTs (EC: 2.1.1.63; alternative names: MGMT or OGT) are small suicidal proteins that perform the stoichiometric transfer of the alkyl group from the O^6 position of an alkylated-

Abbreviations: OGT/AGT, O^6-alkylguanine-DNA alkyltransferase; hAGT, human O^6-alkylguanine-DNA alkyltransferase; BG, benzylguanine; SVG, SNAP-Vista Green™ reagent; HTH, Helix-Turn-Helix.

* Corresponding author. DSF-Dipartimento di Scienze del Farmaco, University of Piemonte Orientale, Via Bovio 6, 28100 Novara, Italy.
** Corresponding author.
E-mail addresses: franca.rossi@uniupo.it (F. Rossi), castrese.morrone@uniupo.it (C. Morrone), alberto.massarotti@uniupo.it (A. Massarotti), davide.ferraris@uniupo.it (D.M. Ferraris), anna.valenti@ibbr.cnr.it (A. Valenti), giuseppe.perugino@ibbr.cnr.it (G. Perugino), riccardo.miggiano@uniupo.it (R. Miggiano).

https://doi.org/10.1016/j.bbrc.2018.04.139
0006-291X/

guanine (or from the O^4 position of an alkylated-thymine) to the catalytic cysteine residue of the protein active site [11]. The single-step trans-alkylation reaction brings the base back to the original state, while leaving the protein permanently inactivated and more prone to destabilization and degradation [12,13]. The SNAP-tag® was obtained through a directed protein evolution approach, by introducing 19 amino acid substitutions in the hAGT sequence and by deleting the 25 residues at the C-terminus of the protein [9,10,14]. This engineered protein reacts with BG derivatives bearing a fluorescent substituting group at their O^6 position, such as the SVG reagent [9]. Compared to the parental hAGT the SNAP-tag® displays an enhanced activity and a remarkably improved stability both *in vitro* and when expressed inside the cell. Conversely, the SNAP-tag® does not possess any significant DNA-binding or alkylated-DNA repairing activity, thus satisfying a mandatory requirement to avoid an improper localization or an altered function of the SNAP-tag-fused protein of interest [10].

The SNAP-tag® is generally used for studying mesophilic species [10], whereas a thermostable OGT-based protein-tag could offer a greater exploitability to characterize organisms which are exposed to harsh growth conditions, and a superior intrinsic robustness in *in vitro* experiments requiring a wide range of temperature, pH, ionic strength and the presence of common denaturing agents [15,16]. To this end, an engineered variant of the *Sulfolobus solfataricus* OGT was previously prepared (hereon indicated as *Ss*OGT-H^5), which is characterized by the presence of a number of amino acid substitutions in the catalytic C-terminal domain of the protein. In detail, 5 mutated residues map at the DNA HTH motif (*i.e.* S100A, R102A, G105K, M106T and K110E), and on residue maps at the C-side of the active site loop (S132E), thus only partially matching the point-mutation profile of the hAGT-derived SNAP-tag® [10,15,16]. Nonetheless, thus mutagenesis approach completely abolished the *Ss*OGT-H^5 variant capability to bind the DNA, while preserved the protein self-labeling efficiency in the presence of the SVG probe [15,16]. The SsOGT-H5 properly works as a self-labeling protein-tag both in the mesophilic organism *E. coli*, as well as in thermophilic species as the bacterium *Thermus thermophilus* and the archaeaon *Sulfolobus islandicus* [16,17].

Although the SNAP-tag® and the *Ss*OGT-H^5 proteins have been extensively characterized, a structural snapshot of the protein architecture upon the trans-alkylation reaction with a synthetic probe was missing. We solved and report here the crystal structure of the *Ss*OGT-H^5 protein in covalent complex with the O^6-substituting group of the SVG reagent (abbreviated as *Ss*OGT-H^5-SVG). Beside representing the first crystal structure of a AGT-derived self-labeling protein-tag in complex with a fluorescent probe, our data provide a structural overview of the amino acids that participate in the coordination of the bulky adduct in the post-reaction state of the protein, including an unprecedented role of residues belonging to the poorly functionally characterized N-terminal domain. Moreover, by performing molecular dynamics simulations on the novel *Ss*OGT-H^5-SVG structure, we gained insights into how the overall stability of the *Ss*OGT-H^5 protein could be affected by the presence of the fluorescent moiety in the protein active site. Overall, our analysis could be used for further improving the selectivity of the corresponding *Ss*OGT-H^5-tag/fluorescent substrate system.

2. Material and methods

2.1. Chemicals

All reagents were obtained from Sigma-Aldrich unless otherwise specified.

2.2. Protein expression and purification

The *Ss*OGT-H^5 variant was expressed in the *E. coli* ABLE-C strain and purified as previously described [15,16].

2.3. Crystallization and data collection

In order to prepare the self-labeled *Ss*OGT-H^5 protein to be used in crystallization experiments, a solution of the commercial SVG reagent (250 μM) was mixed with the *Ss*OGT-H^5 protein solution (10 mg/mL in 20 mM phosphate buffer, pH 7.3 and 150 mM NaCl) at a protein:SVG molar ratio of 1:1, and incubated overnight at 4 °C before crystallization trials. The initial crystallization conditions were identified by means of a robot-assisted (Oryx4; Douglas Instruments) sitting-drop-based spare-matrix strategy, using screen kits from Hampton Research and Qiagen, by the vapor diffusion technique.

The *Ss*OGT-H^5-SVG crystals grew to their maximum dimensions in two months at 4 °C in a drop obtained by mixing equal volumes of a 7 mg/mL labeled-protein solution and 2.0 M ammonium sulfate, in a final droplet volume of 1 μL. Single crystals were cryo-protected in the precipitant solution containing 12% glycerol, mounted in a cryo-loop, and flash frozen in liquid nitrogen at 100 K for subsequent X-ray diffraction analysis. The best crystal diffracted at 2.0 Å of resolution at the ID23-2 synchrotron radiation ($\lambda = 0.87$ Å) (European Synchrotron Radiation Facility [ESRF], Grenoble, France). The diffraction data were indexed with XDS program [18], whose indexing score assigned the crystal to the hexagonal space-group P6322 with the cell dimension a = 63.78 Å, b = 63.78 Å, c = 159.04 Å, containing 1 molecule per asymmetric unit, with a corresponding solvent content of 50% and a Matthews coefficient of 2.46 [19]. Further data analysis was carried out using the CCP4 program suite [20]; in particular, the diffraction intensities were scaled and merged with the programs SCALA and TRUNCATE, respectively.

The structure of the *Ss*OGT-H^5-SVG protein was solved by molecular replacement using the program Phaser of the PHENIX software suite [21,22] and the structure of the wild type *Ss*OGT protein as the search model (Protein Data Bank [PDB] accession number: 4ZYE). The SVG scaffold to be fitted in the correspondent electron density was drawn with ACD/ChemSketch chemical drawing package, whereas the ligand coordinates and restraints were generated using ELBOW of the PHENIX software suite [22]. The resulting electron density map was of good quality, allowing manual model building, using the program Coot [23]. The programs PHENIX and Refmac [20,22] were used for crystallographic refinement and to add water molecules. Data collection and refinement statistics are summarized in Table 1. The atomic coordinates and structure factors of the H^5-SVG have been deposited in the PDB (http://www.rcsb.org/) under the accession number 6GA0.

All figures have been generated using PyMol (https://www.pymol.org).

2.4. Molecular dynamics (MD)

The atomic coordinates of the *Ss*OGT-H^5-SVG protein were used for the initial state of the MD simulations, either including or omitting the SVG molecule coordinates. The MD protocol was performed as previously reported [24]. Two series of simulations were carried out: at a constant temperature of 353 K, since the *Ss*OGT-H^5 is a hyperthermophilic protein, and at a constant temperature of 500 K, to analyze the protein unfolding.