BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

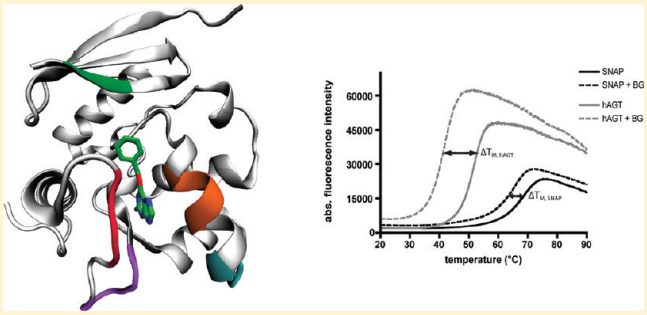
Directed Evolution of the Suicide Protein O^6-Alkylguanine-DNA Alkyltransferase for Increased Reactivity Results in an Alkylated Protein with Exceptional Stability

Birgit Mollwitz,[†] Elizabeth Brunk,[‡] Simone Schmitt,[†] Florence Pojer,[§] Michael Bannwarth,[†] Marc Schiltz,[∥] Ursula Rothlisberger,[‡] and Kai Johnsson*[,†]

[†]Laboratory of Protein Engineering, Institute of Chemical Sciences and Engineering, [‡]Laboratory of Computational Chemistry and Biochemistry, Institute of Chemical Sciences and Engineering, [§]Chair of Microbial Pathogenesis, Global Health Institute, School of Life Sciences, and [∥]Laboratory of Crystallography, Institute of Physics of Biological Systems, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Supporting Information

ABSTRACT: Here we present a biophysical, structural, and computational analysis of the directed evolution of the human DNA repair protein O^6-alkylguanine-DNA alkyltransferase (hAGT) into SNAP-tag, a self-labeling protein tag. Evolution of hAGT led not only to increased protein activity but also to higher stability, especially of the alkylated protein, suggesting that the reactivity of the suicide enzyme can be influenced by stabilizing the product of the irreversible reaction. Whereas wild-type hAGT is rapidly degraded in cells after alkyl transfer, the high stability of benzylated SNAP-tag prevents proteolytic degradation. Our data indicate that the intrinstic stability of a key α helix is an important factor in triggering the unfolding and degradation of wild-type hAGT upon alkyl transfer, providing new insights into the structure−function relationship of the DNA repair protein.



The specific labeling of proteins with synthetic probes is a powerful approach for studying protein function. One way to achieve such a specific labeling is based on so-called self-labeling protein tags.[1] In this approach, the protein of interest is expressed as a fusion protein with a peptide or protein (i.e., tag) whose role is to specifically bind to a synthetic probe in vitro or in vivo. A well-established example of a self-labeling protein tag is SNAP-tag.[2] SNAP-tag specifically reacts with substituted O^6-benzylguanine derivatives and thereby permits the labeling of SNAP-tag fusion proteins with a wide variety of different synthetic probes. Recent applications include its use for the analysis of protein complexes,[3] super-resolution microscopy,[4] the identification of protein−protein interactions,[5] drug target identification,[6] and the determination of protein half-life in animals.[7] The appeal of self-labeling tags such as SNAP-tag is the ease with which fusion proteins can be labeled with synthetic probes even in living cells. A conceptual limitation of the approach is the fact that the tag can affect the properties of its fusion partner. It is therefore important that the properties of the tag be as thoroughly characterized as possible.

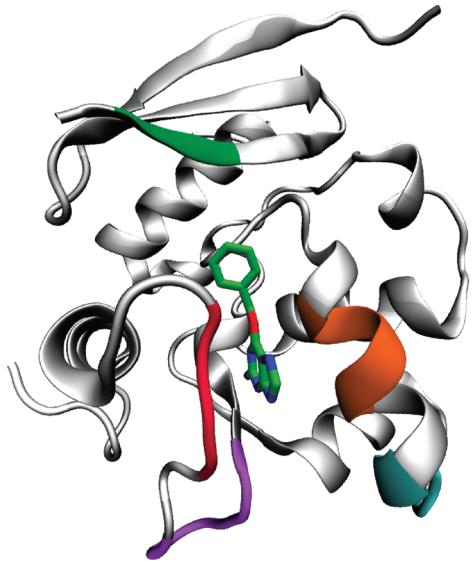
SNAP-tag was generated in a stepwise manner from human O^6-alkylguanine-DNA alkyltransferase (hAGT) by introduction of a total of 19 point mutations (Figure 1) and deletion of 25 C-terminal residues. Saturation mutagenesis of four active-site residues followed by phage display and selection for activity against BG derivatives resulted in ^{GE}AGT, a mutant with 20-fold increased activity toward such substrates (Figure 1B).[8] Subsequent saturation mutagenesis of four additional residues involved in substrate binding followed by phage selection resulted in AGT-54, a mutant with 1.5-fold higher activity than ^{GE}AGT. To further optimize the protein for applications in protein labeling, mutations were introduced to suppress DNA binding and reactivity toward nucleosides, to remove nonessential cysteines, and to truncate the last 25 residues.[9] The resulting mutant ^{M}AGT displayed relatively low activity toward BG derivatives (Figure 1B) (5-fold higher than that of hAGT). To rescue the activity of ^{M}AGT against BG derivatives, an additional round of saturation mutagenesis (residues 150−154 and 32 and 33) followed by phage display was performed, resulting in SNAP-tag.[10] Relative to hAGT, SNAP-tag possesses a 52-fold higher reactivity toward BG derivatives, does not bind to DNA, and is expressed as well in cells as on cell surfaces. However, our understanding of the structure−function relationship of SNAP-tag and how the introduced mutations affect activity is poor.

A unique feature of O^6-alkylguanine-DNA alkyltransferases is that the protein is not regenerated but degraded after DNA

Received: November 2, 2011
Revised: December 22, 2011
Published: January 17, 2012

   dx.doi.org/10.1021/bi2016537 | *Biochemistry* 2012, 51, 986−994

A



B

mutant	reactivity [1]
hAGT	set to 1
^{GE}AGT	20-fold
AGT54	30-fold
^{M}AGT	5-fold
SNAP	52-fold

[1] = relative reactivity with BG-Cy3

C

```
hAGT     MDKDCEMKRT TLDSPLGKLE LSGCEQGLHE IKLLGKGTSA ADAVEVPAPA   50
GEAGT    MDKDCEMKRT TLDSPLGKLE LSGCEQGLHE IKLLGKGTSA ADAVEVPAPA   50
AGT54    MDKDCEMKRT TLDSPLGKLE LSGCEQGLHE IKLLGKGTSA ADAVEVPAPA   50
MAGT     MDKDCEMKRT TLDSPLGKLE LSGCEQGLHE IKLLGKGTSA ADAVEVPAPA   50
SNAP     MDKDCEMKRT TLDSPLGKLE LSGCEQGLHE IIPLGKGTSA ADAVEVPAPA   50

hAGT     AVLGGPEPLM QCTAWLNAYF HQPEAIEEFP VPALHHPVFQ QESFTRQVLW   100
GEAGT    AVLGGPEPLM QCTAWLNAYF HQPEAIEEFP VPALHHPVFQ QESFTRQVLW   100
AGT54    AVLGGPEPLM QCTAWLNAYF HQPEAIEEFP VPALHHPVFQ QESFTRQVLW   100
MAGT     AVLGGPEPLM QATAWLNAYF HQPEAIEEFP VPALHHPVFQ QESFTRQVLW   100
SNAP     AVLGGPEPLM QATAWLNAYF HQPEAIEEFP VPALHHPVFQ QESFTRQVLW   100

hAGT     KLLKVVKFGE VISYQQLAAL AGNPKAARAV GGAMRGNPVP ILIPCHRVVC   150
GEAGT    KLLKVVKFGE VISYQQLAAL AGNPKAARAV GGAMRGNPVP ILIPCHRVVC   150
AGT54    KLLKVVKFGE VISYQQLAAL AGNPKAARAV KTALSGNPVP ILIPCHRVVC   150
MAGT     KLLKVVKFGE VISYSHLAAL AGNPAATAAV KTALSGNPVP ILIPCHRVVN   150
SNAP     KLLKVVKFGE VISYSHLAAL AGNPAATAAV KTALSGNPVP ILIPCHRVVQ   150

hAGT     SSGAVGNYSG GLAVKEWLLA HEGHRLGKPGLG*                     182
GEAGT    SSGAVGCYEG GLAVKEWLLA HEGHRLGKPGLG*                     182
AGT54    SSGAVGCYEG GLAVKEWLLA HEGHRLGKPGLG*                     182
MAGT     INGAVGCYEG GLAVKEWLLA HEGHRLGKPGLG*                     182
SNAP     GDLDVGCYEG GLAVKEWLLA HEGHRLGKPGLG*                     182
```

Figure 1. Overview of the directed evolution of hAGT to SNAP-tag. (A) Crystal structures of the SNAP-tag C145A mutant cocrystallized with BG. Mutagenized parts of the protein are highlighted in color. (B) Relative reactivities of hAGT and mutants with BG-Cy3. (C) Sequence alignment of hAGT with intermediate mutants and SNAP-tag. Colors that highlight mutations correspond to those used in panel A.

repair.[11] In mammalian cells, it is believed that alkyl transfer triggers a conformational change in hAGT, leading to ubiquitination and degradation of the alkylated protein.[12] Supporting this hypothesis is the observation that alkylation of hAGT increases its sensitivity toward proteolysis in vitro.[13] Furthermore, structural analysis of hAGT before and after alkylation revealed that alkylation leads to sterically unfavorable interactions that result in partial unfolding of the protein.[14] Interestingly, an increased level of degradation of SNAP-tag fusion proteins after labeling has not been observed.[7] However, a thorough analysis of how the stability of SNAP-tag is affected by labeling is missing.

Here we report a detailed biochemical and structural analysis of (i) the interaction of SNAP-tag with its substrate, (ii) the factors responsible for its increased reactivity, and (iii) how the labeling affects the stability of the protein. These results will facilitate future applications and improvements of SNAP-tag and shed light on the factors that balance the stability of wild-type hAGT before and after DNA repair.

dx.doi.org/10.1021/bi2016537 | *Biochemistry* 2012, 51, 986−994

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. hAGT and mutants were cloned into a pRSET vector carrying an N-terminal His tag followed by a PreScission Protease cleavage site. The lower primer of SNAP-tag gave rise to two different clones, where one clone contained the point mutation of Pro179 to Arg (CCA was converted to CGA). This mutation eventually facilitated protein crystallization, and this specific clone was used for structure determination. For all other experiments, the Pro179 clone was expressed and purified separately. The plasmids were transformed into *Escherichia coli* BL21 DE3 and directly used to inoculate a 20 mL preculture in LB containing 100 μg/mL ampicillin. The preculture was diluted 50-fold, and cells were grown at 37 °C until an OD of 0.6−0.8 was reached. At this cell density, protein expression was induced by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After 16 h at 18 °C, the cultures were harvested, and the pellets were taken up in PBS supplemented with protease inhibitor cocktail and lysed by sonication. The soluble fraction of the lysate was used for Ni-NTA purification. Protein was eluted in 50 mM K_2HPO_4, 150 mM imidazole, and 10% glycerol (pH 8.0). The eluted protein was incubated overnight with PreScission Protease and 10 mM DTT at 4 °C. Subsequently, it was subjected to a ResQ 6 mL column (GE Healthcare) using buffer A [20 mM Tris and 4 mM DTT (pH 8.0)] and buffer B [20 mM Tris, 4 mM DTT, and 1 M NaCl (pH 8.0)]. The corresponding elution fractions were pooled and concentrated. Homogeneous protein was obtained by size exclusion chromatography (Superdex 200 column, GE Healthcare) using a buffer containing 20 mM Tris, 4 mM DTT, and 200 mM NaCl (pH 8.0). The main fractions of the elution peak were pooled and concentrated.

SNAP-tag Crystallization. Crystallization trials of SNAP-tag were performed with either 8 or 18 mg/mL protein. First, crystals were found with crystallization screen Qiagen/Nextal Classics II Suite [buffer 78, which consists of 0.2 M ammonium acetate, 0.1 M Bis-Tris (pH 5.5), and 25% PEG 3350]. Crystals were optimized and finally produced in a hanging drop with 2 μL of a 12 mg/mL protein solution mixed with 2 μL of buffer [for SNAP-tag, 34% PEG 3350, 100 mM Bis-Tris (pH 5.5), and 500 mM NaCl; for SNAP-tag C145A bound to BG, 32% PEG 3350, 100 mM Bis-Tris (pH 6.5), and 400 mM NaCl; for SNAP-tag bound to BG, 34% PEG 3350, 100 mM Bis-Tris (pH 6.5), and 400 mM NaCl]. Crystals were grown at 18 °C and appeared after 48 h. Diffraction data were produced by the ESRF synchrotron (Grenoble, beamlines ID14EH3 and BM30A). Prior to freezing, crystals were briefly soaked in crystallization buffer and 20% glycerol. Data were processed using Scala[15] (SNAP-tag) and XDS[16](SNAP-tag C145A bound to BG; SNAP-tag benzylated), and structures were determined using Phaser[17](CCP4 Suite) and hAGT [Protein Data Bank (PDB) entry 1EH6] as a search model. Structures were refined using Refmac5.[18] Model visualization and building were performed using Coot.[19] The final structures were verified using PROCHECK and Molprobity.[20,21] Figures were generated using Visual Molecular Dynamics (VMD)[22] and CCP4 Molecular Graphics Program.[23] Coordinates and structure factors for SNAP-tag, benzylated SNAP-tag, and SNAP-tag bound to its substrate BG have been deposited in the PDB as entries 3KZY, 3LOO, and 3KZZ, respectively.

Proteolysis Experiment. Protein solutions of hAGT and mutants were diluted to 1 mg/mL in reaction buffer [20 mM Tris, 4 mM DTT, and 200 mM NaCl (pH 8.0)] and split into two aliquots. One aliquot of each protein was incubated with a 2-fold molar excess of BG-fluorescein (in DMSO) for 1 h at room temperature; to the second aliquot was added the corresponding amount of DMSO. To control if the labeling reaction was complete, a small aliquot was taken and the reaction quenched with a 40-fold excess of BG-Cy5. Samples were analyzed via sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) to confirm complete labeling of the proteins prior to trypsinization (data not shown). Per reaction, 50 μL (50 μg) of protein were mixed with 50 μL of different trypsin dilutions [trypsin stocks of 1.0 mg/mL (1:1 dilution), 100 μg/mL (1:10), 10 μg/mL (1:100), 1.0 μg/mL (1:1000), 0.1 μg/mL (1:10000), and 0 μg/mL]. Samples were incubated for 1 h at room temperature and reactions quenched by the addition of a 6-fold excess of sample buffer and heating for 3 min at 95 °C. Samples were analyzed via SDS−PAGE followed by fluorescence-in-gel scanning on a Pharos FX Molecular Imager (Bio-Rad) and by Coomassie staining.

Thermal Denaturation Assay. Protein was diluted to 25 μM in reaction buffer [20 mM Tris, 4 mM DTT, and 200 mM NaCl (pH 8.0)]. Ten microliters of the solution was mixed with 10 μL of a 30× SYPRO Orange (Invitrogen) stock solution. The experiment was performed in a 96-well plate using MicroAmp Fast 96-Well Reaction Plates (Applied Biosystems) on a 7900 HT Real Time PCR System (Applied Biosystems). Samples were heated from 20 to 95 °C at a ramping of 1%. Fluorescence changes were monitored simultaneously with a charge-coupled device (CDD) camera. Dissociation curves were analyzed by plotting the first derivative of the fluorescence curve (slope) versus temperature ($dF/dT°$) averaged over 4 °C increments. The maximal value of these curves corresponds to the melting temperature (T_M) of the individual proteins.

Intracellular Half-Life Determination. HEK 293 cells were cultured in EX-CELL 293 medium (Sigma) and transfected with the corresponding plasmids according to the standard transfection procedure using polyethyleneimine (PEI). After transfection, cells were cultured for 24 h prior to the experiment. Cells expressing hAGT, the intermediate mutants, or SNAP-tag were labeled for 15 min with 0.65 μM CP-TMRstar (Covalys) and blocked with 500 μM BG. After being extensively washed with BG-containing media (first wash with 500 μM BG for 30 min and later with 100 μM BG in a 10-fold larger volume for 3 h), samples were taken at different time points for protein extraction and SDS−PAGE analysis followed by fluorescence-in-gel scanning (Pharos FX Molecular Imager, Bio-Rad).

Molecular Dynamics Simulations. Computational models of SNAP-tag and wild-type hAGT were constructed, starting from crystallographic nonbenzylated (PDB entries 3KZY and 1EH6, respectively) and benzylated structures (PDB entries 3L00 and 1EH8, respectively) in which the unresolved loop region (residues 36−44 in hAGT and residues 36−49 in SNAP-tag) was added manually. AMBER parm99SB charges and atom types were used to build the topologies for each of the structures with zinc parameters, where distance restraints were used for the coordination of the zinc ion to the cysteine residues in the N-terminal domain. The structures were solvated in a periodically repeated TIP3P water box with dimensions of 81 Å × 80 Å × 70 Å (corresponding to a 15 Å solvation shell around the protein). All structures were minimized, heated to 300 K under constant volume conditions with 5.0 kcal mol^{-1} positional restraints on the protein (except

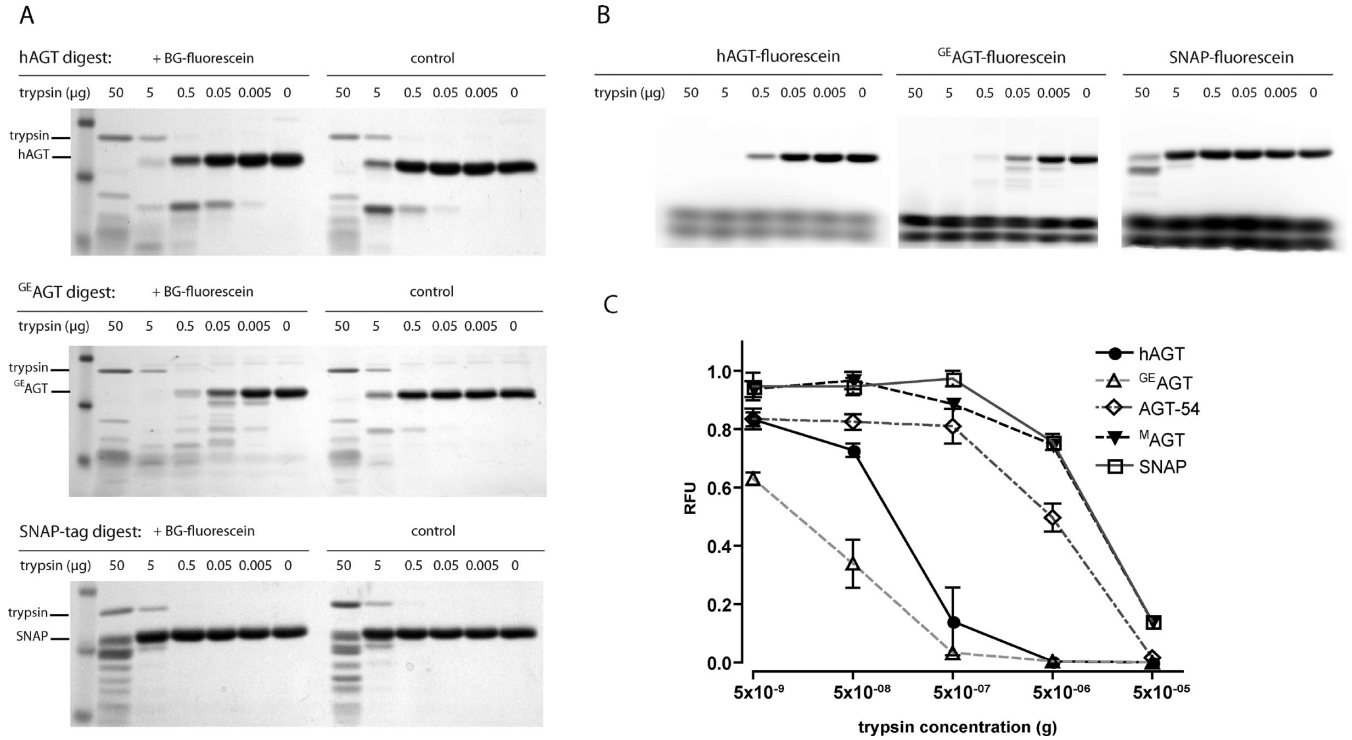


Figure 2. Proteolysis experiment with trypsin. Before and after being labeled with BG-fluorescein (control DMSO), proteins were digested for 1 h with increasing amounts of trypsin. (A) SDS−PAGE of hAGT, ^{GE}AGT, and SNAP-tag. (B) Fluorescence scan of gels used for quantification in panel C. The high-mobility bands correspond to free BG-fluorescein. Different intensities are due to diffusion out of the gel prior to analysis. (C) Quantification of fluorescence intensities of bands in gels of all mutants at different trypsin concentrations.

the unresolved loop), and then equilibrated under constant-temperature and -pressure conditions, slowly releasing the restraints over 4 ns. Data were collected from production phase simulations, in which molecular dynamics (MD) trajectories were run for 30−50 ns using the particle mesh Ewald (PME) MD module in AMBER version 10. During the simulations, the PME method was used with a cutoff of 8.0 Å for nonbonding interactions. Constant-pressure periodic boundary conditions were maintained with a pressure relaxation time of 2 ps. The SHAKE algorithm was used to constrain all bonds that involve hydrogen atoms. The Langevin method for temperature control was used with a collision frequency of 1 ps^{-1}.

Free energy perturbation methods, in combination with thermodynamic integration (FEP/TI), were used within the Simulated Annealing with NMR-derived Energy Restraints (SANDER) module of AMBER version 10. Using a dual-topology paradigm, two topologies (state 0 and state 1) for each transformation were constructed by manually imposing point mutations on the equilibrated MD structures of the free enzymes, SNAP-tag, and hAGT. The electrostatic and Lennard-Jones terms were decoupled by performing three separate alchemical transformations: (i) decharging state 0 using 30−40 λ points, (ii) transforming the atoms of state 0 into those of state 1 using 20 λ points, and (iii) recharging state 1 using 30−40 λ points. Convergence was examined by extending the duration of the MD run and by increasing the number of λ points. In each transformation, all λ points were individually minimized and equilibrated, and data were collected during a production phase of 1 ns. The change in the potential energy as a result of the perturbation was integrated over the λ values to produce the ΔG for each mutation. A per residue decomposition for the charging and decharging steps was performed to indicate the primary points in the protein that are affected by the mutation.

RESULTS

Proteolysis Experiments. It has been reported previously that alkylation of wild-type hAGT increases the susceptibility of the protein to proteases such as trypsin.[13] The increased protease susceptibility can be interpreted as a decreased stability of the alkylated protein. As a first measure of how the stability of hAGT changed in the course of its directed evolution into SNAP-tag, we therefore measured the susceptibility of the different hAGT mutants to trypsin before and after labeling with BG-fluorescein. Proteins were titrated with increasing amounts of trypsin and analyzed via SDS−PAGE by Coomassie staining (Figure 2A) and, in the case of fluorescein-labeled proteins, by fluorescence-in-gel scanning (Figure 2B). As reported previously, hAGT showed increased susceptibility to trypsin, especially upon being labeled with BG-fluorescein (Figure 2A, top). In contrast, no increased sensitivity of SNAP-tag could be observed (Figure 2A, bottom). Furthermore, an almost 100-fold higher trypsin concentration had to be used to degrade SNAP-tag to an extent similar to that of hAGT, indicating an increased stability of SNAP-tag. Mutations introduced into ^{GE}AGT (Asn157Gly and Ser159-Glu) had a destabilizing effect, especially on the labeled protein (Figure 2A, middle). However, the trypsin resistance significantly increased for AGT-54 and ^{M}AGT (Figure 2C). These data showed that during the directed evolution of hAGT for higher reactivity toward BG the stability of the protein toward proteolytic digest, especially after benzylation, increased drastically. It should be noted that the mutations introduced throughout the directed evolution of hAGT might affect the

dx.doi.org/10.1021/bi2016537 | *Biochemistry* 2012, 51, 986−994

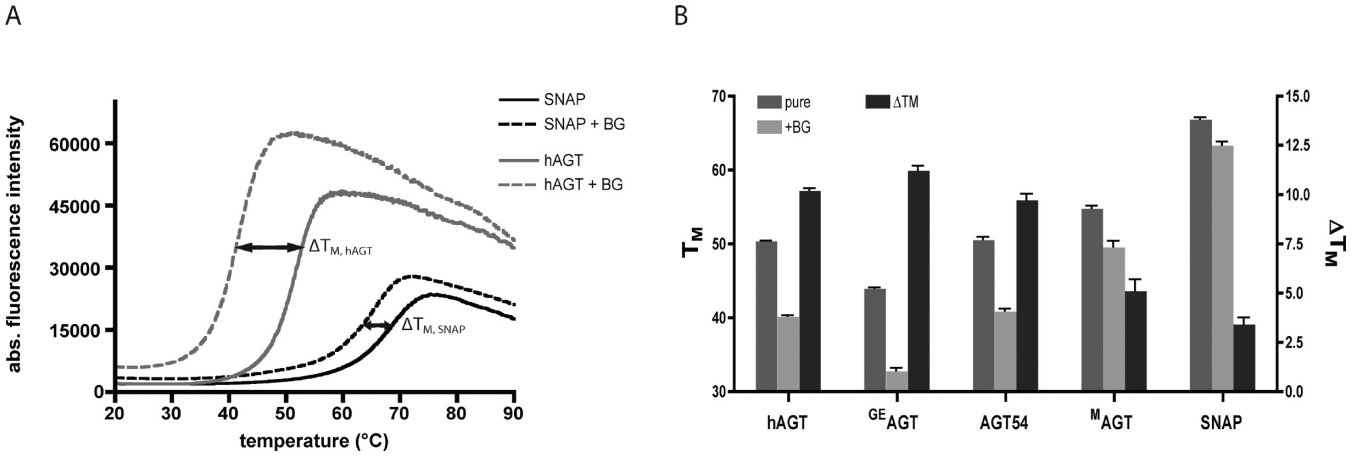


Figure 3. Melting point (T_M) measurement for hAGT, SNAP-tag, and mutants. (A) Raw data of the thermofluor assay presented for hAGT and SNAP-tag before and after reaction with BG. Differences in ΔT_M before and after benzylation are highlighted by arrows. (B) Melting point analysis. Shown are T_M values of pure protein (gray) and benzylated protein (light gray) as well as the ΔT_M before and after benzylation (dark gray).

susceptibility to trypsin not only by changing the stability of the protein but also through the introduction or removal of trypsin cleavage sites. We therefore sought to confirm these results through an independent assay.

Thermal Denaturation Assay. To independently confirm the results obtained in the proteolysis experiment, we evaluated the stability of hAGT and mutants by thermal denaturation.[24] The melting temperatures (T_M) were determined with the help of a fluorogenic molecule (SYPRO Orange) that shows a strong increase in fluorescence upon binding to hydrophobic regions of a protein. The fluorescence intensity reached a maximum and then started to decrease, probably because of precipitation of the complex of the fluorescent probe and the denatured protein.[25] We determined the melting temperatures of all proteins in the presence and absence of BG to assess the impact of benzylation on protein stability (Figure 3). We observed that the T_M of SNAP-tag was increased by 17 °C relative to that of hAGT. The directed evolution for higher reactivity toward BG derivatives led at the same time to an increased stability of the labeled protein relative to the unmodified protein (see ΔT_M in Figure 3A,B). A 10.2 °C decrease in melting temperature was observed for benzylated hAGT, whereas the T_M of SNAP-tag was reduced by only 3.5 °C upon benzylation. These results were in agreement with the data obtained from the proteolysis experiments and could confirm the increased stability throughout protein evolution.

Stability of Labeled hAGT and SNAP-tag in Living Cells. Our data for the sensitivity to proteases and thermal denaturation showed a stability of SNAP-tag significantly increased relative to that of wild-type hAGT, in particular for the labeled protein. As hAGT is known to be rapidly degraded upon alkyl transfer in living cells,[11] we investigated the extent to which the increased in vitro stability of SNAP-tag would also translate into an increased intracellular half-life of the labeled protein. Therefore, we performed pulse−chase experiments with all mutants (Figure 4). Suspension cultures of HEK 293 cells expressing the different mutants were incubated for 15 min with 0.65 μM CP-TMRstar (NEB), and the reaction was quenched with 100 μM BG. Aliquots were taken at distinct time points and analyzed via SDS−PAGE and fluorescence-in-gel scanning (Figure 4A). Labeled hAGT showed a half-life of ~3 h, whereas the apparent half-life of labeled SNAP-tag was determined to be approximately 42 h (Figure 4B). It should be noted that the actual half-life of SNAP-tag should be even higher, since the continuous cell growth results in the dilution of labeled protein, causing a decrease in signal. With respect to the intermediate mutants, ^{GE}AGT showed stability similar to

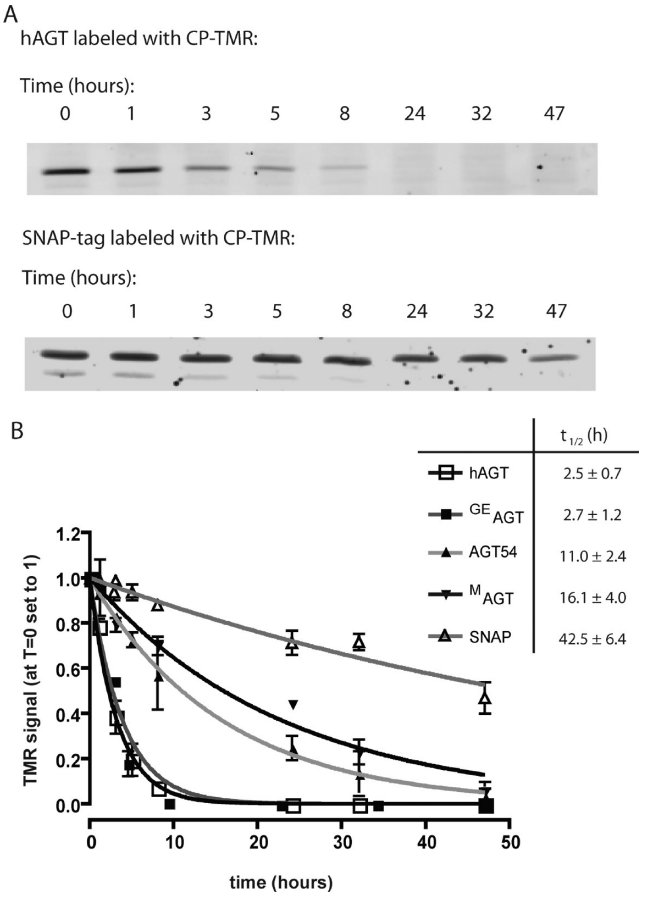


Figure 4. Pulse−chase experiment with hAGT, SNAP-tag, and mutants in HEK 293 cells. (A) Fluorescence-in-gel scan of SDS−PAGE of hAGT and SNAP-tag pulse−chase experiments. Samples were analyzed at indicated time points for the amount of fluorescence-labeled protein present in each sample. Slight proteolytic degradation during sample workup gave rise to an additional minor band at a lower molecular weight. (B) Plot of the relative fluorescence signal intensities for analyzed proteins as a function of time.

dx.doi.org/10.1021/bi2016537 | Biochemistry 2012, 51, 986−994

that of wild-type hAGT (half-life of 2.7 h), AGT-54 showed an increased half-life of 11 h and mutations introduced into ^{M}AGT resulted in a further prolonged half-life of 16 h (Figure 4B). These data clearly show that labeled SNAP-tag is not degraded to a significant degree in living cells and that the directed evolution for higher reactivity resulted in the generation of mutants with increased in vitro stability that translates into an increased intracellular protein half-life.

Structural Analysis of Substrate Binding and Increased Stability. To obtain further insights into how the mutations influenced the reactivity and stability of SNAP-tag, we obtained the crystal structures of (i) SNAP-tag, (ii) SNAP-tag mutant Cys145Ala with BG bound, and (iii) benzylated SNAP-tag. These structures were determined at 1.9, 1.9, and 1.7 Å resolution, respectively (Table S1 of the Supporting Information). The C145A mutation of the active-site cysteine was necessary to prevent the reaction of the protein with its substrate.

Like hAGT, SNAP-tag consists of two domains, and in all three SNAP-tag structures, a Zn^{2+} ion was located in the N-terminal domain. As observed in the hAGT structure, a large flexible loop within the N-terminal domain of SNAP-tag was not resolved in the electron density maps (SNAP-tag, Lys36–Pro49; benzylated SNAP-tag, Gly35–Leu53; SNAP-BG, Gly35–Leu53). A comparison of the backbone structures of hAGT (PDB entry 1EH6) and SNAP-tag (PDB entry 3KZY) gave rise to a root-mean-square deviation (rmsd) of 0.794 Å calculated between Cα positions for residues 1–182 of SNAP-tag and hAGT (EMBL-EBI, Secondary Structure Matching).

Substrate Binding. A comparison of the crystal structures of hAGT and SNAP-tag demonstrates that key features for substrate binding have been conserved during the directed evolution of hAGT to SNAP-tag: Tyr114 and Val148 form hydrogen bonds to the purine, thereby positioning the alkylated base for efficient alkyl transfer in the active site and stabilizing the leaving group during alkyl transfer (Figure 5).[14] As observed for hAGT, an overlay of SNAP-tag with SNAP-tag Cys145Ala cocrystallized with BG shows that the thiol of Cys145 is 3.1 Å from the CH_2 group of the benzyl ring and ideally positioned for alkyl transfer.

However, there are additional interactions in the crystal structure of SNAP C145A with bound BG that explain the improved substrate binding and alkyl transfer in SNAP-tag relative to those of hAGT. Mutations introduced at positions 157 and 159 (see the red segment in Figure 1A,C) increased significantly the reactivity toward BG (see ^{GE}AGT). We speculated previously that Glu159 would form a hydrogen bond with N7 of BG.[8] Indeed, in the crystal structure of SNAP C145A with bound BG, one of the carboxylate oxygens of Glu159 is within 2.7 Å of N7 of BG. In addition, Cβ and Cγ of Glu159 make hydrophobic contacts with the benzyl ring of BG (distance of ~3.7 Å), which is located between the glutamate side chain and Pro140 with an additional edge-on hydrophobic interaction with Tyr158. Further, it is noteworthy that in the structure of SNAP-tag bound with BG the side chain of Glu159 adopts a conformation that is not present in the structure of free SNAP-tag, suggesting that binding of BG fixes the conformation of this residue (Figure 5).

Increased Stability. It has been suggested that after alkyl transfer in hAGT, sterically unfavorable interactions between the benzylic CH_2 group and the carbonyl oxygen of Met134 cause a displacement of α helix residues 127–136.[11] This displacement of the α helix triggers unfolding of the alkylated protein. In the structures of hAGT and benzylated hAGT, the displacement of helix residues 127–136 upon alkylation also manifests itself in an increased distance from the sulfur of Cys145 and benzylated Cys145 to the carbonyl oxygen of Met134 [from 4.2 to 4.6 Å (Figure 6B,D)]. In SNAP-tag, structural changes in the active site prevent unfavorable steric interactions upon benzylation (Figure 6A,C). When the structures of SNAP-tag and benzylated SNAP-tag are overlaid, a displacement of α helix residues 127–136 cannot be detected.

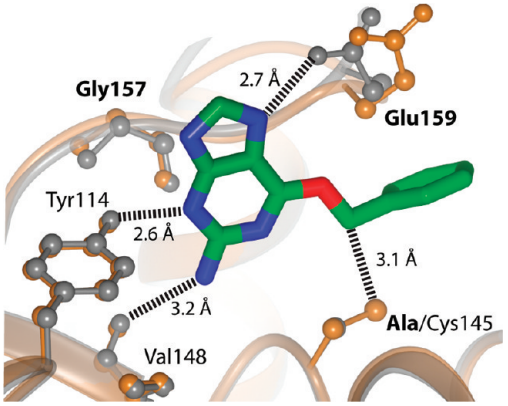


Figure 5. Active site of the SNAP-tag C145A mutant (gray) cocrystallized with BG. Free SNAP-tag (gold) is overlaid to highlight the positioning of Cys145 with respect to BG. Ala145, Gly157, and Glu159 (bold) have been mutated compared to the residues of wild-type hAGT. Tyr114, Val148, and Glu159 form hydrogen bonds with the substrate. In the presence of BG, Glu159 undergoes a change in position and forms a hydrogen bond interaction with BG, and Cβ and Cγ of Glu159 make hydrophobic contacts with the benzyl ring of BG. Cys145 is 3.1 Å from the CH_2 group of the benzyl ring and ideally positioned to perform the alkyl transfer.

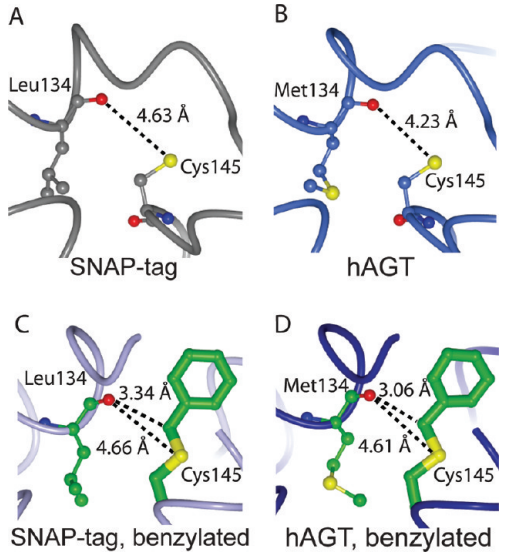


Figure 6. Distances between the carbonyl oxygen of Met134 (in hAGT) or Leu134 (in SNAP-tag) and the sulfur of Cys145 (A and B) or the sulfur of benzylated Cys145 and the CH_2 group of the benzyl ring (C and D). In SNAP-tag, distances upon benzylation remained virtually unchanged as opposed to those of hAGT where the distance increases because of the movement of α helix residues 127–136 and the proximity especially to the Met134 carbonyl group.

dx.doi.org/10.1021/bi2016537 | *Biochemistry* 2012, 51, 986–994

Also, the distance between the sulfur atom of Cys145 and the carbonyl oxygen of Leu134 is 4.6 Å in SNAP-tag, and this distance remains virtually unchanged upon benzylation (Figure 6A,C). The absence of structural information about single-point mutants makes an unambiguous identification of the key residues responsible for this structural change difficult, but we believe that particular mutations in α helix residues 127−136 are important. Mutations introduced at this position of SNAP-tag shorten hydrogen bonds within the α helix considerably (Table S2 of the Supporting Information): the distance between O131 and N135 is 3.0 Å in SNAP-tag but 3.8 Å in hAGT. Glycine residues are known to destabilize α helices,[26,27] and the mutations in SNAP-tag result in a more compact α helix and increased available space in the active site. Not only Gly131 but also Gly132 is well conserved among O^6-alkylguanine-DNA alkyltransferases,[28] and it has been suggested that for steric reasons the flipping of alkylated bases out of double-stranded DNA requires glycines at these positions.[14,29] Our results suggest that Gly131 and Gly132 in hAGT also play an important role in triggering protein unfolding and degradation upon alkyl transfer.

Molecular Dynamics Simulations. To further understand the impact of specific mutations introduced into SNAP-tag on protein stability and reactivity, we used MD methods to study both the benzylated and nonbenzylated systems. For this, we used a method using free energy perturbation in conjugation with thermodynamic integration (FEP/TI) to address specific stabilizing versus destabilizing interactions by determining the contribution of a mutation to the overall change in free energy for each residue. In particular, we focused on the specific mutations at positions (i) 157/159, (ii) 32/33, and (iii) 150−154 (see Figure 1 for the various sites of mutagenesis). The first mutations investigated were Asn157Gly and Ser159Glu. To study the effect of these mutations, we alchemically transformed hAGT into ^{GE}AGT using FEP/TI simulations. We found that this transformation has a destabilizing effect (ΔG of 18 kcal mol^{-1}), which is in line with the experimentally observed decrease in stability in ^{GE}AGT versus hAGT. A per residue decomposition was used to indicate the specific molecular interactions that were most affected as a result of the transformation. The results of the decomposition show that the destabilization is mainly due to a loss of hydrogen bonding interactions between the α helix (residues 127−136) and the loop (residues 157 and 159). The mutation of Ser159 to Glu introduces an interaction with Lys32, situated in the N-terminal β sheet. Because of this interaction, Glu159 in our ^{GE}AGT model is oriented toward the β sheet, as illustrated in Figure 7B. It is noteworthy that this conformation would prevent hydrogen bonding of the glutamate carboxylate with BG. However, Lys32 is mutated to Ile during the directed evolution from ^{M}AGT to SNAP-tag. FEP/TI simulations of this transformation show that this mutation induces a destabilization of the SNAP-tag (by a ΔG of 29 kcal mol^{-1}). The per residue decomposition of the ΔG indicates that the Lys32Ile mutation strongly affects Glu159, because of the loss of interaction toward Lys32 (see Figure 7A and Figure S1 of the Supporting Information for more details on the per residue decomposition). Further, the Lys32Ile mutation strongly affects the neighboring (charged) residues within the N-terminal β sheet. Overall, the picture that emerges from our studies is that the Ser159Glu mutation in ^{GE}AGT increases the reactivity of SNAP-tag because of hydrogen bonding between Glu159 and BG. However, interactions with Lys32 keep Glu159 at least partially in an inactive conformation. The Lys32Ile mutation then increases the reactivity of SNAP-tag by disfavoring the inactive conformation of Glu159.

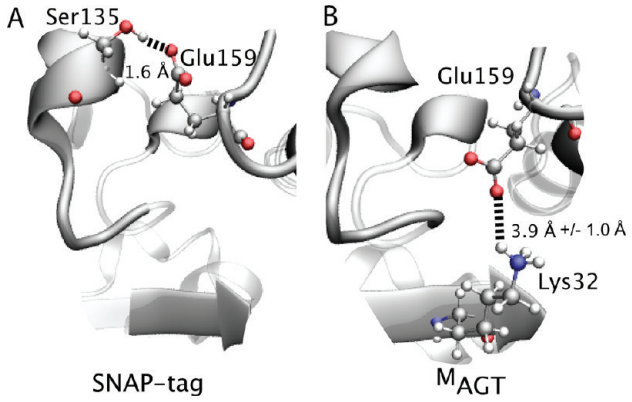


Figure 7. Snapshots of two equilibrated MD trajectories from a FEP/TI simulation for the free protein SNAP-tag in which the residues 32 and 33 have been perturbed. The point mutation at this position represents one of several mutations that are introduced in the directed evolution step from ^{M}AGT to SNAP-tag. (A) In SNAP-tag, Glu159 is oriented toward the α helix and is within hydrogen bonding distance of Ser135, whereas in ^{M}AGT (B), Glu159 is oriented toward the N-terminal β sheet with a possibility for salt bridge formation with Lys32. All distances were averaged during the last 20 ns of the MD trajectory. The distance given in panel B corresponds to the average distance plus its thermal fluctuations measured by the standard deviation of 1 Å.

Analogous studies of the transformation of residues 150−154 demonstrate that SNAP-tag is strongly stabilized over ^{M}AGT (a ΔG of 87 kcal mol^{-1}), compensating for the destabilizing effect of the Lys32Ile mutation, in agreement with the experimentally determined melting temperatures. We performed a per residue decomposition to evaluate the residues that are the strongest contributors to the stabilization of SNAP-tag and observed that the Asn150Gln, Asn152Asp, and Ala154Asp mutations introduced into SNAP-tag contribute the most to the free energy difference. They create a highly structured hydrogen bonding network (see Figure 8) that extends from the α helix (Lys131 hydrogen bonding to Gln150) to the tip of the loop (Gln150 hydrogen bonding to Asp154) to the C-terminus (Asp152 hydrogen bonding to Arg175). As opposed to hAGT, in which only Ser152 is directly involved in the binding to the DNA minor groove, this engineered network has a strong effect on SNAP-tag stability, especially after protein alkylation. The interactions within this hydrogen bonding network are also observed during the MD trajectories and are preserved for 50 ns for both free and benzylated SNAP-tag structures.

■ DISCUSSION

Here we demonstrate that the directed evolution of wild-type hAGT into SNAP-tag not only increased the reactivity of the protein toward BG but also affected the stability of unlabeled and, in particular, labeled SNAP-tag. The increase in stability is remarkable because our directed evolution experiments selected for increased reactivity, but not higher stability. Generally, the evolution of proteins favors the selection of marginal stability with improvements in functionality.[30] It is also known from other directed evolution experiments that activating mutations usually come at the cost of stability.[31,32] As most mutations are destabilizing and because evolution favors the most likely solutions over less likely ones, (directed) evolution generally

dx.doi.org/10.1021/bi2016537 | Biochemistry 2012, 51, 986−994

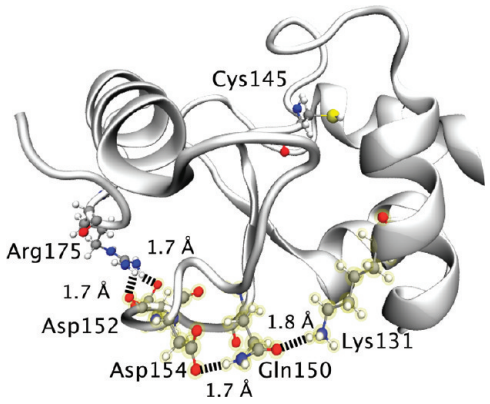


Figure 8. Displayed is a snapshot of the free protein structure of SNAP-tag from an equilibrated MD trajectory in which hydrogen bonding distances were averaged over the last 20 ns of the trajectory. An extended hydrogen bonding network forms as a result of several point mutations, involving residues 131, 150, 152, and 154. The hydrogen bonding network extends from the C-terminal α helix (Lys131) to the loop region (Gln150, Asp152, and Asp154) and finally to the C-terminal Arg175. Such a network strongly reinforces the active site by providing stability to the α helix–loop domain.

favors mutants with marginal stability. We believe that in the case of a suicide protein such as hAGT, the reason for the isolation of mutants with improved stability occurs because some mutations increased reactivity by stabilizing the product of the reaction, the alkylated protein. The demonstrated stability of (labeled) SNAP-tag is an important feature for its use in protein labeling. The observed in vitro stability translates well into increased protein half-life in mammalian cells as no significant SNAP-tag degradation was detected upon labeling. This is in direct contrast to the behavior of wild-type hAGT, which becomes rapidly degraded upon alkylation.

Finally, our analysis of SNAP-tag also provides new insights into the structure–function relationship of AGTs. Gly131 and Gly132 are two highly conserved residues among AGTs residing in an α helix that plays an important role in substrate binding (the so-called recognition helix).[28,29] It has been argued that the role of the glycine residue at this position is to accommodate the natural substrate, an alkylated guanine flipped out of double-stranded DNA. However, mutations of residues 127–136 stabilize the α helix, as evidenced by shortened hydrogen bonds (Table S2 of the Supporting Information). Presumably, these mutations relieve some of the unfavorable interactions occurring upon alkyl transfer and stabilize the benzylated protein. We therefore believe that α helix residues 127–136, especially residues Gly131 and Gly132, play a dual role in AGTs: accommodation of the substrate and contribution to unfolding and degradation of AGT upon alkyl transfer.

In conclusion, our analysis of the directed evolution of wild-type hAGT to SNAP-tag reveals a correlation between the reactivity and stability of this suicide protein. Our insights into the structure–function relationship of SNAP-tag should facilitate future engineering experiments to further improve reactivity or change the specificity of the tag. In addition, our studies provide new insights into the structure–function relationship of AGTs in general.

■ ASSOCIATED CONTENT

Supporting Information

Data collection and refinement statistics of SNAP-tag crystals (Table S1), comparison of hydrogen bonding length in α helix residues 126–137 for hAGT and SNAP in benzylated and nonbenzylated states (Table S2), and a graphical representation of the decomposition of $\Delta\Delta G$ energies (Figure S1). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +41 (0)21 693 93 56. Fax: +41 (0)21 693 93 55. E-mail: kai.johnsson@epfl.ch.

Funding

Supported by the European Commision CORDIS Framework Programme 6, Grant 41455, a Marie Curie Intra Fellowship (S.S.), the Marie Heim-Vögtlin Programme (F.P.), and the Swiss National Science Foundation.

■ ACKNOWLEDGMENTS

We thank Prof. Dr. Michael Groll for the generous gift of expression vectors for His tag fusion proteins and PreScission Protease.

■ ABBREVIATIONS

hAGT, human O^6-alkylguanine-DNA alkyltransferase; BG, O^6-benzylguanine; CP-TMRstar, benzylchloropyrimidine-tetramethylrhodamine; rmsd, root-mean-square deviation.

■ REFERENCES

(1) Hinner, M. J., and Johnsson, K. (2010) How to obtain labeled proteins and what to do with them. *Curr. Opin. Biotechnol. 21*, 766–776.

(2) Gendreizig, S., Kindermann, M., and Johnsson, K. (2003) Induced protein dimerization in vivo through covalent labeling. *J. Am. Chem. Soc. 125*, 14970–14971.

(3) Hoskins, A. A., Friedman, L. J., Gallagher, S. S., Crawford, D. J., Anderson, E. G., Wombacher, R., Ramirez, N., Cornish, V. W., Gelles, J., and Moore, M. J. (2011) Ordered and dynamic assembly of single spliceosomes. *Science 331*, 1289–1295.

(4) Klein, T., Löschberger, A., Proppert, S., Wolter, S., van de Linde, S., and Sauer, M. (2011) Live-cell dSTORM with SNAP-tag fusion proteins. *Nat. Methods 8*, 7–9.

(5) Gautier, A., Nakata, E., Lukinavicius, G., Tan, K.-T., and Johnsson, K. (2009) Selective cross-linking of interacting proteins using self-labeling tags. *J. Am. Chem. Soc. 131*, 17954–17962.

(6) Chidley, C., Haruki, H., Pedersen, M. G., Muller, E., and Johnsson, K. (2011) A yeast-based screen reveals that sulfasalazine inhibits tetrahydrobiopterin biosynthesis. *Nat. Chem. Biol. 7*, 375–383.

(7) Bojkowska, K., Santoni de Sio, F., Barde, I., Offner, S., Verp, S., Heinis, C., Johnsson, K., and Trono, D. (2011) Measuring in vivo protein half-life. *Chem. Biol. 18*, 805–815.

(8) Juillerat, A., Gronemeyer, T., Keppler, A., Gendreizig, S., Pick, H., Vogel, H., and Johnsson, K. (2003) Directed evolution of O6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. *Chem. Biol. 10*, 313–317.

(9) Juillerat, A., Heinis, C., Sielaff, I., Barnikow, J., Jaccard, H., Kunz, B., Terskikh, A., and Johnsson, K. (2005) Engineering substrate specificity of O6-alkylguanine-DNA alkyltransferase for specific protein labeling in living cells. *ChemBioChem 6*, 1263–1269.

(10) Gronemeyer, T., Chidley, C., Juillerat, A., Heinis, C., and Johnsson, K. (2006) Directed evolution of O-6-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Eng., Des. Sel. 19*, 309–316.

(11) Pegg, A. E., Wiest, L., Mummert, C., Stine, L., Moschel, R. C., and Dolan, M. E. (1991) Use of antibodies to human O6-alkylguanine-DNA alkyltransferase to study the content of this protein in cells treated with O6-benzylguanine or N-methyl-N′-nitro-N-nitrosoguanidine. *Carcinogenesis 12*, 1679−1683.
(12) Srivenugopal, K. S., Yuan, X. H., Friedman, H. S., and Ali-Osman, F. (1996) Ubiquitination-dependent proteolysis of O6-methylguanine-DNA methyltransferase in human and murine tumor cells following inactivation with O6-benzylguanine or 1,3-bis(2-chloroethyl)-1-nitrosourea. *Biochemistry 35*, 1328−1334.
(13) Kanugula, S., Goodtzova, K., and Pegg, A. E. (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.
(14) Daniels, D. S., Mol, C. D., Arvai, A. S., Kanugula, S., Pegg, A. E., and Tainer, J. A. (2000) Active and alkylated human AGT structures: A novel zinc site, inhibitor and extrahelical base binding. *EMBO J. 19*, 1719−1730.
(15) Collaborative Computational Project Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D50*, 760−763.
(16) Kabsch, W. (2010) XDS. *Acta Crystallogr. D66*, 125−132.
(17) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658−674.
(18) Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D67*, 355−367.
(19) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126−2132.
(20) Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D66*, 12−21.
(21) Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res. 35*, W375−W383.
(22) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics 14*, 27−38.
(23) McNicholas, S., Potterton, E., Wilson, K. S., and Noble, M. E. M. (2011) Presenting your structures: The CCP4mg molecular-graphics software. *Acta Crystallogr. D67*, 386−394.
(24) Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P., and Salemme, F. R. (2001) High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J. Biomol. Screening 6*, 429−440.
(25) Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N., and Nordlund, P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem. 357*, 289−298.
(26) Javadpour, M. M., Eilers, M., Groesbeek, M., and Smith, S. O. (1999) Helix packing in polytopic membrane proteins: Role of glycine in transmembrane helix association. *Biophys. J. 77*, 1609−1618.
(27) Pace, C. N., and Scholtz, J. M. (1998) A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J. 75*, 422−427.
(28) Loktionova, N. A., and Pegg, A. E. (2002) Interaction of mammalian O(6)-alkylguanine-DNA alkyltransferases with O(6)-benzylguanine. *Biochem. Pharmacol. 63*, 1431−1442.
(29) Daniels, D. S., Woo, T. T., Luu, K. X., Noll, D. M., Clarke, N. D., Pegg, A. E., and Tainer, J. A. (2004) DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat. Struct. Mol. Biol. 11*, 714−720.
(30) Taverna, D. M., and Goldstein, R. A. (2002) Why are proteins marginally stable? *Proteins 46*, 105−109.
(31) Tokuriki, N., Stricher, F., Serrano, L., and Tawfik, D. S. (2008) How Protein Stability and New Functions Trade Off. *PLoS Comput. Biol. 4*, e1000002.
(32) Romero, P. A., and Arnold, F. H. (2009) Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol. 10*, 866−876.