

Xiaofeng Zhu, Xuan Yan,  
Lester G. Carter, Huanting Liu,  
Shirley Graham, Peter J. Coote  
and James Naismith\*

The BSRC, University of St Andrews, North  
Haugh, St Andrews KY16 9ST, Scotland

Correspondence e-mail: naismith@st-and.ac.uk

Received 29 March 2012

Accepted 15 April 2012

**PDB References:** 3-methyladenine DNA  
glycosylase I–3-MeA complex, 4aia; Y16F  
mutant, 4ai5.

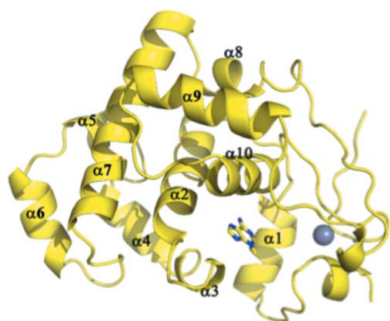
## A model for 3-methyladenine recognition by 3-methyladenine DNA glycosylase I (TAG) from *Staphylococcus aureus*

The removal of chemically damaged DNA bases such as 3-methyladenine (3-MeA) is an essential process in all living organisms and is catalyzed by the enzyme 3-MeA DNA glycosylase I. A key question is how the enzyme selectively recognizes the alkylated 3-MeA over the much more abundant adenine. The crystal structures of native and Y16F-mutant 3-MeA DNA glycosylase I from *Staphylococcus aureus* in complex with 3-MeA are reported to 1.8 and 2.2 Å resolution, respectively. Isothermal titration calorimetry shows that protonation of 3-MeA decreases its binding affinity, confirming previous fluorescence studies that show that charge–charge recognition is not critical for the selection of 3-MeA over adenine. It is hypothesized that the hydrogen-bonding pattern of Glu38 and Tyr16 of 3-MeA DNA glycosylase I with a particular tautomer unique to 3-MeA contributes to recognition and selection.

### 1. Introduction

Bacterial 3-methyladenine DNA glycosylase I (TAG; Forsyth *et al.*, 2002; Ji *et al.*, 2001) is ubiquitous in eubacteria (Supplementary Fig. S1<sup>1</sup>; Drohat *et al.*, 2002) but shows no sequence or structural similarity to mammalian 3-methyladenine DNA glycosylase (AAG; Lau *et al.*, 2000). TAG belongs to the alkylpurine DNA glycosylase superfamily and hydrolyzes the N9–C1' glycosylic bond between a 3-methyladenosine (3-MeA) nucleobase lesion and the deoxyribose ring (Riazuddin & Lindahl, 1978; Bjelland *et al.*, 1993; Fig. 1a). 3-Methylation of adenine does not influence base pairing (Sedgwick *et al.*, 2007); rather, the methyl group blocks replication by interfering with the interactions of DNA polymerase (Sedgwick *et al.*, 2007; Engelward *et al.*, 1996). Like the 8-oxoguanylate DNA glycosylases MutM and hOGG1 (Banerjee *et al.*, 2005, 2006; Banerjee & Verdine, 2006; Blainey *et al.*, 2006), TAG is thought to slide along the duplex until it encounters a lesion. TAG binds flipped-out 3-MeA and then cleaves the damaged base from the ribose. TAG from *Staphylococcus aureus* shares around 40% amino-acid sequence identity with the structurally characterized TAG enzymes from *Salmonella typhi* (Metz *et al.*, 2007) and *Escherichia coli* (Drohat *et al.*, 2002). The crystal structure of the *S. typhi* enzyme complexed with 3-MeA and abasic DNA (Metz *et al.*, 2007) and an NMR structure of the *E. coli* enzyme complexed with 3-MeA (Cao *et al.*, 2003) have been reported. Two absolutely conserved residues, Tyr16 and Glu38, were identified to form hydrogen bonds with 3-MeA and Trp46 stacks with 3-MeA (Cao *et al.*, 2003; Metz *et al.*, 2007). The methyl group does not appear to make extensive contacts. The crystal structure of the apo *S. aureus* enzyme has been reported (Oke *et al.*, 2010). We wished to probe the basis of the discrimination between adenine and 3-MeA in the *S. aureus* enzyme.

<sup>1</sup> Supplementary material has been deposited in the IUCr electronic archive (Reference: GX5204).

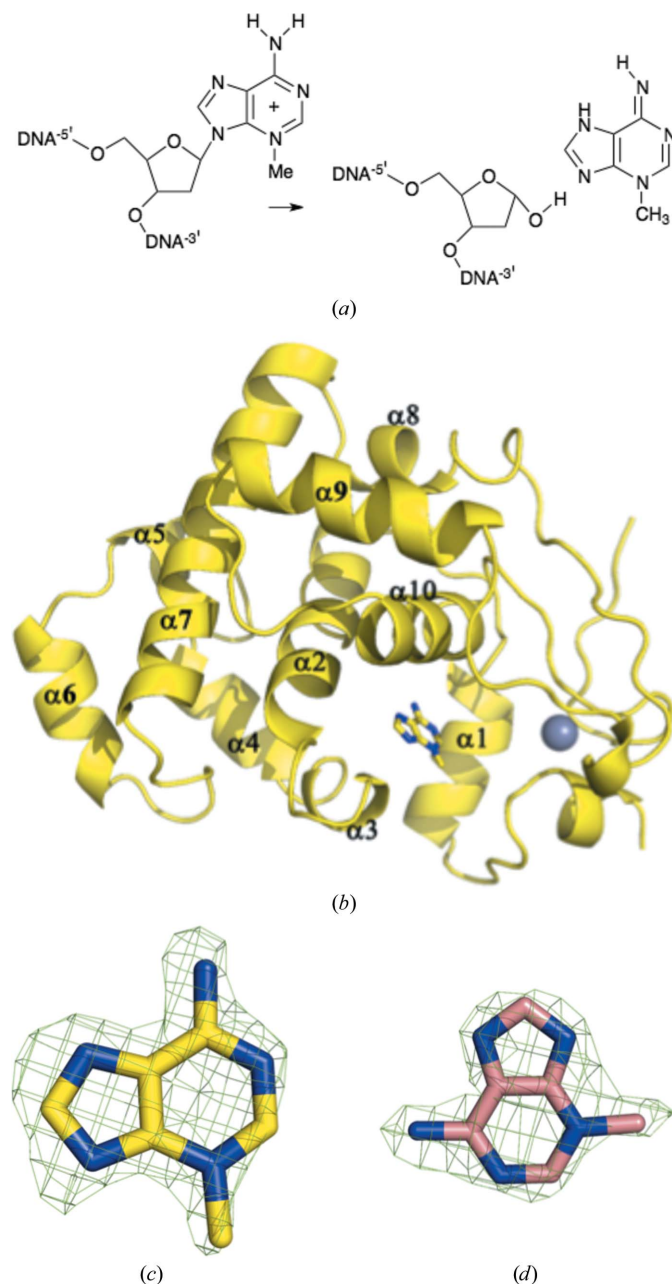


## 2. Materials and methods

### 2.1. Protein production

Native and mutant protein were purified as described by Oke *et al.* (2010). Y16F and E38Q mutations were introduced using Quik-Change (Stratagene); primers are listed in Table 1.

Fluorescence binding measurements were performed as described by Cao *et al.* (2003) and Drohat *et al.* (2002). 2  $\mu\text{M}$  TAG was titrated with 10–650  $\mu\text{M}$  3-MeA or adenine in 20 mM phosphate buffer pH 7.8 and 5.8; Figs. 2*a* and 2*b*). Isothermal titration calorimetry (ITC) experiments were carried out using a VP-ITC device (MicroCal)



**Figure 1**  
(*a*) The reaction catalyzed by TAG. (*b*) TAG is mainly  $\alpha$ -helical; a structural zinc ion (grey sphere) is found in all homologues of the enzyme. 3-MeA is shown in stick representation, with C atoms coloured yellow, N atoms coloured blue and O atoms coloured red. (*c*) Difference  $F_o - F_c$  electron density contoured at  $3\sigma$  for 3-MeA in the active site of TAG. (*d*) Difference  $F_o - F_c$  electron density contoured at  $3\sigma$  for 3-MeA in the active site of Y16F-mutant TAG; C atoms are coloured pink. 3-MeA binds in a different orientation in the Y16F mutant.

**Table 1**

Macromolecule-production information.

The following primers were used to create the mutations: Y16F, 5'-GTACTAAAGATC-CAGTCTACTTAAACTTTCATGATCATGTATGGG-3' and 5'-CCATACATGATC-ATGAAAGTTTAAAGTAGACTGGATCTTTAGTAC-3'; E38Q, 5'-GCAAGGCATTG-TTTAAACTTTTAGCATTACAGTCACAACATGCTGGG-3' and 5'-CCCAGCATG-TTGTGACTGTAATGCTAAAAGTTTAAACAATGCCTTGC-3'. Mutation sites are shown in bold.

|  |   |
|--|---|
| Source organism  | <i>S. aureus</i> strain MSSA476   |
| Expression vector                                      | pHis-TEV  |
| Expression host  | <i>E. coli</i>  |
| Complete amino-acid sequence of the construct produced | GAMNECAFGTKDPVYLYNHHDHVVWQPLYDSK-ALFKLLALESQHAAGLSWLTLKKKEAYEEAF-YDFEPEKVAQMTAQDIDR LMTFNPVHHRK-KLEAIVNQAQGYLKIEQAYGSFKLWSYVN-GKPKDLQYEHASDRITVDDTATQLSKDLKQ-YGFKFLGPVTVFSFLEAAGLYDAHLKDCPSK-PKHN |

**Table 2**

Data-collection and processing statistics.

Values in parentheses are for the last shell.

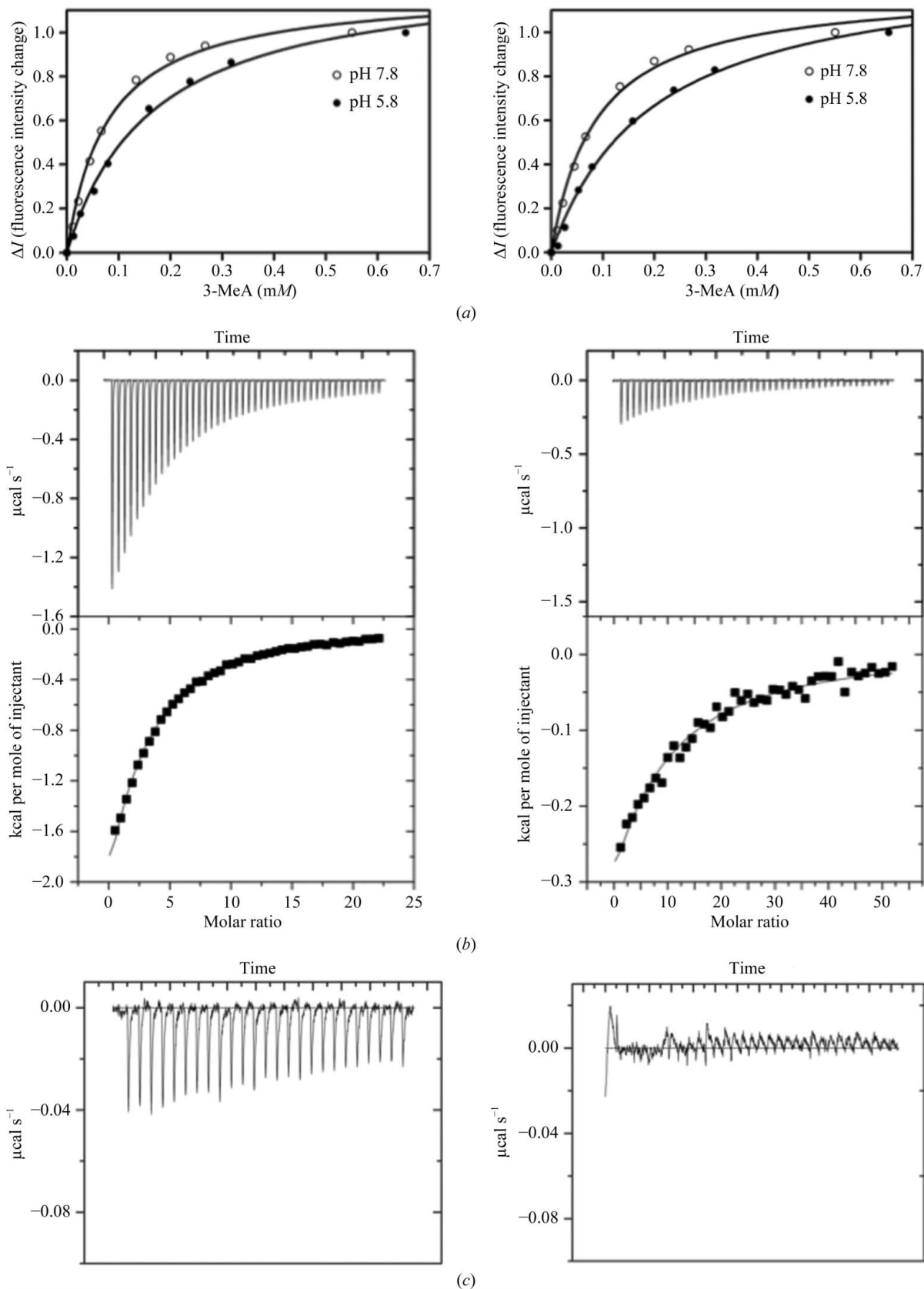
| Protein  | Native, 3-MeA complex  | Y16F, 3-MeA complex |
|--|------------------------|---------------------|
| Diffraction source   | ESRF beamline ID14-2   | Rotating anode      |
| Wavelength (Å)   | 0.933                  | 1.54                |
| Temperature (K)  | 100                    | 100                 |
| Detector   | ADSC Quantum 4 CCD     | Saturn CCD          |
| Crystal-to-detector distance (mm)                          | 203                    | 55                  |
| Rotation range per image (°)                               | 0.2                    | 0.5                 |
| Total rotation range (°)                                   | 108                    | 180                 |
| Exposure time per image (s)                                | 5                      | 5                   |
| Space group  | C2                     | C2                  |
| Unit-cell parameters                                       |                        |                     |
| <i>a</i> , <i>b</i> , <i>c</i> (Å)                         | 73.00, 78.59, 179.81   | 72.3, 78.8, 179.3   |
| $\alpha$ , $\beta$ , $\gamma$ (°)                          | 90, 90.56, 90          | 90, 90.5, 90        |
| Mosaicity (°)  | 0.3                    | 0.56                |
| Resolution range (Å)                                       | 29.60–1.80 (1.85–1.80) | 50–2.2 (2.28–2.20)  |
| Total No. of reflections                                   | 341926                 | 118143              |
| No. of unique reflections                                  | 92544 (5876)           | 47714 (3209)        |
| Completeness (%)   | 98.4 (91.6)            | 95.5 (89.1)         |
| Redundancy   | 3.7 (3.1)              | 2.6 (2.3)           |
| $\langle I/\sigma(I) \rangle$                              | 17.50 (3.9)            | 28.2 (10.9)         |
| $R_{r.i.m.}^\dagger$                                       | 0.059 (0.292)          | 0.04 (0.11)         |
| Overall <i>B</i> factor from Wilson plot (Å <sup>2</sup> ) | 18                     | 24.2                |

$^\dagger$  Estimated  $R_{r.i.m.} = R_{merge}[N/(N-1)]^{1/2}$ , where *N* is the data multiplicity.

in the same buffer. 5 mM 3-MeA or 1.5 mM adenine solution was injected at 298 K into a sample cell containing ~1.4 ml protein solution at 30–40  $\mu\text{M}$ . Each titration consisted of a first 1  $\mu\text{l}$  injection followed by up to 25 subsequent 10  $\mu\text{l}$  injections or 48 subsequent 5  $\mu\text{l}$  injections of the ligand as indicated. Calorimetric data were analyzed using the MicroCal *ORIGIN* software, fixing the stoichiometry as *N* = 1 (Figs. 2*c* and 2*d*; Supplementary Table S1).

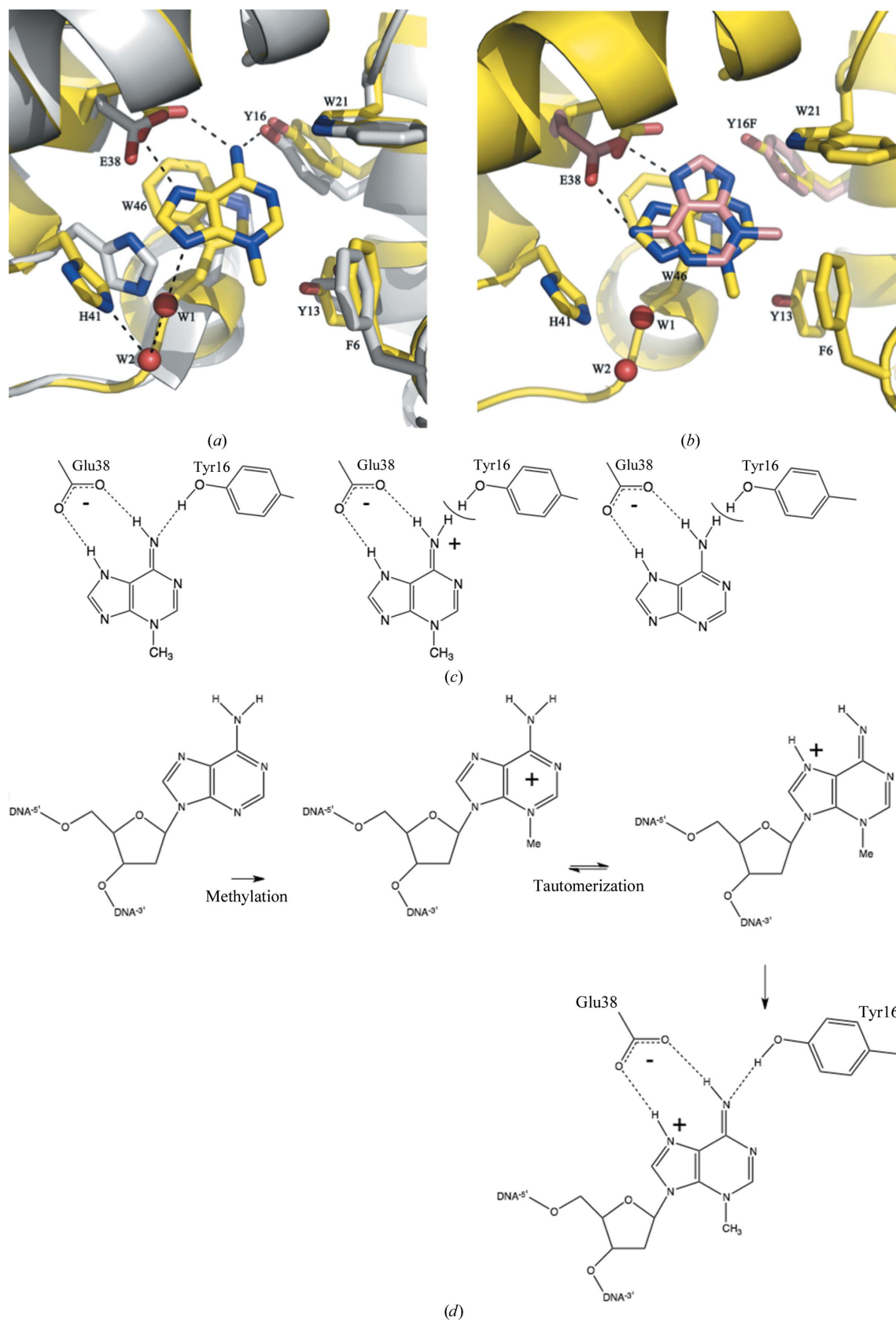
### 2.2. Crystallization

Sitting-drop vapour-diffusion crystallization trials (1  $\mu\text{l}$  protein solution plus 1  $\mu\text{l}$  precipitant solution) were set up using a Cartesian Honeybee nanodrop crystallization robot which was integrated in a Hamilton-Thermo Rhombix system. The 3-MeA complexes of native and Y16F TAG were obtained by incubating TAG with 10 mM 3-MeA for 6 h before crystallization at 277 K. The complex crystals grew using a precipitant solution consisting of 0.1 M Tris-HCl pH 8.5, 1.8 M ammonium sulfate, 0.2 M Li<sub>2</sub>SO<sub>4</sub> at 293 K as thin plates and grew to full size (0.2  $\times$  0.2  $\times$  < 0.05 mm) in two to three weeks. Cryoprotectant solution was made by supplementing the crystallization precipitant solution with 20% glycerol. Crystals were mounted



**Figure 2**

(a) Measurement of the binding of 3-MeA to *S. aureus* TAG using intrinsic fluorescence quenching at pH 5.8 ( $K_d = 165 \mu\text{M}$ ) and pH 7.8 ( $K_d = 78 \mu\text{M}$ ); the results are similar to those previously reported for the *E. coli* enzyme (Cao *et al.*, 2003). (b) Fluorescence quenching of 3-MeA with E38Q-mutant *S. aureus* TAG at pH 5.8 and 7.8. The small reduction in the binding constant was inconsistent with structural and previous functional data (Cao *et al.*, 2003). This indicated that the fluorescence was unreliable for the *S. aureus* enzyme. (c) ITC measurement of the binding of 3-MeA to *S. aureus* TAG at pH 7.8 ( $K_d = 220 \mu\text{M}$ ) and pH 5.8 ( $K_d = 470 \mu\text{M}$ ). Adenosine does not bind. (d) ITC measurement of the binding of 3-MeA to Y16F-mutant ( $K_d = 1.2 \text{ mM}$ ; left) and E38Q-mutant (no binding; right) *S. aureus* TAG at pH 7.8. 1 cal = 4.186 kJ.


**Figure 3**

(a) Structure of the 3-MeA–TAG complex (C atoms, yellow; N atoms, blue; O atoms, red) showing the key interactions. The apo structure is shown with C atoms in white. (b) Structure of the 3-MeA–Y16F TAG complex (C atoms shown in pink); the 3-MeA ring adopts a different orientation in the mutant. The 3-MeA in the native protein is also shown. (c) The most common tautomer of 3-MeA could be recognized by a specific hydrogen-bond arrangement of Tyr16 and Glu38. The predominant tautomer of protonated 3-MeA and adenosine would not match this hydrogen-bonding arrangement. (d) DNA damage leads to formation of the positively charged tautomer that is optimal for recognition by TAG; in addition, the highly electron-deficient ring would interact favourably with the TAG active site.



in Hampton Research cryoloops and rapidly cooled to 100 K prior to data collection.

### 2.3. Data collection and processing

Data for the native TAG–3-MeA complex were collected from a single crystal using  $0.2^\circ$  oscillations at a wavelength of  $0.933 \text{ \AA}$  (ESRF beamline ID14-2) and were reduced using *XDS* (Kabsch, 2010). Data were collected from a single crystal of the Y16F TAG–3-MeA complex using an in-house Rigaku MicroMax-007 HF rotating-anode generator and Saturn 944 CCD detector. Data were reduced using *HKL-2000* (Otwinowski & Minor, 1997) and *POINTLESS* (Evans, 2006; Potterton *et al.*, 2003; Winn *et al.*, 2011). Full details are given in Table 2. The E38Q mutant was also crystallized, but as no 3-MeA was located in the active site the structure is not described here; however, the structure has been deposited (PDB entry 4ai4).

### 2.4. Structure solution and refinement

The structures were solved with *Phaser* (McCoy *et al.*, 2007) using the native apo structure (Oke *et al.*, 2010; PDB entry 2jg6) as a search model. As the complex crystals grew in a different space group to the native crystals, a new free set of reflections was assigned for refinement. All structures were refined with *REFMAC* v.5.6.0117 (Murshudov *et al.*, 2011); manual intervention employed *Coot* (Emsley & Cowtan, 2004). 3-MeA was added to the models when the  $F_o - F_c$  density was clear (Figs. 1c and 1d). *MolProbity* (Chen *et al.*, 2010) was used for structure validation and Ramachandran analysis. TLS parameters were used in refinement. TLS groups were assigned using the *TLSMD* server (Painter & Merritt, 2006). Details of the refinement are given in Table 3.

## 3. Results and discussion

The structure of the *S. aureus* TAG–3-MeA complex was determined to  $1.8 \text{ \AA}$  resolution and that of the Y16F TAG–3-MeA complex to  $2.22 \text{ \AA}$  resolution. The structure of the native 3-MeA complex is very similar to the crystal structure of the *S. typhi* TAG–3-MeA–abasic DNA complex (Metz *et al.*, 2007) and the NMR structure of the *E. coli* TAG–3-MeA complex (Cao *et al.*, 2003). Relative to apo TAG (Oke *et al.*, 2010), Glu38 has rotated to make  $2.7 \text{ \AA}$  contacts with the exocyclic N atom and N7 of 3-MeA. Tyr16 moves to make a  $2.8 \text{ \AA}$  contact with the exocyclic N atom of 3-MeA (Fig. 3a). Trp46 stacks with the bound purine ring of 3-MeA, while Phe6, Tyr13 and Tyr21 make edge-on contacts. His41 rotates  $80^\circ$  to create space for 3-MeA to bind. The Y16F-mutant complex revealed that 3-MeA adopts a different orientation, although it preserves a bidentate hydrogen bond to Glu38 and a stacking interaction with Trp46 (Fig. 3b). This conformation is unlikely to be physiologically relevant, as it would require a very different orientation of the DNA to that observed in the *S. typhi* complex (Metz *et al.*, 2007). Using a fluorescence assay, we measured 3-MeA binding (Fig. 2a), obtaining a similar result at pH 7.8 ( $K_d = 78 \mu\text{M}$ ) to that for the *E. coli* enzyme at pH 7.5 ( $K_d = 42 \mu\text{M}$ ; Cao *et al.*, 2003). However, the assay is flawed for the *S. aureus* enzyme as the E38Q mutant gave the same result as for the native protein (Fig. 2b), which is physically unreasonable. ITC (Figs. 2c and 2d) showed clear differences between the native and mutant *S. aureus* enzymes (Y16F,  $K_d = 1.2 \text{ mM}$ ; E38Q, no binding) and gave  $K_d$  values of  $220 \mu\text{M}$  at pH 7.8 and  $471 \mu\text{M}$  at pH 5.8 for the native enzyme. We did not detect adenine binding.

3-Methyldeoxyadenosine is positively charged in DNA, whilst deoxyadenosine is neutral; simple charge–charge recognition was therefore the original explanation for the specificity of TAG (Labahn

**Table 3**  
Structure refinement.

Values in parentheses are for the last shell.

| Protein                                | Native, 3-MeA complex (PDB entry 4aia) | Y16F, 3-MeA complex (PDB entry 4ai5) |
|--|--|--------------------------------------|
| Resolution range ( $\text{\AA}$ )      | 28.19–1.80 (1.847–1.800)               | 179.29–2.22 (2.276–2.218)            |
| Completeness (%)                       | 98.2                                   | 95.3                                 |
| $\sigma$ cutoff                        | 0                                      | 0                                    |
| No. of reflections, working set        | 87884 (5568)                           | 45350 (3043)                         |
| No. of reflections, test set           | 4654 (308)                             | 2364 (166)                           |
| Final $R_{\text{cryst}}$               | 0.179 (0.233)                          | 0.183 (0.193)                        |
| Final $R_{\text{free}}$                | 0.218 (0.289)                          | 0.216 (0.244)                        |
| No. of non-H atoms                     |  |                                      |
| Protein                                | 7598                                   | 7602                                 |
| Ion                                    | 25                                     | 25                                   |
| Ligand                                 | 55                                     | 55                                   |
| Water                                  | 927                                    | 486                                  |
| Total                                  | 8605                                   | 8168                                 |
| R.m.s. deviations                      |  |                                      |
| Bonds ( $\text{\AA}$ )                 | 0.009                                  | 0.015                                |
| Angles ( $^\circ$ )                    | 1.189                                  | 1.550                                |
| Average $B$ factors ( $\text{\AA}^2$ ) |  |                                      |
| Protein                                | 22.2                                   | 21.7                                 |
| Ion                                    | 29.9                                   | 29.2                                 |
| Ligand                                 | 15.4                                   | 17.3                                 |
| Water                                  | 25.5                                   | 22.1                                 |
| Ramachandran plot                      |  |                                      |
| Favoured regions (%)                   | 98.5                                   | 98.4                                 |
| Additionally allowed (%)               | 1.4                                    | 1.5                                  |

*et al.*, 1996; Lau *et al.*, 2000; Hollis *et al.*, 2000). However, it has been shown that *E. coli* TAG binds 3-MeA but not adenine and binds protonated 3-MeA (pH 5.7) more weakly than neutral 3-MeA (pH 7.5) (Cao *et al.*, 2003; Drohat *et al.*, 2002), establishing that charge–charge recognition is not the sole explanation (Cao *et al.*, 2003). We suggest that a particular hydrogen-bond pattern contributes to the selection of a specific but favoured (Sharma & Lee, 2002) neutral tautomer of 3-MeA (Fig. 3c) that is not available to adenosine (Fig. 3c) and that is disfavoured for protonated 3-MeA (Fig. 3c). Our hypothesis implies that there is an energetic penalty in reorganizing the hydrogen-bond network around Tyr16 to avoid a van der Waals clash (Fig. 3c). In DNA, 3-methyldeoxyadenosine can adopt a tautomer that has the same hydrogen arrangement as neutral 3-MeA and has positive charge (Fig. 3d), which is favoured at the active site (Metz *et al.*, 2007). A clash of H atoms was observed between the amide of His136 and the amino group of adenine in human AAG and is used to preferentially select the damaged purine base (O'Brien & Ellenberger, 2004). Higher resolution data or neutron diffraction are required to further test the hypothesis for the TAG enzyme.

The work was funded by the BBSRC SPoRT initiative (BB\_BBS/B/14426).

## References

- Banerjee, A., Santos, W. L. & Verdine, G. L. (2006). *Science*, **311**, 1153–1157.  
 Banerjee, A. & Verdine, G. L. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 15020–15025.  
 Banerjee, A., Yang, W., Karplus, M. & Verdine, G. L. (2005). *Nature (London)*, **434**, 612–618.  
 Bjelland, S., Bjørås, M. & Seeberg, E. (1993). *Nucleic Acids Res.* **21**, 2045–2049.  
 Blainey, P. C., van Oijen, A. M., Banerjee, A., Verdine, G. L. & Xie, X. S. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 5752–5757.  
 Cao, C., Kwon, K., Jiang, Y. L., Drohat, A. C. & Stivers, J. T. (2003). *J. Biol. Chem.* **278**, 48012–48020.  
 Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst. D* **66**, 12–21.

- Drohat, A. C., Kwon, K., Krosky, D. J. & Stivers, J. T. (2002). *Nature Struct. Biol.* **9**, 659–664.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Engelward, B. P., Dreslin, A., Christensen, J., Huszar, D., Kurahara, C. & Samson, L. (1996). *EMBO J.* **15**, 945–952.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Forsyth, R. A. *et al.* (2002). *Mol. Microbiol.* **43**, 1387–1400.
- Hollis, T., Ichikawa, Y. & Ellenberger, T. (2000). *EMBO J.* **19**, 758–766.
- Ji, Y., Zhang, B., Van Horn, S. F., Warren, P., Woodnutt, G., Burnham, M. K. & Rosenberg, M. (2001). *Science*, **293**, 2266–2269.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Labahn, J., Schärer, O. D., Long, A., Ezaz-Nikpay, K., Verdine, G. L. & Ellenberger, T. E. (1996). *Cell*, **86**, 321–329.
- Lau, A. Y., Wyatt, M. D., Glassner, B. J., Samson, L. D. & Ellenberger, T. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 13573–13578.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Metz, A. H., Hollis, T. & Eichman, B. F. (2007). *EMBO J.* **26**, 2411–2420.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- O'Brien, P. J. & Ellenberger, T. (2004). *J. Biol. Chem.* **279**, 9750–9757.
- Oke, M. *et al.* (2010). *J. Struct. Funct. Genomics*, **11**, 167–180.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Painter, J. & Merritt, E. A. (2006). *J. Appl. Cryst.* **39**, 109–111.
- Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. (2003). *Acta Cryst.* **D59**, 1131–1137.
- Riazuddin, S. & Lindahl, T. (1978). *Biochemistry*, **17**, 2110–2118.
- Sedgwick, B., Bates, P. A., Paik, J., Jacobs, S. C. & Lindahl, T. (2007). *DNA Repair*, **6**, 429–442.
- Sharma, S. & Lee, J. K. (2002). *J. Org. Chem.* **67**, 8360–8365.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.