A Novel Zinc Snap Motif Conveys Structural Stability to 3-Methyladenine DNA Glycosylase I*

Received for publication, January 28, 2003, and in revised form, March 19, 2003
Published, JBC Papers in Press, March 24, 2003, DOI 10.1074/jbc.M300934200

Keehwan Kwon, Chunyang Cao, and James T. Stivers‡
From the Department of Pharmacology and Molecular Sciences, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205-2185

The Escherichia coli 3-methyladenine DNA glycosylase I (TAG) is a DNA repair enzyme that excises 3-methyladenine in DNA and is the smallest member of the helix-hairpin-helix (HhH) superfamily of DNA glycosylases. Despite many studies over the last 25 years, there has been no suggestion that TAG was a metalloprotein. However, here we establish by heteronuclear NMR and other spectroscopic methods that TAG binds 1 eq of Zn$^{2+}$ extremely tightly. A family of refined NMR structures shows that 4 conserved residues contributed from the amino- and carboxyl-terminal regions of TAG (Cys$^4$, His$^{17}$, His$^{175}$, and Cys$^{179}$) form a Zn$^{2+}$ binding site. The Zn$^{2+}$ ion serves to tether the otherwise unstructured amino- and carboxyl-terminal regions of TAG. We propose that this unexpected “zinc snap” motif in the TAG family (CX$_{12-17}$HX$_{150}$HX$_C$) serves to stabilize the HhH domain thereby mimicking the functional role of protein-protein interactions in larger HhH superfamily members.

Excision repair of damaged bases in DNA is one pathway that cells use to protect the genome from the damaging effects of ionizing radiation, reactive oxygen species, and alkylating agents (1). Lesion repair begins by enzymatic hydrolysis of the glycosidic bond, the critical initiating step of the DNA base excision repair pathway. There are many different types of DNA glycosylases in cells. In general, each DNA glycosylase exhibits specificity for a cognate damaged base while ignoring undamaged bases in DNA. One alkylated base, 3-methyladenine, is highly toxic, because DNA replication is blocked at this lesion site (2). In Escherichia coli, the glycosidic bond of 3-methyladenine is hydrolyzed by two enzymes, 3-methyladenine DNA glycosylase I (TAG)$^1$ and II (AlkA) (3).

It has been recently discovered that TAG is a member of the helix-hairpin-helix (HhH) superfamily of DNA glycosylases (4).

* This work was supported by National Institutes of Health Grant GM46835. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1NKU) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‡ To whom correspondence should be addressed: Dept. of Pharmacology and Molecular Sciences, The Johns Hopkins University School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205-2185. Tel.: 410-502-2758; Fax: 410-955-3023; E-mail: jsstivers@jhmi.edu.

1 The abbreviations used are: TAG, 3-methyladenine DNA glycosylase I; HhH, helix-hairpin-helix; AAS, atomic absorption spectroscopy; NOE, nuclear Overhauser effect; r.m.s.d., root-mean-square deviation; MES, 4-morpholineethanesulfonic acid; NOES, nuclear Overhauser effect; HSQC, heteronuclear single quantum coherence.

The HhH motif is a sequence-nonspecific DNA binding module found in DNA polymerases, NAD$^+$-dependent DNA ligases, and some DNA glycosylases (5). This motif consists of two α-helices connected by a consensus hairpin loop that interacts nonspecifically with the DNA backbone. Unlike other HhH family members, TAG does not possess the consensus hairpin sequence, (L/F)PG(V/I)G, nor does it contain the conserved aspartate group that was previously believed to be required for catalysis as a water activating group or, alternatively, for stabilization of a transition state with glycosyl cation character. Thus, TAG appears to be unique with respect to structure and mechanism within this superfamily (4).

Although it has been concluded previously that TAG does not require a metal ion for activity (6), the possibility of a metal binding site was suggested by our recent NMR structure and a coincident bioinformatics study (4, 7). According to the solution structure, two conserved sulfur and histidine ligands at opposing ends of the linear sequence are closely positioned in three-dimensional space (Cys$^4$, His$^{17}$, His$^{175}$, and Cys$^{179}$) (4, 7), but given the limitations of the NMR method, the presence of a metal ion was not established. As shown in Fig. 1, the sequence alignment of the TAG family shows that these potential metal ligands are also completely conserved across all species. Taken together, these observations clearly suggested the presence of a metal ion binding site. Here we establish this hypothesis and present the structure of a long overlooked Zn$^{2+}$ binding site in TAG. The metal site stabilizes the structure of TAG by “snapping” together the largely unstructured N- and C-terminal regions. This “zinc snap” motif (CX$_{12-17}$HX$_{150}$HX$_C$) is distinct from that of zinc finger proteins or the zinc binding sites in other DNA repair proteins such as formamidopyrimidine DNA glycosylase, poly(ADP-ribose) polymerase, and UvrA (8–11). We propose that the minimal HhH domain is intrinsically unstable and that the zinc snap provides the stabilizing forces that are assumed by intramolecular protein-protein contacts in other larger HhH family members.

EXPERIMENTAL PROCEDURES

Enzyme Purification—Wild-type TAG was overexpressed from pET21a(TAG) in BL21(DE3) and was purified by two sequential chromatography steps as described previously by Drohat et al. (4). A His$_{6}$-tagged version was overexpressed from pET28a(TAG) in BL21(DE3)pLysS and purified using a nickel-nitrilotriacetic acid-agarose column. Final purification was achieved using a Poros 20-S HPLC column (Applied Biosystems, Foster City, CA) using a linear gradient of 0.1–1 M NaCl. The purity of TAG was >95% by both procedures.

Atomic Absorption Spectroscopy (AAS)—The presence of zinc in the holoenzyme was determined using a PerkinElmer model 370 atomic absorption spectrophotometer. The purified enzyme was dialyzed extensively against buffer A (50 mM Tris-HCl, 200 mM NaCl, pH 7.5) containing 4 ml of hydrated chelex-100 resin (Sigma)/500 ml of buffer.

This paper is available on line at http://www.jbc.org
NMR Spectroscopy—The NMR sample, prepared as described in a previous report (4), contained 0.5 mM TAG, 10 mM phosphate buffer (pH 6.6), 100 mM NaCl, 3 mM dithiothreitol, 0.34 mM NaN₃, and 10% D₂O in a total volume of 300 μl. The two-dimensional ¹H-¹⁵N LR HSQC experiment was conducted at 20 °C on a Varian Unity Plus 600-MHz spectrometer equipped with four frequency channels and pulse-field gradients as described (12). This experiment was simply a conventional HSQC used for backbone amide correlations collected with an optimized two-bond ²J_HH value of 22 Hz in order to observe signals from the weak two-bond couplings in the histidine rings and suppress the signals from the one-bond J_HN amide couplings. The ¹N dimension was collected with 256 complex points, a 120-ppm sweep width, 128 scans, and the ³¹P carrier set at 205 ppm. The ¹H dimension was collected with 1024 complex points and a 13.3-ppm sweep width centered at 4.82 ppm. Decoupling of ³¹P was accomplished with the WALTZ16 sequence (13) using a 1.1-kHz field. 

NRQ Structural Calculations—Starting from the published NMR structure of TAG (Protein Data Bank code: 1LMZ), which was originally refined without any constraints to Zn²⁺, a set of 100 structures that included restraints from the enzyme ligands to Zn²⁺ were calculated using the program XPLOR-NIH, version 2.0.2 (14). These calculations employed the same NOE constraints, dihedral angle, and H-bond restraints as the original structures. Additionally, because of new assignments obtained from the ¹H-¹⁵N LR-HSQC experiment, a total of 11 new NOEs and three new hydrogen bonds were used as restraints in the calculation. From this set, 25 low energy structures were selected with constraints violations less than 5°. Constraints between the protein ligands and the zinc ion were added using the procedure of Neuhaus and the zinc ion was added using the procedure of Neuhaus (15). The constraint values were based on tetrahedral coordination, using literature values for the HisN-²-Zn and CysS-²-Zn bond lengths of 2.0 Å.
and 2.3 Å, respectively, which were obtained from EXAFS (extended x-ray absorption fine structure) measurements (16). The following additional distance constraints were used (15): His\(^{17}\)N\(^{-}\) to His\(^{17}\)N\(^{2}\), 3.1 Å < d < 3.5 Å; His\(^{17}\)N\(^{-}\) to Cys\(^{4}\) or Cys\(^{37}\)S\(^{-}\) and His\(^{17}\)N\(^{-}\) to Cys\(^{4}\) or Cys\(^{37}\)S\(^{-}\); 3.4 Å < d < 3.8 Å; His\(^{17}\)N\(^{-}\) to Cys\(^{37}\)S\(^{-}\); 3.4 Å < d < 3.9 Å; and Cys\(^{37}\)S\(^{-}\) to Cys\(^{17}\)S\(^{-}\). The following ligand-zinc bond angles were used as constraints: His\(^{17}\)N\(^{-}\)-Zn-His\(^{17}\)N\(^{2}\) = 109.5°, and Zn-His\(^{17}\)N\(^{-}\)-His\(^{17}\)N\(^{-}\) = 126°. The improper dihedral angle constraints were (Zn-His\(^{17}\)N\(^{-}\)-His\(^{17}\)N\(^{-}\)-His\(^{17}\)N\(^{-}\)-His\(^{17}\)N\(^{-}\) = 180°. The zinc atom was given a charge of +2 in the calculations, and the default values for the bond and angle force constants in X-PLOR were employed (500 kcal/mol Å\(^{-}\) and 70 kcal/mol Å\(^{-}\) Å\(^{-}\), respectively).

RESULTS AND DISCUSSION

TAG Has One Tight Zn\(^{2+}\) Binding Site—To determine whether TAG possesses a tight metal binding site, AAS measurements were performed (Table I). Two independently purified samples of TAG that had been dialyzed extensively against a neutral pH buffer containing excess chelating resin showed the presence of 1 metal of Zn\(^{2+}\) by AAS analysis (Table I). It was found that exhaustive dialysis of these same samples against a buffer that contained 40 \(\mu\)M CoCl\(_2\) (pH 7.5) still retained 1 eq of Zn\(^{2+}\), indicating that the site was inert to exchange under these conditions. Surprisingly, even extensive dialysis of the enzyme against a low pH buffer (pH 5.5) that contained 50 mM EDTA removed only about 0.8 eq of the bound metal. In fact, complete removal of Zn\(^{2+}\) required denaturation of the enzyme using 0.1 M HCl (Table I). These results indicate that previous reports of TAG not requiring a metal ion for activity were due to the very high binding affinity of Zn\(^{2+}\) for this site, which cannot be removed by standard chelating agents that are included in the purification buffers (6, 14). The previously reported dramatic loss in TAG activity below pH 6 may be attributed to metal ion loss and partial protein unfolding (17). We found that the apoTAG was unstable and aggregated extensively at pH 7.5 (data not shown).

Geometry and Ligands of the Zn\(^{2+}\) Site—One valuable method for determining the ligands and coordination geometry for a Zn\(^{2+}\) binding site in enzymes is to perform metal replacement with Co\(^{2+}\), a transition metal that gives diagnostic features in its UV-visible spectra (18–20). For example, thiolate-Co\(^{2+}\) bonds show a diagnostic charge transfer band at about 350 nm with a typical extinction coefficient of 900–1300 m\(^{-1}\)cm\(^{-1}\)/thiolate ligand (21, 22). Additional bands in the 576–670 nm range are indicative of the coordination geometry; tetrahedral coordination gives a strong signal (\(E_{350}\) >300 m\(^{-1}\)cm\(^{-1}\)), whereas pentacoordinate and octahedral sites have much lower molar absorptivities (50 \(< E < 225\) m\(^{-1}\)cm\(^{-1}\) and \(< 30\) m\(^{-1}\)cm\(^{-1}\), respectively) (18, 23–25).

The UV-visible spectrum of Co\(^{2+}\)-TAG (Fig. 2) is consistent with two sulfur ligands and a tetrahedral coordination geometry. Because of the extensive aggregation of apoTAG during the Co\(^{2+}\) spectroscopy, the concentration of the soluble protein in the cuvette could not be known precisely. However, comparing the ratio (R) of extinction coefficients of the 350 nm and 617 nm bands (\(R = E_{350}/E_{617}\)) after base-line correction still provides quantitative information on the number of sulfur ligands and the coordination geometry. The expected ratio for tetrahedral geometry is 1 < R < 4 for one thiolate ligand, and 2 < R < 9 for two thiolate ligands. The latter value encompasses the E\(_{350}/E_{617}\) = 4 measured for TAG. This value differs considerably from the expected ratios for pentacoordinate and octahedral geometry for two thiolate ligands, which would be 8 < R < 52 and R > 60, respectively. Although these estimates have some uncertainty, they are consistent with the strict conservation of 2 Cys and 2 His residues in the TAG family (Fig. 1), and the structural results described below.

Electronic Properties of the Zn\(^{2+}\) Ligands Determined by Heteronuclear NMR Spectroscopy—Although the Co\(^{2+}\) replacement studies described above indicate the ligation of two sulfur atoms to the Zn\(^{2+}\), the identities of the remaining two ligands are not revealed by this method. Because our previous NMR structure and the homology comparison suggested that His\(^{17}\) and His\(^{17}\) were the remaining ligands (Fig. 1), we performed a 1H-15N LR-HSQC NMR experiment to investigate the electronic properties of these histidines (Fig. 3A); the histidine spin systems in Fig. 3A were obtained from three-dimensional 15N-edited aromatic and aliphatic NOESY experiments). This simple LR-HSQC experiment correlates the carbon-bond protons of the histidine rings with the imidazole nitrogen atoms and can unambiguously establish the tautomeric and protonation states of histidines in proteins (12, 26). The characteristic upside-down and sideways L-shaped patterns for the peaks in the two-dimensional spectrum for His\(^{17}\) and His\(^{17}\) and their well separated 15N chemical shifts indicate that these histidines are...
in the Nβ2-H and Nγ1-H tautomeric forms, respectively, and that both are neutral (Fig. 3, A and B). Thus the nitrogens that are available to coordinate the zinc are the Nε1 of His175 and the Nγ2 of His177. The 15Nε1 shift for His175 is by far the most deshielded nitrogen of all histidines in TAG, consistent with strong chelation to an electropositive Zn2⁺.

The 15Nγ2 chemical shift of the ligating nitrogen of His177 is also the most deshielded of the five histidines, also indicating strong chelation to the metal from this position. Similarly, the 13Cα chemical shifts of Cys4 and Cys179 are the most deshielded of the 8 cysteine residues of TAG, consistent with thiolate metal coordination from these side-chains (Fig. 3C).

Structure of Zn2⁺ Binding Site—Because a zinc binding site was entirely unanticipated in TAG, our previous NMR structure did not include any restraints to a Zn2⁺ atom (Protein Data Bank code: 1LMZ). Therefore, with unambiguous evidence for a Zn2⁺ site in hand, we further refined the solution structure using the same constraints employed previously but including distance and angle constraints between the Zn2⁺ and Cys4, His175, His177, and Cys179 (see “Experimental Procedures”). Inclusion of these constraints was justified on the basis of the UV-visible spectroscopy measurements, which indicated two sulfur ligands with tetrahedral geometry, and the LR-HSQC experiments, which indicated coordination of His177Nε and His175Nγ to the zinc. The addition of the zinc ion and these new constraints did not increase the Lennard-Jones potential energy for the ensemble of 25 lowest energy structures, nor did it introduce new NOE violations. For the new ensemble, about 85% of the dihedral angles are in the most favored region, with 11% in the additionally allowed region. These statistics are within error of those reported previously; further structural statistics are reported in Table II.

The new NMR structure containing Zn2⁺ is depicted as a ribbon diagram in Fig. 4A and confirms that the Zn2⁺ binding site tethers the amino and carboxyl termini of TAG. This structure is nearly identical to the deposited TAG structure with r.m.s.d. values of 0.66 Å for backbone atoms and 0.86 Å for all heavy atoms for residues 11–174. The 3-methyladenine binding site, marked with an asterisk in Fig. 4A, is removed from the metal site, precluding any direct involvement of the metal in catalysis. However one metal ligand, His175, is involved in a strong H-bond to the carboxyl oxygen of the active site group, Trp21 (His175 δ(NH2) = 14.75 ppm), suggesting that the zinc could indirectly organize the structure of the active site. The detailed structure of the Zn2⁺ binding site is shown in Fig. 4B (27).

Role of the Zinc Snap Motif in the Helix-Hairpin-Helix Superfamily—The Zn2⁺ binding site in TAG is distinct from zinc finger motifs found in bacteria to eukaryotes (28), as well as the zinc binding motifs found in the bacterial 8-oxoguanine glycosylase, MutM (8, 9). The four Cys zinc finger motif in MutM is in the arrangement CX5CX16CX2C and serves to stabilize two β-strands that deliver three polar side chains to their interac-

---

**Table II**

| Parameter | (25) |
|-----------|------|
| r.m.s.d. with respect to the mean structure | 0.73 ± 0.11 |
| Backbone atoms residues 11-174 | 1.31 ± 0.13 |
| All heavy atoms residues (1-187) | 0.93 ± 0.16 |
| All heavy atoms for all residues (1-187) | 1.48 ± 0.16 |
| r.m.s.d. from experimental distance restraints (Å) | |
| Intraresidue (300) | 0.045 ± 0.003 |
| Sequential (477) | 0.043 ± 0.005 |
| Medium range (465) | 0.029 ± 0.003 |
| Long range (520) | 0.026 ± 0.005 |
| Hydrogen bonds (118) | 0.033 ± 0.006 |
| r.m.s.d. from ϕ/ψ dihedral angle restraints (255) (°) | 0.31 ± 0.06 |
| r.m.s.d. from idealized covalent geometry | |
| Bonds (Å) | 0.0019 ± 0.0001 |
| Angles (°) | 0.387 ± 0.024 |
| Impropers (°) | 0.337 ± 0.027 |
| Lennard-Jones potential energy (kcal · mol⁻¹) | -704.9 ± 17.8 |
| Bad contacts/100 residues | 16.7 ± 3.0 |

---

* (25) is the ensemble of 25 structures calculated as described in the text.

† None of the structures has a distance violation >0.5 Å.

‡ The dihedral restraints were derived using TALOS as described in the text. None of the structures exhibits a dihedral angle violation >5°.

§ Lennard-Jones potential energy calculated using the CHARMM parameters.
ions with the phosphodiester backbone (9). In contrast, the zinc motif in TAG consists of two sets of CX3H submotifs, which are separated by 143–151 amino acid residues. This arrangement differs significantly from zinc finger motifs, which cluster in the region of a turn or a loop. The split zinc binding motif, with ligands donated from both termini of TAG, acts as a “snap” for closing the ends of the protein. The zinc binding site is distant from the previously modeled DNA binding region of TAG (4) and serves to stabilize tertiary structure, but it is not involved directly in DNA binding. Given the small size of the domain, the zinc ion provides a critical stabilizing element. This is supported by our observation that the apoprotein appears to be poorly folded, leading to aggregation and eventual precipitation. Difficulties in preparation of the TAG mutant, H17A, because of low expression levels and aggregation (data not shown), are consistent with the conclusion that the Zn2+ binding motif is important for stabilization of the protein. Structural comparisons between TAG and other HhH glycosylases show that all other family members have appendages to their amino and carboxyl termini that interact extensively and are likely to contribute to stabilizing the HhH fold. Therefore, the newly revealed zinc snap motif allows the TAG family to efficiently stabilize the HhH structure without additional protein scaffolding.

Acknowledgments—We thank Dr. Albert S. Mildvan for assistance with the atomic absorption spectroscopy analysis and Daniel Krosky and Lauren Morgans for helpful comments on the manuscript.

REFERENCES
1. Lindahl, T. (1993) Nature 362, 709–715
2. Boiteux, S., Huisman, O., and Laval, J. (1984) EMBO J. 3, 2569–2573
3. Wyatt, M. D., Allan, J. M., Lau, A. Y., Ellenberger, T. E., and Samson, L. D. (1999) Bioessays 21, 668–676
4. Drobak, A. C., Kwon, K., Krosky, D. J., and Stivers, J. T. (2002) Nat. Struct. Biol. 9, 659–664
5. Nash, H. M., Bruner, S. D., Scharer, O. D., Kawate, T., Addona, T. A., Spooner, E., Lane, W. S., and Verdin, G. L. (1996) Curr. Biol. 6, 986–980
6. Rissanen, S., and Lindahl, T. (1978) Biochemistry 17, 2110–2118
7. Bujnicki, J. M., and Rychlewski, L. (2002) DNA Repair (Amst.) 1, 391–395
8. Fromme, J. C., and Verdin, G. L. (2002) Nat. Struct. Biol. 9, 544–552
9. Gilboa, E., Zharkov, D. O., Golan, G., Fernandes, A. S., Gerchman, S. E., Matz, E., Kycia, J. H., Grollman, A. P., and Shoham, G. (2002) J. Biol. Chem. 277, 19811–19816
10. Narasimhan, S., Myles, G. M., Strange, R. W., and Sancar, A. (1989) J. Biol. Chem. 264, 16067–16071
11. Menissier-de Murcia, J., Molinet, M., Gradwohl, G., Simonin, F., and de Murcia, G. (1989) J. Mol. Biol. 210, 229–233
12. Drobak, A. C., Xiao, G., Tordova, M., Jagadeesh, J., Pankiewicz, K. W., Watanabe, K. A., Gilliland, G. L., and Stivers, J. T. (1999) Biochemistry 38, 11876–11886
13. Shaka, A. J., Koehler, J., Frenkel, T., and Freeman, R. (1983) J. Magn. Reson. 52, 335–338
14. Kuszewski, J., and Clare, G. M. (2000) J. Magn. Reson. 146, 249–254
15. Neuhau, D., Nakaseko, Y., Schwabe, J. W., and Klug, A. (1992) J. Mol. Biol. 228, 637–651
16. Daikun, G. P., Fairall, L., and Klug, A. (1986) Nature 324, 696–699
17. Bjelland, S., and Seeberg, E. (1987) Nucletic Acids Res. 15, 2787–2791
18. Rosenberg, R. C., Root, C. A., and Gray, H. B. (1975) J. Am. Chem. Soc. 97, 21–26
19. Outten, C. E., Tobin, D. A., Penner-Hahn, J. E., and O’Halloran, T. V. (2001) Biochemistry 40, 10417–10425
20. VanZile, M. L., Cosper, N. J., Scott, R. A., and Giedroc, D. P. (2000) Biochemistry 39, 11818–11829
21. May, S. W., and Koo, J. Y. (1978) Biochemistry 17, 3333–3338
22. Lane, R. W., Ibers, J. A., Franke, R. B., Papaefthymiou, G. C., and Holm, R. H. (1977) J. Am. Chem. Soc. 99, 84–98
23. Rosenberg, R. C., Root, C. A., Wang, R. H., Cerdonio, M., and Gray, H. B. (1973) Proc. Natl. Acad. Sci. U. S. A. 70, 161–163
24. Vasak, M., Kagi, J. H., Holmquist, B., and Vallee, B. L. (1981) Biochemistry 20, 6659–6664
25. Bolognese, L., Magliozzo, R. S., Beltrami, M., Salvato, B., and Peisach, J. (1992) Biochemistry 31, 9294–9303
26. Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) Protein Sci. 2, 543–558
27. Borden, K. L., Lally, J. M., Martin, S. R., O’Reilly, N. J., Etkin, L. D., and Freemont, P. S. (1995) EMBO J. 14, 5947–5956
28. Berg, J. M. (1990) J. Biol. Chem. 265, 6513–6516
