JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

A Catalytic Role for C−H/π Interactions in Base Excision Repair by *Bacillus cereus* DNA Glycosylase AlkD

Zachary D. Parsons,[†] Joshua M. Bland,[†] Elwood A. Mullins,[†] and Brandt F. Eichman*[,†]

[†]Department of Biological Sciences and Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37232, United States

Supporting Information

ABSTRACT: DNA glycosylases protect genomic integrity by locating and excising aberrant nucleobases. Substrate recognition and excision usually take place in an extrahelical conformation, which is often stabilized by π-stacking interactions between the lesion nucleobase and aromatic side chains in the glycosylase active site. *Bacillus cereus* AlkD is the only DNA glycosylase known to catalyze base excision without extruding the damaged nucleotide from the DNA helix. Instead of contacting the nucleobase itself, the AlkD active site interacts with the lesion deoxyribose through a series of C−H/π interactions. These interactions are ubiquitous in protein structures, but evidence for their catalytic significance in enzymology is lacking. Here, we show that the C−H/π interactions between AlkD and the lesion deoxyribose participate in catalysis of glycosidic bond cleavage. This is the first demonstration of a catalytic role for C−H/π interactions as intermolecular forces important to DNA repair.

DNA damage arising from reactive cellular metabolites and environmental toxins can corrupt stored genetic information and disrupt normal cellular processes, leading to heritable diseases, aging, and cell death.[1] DNA nucleobases are particularly susceptible to oxidation, methylation, and hydrolytic deamination,[2] giving rise to a broad spectrum of chemical modifications, most of which are eliminated from the genome by the base excision repair (BER) pathway.[3] BER is initiated by DNA glycosylases, which locate and remove damaged nucleobases from the DNA backbone by catalyzing heterolysis of the N-glycosidic bond. DNA glycosylases acting on modified purines typically activate the nucleobase for departure as a leaving group through a base-flipping mechanism, wherein the damaged nucleobase is extruded from the DNA helix and positioned in an active site pocket.[4,5] However, we recently showed that base-flipping is not a requisite for catalysis, as demonstrated by the *Bacillus cereus* DNA glycosylase AlkD, which is specific for cationic lesions.[6] Time-resolved crystal structures of AlkD-mediated base excision revealed that the substrate nucleobase remains stacked in the duplex over the course of the reaction and that three active site residues—Asp113, Trp109, and Trp187—cradle the substrate deoxyribose without contacting the nucleobase itself. Importantly, the indole rings of both tryptophan side chains form a series of C−H/π interactions with the lesion deoxyribose, and computational modeling suggested that these interactions contribute to lowering the barrier to depurination. C−H/π interactions are ubiquitous in protein structure and have been observed to facilitate protein-DNA binding.[7,8] However, despite their importance, to our knowledge there is only one enzymatic system in which these interactions have been shown to be catalytic,[9,10] and no empirical demonstration that C−H/π interactions participate in DNA repair catalysis. Here, we provide direct experimental evidence for the importance of C−H/π interactions in the catalytic mechanism of DNA base excision through a series of structural, biochemical, and computational studies on the AlkD enzyme and a well-defined 7-methyl-2′-deoxyguanosine (d7mG) oligonucleotide substrate.

Our previous AlkD structural analysis focused on repair of the minor groove lesion 3-methyl-2′-deoxyadenosine (d3mA).[6] To verify that the non-base-flipping mechanism and C−H/π interactions important to d3mA excision persist when the lesion present is d7mG, we determined the crystal structure of AlkD in a ternary complex with 7-methylguanine (7mGua) nucleobase and DNA containing 1-aza-2′,4′-dideoxyribose (1aR), an analogue to the key oxocarbenium ion intermediate formed during depurination of d7mG (Figure 1a, Scheme 1). This new AlkD/1aR-DNA/7mGua structure is virtually identical to the corresponding ternary complex with 3-methyladenine (3mAde), with an rmsd of 0.35 Å for all AlkD/1aR-DNA atoms. Specifically, the 7mGua nucleobase remained stacked in the DNA duplex and retained Watson−Crick pairing with the opposing cytosine (Figure 1b), as 3mAde did with thymine in the previous structure. Importantly, the contacts observed in the 3mAde structure presumed to be important for AlkD-mediated glycosidic bond cleavage were again observed in the 7mGua structure, including Asp113-mediated preorganization of a water molecule for nucleophilic attack on the oxocarbenium ion, as well as Trp109- and Trp187-mediated C−H/π interactions to the 1aR (Figure 1b,c).

With this structure in hand, we set out to delineate the roles of Asp113, Trp109, and Trp187 in lesion binding and catalysis using Michaelis−Menten kinetic analyses. We accomplished this by employing alanine mutants against a well-defined d7mG-oligonucleotide substrate. Wild-type or mutant AlkD proteins were added to solutions containing a fixed concentration of oligonucleotide bearing a centrally located d7mG lesion. Aliquots were removed from these reaction mixtures at various times and assayed for product formation by

Received: July 21, 2016
Published: August 29, 2016

DOI: 10.1021/jacs.6b07399
J. Am. Chem. Soc. 2016, 138, 11485−11488

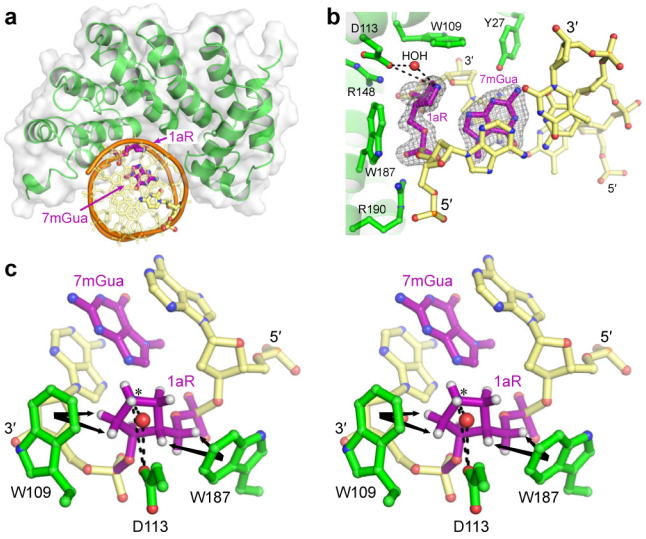
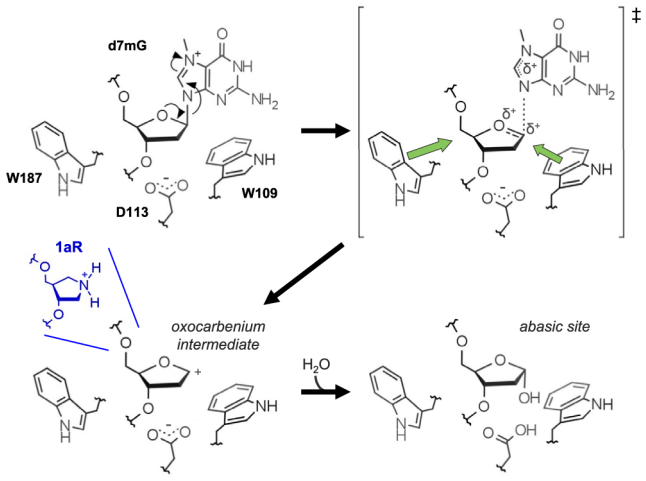


Figure 1. Crystal structure of AlkD bound to 1aR-DNA and 7mGua nucleobase. (a) Overall structure, with the protein in green, the DNA lesion in magenta, and the opposing dC nucleotide in yellow. (b) Close-up of the AlkD active site, with annealed omit electron density for the DNA lesion contoured to 3σ. (c) Stereoview of the AlkD active site. The putative water nucleophile is rendered as a red sphere. Hydrogen bonds are shown as dashed lines, C−H/π interactions are depicted as black arrows, and N1′ of 1aR is labeled with an asterisk (*).

Scheme 1. Mechanism for Depurination of d7mG-DNA by AlkD[a]



[a]C−H/π interactions in the transition state are indicated by green arrows, and structural similarity between the oxocarbenium intermediate and the 1aR isostere (blue) is highlighted. 7mGua was removed following depurination for visual clarity.

alkaline hydrolysis, which selectively cleaves DNA at the abasic site. In accord with the assumptions implicit in Michaelis−Menten analyses, only linear, initial-phase kinetic data were considered in determining reaction velocities (Figure S1). Replots of these initial reaction velocities (slopes) versus protein concentration afforded the expected saturation kinetics (Figure 2). From the data, Michaelis constants (K_m) and first-order rate constants (k_{cat}) were determined (Table 1). The K_m values indicate that all three alanine mutants bound substrate with comparable affinities as the wild-type enzyme (Table 1), suggesting that individually these side chains do not

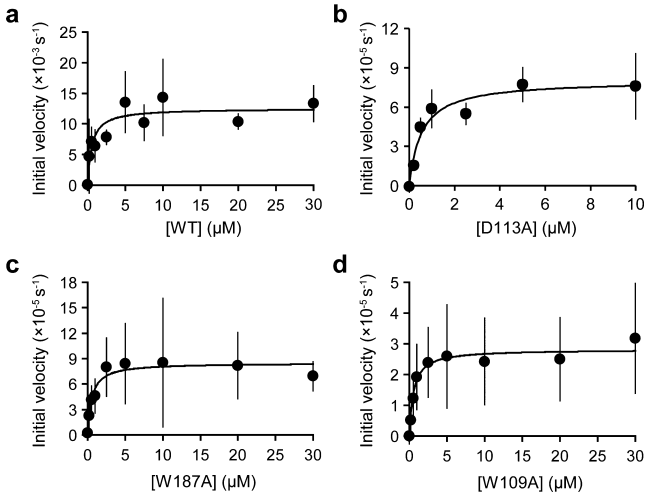


Figure 2. Michaelis plots showing saturation kinetics for depurination of d7mG-DNA by increasing concentrations of AlkD: (a) wild-type, (b) D113A, (c) W187A, (d) W109A. Numerical values of enzyme concentrations are provided in the Supporting Information (SI).

significantly contribute to lesion binding by AlkD. In contrast, comparison of k_{cat} values for AlkD mutants D113A, W109A, and W187A revealed that each catalyzed depurination of d7mG at rates <1% of that of wild-type AlkD (Table 1). We note that the k_{cat} values for wild-type and mutant AlkD determined here are consistent with rates of d7mG excision previously reported by our group for the same reaction at 20 μM AlkD.[11] The full Michaelis−Menten analysis reported here unequivocally indicates that Asp113, Trp109, and Trp187 each play a role in the chemistry of glycosidic bond cleavage.

The fact that tryptophan residues 187 and 109 are comparable to Asp113 in their catalytic importance is striking. Structurally divergent DNA glycosylases frequently contain an active site carboxylate residue that, when mutated, abrogates base excision activity.[4,5,12] Conversely, mutation of active site aromatic residues typically results in far less deleterious effects on catalytsis.[13−15] Aromatic side chains in base-flipping DNA glycosylases usually stabilize the enzyme−substrate complex by forming $\pi-\pi$ stacking interactions with the extruded nucleobase or by intercalating into the corresponding gap in the DNA helix. Conversely, the more than 100-fold attenuation of AlkD's d7mG excision activity upon removal of the aromatic side chains of either Trp109 or Trp187 demonstrates that the corresponding C−H/π contacts to the lesion deoxyribose contribute to glycosylase activity in a fundamentally different fashion than the π-stacking seen in base-flipping glycosylases. Indeed, using the Eyring equation to calculate reaction barriers ($\Delta G^{\ddagger}$) from saturating rate constants for AlkD-catalyzed d7mG depurination, we found that each set of C−H/π interactions contributed ~3−4 kcal mol^{-1} to lowering the barrier to glycosidic bond cleavage (Table 1). For comparison, similar interactions between Trp334 and aliphatic carbenium ions in the terpene biosynthesis enzyme aristolochene synthase were calculated to contribute from 2 to 4 kcal mol^{-1} to catalysis there.[16,10] These observations strongly implicate C−H/π interactions as relevant forces important for the catalytic mechanism of AlkD-mediated excision of cationic alkylpurine lesions.

To gain mechanistic insight into the catalytic role of C−H/π interactions, we developed a reductive computational model of the AlkD active site, from which we generated an

DOI: 10.1021/jacs.6b07399
J. Am. Chem. Soc. 2016, 138, 11485−11488

Table 1. Michaelis Parameters and Experimental and Computational Energetics

	WT	D113A	W187A	W109A
K_m (nM)	700 ± 500	600 ± 300	510 ± 70	1100 ± 900
k_{cat} (s^{-1})	(1.3 ± 0.3) × 10^{-2}	(8 ± 1) × 10^{-5}	(8 ± 4) × 10^{-5}	(3 ± 2) × 10^{-5}
k_{rel} (%)[a]	100	0.6	0.6	0.2
expt. r.e.[b]	3.3 × 10^4	200	200	70
$\Delta\Delta G^{\ddagger}$[c]	−6.2	−3.2	−3.2	−2.6
pred. r.e.[d]	6.9 × 10^4	200	9000	2.7 × 10^5
$\Delta\Delta E^{\ddagger}$[e]	−6.7	−3.2	−5.4	−7.4

[a]k_{cat} values relative to wild-type AlkD. [b]Rate enhancement relative to spontaneous depurination of d7mG (ref 18). [c]Decrease in activation free energy (kcal mol^{-1}) relative to spontaneous depurination. [d]Predicted rate enhancement for depurination of d7mG. [e]Predicted decrease in activation energy (kcal mol^{-1}).

approximation for the transition state to depurination of d7mG similar to that previously described.[6,17] The active site was modeled from the side chains of Asp113, Trp109, and Trp187, truncated at their β-carbons, and the lesion was modeled as the nucleoside (Figure 3). To validate this model, we first

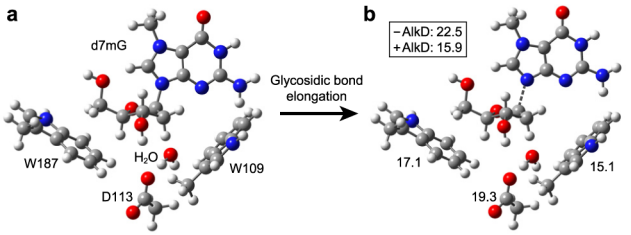


Figure 3. Computational structures and energetics for depurination of d7mG in a model of the AlkD active site. (a) Predicted structure of the GS of bound d7mG lesion in the AlkD active site with catalytic residues labeled. (b) Structure of the TSA to d7mG depurination in the AlkD active site. Predicted barriers to d7mG depurination ($\Delta E^{\ddagger}$, kcal mol^{-1}) in the absence (−AlkD) and presence (+AlkD) of wild-type enzyme are shown at the top left, and those calculated when catalytic residues are replaced by ghost atoms are given next to each residue.

calculated the activation energy to spontaneous depurination of d7mG. The value predicted by this model ($\Delta E^{\ddagger}$ = 22.5 kcal mol^{-1}) agreed very well with the related value calculated using the Eyring equation from the empirical data of Osborne and Phillips ($\Delta H^{\ddagger}$ = 23.4 kcal mol^{-1}) (Table 1).[18] For the enzyme-catalyzed reaction, we calculated from experimental rates of spontaneous and AlkD-catalyzed depurination of d7mG under saturating conditions (4.2 × 10^{-7} s^{-1} and 1.3 × 10^{-2} s^{-1}, respectively) that AlkD lowers the activation free energy to depurination ($\Delta G^{\ddagger}$) by 6.2 kcal mol^{-1}. Our computational model predicted a decrease in the corresponding activation energy ($\Delta E^{\ddagger}$) of 6.6 kcal mol^{-1} (Table 1). The agreement between these values indicates that the conformations considered in both the ground state (GS) and the transition-state approximation (TSA) in our model reasonably recapitulate those in solution.

Having validated this model, we next explored the contribution of each catalytic residue by sequentially replacing the atoms of residues Asp113, Trp109, and Trp187 with the appropriate Gaussian ghost atoms in order to maintain consistent basis. This approach predicted that knocking out Asp113 would decrease catalytic activity to 0.3% of that of wild-type AlkD, recapitulating the experimentally measured contribution of Asp113 to catalysis with excellent accuracy (k_{rel} of D113A = 0.6% wild-type activity) (Table 1). Because the conformations of the GS and TSA were held constant during this analysis, these observations suggest that Asp113 plays strictly an electronic role in AlkD's catalytic mechanism by providing electrostatic stabilization to the incipient positive charge on C1′ of the lesion deoxyribose during depurination. Such a contribution is the same as that postulated for active site carboxylates in other glycosylases.[5] Similarly, our calculations predicted that replacing Trp187 with ghost atoms lowers the relative enzyme activity to 12.1%, compared to the empirically determined 0.6% (versus wild-type enzyme). The fact that our conformationally restrained model predicts much of Trp187's role in catalysis suggests that this residue also facilitates depurination largely in an electrostatic fashion. For example, the electron-rich π-system of Trp187's phenyl ring is within 4 Å of the lesion O4′, which may permit effective donation of electron density from Trp187 to the nascent oxocarbenium ion in a fashion analogous to that postulated to occur from carboxylate residues in other glycosylases.[5] Interestingly, removing Trp109 was predicted to result in a 350% *increase* in catalytic activity relative to wild-type AlkD. The failure of this model to recapitulate the W109A experimental result (k_{rel} = 0.2% of wild-type activity) may be a result of consideration of only a single lesion conformation in the model. For example, Trp109 may be important in establishing the conformation of the lesion deoxyribose observed *in crystallo*. Crystal structures of AlkD bound to DNA that contain a lesion bearing an intact glycosidic bond (e.g., PDB 5CL3) show that the lesion deoxyribose adopts a C3′-endo sugar pucker, such as that observed in A-form DNA. This sugar pucker differs from the C2′-endo pucker observed in B-DNA. Notably, the presence of d7mG in duplex DNA is known to *reinforce* the C2′-endo conformation.[19] However, the C3′-endo conformation is believed to weaken the glycosidic bond by enhancing overlap between the occupied, nonbonding sp^3 orbital on O4′ and the vacant σ^* orbital on C1′, which is antibonding to the glycosidic bond.[5] Indeed, an enzyme utilizing this anomeric effect alone has been predicted to provide between 2 and 5 kcal mol^{-1} toward barrier reduction to depurination. Thus, Trp109's primary role in catalysis may be to establish the reactive C3′-endo conformation at d7mG by forming a strong C−H/π interaction with the highly polarized C−H bond at C2′, which is directly adjacent to the cationic alkylpurine.

The structural, biochemical, and computational data presented here unequivocally show that the C−H/π interactions between the AlkD active site and the lesion deoxyribose directly participate in catalyzing d7mG glycosidic bond hydrolysis. The nature of these interactions in AlkD is distinctly different than the π−π stacking interactions observed in base-flipping enzymes, which provide binding energy to

DOI: 10.1021/jacs.6b07399
J. Am. Chem. Soc. 2016, 138, 11485−11488

stabilize the extrahelical nucleobase lesion. Furthermore, our computational data reveal that these C−H/π interactions likely play a role other than merely providing direct electronic stabilization to the deoxyribose. Indeed, it has been suggested that C−H/π interactions in aristolochene synthase affect the reaction coordinate by modulating aliphatic carbenium ion structure.[20] AlkD's C−H/π interactions may play a similar role to guide conformation, for example, by establishing hyperconjugative overlap between the C2′−H σ-bond and the C1′−N9 σ^* orbitals, further weakening the glycosidic bond. Moreover, one can envision that π-donation from Trp187 and Trp109 might serve to increase the C−H σ-bonding electron density available for hyperconjugative stabilization of the incipient oxocarbenium ion. Undoubtedly, more work will be required to determine the precise nature of these C−H/π interactions' contributions to catalyzing glycosidic bond cleavage.

It is worth noting that carbon-centered cationic intermediates are common in both terpene biosynthetic pathways, where C−H/π interactions are known to be catalytic, and in enzymatic depurination of DNA. This observation lends credence to the assertion that C−H/π interactions may be catalytic in DNA repair beyond just the AlkD system, as there exists a preponderance of C−H/π interactions among DNA-binding proteins.[7,8] Along these lines, the strongest C−H/π interactions occur when π-rich aryl systems function as Lewis bases, donating electron density to positively polarized hydrogen atoms.[21−23] As evinced by proton NMR chemical shift values,[24] endo- and exocyclic oxygen atoms serve to polarize C−H bonds around the deoxyribose ring, making deoxyribose a suitable C−H donor even in an intact nucleotide (i.e., before depurination). Furthermore, the indole moiety of tryptophan residues is an extremely π-rich aryl system, rendering the pairing between tryptophan residues and nucleic acids particularly well-suited to engage in chemically useful C−H/π interactions.[25,26]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b07399.

Crystallographic refinement statistics, biochemical and computational methods, and kinetic data (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*brandt.eichman@vanderbilt.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Sheila David for providing the oligonucleotide containing 1aR. This work was funded by the National Science Foundation (MCB-1517695) and the National Institutes of Health (R01ES019625). Z.D.P. was supported by an NIH postdoctoral fellowship (F32ES027332). Z.D.P. and E.A.M. received financial support from the Vanderbilt Training Program in Environmental Toxicology (NIH T32ES07028). Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy Office of Science by Argonne National Laboratory, was supported by the U.S. Department of Energy (DE-AC02-06CH11357). Use of LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (085P1000817).

■ REFERENCES

(1) Friedberg, E. C.; Walker, G. C.; Siede, W.; Wood, R. D.; Schultz, R. A.; Ellenberger, T. *DNA Repair and Mutagenesis*, 2nd ed.; ASM Press: Washington, D.C., 2006

(2) Gates, K. S. *Chem. Res. Toxicol.* **2009**, *22*, 1747.

(3) Kim, Y.-J.; Wilson, D. M., III *Curr. Mol. Pharmacol.* **2012**, *5*, 3.

(4) Brooks, S. C.; Adhikary, S.; Rubinson, E. H.; Eichman, B. F. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834*, 247.

(5) Stivers, J. T.; Jiang, Y. L. *Chem. Rev.* **2003**, *103*, 2729.

(6) Mullins, E. A.; Shi, R.; Parsons, Z. D.; Yuen, P. K.; David, S. S.; Igarashi, Y.; Eichman, B. F. *Nature* **2015**, *527*, 254.

(7) Wilson, K. A.; Kellie, J. L.; Wetmore, S. D. *Nucleic Acids Res.* **2014**, *42*, 6726.

(8) Wilson, K. A.; Wells, R. A.; Abendong, M. N.; Anderson, C. B.; Kung, R. W.; Wetmore, S. D. *J. Biomol. Struct. Dyn.* **2016**, *34*, 184−200.

(9) Faraldos, J. A.; Antonczak, A. K.; González, V.; Fullerton, R.; Tippmann, E. M.; Allemann, R. K. *J. Am. Chem. Soc.* **2011**, *133*, 13906.

(10) Deligeorgopoulou, A.; Taylor, S. E.; Forcat, S.; Allemann, R. K. *Chem. Commun.* **2003**, 2162.

(11) Mullins, E. A.; Rubinson, E. H.; Eichman, B. F. *DNA Repair* **2014**, *13*, 50.

(12) O'Brien, P. J.; Ellenberger, T. *Biochemistry* **2003**, *42*, 12418.

(13) Hendershot, J. M.; O'Brien, P. J. *Nucleic Acids Res.* **2014**, *42*, 12681.

(14) Klapacz, J.; Lingaraju, G. M.; Guo, H. H.; Shah, D.; Moar-Shoshani, A.; Loeb, L. A.; Samson, L. D. *Mol. Cell* **2010**, *37*, 843.

(15) Livingston, A. L.; Kundu, S.; Henderson Pozzi, M.; Anderson, D. W.; David, S. S. *Biochemistry* **2005**, *44*, 14179.

(16) Faraldos, J. A.; Antonczak, A. K.; González, V.; Fullerton, R.; Tippmann, E. M.; Allemann, R. K. *J. Am. Chem. Soc.* **2011**, *133*, 13906.

(17) Dinner, A. R.; Blackburn, G. M.; Karplus, M. *Nature* **2001**, *413*, 752.

(18) Osborne, M. R.; Phillips, D. H. *Chem. Res. Toxicol.* **2000**, *13*, 257.

(19) Ezaz-Nikpay, K.; Verdine, G. L. *Chem. Biol.* **1994**, *1*, 235.

(20) Hong, Y. J.; Tantillo, D. J. *Chem. Sci.* **2013**, *4*, 2512.

(21) Nepal, B.; Scheiner, S. *J. Phys. Chem. A* **2014**, *118*, 9575.

(22) Zhao, Y.; Cotelle, Y.; Sakai, N.; Matile, S. *J. Am. Chem. Soc.* **2016**, *138*, 4270.

(23) Mishra, B. K.; Deshmukh, M. M.; Venkatnarayan, R. *J. Org. Chem.* **2014**, *79*, 8599.

(24) Wijmenga, S. S.; van Buuren, B. N. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287.

(25) Vandenbussche, S.; Diaz, D.; Fernandez-Alonso, M. C.; Pan, W.; Vincent, S. P.; Cuevas, G.; Canada, F. J.; Jimenez-Barbero, J.; Bartik, K. *Chem. - Eur. J.* **2008**, *14*, 7570.

(26) Mecozzi, S.; West, A. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 10566.

DOI: 10.1021/jacs.6b07399
J. Am. Chem. Soc. 2016, 138, 11485−11488

SUPPORTING INFORMATION FOR

A catalytic role for C-H/π interactions in base excision repair by *Bacillus cereus* DNA glycosylase AlkD

Zachary D. Parsons, Joshua M. Bland, Elwood A. Mullins, and Brandt F. Eichman

CONTENTS

Materials and Methods

Enzyme purification. Wild-type AlkD and all catalytic mutants were expressed in *E. coli* and purified to homogeneity as previously described.[1] All proteins except D113A were concentrated to 15 mg/mL before being flash-frozen in liquid nitrogen and stored at -80 °C. Due to poor solubility, the Asp113A construct was concentrated to 2 mg/mL before flash-freezing.

X-ray crystallography. Oligodeoxynucleotides [d(CCCGA(1aR)AGTCCG) and d(CGGACTCTCGGG), each at 0.54 mM] were annealed in 10 mM MES, pH 6.5, 40 mM NaCl, and ≤5.4 mM 7mGua nucleobase (saturated slurry) by heating to 85°C and slowly cooling to 20°C over several hours. Insoluble 7mGua was removed from the DNA solution by centrifugation. The AlkD/DNA complex was formed by incubating equal volumes of 0.54 mM DNA and 0.45 mM protein at 4°C for 30 min. The complex was crystallized using the hanging drop vapor diffusion technique. Drops were prepared from 2 µL protein-DNA solution [0.22 mM AlkD and 0.27 mM DNA], 2 µL reservoir solution [17% (w/v) PEG 4,000, 42 mM sodium acetate, pH 4.6, 85 mM ammonium acetate, and 5% (v/v) glycerol], and 1 µL seed solution [submicroscopic crystals of AlkD/1mA•T-DNA] and equilibrated at 21°C against an additional 500 µL reservoir solution. After several days, crystals were harvested, briefly soaked in reservoir solution supplemented with 15% (v/v) glycerol, and flash-cooled in liquid nitrogen.

X-ray diffraction data were collected at 100 K on beamline 21-ID-F at the Advanced Photon Source and processed using HKL2000.[2] Data collection statistics are provided in Table S1. Initial phases were determined by molecular replacement from a model of AlkD lacking DNA (PDB accession 3BVS) using the program Phaser.[3] DNA was then manually built in Coot,[4] guided by inspection of 2mF_o−DF_c and mF_o−DF_c electron density maps. The entirety of the 12-mer oligodeoxynucleotide duplex, including 7mGua nucleobase, was readily apparent, as were AlkD residues 1–229 and two non-native residues (−1–0) from the cleaved N-terminal affinity tag. The last eight residues (230–237) at the C-terminus were not apparent in the electron density maps and therefore were not modeled. Hydrogen atoms were placed in riding positions while avoiding steric clashes and were not refined against the X-ray data. For all non-hydrogen atoms, atomic coordinates, anisotropic temperature factors (TLS), and fractional occupancies were refined in PHENIX.[5] The final AlkD/DNA/7mGua model was validated using MolProbity[6] and contained no residues in the disallowed regions of the Ramachandran plot. Refinement and validation statistics are given in Table S1. The model was deposited in the Protein Data Bank under accession number 5KUB.

Structural images were created in PyMOL (https://www.pymol.com). mF_o−DF_c omit electron density maps were calculated using PHENIX after removing the lesion and the opposing cytosine and then performing simulated annealing on the remaining AlkD/DNA complex to minimize model bias. Maps were carved around the omitted atoms with a 2 Å radius and contoured to 3σ.

Substrate (d7mG-DNA) preparation and base excision assay. An HPLC-purified 5'-(6-carboxyfluorescein)-end-labeled oligonucleotide primer [FAM-d(GACCACTACACC)] and a desalted, unlabeled template [d(TTGATGGGAAATCGGTGTAGTGGTC)] were purchased from Integrated DNA Technologies. After reconstitution in ddH_2O, the two strands were diluted into annealing buffer (10 mM MES, 40 mM NaCl, pH 6.5) to final concentrations of 24 μM (primer) and 73 μM (template), heated to 85°C, and slowly cooled to 23 ± 2°C. 7-methyl-2'-deoxyguanosine (d7mG) was enzymatically incorporated into the oligonucleotide using 2.5 μM *E. coli* DNA polymerase I Klenow fragment, 200 μM each of dATP, dTTP, and dCTP, and 1mM 2'-deoxy-7-methylguanosine triphosphate (d7mGTP) in 1 mM DTT, 4 mM $MnCl_2$, 70 mM Tris, and 50 mM KCl, pH 7.6 final). The extension reaction was carried out for 20 min at 37°C and quenched by addition of EDTA to a final concentration of 80 mM. The extended substrate was diluted to 2 μM in water and stored on ice until used.

Glycosylase activity assays were performed as reported previously,[7] and electrophoresis bands were quantified using Gel Analyzer 2010 (http://www.gelanalyzer.com/). Initial reaction velocities (v_o) were extracted from plots of the slopes of fraction product as a function of time and at various protein concentrations (0, 0.2, 0.5, 1.0, 2.5, 5.0, 7.5, 10, 20, and 30 μM, as shown in Fig. S1, except for AlkD D113A, which was limited to 10 μM maximum concentration due to poorer solubility). Experiments were performed in triplicate for statistical analysis. In accord with the assumptions implicit to the Michaelis-Menten formalism, only the regions in which product release was well-described by linear functions were considered in analysis. For AlkD constructs D113A, W109A, and W187A, this corresponded to consideration of fractional product release up to 0.1. Due to the alacrity of the reaction between wild-type AlkD and d7mG-DNA, fractional product release up to 0.2 was considered in order to capture enough data points to adequately perform linear regression.

Computational methods. All electronic structure calculations were carried out using Gaussian09® software[8] on a Unix cluster running Linux; (mixed processors) AMD Opteron/Intel Xeon Westmere/Intel Xeon Sandy Bridge. Cartesian coordinates of protein residues Asp113, Trp109, Trp187, the catalytic water molecule, and the lesion nucleotide were extracted from PDB file 5CL3 (AlkD/3d3mA-DNA complex). The necessary atomic substitutions on 3-deaza-3-methyladenine to afford 7-methylguanine were made, the nucleotide was truncated to the corresponding nucleoside to reduce computational cost, and hydrogen atoms were manually placed to fill all open valences on heavy atoms (Asp113 was assumed to be in carboxylate form). To simulate the crystallographically-observed lack of motion in both DNA and protein during depurination of 3d3mA, the Cartesian coordinates of all AlkD heavy atoms were frozen, as were O3' and O5' of the d7mG nucleoside. The remaining atoms—all hydrogens and the heavy atoms in the lesion deoxyribose and nucleobase—were optimized until reaching a stationary point on the potential energy surface at the M06-2X/6-31+g(d) level of theory. To approximate the transition state to depurination, the glycosidic bond was elongated by 0.5 Å relative to its length in the optimized ground state structure,[9,10] the coordinates for atoms C1' and N9 additionally frozen, and the structure optimized again at the M06-2X/6-31+G(d) level of

theory. To verify that these structures represented minima along the potential energy surface for every allowed degree of freedom, frequency calculations were carried out on each at the same level of theory, holding constant those atoms for which coordinates had been frozen during optimization. Vibrational frequency analyses afforded zero imaginary frequencies in the ground state structure, and a single imaginary frequency in the transition state approximation structure. Animation of this imaginary frequency revealed rotation about the C-N bond of the 7-methyl substituent, occurring with a frequency of 17.4i cm^{-1}. Such a low-lying imaginary vibrational frequency was deemed negligible for subsequent computation of energetics, and this geometry was utilized for approximating the transition state to depurination.

Bond heterolysis energetics and the contributions of individual AlkD residues to catalysis of d7mG depurination were computed by single-point evaluations at the M06-2X/6-311++G(3df,2p) level of theory on the optimized structures of the ground state and transition state approximation complexes. These calculations were repeated, maintaining a consistent basis, by substituting for AlkD residue atoms the appropriate Gaussian ghost atoms (Table S2). Differences in single-point energies derived from these calculations were used to determine each protein residue's contribution to catalyzing depurination. For example, subtracting the single-point energy of the ground state complex, in which atoms defining Asp113 had been replaced by ghost atoms, from the single-point energy of the transition state approximation complex, in which Asp113 had been replaced by ghost atoms, afforded the barrier to depurination associated with knockout of Asp113 in AlkD. The same arithmetic calculation, but where no atoms had been replaced by ghost atoms, was taken as the barrier to wild-type AlkD-mediated depurination of d7mG. The difference in these two barriers was taken as the contribution of Asp113 to catalysis. The contributions made by Trp109 and Trp187 were determined in the same way.

The program Marvin (version 15.7.13) was used for drawing chemical structures (ChemAxon; www.chemaxon.com).

References

(1) Mullins, E. A.; Rubinson, E. H.; Eichman, B. F. *DNA Repair* **2014**, *13*, 50.

(2) Otwinowski, Z.; Minor, W. Enzymology, B.-M. in, Ed.; Macromolecular Crystallography Part A; Academic Press, 1997; Vol. 276, pp. 307–326.

(3) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658.

(4) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486.

(5) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213.

(6) 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.; Richardson, D. C. *Nucleic Acids Res.* **2007**, *35*, W375.

(7) Mullins, E. A.; Rubinson, E. H.; Pereira, K. N.; Calcutt, M. W.; Christov, P. P.; Eichman, B. F. *Methods* **2013**, *64*, 59.

(8) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, USA, 2009.

(9) Mullins, E. A.; Shi, R.; Parsons, Z. D.; Yuen, P. K.; David, S. S.; Igarashi, Y.; Eichman, B. F. *Nature* **2015**, *527*, 254.

(10) Dinner, A. R.; Blackburn, G. M.; Karplus, M. *Nature* **2001**, *413*, 752.

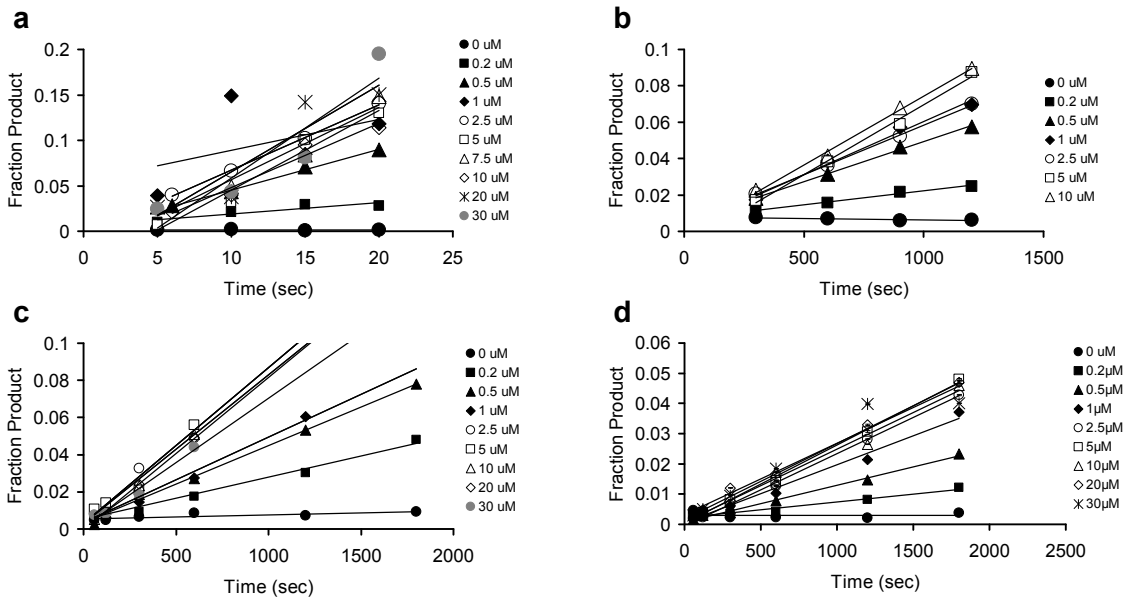


Figure S1. Initial reaction velocities for depurination of d7mG-DNA by wild-type and mutant AlkD proteins. (**a**) Wild-type, (**b**) D113A, (**c**) W187A, (**d**) W109A. All plots except **a** are derived from the average fraction product developed at various time points from three independent experiments. Plot **a** is from a single experiment, owing to the reaction rate precluding reliable collection of data at consistent time points during triplication. Error bars are omitted from **b** – **d** for visual clarity.

Table S1. X-ray data collection and refinement statistics

	AlkD/1aR•C-DNA/7mGua
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	38.45, 93.60, 47.99
α, *β*, *γ* (°)	90.00, 112.54, 90.00
Resolution (Å)	50.00–1.73 (1.79–1.73)[a]
R_{sym}	0.061 (0.215)
Avg. $I/\sigma I$	23.6 (9.0)
Completeness (%)	97.7 (96.2)
Redundancy	6.3 (6.0)
Wilson *B*-factor (Å^2)	19.0
Refinement	
Resolution (Å)	35.04–1.73 (1.79–1.73)
No. reflections	31,900 (2,809)
R_{work}	0.141 (0.160)
R_{free}[b]	0.180 (0.198)
No. atoms[c]	
Protein	1,965
DNA	487
Water	329
Avg. *B*-factors[c,d] (Å^2)	
Protein	23.9
DNA	42.2
Water	38.1
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.028
Ramachandran distribution (%)	
Favored	97.8
Allowed	2.2
Disallowed	0.0

[a] Statistics for the highest resolution shell are shown in parentheses.
[b] R_{free} was determined from the 5% of reflections excluded from refinement.
[c] Riding hydrogens were not included in no. atoms or avg. *B*-factors.
[d] Equivalent isotropic *B*-factors were calculated in conjunction with TLS-derived anisotropic *B*-factors.

Table S2. Zero point-uncorrected electronic energies of ground state and transition state approximation structures.

AlkD residues replaced by ghost atoms	Ground State[a]	TSA[a,b]	ΔE (kcal mol^{-1})[c]
None	-2114.521809	-2114.496531	15.9
All residues (AlkD, HOH)	-1003.240274	-1003.204354	22.5
Asp113	-1885.870471	-1885.839679	19.3
Trp109	-1711.420723	-1711.39662	15.1
Trp187	-1711.418309	-1711.391041	17.1

[a] Electronic energies computed at M06-2X/6-311++G(3df,2p), reported in Hartrees

[b] TSA = Transition state approximation

[c] $(E_{TSA} - E_{GS}) * 627.5095$

Table S3. Optimized Cartesian coordinates for computational models

AlkD/d7mG ground state structure
Charge = 0
Multiplicity = 1

C	-1.95763000	-5.07066000	-1.37684600
C	-3.19600200	-4.23337600	-1.57582900
C	-4.10042800	-4.34458600	-2.58523500
C	-3.65772000	-3.16353100	-0.73799200
C	-4.85314300	-2.68189900	-1.30225500
C	-3.17399000	-2.56747300	0.42971400
N	-5.11127600	-3.42177400	-2.42602800
C	-5.57510300	-1.62577500	-0.74013900
C	-3.90139700	-1.53255700	0.99672100
C	-5.08223900	-1.06856800	0.40551300
C	1.02290600	-4.14599100	3.20234400
C	1.30855600	-4.11127100	1.71175700
O	1.97779900	-5.03415600	1.20398300
O	0.86184300	-3.14557800	1.05286000
C	7.25007800	-0.03100600	-2.37456700
C	6.79107500	0.45057300	-1.04608200
C	7.21990300	1.54960100	-0.34926200
C	5.78281600	-0.16506600	-0.24531700
C	5.64666000	0.60289300	0.92976200
C	4.99289100	-1.30980000	-0.40188400
N	6.52748400	1.65069000	0.84492300
C	4.74614500	0.26941100	1.93650600
C	4.08617800	-1.63332200	0.59966300
C	3.96998200	-0.84579700	1.75168100
C	2.71825400	0.90892800	-1.49159500
O	2.30316400	2.14288900	-2.09623000
C	1.58980400	0.28477300	-0.70636900
O	1.24454600	1.13651500	0.40607200
C	0.25981100	0.03704300	-1.42823700
O	0.26069100	-1.09432200	-2.24679500
C	-0.69702700	-0.03747000	-0.23884200
C	-0.11546200	1.01681200	0.68893100
N	-4.28839500	3.93865000	0.97277000
C	-4.00418800	2.73006900	1.57143200
C	-1.98173900	2.77858600	0.66939400
C	-2.16931700	3.98612700	0.02752300
C	-3.41626600	4.68641900	0.14913000
N	-1.01294700	4.26238200	-0.67921600
C	-0.16029100	3.25914800	-0.46997700
N	-0.71762900	2.35455100	0.34482500
O	-0.35315700	-1.30163600	2.48504900
H	8.01459900	0.62869700	-2.79646700
H	7.67802000	-1.03812100	-2.30700400
H	6.42089700	-0.08340900	-3.09100400
H	6.69423400	2.33839300	1.56242700
H	5.08142500	-1.93439400	-1.28796300
H	3.45336200	-2.51490200	0.50172400
H	3.23560700	-1.12748000	2.50172100
H	4.65535300	0.87861400	2.83281500
H	7.97846000	2.27692700	-0.60852200
H	-1.06139800	-4.45074800	-1.26976300
H	-2.02947200	-5.67571900	-0.46659800
H	-1.80793400	-5.74745000	-2.22270500
H	-4.10988300	-5.03008400	-3.42255500
H	-5.89546800	-3.29658100	-3.04604600

H	-2.24211400	-2.89991800	0.88501000
H	-3.51104700	-1.05881900	1.89485500
H	-5.62842200	-0.24842100	0.86882600
H	-6.49532900	-1.26908000	-1.19677500
H	-0.30171200	0.83207000	1.74758600
H	-0.60224100	-1.01278100	0.24997700
H	-1.75215200	0.12304700	-0.47636700
H	0.02540400	0.88182100	-2.08882200
H	3.56199600	1.08403500	-0.81473600
H	3.02578300	0.19033700	-2.26231200
H	0.82952800	3.14877100	-0.91024700
H	1.49426300	-3.28589100	3.69537100
H	-0.05700400	-4.08579100	3.38637300
H	1.41319000	-5.06302300	3.64581200
H	-0.33795300	-1.60514900	3.40172800
H	0.12602000	-2.02841400	1.96024000
H	3.05264800	2.49941900	-2.59100000
H	0.39749100	-1.88221200	-1.68934200
H	1.95037100	-0.67345400	-0.30729800
N	-2.85697800	2.11216500	1.44473100
N	-4.96159500	2.19815300	2.37383000
H	-5.93514400	2.41147400	2.20507200
H	-4.77003000	1.24951600	2.67607700
C	-0.77750900	5.44364900	-1.50723000
H	-1.56558100	5.51788600	-2.25698100
H	-0.78984900	6.33575900	-0.88007700
H	0.19387300	5.33027500	-1.98837500
O	-3.75148300	5.74816900	-0.33811100
H	-5.18131900	4.39072200	1.14540300

AlkD/d7mG transition state approximation structure
Charge = 0
Multiplicity = 1

C	1.33634300	5.37361200	-1.42601600
C	2.66459400	4.67779300	-1.58660000
C	3.56623200	4.86537200	-2.58715900
C	3.23146700	3.68910000	-0.71432500
C	4.48244100	3.33231500	-1.25020600
C	2.80024400	3.07185000	0.46290700
N	4.67255700	4.06786200	-2.39026700
C	5.31059600	2.37981700	-0.65061800
C	3.63133500	2.14116400	1.06687600
C	4.86624300	1.79955600	0.50340200
C	-1.59129300	4.23299800	3.13858000
C	-1.84786500	4.12804100	1.64599200
O	-2.60939000	4.95558800	1.10500100
O	-1.28433100	3.20292900	1.01920500
C	-7.22407300	-0.70425000	-2.40290700
C	-6.73423700	-1.09649000	-1.05617600
C	-7.04655200	-2.21903700	-0.33530000
C	-5.81488300	-0.34997600	-0.25973000
C	-5.61091700	-1.06727000	0.93735000
C	-5.15745800	0.87279900	-0.43727700
N	-6.36586400	-2.21025800	0.86982600
C	-4.76983300	-0.60799300	1.94596900
C	-4.30900700	1.32275600	0.56649600
C	-4.12229100	0.58317200	1.74080200
C	-2.61313300	-1.12932700	-1.38476400
O	-2.06758200	-2.29366300	-2.00338400

C	-1.54779100	-0.35747600	-0.65172900
O	-1.08430800	-1.11755600	0.51919600
C	-0.25478100	0.02243500	-1.38821400
O	-0.40311100	1.14992600	-2.21582500
C	0.71322900	0.16438600	-0.19945300
C	0.16404000	-0.82922500	0.78053600
N	4.82950200	-3.96147300	0.88565700
C	4.49651600	-2.78540400	1.52486700
C	2.45342800	-2.89251500	0.64988200
C	2.70669900	-4.06882100	-0.03722300
C	3.98171600	-4.71781800	0.04754700
N	1.55660500	-4.38481600	-0.73214500
C	0.66458200	-3.42047800	-0.45301400
N	1.16985300	-2.51919100	0.38084100
O	0.10933600	1.54624000	2.51602600
H	-7.91981000	-1.44589300	-2.80707900
H	-7.74572300	0.25950300	-2.37124800
H	-6.39765300	-0.59713900	-3.11638200
H	-6.46125200	-2.89647000	1.60181900
H	-5.30625500	1.46189600	-1.33944200
H	-3.78209400	2.26992200	0.45366500
H	-3.44126000	0.96911000	2.49479200
H	-4.62627000	-1.17737300	2.86128900
H	-7.71245200	-3.03482800	-0.58495600
H	0.51732600	4.65849000	-1.29637200
H	1.33054400	6.01928200	-0.54124800
H	1.11517800	5.99317700	-2.29957100
H	3.51202600	5.52689300	-3.44195400
H	5.47473200	4.01517800	-2.99755000
H	1.83259900	3.31344000	0.90107000
H	3.29536900	1.65262200	1.97817000
H	5.49041000	1.05948700	0.99994400
H	6.27231700	2.11909600	-1.08629700
H	0.45902100	-0.85742900	1.82239600
H	0.58586600	1.14774700	0.28078400
H	1.77450100	0.03364500	-0.41585200
H	0.05292800	-0.79437100	-2.05081100
H	-3.40419600	-1.39758400	-0.67424700
H	-3.03629400	-0.44885000	-2.13506700
H	-0.33569400	-3.36603400	-0.87837700
H	-1.95036200	3.32907900	3.64739100
H	-0.51542400	4.32141800	3.33510500
H	-2.10431100	5.10015700	3.55623800
H	0.08155300	1.87406000	3.42362600
H	-0.43914800	2.20757000	1.97178400
H	-2.77064900	-2.72377800	-2.50813000
H	-0.60943800	1.92548400	-1.66112700
H	-2.00842600	0.55214800	-0.24218400
N	3.32709600	-2.21522200	1.43660600
N	5.45595900	-2.24514400	2.33437400
H	6.42800500	-2.37267000	2.08579600
H	5.22609700	-1.31546200	2.66531500
C	1.36651900	-5.54242300	-1.59931100
H	2.14479600	-5.55435200	-2.36336300
H	1.42920100	-6.45849700	-1.01004300
H	0.38384600	-5.45960500	-2.06452900
O	4.35820000	-5.75070200	-0.47582200
H	5.73725500	-4.38546000	1.05012600