Articles

Implications of Alternative Substrate Binding Modes for Catalysis by Uracil-DNA Glycosylase: An Apparent Discrepancy Resolved†

Ao Ma,‡ Jie Hu,‡ Martin Karplus,*,§ and Aaron R. Dinner*,‡

Department of Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, and Laboratoire de Chimie Biophysique, ISIS, Université Louis Pasteur, 67000 Strasbourg, France

Received May 28, 2006; Revised Manuscript Received August 25, 2006

ABSTRACT: A theoretical study showed that the base excision repair enzyme uracil-DNA glycosylase (UDG) exploits electrostatic interactions with backbone phosphate groups in the substrate for catalysis. Although experiments performed to test the calculated results confirmed the predicted importance of the −2, −1, and +1 phosphate groups, there was an apparent disagreement with regard to the +2 phosphate group. The calculations indicated that it made an important contribution, while experimentally, the effect of its deletion or neutralization was small. The +2 phosphate group interacts directly with an active site histidine (H148 in humans) in the crystal structure of UDG in complex with double-stranded (ds) DNA. We previously calculated that H148 has a strong anticatalytic effect due to its protonation, and here we use alchemical free energy simulations to estimate its site-specific pK_a. We find that it is positively charged over the entire experimental pH range (4−10), so its deprotonation cannot compensate for deletion or neutralization of the +2 phosphate group. The free energy simulations are facilitated by an efficient charge-scaling procedure that allows quantitative correction for the implicit treatment of solvent far from the active site; improvements are made to that method to account carefully for differences in the truncation of electrostatic interactions in the contributing molecular-mechanical and continuum-electrostatic approaches. Additional simulations are used to demonstrate that the +2 phosphate group is fully solvent exposed in complexes with single-stranded DNA substrates like those used in the experiments. In contrast, it is well-structured and buried in the dsDNA complex used in the original simulations. Differences in solvent shielding thus account for the apparent lack of an effect observed experimentally upon neutralization or deletion of this group.

Enzymes are remarkable catalysts capable of 10^{18}-fold rate enhancements relative to the corresponding reaction rates in solution (*1*, *2*). Most enzymes act primarily through specific functional groups presented in their active sites (*3*). However, an unusual example in which the substrate itself makes an important contribution to catalysis was identified recently by hybrid quantum-mechanical/molecular-mechanical (QM/MM) calculations (*4*). The base excision repair (BER) enzyme uracil-DNA glycosylase (UDG) cleaves the C1′−N1 bond of deoxyuridine (dU) through an oxacarbenium cation intermediate stabilized by phosphate groups in the substrate DNA backbone buried by the enzyme. This mechanism accounts for the activity of UDG mutants that lack key residues (*5*−*7*), an understanding of which is in turn important for interpreting recent experiments that seek to determine whether UDG plays a significant role in immunoglobulin class switching (*8*, *9*).

Given that the proposed "substrate autocatalysis" was based on simulations, Stivers and co-workers, who had independently suggested that the phosphate groups were important (*10*), undertook experiments (*11*) to test the results of the calculations. They determined the effect on catalysis of introducing methylphosphonate (MeP) substitutions at specific positions (*11*) of single-stranded DNA (ssDNA) oligomers. Those measurements confirmed most of the predictions of the QM/MM calculations (*4*), which were based on the X-ray crystallographic structure with double-stranded (ds) DNA (*12*). For the −2, −1, and +1 phosphate groups, the total contributions to lowering the activation energy ranged from 6.9 to 13.6 kcal/mol in the experiments, somewhat smaller than but in general agreement with the calculated values (12.5−16.9 kcal/mol). There was, however, a substantial discrepancy at the +2 position. The calculated contribution is ∼5 kcal/mol, but the experiments show a very small change in k_{cat} upon deletion (*10*) or neutralization (*11*).

† A.M. and A.R.D. were supported in part by a Burroughs Wellcome Fund Hitchings-Elion Fellowship. Part of the research at Harvard University was supported by a grant from the National Institutes of Health.

* To whom correspondence should be addressed. M.K.: e-mail, marci@tammy.harvard.edu; fax, (617) 496-3204. A.R.D.: e-mail, dinner@uchicago.edu; telephone, (773) 702-2330; fax, (773) 834-5250.

‡ The University of Chicago.

§ Harvard University and Université Louis Pasteur.

10.1021/bi061061y CCC: $33.50
Published on Web 10/28/2006

Since a full understanding of the mechanism is important, we have conducted additional calculations to investigate the origin of the difference at the +2 position. The +2 phosphate is the terminal one in the ssDNA experiments (*11*), so it is likely to be more exposed to solvent than in the dsDNA structure (*12*). Consequently, solvent shielding could limit the effects observed upon deletion or neutralization. An alternative explanation centers on H148 (in the human enzyme or H67 in *Escherichia coli*). H148 bridges the +2 phosphate and the attacking water molecule. In ref *4*, H148 was taken to be doubly protonated on the basis of continuum electrostatics calculations and, as a result, was found to have a large anticatalytic effect (6.0−8.2 kcal/mol toward destabilizing the rate-limiting transition state). Deletion of the +2 phosphate group and MeP substitution are expected to lower the pK_a for this group relative to that in the presence of dsDNA. If this led to neutralization of H148, the loss of the anticatalytic effect could compensate for that of the +2 phosphate substrate electrostatic interactions in catalysis and, in turn, result in little observed effect.

In this paper, we investigate these two scenarios by means of free energy difference simulations (*13*, *14*). We first establish the site-specific pK_a of H148 in the presence of ssDNA and dsDNA by "alchemical" simulations, in which we transform one chemical species (doubly protonated histidine) into another (singly protonated histidine) (*15*). These calculations support the conclusion that H148 is doubly protonated over the entire experimentally studied pH range [4−10 (*10*)] and thus plays an anticatalytic role in the removal of uracil from both ss- and dsDNA. Moreover, the structures from the simulations indicate that the position of the +2 phosphate of the ssDNA oligomer differs significantly from its position in the dsDNA complex such that it is almost fully solvent exposed. To explore the conformational preferences of the +2 phosphate group more quantitatively, we determine the potential of mean force along a coordinate that characterizes its position relative to the protein. The free energy profile confirms that the +2 phosphate is solvent-exposed at equilibrium and thus suggests that enhanced solvent shielding in ssDNA relative to dsDNA is the origin of the apparent difference between the earlier calculations (*4*) and experiments (*11*). The relation of these results to additional UDG data in the literature is discussed. An improved method for efficiently treating solvent in free energy simulations (*4*, *16*, *17*) is also introduced.

METHODS

Site-Specific pK_a Calculations. Absolute pK_as are difficult to estimate because electronic structure methods are limited by the prohibitive computational cost of sampling the Born−Oppenheimer surface with sufficient accuracy. Consequently, we obtain the pK_a of H148 in the protein environment (pK_a = $\Delta G_a^{pr}/2.3RT$) by combining molecular-mechanical free energy difference simulations with empirical data. The basis for this approach (*18*) is the thermodynamic cycle shown in Figure 1. Because the desired quantity (ΔG_a^{pr}) is a state function, we can follow a hypothetical path along which proton dissociation takes place in aqueous solution (ΔG_a^{aq}, which is known from experiment). To this end, we remove the proton from the acid in the protein environment (ΔG_c^{pr}), transfer the resulting conjugate base to aqueous solution (ΔG_t), and restore the proton ($-\Delta G_c^{aq}$). We then allow the dissociation to take place (ΔG_a^{aq}) and transfer the conjugate base back to the protein environment. Because the same species (the conjugate base) is transferred between the protein and aqueous solution environments, we can assume that the free energies of transfer (ΔG_t and $-\Delta G_t'$) cancel and that we need not evaluate them explicitly. As already mentioned, ΔG_a^{aq} is known experimentally. The remaining two required free energy differences (ΔG_c^{pr} and ΔG_c^{aq}) can be obtained from alchemical methods (*15*), as detailed in the following section. Adding the various contributions along the hypothetical path and incorporating constant factors, we have for the site-specific pK_a of H148

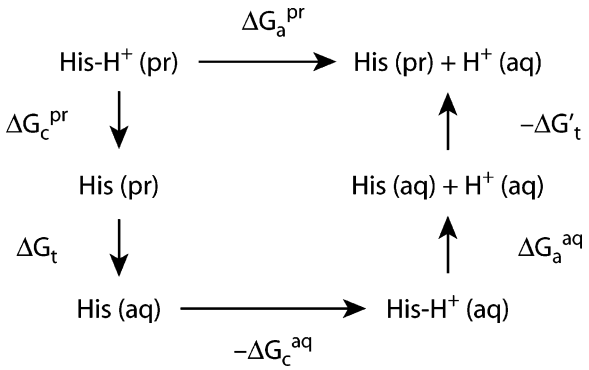

FIGURE 1: Thermodynamic cycle that serves as the basis for the pK_a calculation. ΔG_a^{pr} is the desired free energy difference that gives the pK_a in the protein. ΔG_c^{pr} and ΔG_c^{aq} are the free energy differences between the protonated and neutral forms of histidine in the protein and solution environments, respectively. ΔG_a^{aq} is the free energy for removal of the proton in solution, which can be obtained from experimental data. ΔG_t and $\Delta G_t'$ are the free energies for transferring the histidine from the protein to the aqueous peptide environment; these are assumed to be the same regardless of whether the excess proton is present in the solution, and they are thus neglected in the calculations.

$$pK_a^{pr} = (\Delta G_c^{pr} - \Delta G_c^{aq})/2.3RT + pK_a^{aq} \qquad (1)$$

where pK_a^{aq} is the experimentally measured pK_a value for histidine in solution [6.04 (*19*)], R is the ideal gas constant, and T is the temperature. Because there are two neutral tautomers of histidine (one with the remaining proton on N_δ, denoted HSD, and one with it on N_ϵ, denoted HSE) and we are interested in comparing ss- and dsDNA ligands, we carry out this procedure for four cases: HSP → HSD/ds, HSP → HSE/ds, HSP → HSD/ss, and HSP → HSE/ss.

Alchemical Free Energy Simulations. To calculate the relative free energies of two molecules (A and B), we employ thermodynamic integration (*20*−*23*). To this end, we change A into B by varying a coupling parameter ($0 \le \lambda \le 1$) in a potential energy function of the form

$$U(\lambda) = (1 - \lambda)U_A + \lambda U_B + U_C \qquad (2)$$

In the expression above, U_A (U_B) contains all the terms involving molecule A (B) and U_C contains the remainder of the system, which is common to both end points. For the linear coupling used in eq 2, the corresponding free energy difference (ΔG) is

$$\Delta G = G_B - G_A = \int_0^1 d\lambda \frac{\partial G}{\partial \lambda} = \int_0^1 d\lambda \langle U_B - U_A \rangle_\lambda \qquad (3)$$

where the $\langle U_B - U_A \rangle_\lambda$ denotes an average over the ensemble with the given λ value. In this study, A and B represent doubly and singly protonated forms of histidine, respectively. Bond length and angle terms were not included in U_A or U_B (in other words, these terms were not scaled by the coupling parameter).

For the linear coupling in eq 2 and the empirical energy function that was used (*24*, *25*) (specifically, the repulsive r^{-12} term in the Lennard-Jones potential), the derivative of the free energy in eq 3 is expected to be singular at $\lambda = 0$ and $\lambda = 1$ [$\partial G/\partial\lambda \sim \lambda^{-3/4}$ and $(1 - \lambda)^{-3/4}$, respectively] (*26*). The physical basis for this behavior, which can be derived by performing an expansion of $\partial G/\partial\lambda$ in powers of the density for a homogeneous system of soft spheres (*26*), is that the ensemble average in the integrand involves structures in which environment atoms overlap sterically with the atoms being introduced or removed. Consequently, in this study, the windows were spaced at intervals of $\Delta\lambda = 0.1$ between $\lambda = 0.1$ and $\lambda = 0.9$, and the contribution from this region was integrated numerically using Simpson's rule. When $0.01 \leq \lambda \leq 0.04$ and $0.96 \leq \lambda \leq 0.99$, the windows were spaced with $\Delta\lambda = 0.01$, and the $\langle U_B - U_A \rangle_\lambda$ values in each of these two regions were fit to the known functional forms (*26*): $\lambda^{-3/4}$ and $(1 - \lambda)^{-3/4}$, respectively. The contributions obtained by analytically integrating these functions were then added to that from the well-behaved intermediate region.

As discussed below, free energy simulations can converge quite slowly, particularly for protein systems, which are relatively inhomogeneous. It is of interest, therefore, to represent the system as economically as possible. A number of studies have shown that it is reasonable to treat solvent explicitly close to the area of alchemical change and implicitly elsewhere in molecular dynamics free energy simulations (*16*). The method was subsequently extended in the context of QM/MM calculations (*17*) and applied to UDG (*4*). The essential features are that charges of ionic groups farther than a certain distance from the alchemical change are scaled to enable conformational sampling in a vacuum and that continuum electrostatics methods are used subsequently to calculate corrections to the free energy differences. Below, we first describe the molecular dynamics simulations in the presence of the scaled charges and then the correction procedure. Importantly, we introduce an additional step in the procedure relative to earlier studies (*4*, *16*, *17*) to account for truncation effects that are important for the specific choice of charge scale factors used in this study (Tables S1 and S2 of the Supporting Information).

System. For the alchemical changes in the protein−DNA environment, the system studied was human UDG in the presence of either ss- or dsDNA (Figure 2). These molecules were represented by the CHARMM all-hydrogen topology and parameter sets (c22 for the protein and c27 for the nucleic acid) (*24*, *25*, *27*, *28*); the calculations were performed with CHARMM version c30b2 (*24*). In the case of the dsDNA simulations, an initial structure was obtained by swapping the C1 and N5 atoms of the pseudouridine inhibitor in PDB entry 1EMH (*12*). In the case of the ssDNA simulations, we went on to delete one of the DNA strands as well as one and three nucleotides from the 5′ and 3′ ends, respectively, of the remaining strand. The resulting system contains a five-nucleotide ssDNA oligomer (GpTpdUpApT) representative of the ligands studied by Jiang and Stivers (*11*). In both the ss- and dsDNA cases, hydrogen atoms were added using the HBUILD command in CHARMM (*24*).

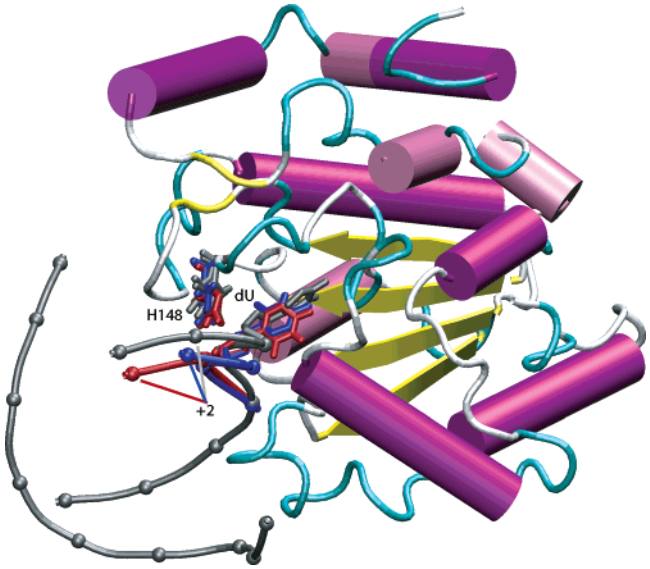

FIGURE 2: Comparison of nucleic acid structures for which $\lambda = 0.01$. The ligand and H148 from the HSE/ss simulation are colored red; those from the HSD/ss simulation are colored blue, and those from the HSD/ds simulation are colored gray (the HSE/ds simulation and starting structure are very similar to the HSD/ds simulation). The protein conformation is for the HSD/ss simulation, which is representative of that part of the system for all the simulations.

The complexes were solvated by centering a 22 Å sphere of pre-equilibrated TIP3P water (*29*, *30*) on H148 and deleting molecules that were within 2.8 Å of protein or DNA heavy atoms. On the basis of continuum electrostatics calculations (*4*), all of the histidine residues other than H148 were taken to be in the HSE neutral form. Side chains of the remaining ionizable residues (arginine, lysine, aspartate, and glutamate) and phosphate groups with any atoms more than 22 Å from H148 were scaled to mimic solvent shielding. The scale factors were determined using the electrostatic potential-based method introduced in ref *16*. In this method, the charges are scaled by the ratio of the potential at the origin from a specific residue in a vacuum ($\epsilon = 1$ everywhere) and that in aqueous solution ($\epsilon = 1$ for the protein and $\epsilon = 80$ for the solution). Here, the origin was taken to be the grid site closest to the N_ϵ atom of H148. The resulting scale factors are given in Tables S1 and S2. The charges of ionizable residues and phosphate groups wholly in the explicit solvent region initially were not scaled. All continuum electrostatic calculations were performed with the PBEQ module in CHARMM (*31*).

The resulting systems were divided into three parts. Protein and nucleic acid atoms more than 22 Å from H148 were fixed in place (FIX, 2448 atoms); those between 17 and 22 Å were harmonically restrained with force constants chosen to reproduce reasonable extents of thermal motion as discussed below (HARM, 791 atoms), and the remainder were completely free (FREE, 833 atoms). Throughout the study, the water molecules were restrained to remain in the 22 Å sphere with the stochastic boundary potential introduced by Brooks and Karplus (*32*). A switch function was applied to the van der Waals interactions over 8−12 Å, and a force shift function was applied to the electrostatic interactions to truncate them at 12 Å (*24*, *33*) (see ref *34* for a discussion

of different cutoff schemes in the context of free energy simulations).

The system configurations were then relaxed with the H148 residue in the doubly protonated form. First, the hydrogen atom positions were relaxed by energy minimization with fixed heavy atom positions [1000 steps of the steepest descent algorithm (SD) followed by 1000 steps of the adopted basis Newton−Raphson algorithm (ABNR) (*24*)]. Then, with the protein and DNA fixed, the water molecules were relaxed by applying the same minimization protocol, heating them from 0 to 300 K over 20 ps, and simulating their dynamics for an additional 20 ps at 300 K with the leap frog algorithm (*35*). The HARM and FREE atoms were then restrained harmonically, and their positions were adjusted by repeatedly applying the same minimization protocol with progressively weaker force constants (from 128 to 1 kcal mol^{-1} Å^{-1}, decreasing each time by a factor of 2). During the production run, the main chain and side chain HARM atoms were restrained to their minimized positions with force constants of 8.78 and 4.92 kcal mol^{-1} Å^{-1} scaled by eq 5 of ref *36*. These values were chosen to yield reasonable mean square displacements (0.410 and 0.732 Å^2 as in Table 7 of ref *36*).

For the free energy difference simulations to obtain ΔG_c^{aq} (Figure 1), the system was composed of a histidine with acetyl and *N*-methylamide blocking groups (to make the N- and C-termini neutral) in a sphere of the same radius as that used above (22 Å). The starting structure of the peptide was constructed from ideal internal coordinate values in the CHARMM (c22) topology and parameter sets (*25*). The waters were added to the doubly protonated system as described above, and their positions were relaxed by the same minimization and heating protocol used for the water molecules in the presence of the protein. Because the same solution calculations are used in the thermodynamic cycles for the ss- and dsDNA ΔG_a^{pr}, only two free energy simulations are needed in this case: one for the N_δ (HSP → HSD) and one for the N_ϵ (HSP → HSE) tautomer.

Molecular Dynamics (MD). Dual-topology free energy perturbation simulations were carried out using the BLOCK module of CHARMM (*37*). In this protocol, the side chains of the singly and doubly protonated tautomers of H148 are represented by two separate sets of atoms. All of the side chain atoms (from C_β outward) were duplicated; the protonated tautomer was represented by the HSP residue and the neutral tautomer by either the HSD or HSE residue in the CHARMM (c22) topology and parameter sets. Nonbonded and dihedral angle interactions between the duplicated side chains and the rest of the system were scaled by the control parameter (λ and $1 - \lambda$ in eq 2). As mentioned above, the initial structures were constructed by energy minimization with the protonated tautomer; the initial structures of the neutral tautomers were obtained by duplicating these coordinates and deleting the appropriate hydrogen atoms. We then set λ equal to 0.01 (protonated end point) and used Langevin dynamics to heat the system in steps of 20 K every 10 ps until 300 K was reached. The SHAKE algorithm (*38*) was used to constrain the lengths of the bonds to hydrogen atoms, and the integration time step was 1 fs. The random force was applied only to atoms more than 17 Å from H148. The friction coefficients for water oxygen atoms in this region were 31 ps^{-1} (*36*). Those for protein and nucleic acid heavy atoms were 500 ps^{-1} scaled by eq 5 of ref *36*.

One important issue concerning successful application of the thermodynamic integration procedure is that the system must reach equilibrium in each window. Recently, it was pointed out in a detailed analysis of thermodynamic integration that convergence can be assessed by evaluating whether the fluctuations in the derivative of the free energy (the integrand in eq 3) exhibit Gaussian statistics (*39*). For the system presented here, we found that this criterion was typically met only after simulation for ∼5 ns at each λ value. However, when we compared the numerical integration of the free energy derivatives from the $\lambda = 0.01$ to $\lambda = 0.04$ windows using the fully converged simulations and ones limited to 2 ns (using data only from the second half of each trajectory), the differences were typically on the order of 0.01 kcal/mol. These results suggest that the Gaussian statistics test is stricter than necessary for achieving errors routinely considered acceptable (∼1 kcal/mol). The reverse cumulative averaging procedure introduced in the same paper (*39*) was used and found to be very effective. At most λ values, we equilibrated the system for 1 ns and sampled for 1 ns; to ensure that there was no incremental error as λ was increased, we equilibrated the system until it exhibited Gaussian fluctuations in the free energy derivative at $\lambda =$ 0.04, 0.5, and 0.96.

Electrostatic Corrections. As mentioned above, continuum electrostatics methods are used to correct for scaling the charges of exposed ionizable groups. In the scheme introduced in ref *16* and refined in ref *17*, the free energy difference in the presence of the scaled charges in a vacuum is determined using thermodynamic integration (step I with contribution ΔG_I), the energy associated with restoring the scaled charges to their full values in a vacuum is calculated (step II with contribution ΔW_{II}), and the system is then transferred from vacuum to bulk solvent (step III with contribution ΔW_{III}). For step II, we use the CHARMM Coulomb electrostatic energies without any truncation, which was shown to provide good agreement with grid-based vacuum calculations in ref *17*. For step III, we use the PBEQ module in CHARMM (*31*). In all the Poisson calculations, no salt was included, and the water dielectric (ϵ) was taken to be 80; an ϵ of 1 was used otherwise (in a vacuum and in the protein interior). The calculations were initially conducted on a 75 × 75 × 75 grid with 2.0 Å spacing and focused to a 140 × 140 × 140 grid with 0.5 Å spacing.

In this study, we added a step to the procedure between step I and step II. Namely, we calculate the contribution to the (free) energy associated with extending the cutoff for the electrostatic interactions with the scaled charges from that used in the molecular dynamics simulation to infinity (step II′ with contribution $\Delta W_{II'}$). Neglecting this step for the systems considered in this study resulted in large corrections (−27.2 to −74.2 kcal/mol in Table 1), in contrast to expectation when the potential-based charge scaling procedure described above is used (*16*). Including step II′ reduced the magnitudes of these values to 2.4−20.4 kcal/mol. The contribution from this step was not important in refs *4* and *17* for two reasons. (1) Entropic contributions to enzyme activation free energies are small in general. In our earlier work on the chemical mechanism of UDG (*4*), we

Table 1: Calculated Free Energy Differences for H148[a]

system	ΔG_I	$\Delta\Delta W_{II,III}$	$\Delta\Delta W_{II',II,III}$	ΔG_c	$\Delta\Delta G$	pK_a
HSP → HSD/ds	39.1	−27.2	20.4	59.5	52.3	44.2
HSP → HSE/ds	20.9	−74.2	−4.0	16.9	8.5	12.2
HSP → HSD/ss	29.4	−63.3	2.4	31.8	24.6	24.0
HSP → HSE/ss	23.1	−65.7	3.3	26.4	18.0	19.2
HSP → HSD/pep	2.0		5.2	7.2		
HSP → HSE/pep	9.7		−1.3	8.4		

[a] $\Delta G_c = \Delta G_I + \Delta W_{II',II,III}$. $\Delta\Delta G = \Delta G_c^{pr} - \Delta G_c^{aq}$. p$K_a^{pr}$ = $\Delta\Delta G/2.3RT$ + pK_a^{aq}. $\Delta W_{II,III}$ is the total electrostatic correction obtained without step II′, and $\Delta W_{II',II,III}$ is that obtained with it.

thus explored only the effective energy (potential energy and implicit solvent free energy) surface with constrained minimization. Here, we are interested in protonation and binding, which both typically involve significant entropic contributions. Consequently, MD simulations at 300 K were used to obtain relative free energies. The MD simulations access a larger portion of the accessible conformation space of the enzyme than does minimization. Due to the greater heterogeneity of the sampled structures, errors are less likely to cancel. (2) The step II calculations in ref *4* were performed entirely with PBEQ even though $\epsilon = 1$ everywhere in that part of the procedure, so the truncation was treated consistently within each step. Here, the energies for increasing the charges to their full magnitudes in a vacuum were calculated with the Coulomb electrostatic term in the CHARMM force field, to which a force shift function is applied (*24*, *33*).

The total free energy differences reported are sums of the energy contributions from steps I, II′, II, and III for each of the four simulations (Table 1). For the blocked histidine in solution, there are no scaled charges, so the free energy differences are simply the sums of steps I and III. That is, we add the free energy change of transferring the blocked histidine and the sphere of water surrounding it into the bulk solution ($\epsilon = 80$) to the thermodynamic integration result.

Decomposition Analysis. Decomposition of the free energy into residue-by-residue contributions can provide valuable insight into the physical basis for the estimated pK_a values (*15*). For residue *i* in C in eq 2

$$\frac{\partial G_i^I}{\partial \lambda} = \sum_{j\in i}\frac{\partial G_j^I}{\partial \lambda} = \sum_{j\in i}\langle U_{jA} - U_{jB}\rangle_\lambda \qquad (4)$$

where G_i^I is the step I contribution from *i* to a calculated free energy difference (ΔG_c^{pr}), U_{jX} is the interaction (with the scaled charges) between atom *j* and block X (X = A or B), and the sums run over the atoms in *i*. The contributions reported include the solvent corrections and are

$$\Delta G_i = \int_0^1 d\lambda \frac{\partial G_i^I}{\partial \lambda} + \Delta W_{iX}^{II'} + \Delta W_{iX}^{II} + \Delta W_{iX}^{III} \qquad (5)$$

$W_{iX}^{II'}$ and W_{iX}^{II} are the average energies associated with extending the cutoffs and restoring the charges of residue *i* to their full magnitudes in a vacuum at a particular value of λ, respectively, and $\Delta W_{iX}^{II'}$ and ΔW_{iX}^{II} are the differences between the $\lambda = 1$ and $\lambda = 0$ end points. These quantities are straightforward to calculate since they are each based on a sum of pairwise terms. For ΔW_{iX}^{III}, we again take a difference between end point averages, in this case, of

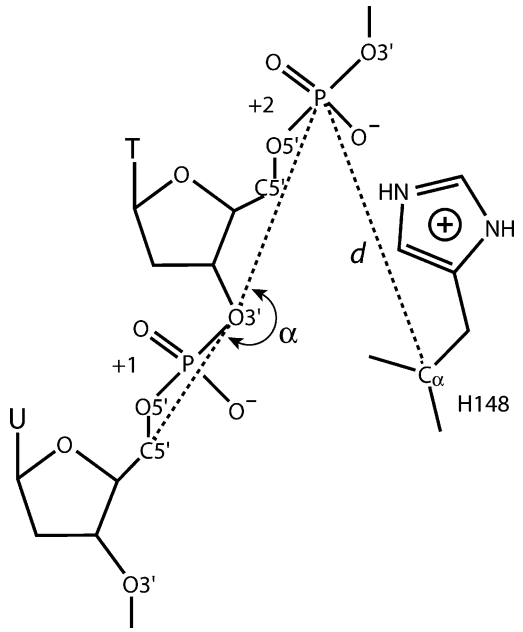

FIGURE 3: Schematic illustrating the coordinates (*d* and α) used to characterize the position of the +2 phosphate group.

$$W_{iX}^{III} = \sum_{j\in i}[\Phi_X(\epsilon{=}80) - \Phi_X(\epsilon{=}1)]q_j \qquad (6)$$

where Φ_X is the electrostatic potential obtained with the partial charges for H148 in state X with the dielectric boundaries determined by the full protein system (all other partial charges are set to zero). The electrostatic corrections were averaged over only every tenth saved structure (1 ps intervals) to limit computational cost.

Potential of Mean Force Calculations. To probe the conformational preferences of the +2 phosphate, which plays an essential role in this analysis, we determined the free energy as a function of an order parameter characterizing the position of the +2 phosphate group relative to the protein (*q*). The choice of *q* is discussed in the Results. In brief, it is a linear combination of *d*, the distance between the P of the +2 phosphate group and C_α of H148, and α, the angle spanned by the P of the +2 phosphate group, O3′ of the +1 phosphate group, and C5′ immediately 3′ of the +1 phosphate group (in other words, C5′ of the dU sugar) (Figure 3); specifically, $q = 0.9985d/\text{angstroms} + 0.0555\alpha/\text{degrees}$. The system was the same as that for the alchemical free energy simulations except that only the doubly protonated tautomer of H148 was included (corresponding to the $\lambda = 0$ end point), since the pK_a calculations make clear that this state is the most populated in all the cases that are considered. The potential of mean force along *q* was evaluated with umbrella sampling (*23*, *40*). A harmonic form was used for the biasing restraints; the force constants and minima of the restraints are detailed in the legend of Figure 6. For each window (each of which corresponded to a different restraint), we equilibrated the system for 500 ps and then sampled for an additional 500 ps. Although the charges were scaled as in step I above, the electrostatic corrections were not applied since our goal was to understand the forces that stabilize the structures observed sampled in the molecular dynamics simulations with the scaled charges

Table 2: Residues That Contribute at Least 2 kcal/mol to ΔG_c^{pr} and 1 kcal/mol to ΔG_I in One or More of the Four Simulations[a]

		ΔΔ*G* (HSD/ds)			ΔΔ*G* (HSE/ds)			ΔΔ*G* (HSD/ss)			ΔΔ*G* (HSE/ss)		
index	type	MD	Ele	Tot	MD	Ele	Tot	MD	Ele	Tot	MD	Ele	Tot
144	Gln	−1.3	−2.4	−3.7	−1.2	−1.2	−2.4	−2.4	−1.0	−3.4			
145	Asp	2.6	4.2	6.8	2.6	4.6	7.2	3.1	7.7	10.8	1.5	7.7	9.2
146	Pro	2.3	1.0	3.3	3.1	1.3	4.4	3.1	0.5	3.6	1.8	0.5	2.3
147	Tyr	−3.3	−1.8	−5.1	−4.0	−0.6	−4.6	−4.2	−0.3	−4.5	−2.6	−0.3	−2.9
210	Arg	2.6	−8.2	−5.6	3.7	−6.6	−2.9						
212	His	6.8	0.0	6.8				2.7	1.0	3.7	3.3	1.0	4.3
213	Gln	4.0	1.1	5.1	7.8	−0.7	7.1				3.3	−0.5	2.8
214	Ala				2.0	2.0	4.0	−2.3	−0.9	−3.2			
+3P		4.2	7.0	11.2	2.3	4.6	6.9						
+2P		52.6	10.6	63.2	39.0	8.3	47.3	39.1	0.1	39.2	17.4	12.2	29.6
+1P		10.5	27.6	38.1	11.3	23.3	34.6	20.4	8.2	28.6	40.2	0.0	40.2
−1P		11.1	4.6	15.7	9.3	4.1	13.4	15.2	9.9	25.1	4.7	11.5	16.2

[a] MD is the contribution to ΔG_I, Ele the contribution to $\Delta W_{II',II,III}$, and Tot their sum. A number of crystallographic water molecules also made contributions greater than 2 kcal/mol but are not listed for clarity. DNA nucleotides are numbered according to their phosphate groups, which dominate the contributions; the ssDNA oligomer spans positions +3 to −2. Non-ionic residues can have non-zero electrostatic corrections due to the contributions from steps II′ and III.

in a vacuum. Data from different windows were combined with the weighted histogram method (WHAM) (*41*).

RESULTS

In this section, we show that H148 is protonated in the presence of ss- and dsDNA. In the former case, the +2 phosphate is fully solvated, which accounts for the small effects observed upon neutralization or deletion in experiments with short ssDNA oligomers.

Protonation State of H148. The calculated free energy differences between the singly and doubly protonated states of H148 are given in Table 1. Both in the presence of the protein and in solution, removing a proton is unfavorable ($\Delta G_c > 0$) due to the loss of electrostatic interactions. However, the free energy cost is greater in the protein−DNA environment in all cases (6.8 kcal/mol ≤ ΔΔ*G* ≤ 51.1 kcal/mol). The ΔΔ*G* values correspond to pK_a shifts relative to the solution value (6.04), and the resulting absolute pK_a values (12.4−44.7) indicate that H148 is expected to be protonated over the experimentally studied pH range [4−10 (*10*)]. The results thus support the earlier assignment of H148 (*4*). Moreover, they are consistent with the observation that there is no change in activity over the pH range of 7.5−10.5 for the *E. coli* enzyme (*42*), which was cited as evidence for a neutral histidine at the corresponding position in ref *11*.

To gain physical insight into the basis for the estimated pK_as, we identified the residues that make the largest contributions to the ΔΔ*G* values using free energy decomposition analysis (Table 2). The sign of the free energy contributions is such that a positive number indicates that the doubly protonated state is preferred. As is reasonable, protein residues that are close to H148 play the most significant roles. These include ones that are close in sequence (Q144−Y147) and those in a loop that packs against H148 (R210−A214) (Figure 4). However, consistent with our earlier findings (*4*), the largest contributions overall come from phosphates in the DNA backbone, in particular the +2 phosphate that interacts directly with H148 in the X-ray crystallographic structure (*12*).

Binding Mode of Single-Stranded DNA. As described in Methods, systematic criteria were employed to ensure that the simulations were adequately equilibrated. Consequently,

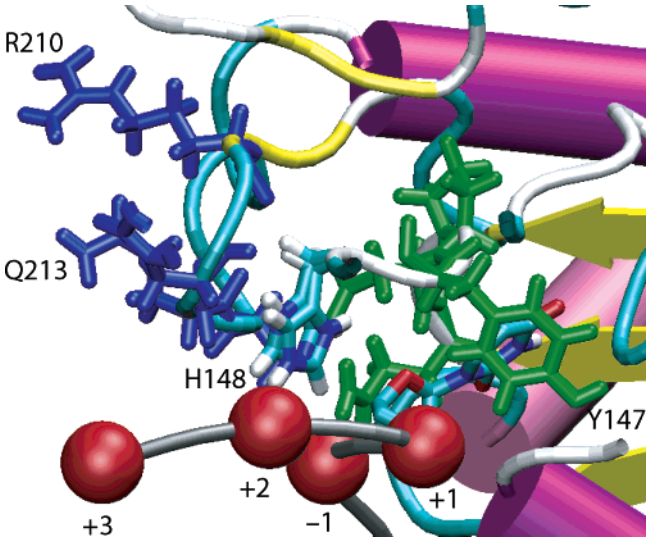

FIGURE 4: Residues contributing significantly to the determination of the site-specific pK_a of H148 (Table 2). Q144, D145, P146, and Y147 are colored green; R210, H212, Q213, and A214 are colored blue, and P atoms of −1 to +3 phosphate groups are colored red. The structure is from the HSD/ds simulation oriented as in Figure 2. H148 is as marked (colored according to atom type with cyan for C, blue for N, white for H, and red for O); the base of dU can be seen behind Y147 and the sugar behind the +1 phosphate group.

we can obtain meaningful information about the conformational space accessible to the molecule by examining the structures sampled at the end points of the free energy simulations ($\lambda = 0.01$ and $\lambda = 0.99$). In particular, in the HSE/ss simulation at the doubly protonated end point ($\lambda = 0.01$), there was a large movement of the ligand atoms 5′ of the deoxyuridine such that the +2 phosphate became more solvent exposed (~6 Å as can be seen by comparing the red and gray structures in Figures 2 and 5). To quantify the extent of solvation, we calculated the surface accessible area (SAA) of the +2 phosphate for the structures sampled in the last nanosecond of this simulation ($\lambda = 0.01$ for HSE/ss) and found that it generally was the same as the SAA in the absence of the protein (~60 Å^2). In contrast, neglecting the protein changes the SAA by ~67 Å^2 on average (from 87 to 20 Å^2) in the dsDNA case.

Consistent with the shift in position of the +2 phosphate, its contribution to ΔΔ*G* for deprotonating H148 in the HSE/

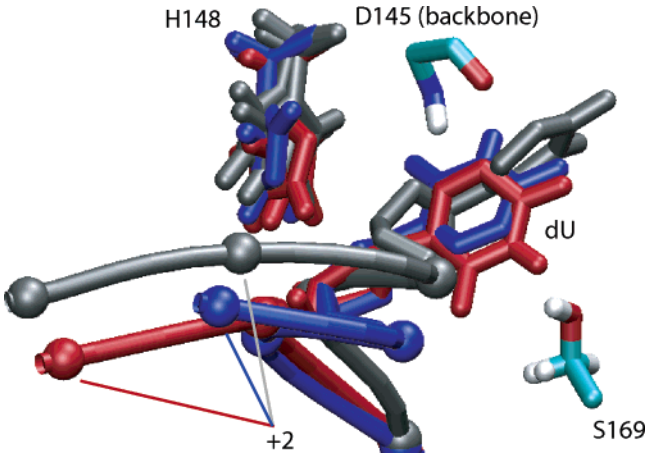

FIGURE 5: Orientation of the uracil base at the $\lambda = 0.01$ end point. Colors and orientation are the same as in Figure 2. The D145 backbone is from the HSD/ss structure, and the S169 side chain is from the HSE/ss structure to show the interactions that stabilize the different conformations of the uracil base in the ssDNA simulations.

ss simulation is considerably smaller than in the others (29.6 kcal/mol compared with 39.2−63.2 kcal/mol in Table 2). To determine if water compensated for this change, we calculated the interaction energy between the solvent molecules and H148 in each of its protonation states. The difference between the neutral and protonated forms is −6.3 kcal/mol for HSD/ss and 15.4 kcal/mol for HSE/ss. These balance the +2 phosphate contributions from the molecular dynamics simulations (39.1 and 17.4 kcal/mol, respectively, in Table 2). In other words, the sums of the +2 phosphate and solvent $\Delta\Delta G$ values are essentially the same for the two ssDNA simulations (32.6 and 32.8 kcal/mol for HSD/ss and HSE/ss, respectively) even though the +2 phosphate group becomes fully solvated in only the HSE/ss case. That the +2 phosphate of the ssDNA oligomer is solvated supports the idea that electrostatic interactions between the +2 phosphate and the active site are shielded in the ssDNA case, which would account for the lack of an effect observed upon deletion or neutralization of that group. The alternative binding mode is not inconsistent with the calculated pK_as. In the HSE/ss simulation, the whole DNA backbone shifts and the +1 phosphate occupies roughly the position of the +2 phosphate in the starting structure (Figure 2).

Comparison of the HSD/ss and HSE/ss structures at the $\lambda = 0.01$ end point, where both simulations correspond to the doubly protonated histidine, shows that the two structures are similar. This suggests that the protein is well-behaved. However, the uracil ring is oriented differently in the active site in the two simulations (Figure 5). In the HSD/ss simulation, the uracil ring rotates approximately +20° from the starting structure such that the C2═O2 and N3−H3 groups hydrogen bond with the backbone amide and carbonyl groups of D145, respectively. In contrast, in the HSE/ss simulation, the uracil ring rotates approximately −150° such that these groups interact with the side chain of S169 (the C2═O2 group with the C_β−H_β group and the N3−H3 group with O_γ). In both cases, the C4═O4 group hydrogen bonds with the backbone amide group of F158; these interactions stabilize the uracil ring in a shallower position in the active site pocket relative to the starting structure. This shift in the position of the uracil base is consistent with the increased

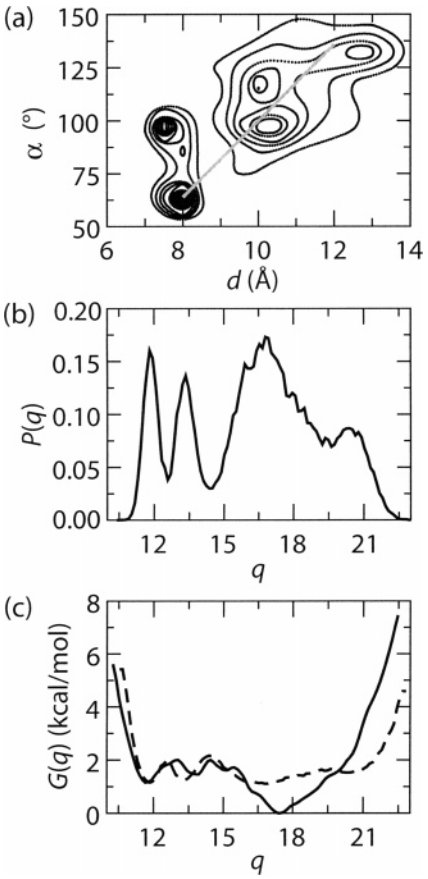

FIGURE 6: Projection of the structures observed in the alchemical free energy simulations. (a) Contour plot of the frequency of observing configurations with different values of d and α (defined in the text) during the ssDNA/HSE $\lambda = 0.01$ simulation. The dimensionless combined coordinate q is defined so that α/degrees = −17.9910d/angstroms + q/0.0555; each value of q/0.0555 thus corresponds to the intercept on the α axis of a line perpendicular to that shown. (b) Histogram of the data in panel a projected onto the combined coordinate q. (c) Potential of mean force along the combined coordinate calculated from independent umbrella sampling simulations (—) and comparison with the data in panel b plotted as k_BT ln $P(q)$ (− − −). The positions of the minima in dimensionless variable q (force constants in kilocalories per mole) of the harmonic biasing potentials used in the simulation are 10.0 (12.0), 10.65 (6.0), 10.8 (10.0), 11.2 (12.0), 11.65 (3.0), 12.3 (10.0), 12.95 (6.0), 13.1 (10.0), 13.25 (10.0), 13.6 (6.0), 14.25 (6.0), 14.6 (12.0), 14.9 (9.0), 15.25 (6.0), 15.65 (9.0), 15.9 (12.0), 16.35 (9.0), 16.55 (3.0), 17.2 (5.0), 17.85 (3.0), 18.4 (5.0), 19.15 (3.0), 19.7 (8.0), 20.45 (6.0), 21.0 (6.0), 21.75 (6.0), 22.25 (6.0), and 23.05 (6.0).

level of solvation of the backbone to which it is covalently linked.

We thus sought to identify a suitable coordinate for characterizing the position of the +2 phosphate and serving as the basis for umbrella sampling simulations. Good separation between structures with the phosphate close to its position in dsDNA and those in the HSE/ss alchemical simulations was obtained in projections onto d and α as defined in Methods (Figures 3 and 6a). To reduce the computational cost of the potential of mean force calculation, which increases exponentially with the number of independent coordinates, we linearly combined these two variables into a single one: $q = 0.9985d$/angstroms + 0.0555α/degrees (Figure 6b). Up to a constant factor, this mapping takes points in Figure 6a to the intercepts of lines orthogonal to that shown and in this way measures the projection onto the

Table 3: Electrostatic Interaction between the +2 Phosphate Group Atoms (P, O1P, O2P, O5′, and O3′) and the Deoxyuridine Sugar Atoms (C1′, H1′, C2′, H2′, H2″, C3′, H3′, O3′, C4′, H4′, and O4′)[a]

system	$\lambda = 0$	$\lambda = 1$	system	$\lambda = 0$	$\lambda = 1$
HSP → HSD/ds	2.2	2.5	HSP → HSD/ss	2.5	2.5
HSP → HSE/ds	2.6	2.5	HSP → HSE/ss	0.7	0.7

[a] Values (in kilocalories per mole) are averages over 50 configurations, spaced 10 ps apart, for the protonated ($\lambda = 0$) and neutral ($\lambda = 1$) end points.

indicated diagonal (see the legend of Figure 6 for further details). The potential of mean force along q is displayed in Figure 6c. There is a rough double well, and the basin corresponding to the solvated state ($q > 15$) is ~1.1 kcal/mol lower in free energy than that corresponding to the bound state ($q < 15$). Figure 6c does not contain information about the rate of diffusion along q, and thus, interconversion between the basins could be relatively slow despite the relatively small barrier height (see refs *43* and *44* for a general discussion). Nevertheless, the potential of mean force is adequate for demonstrating that, at equilibrium, the preferred binding mode of the +2 phosphate of the ssDNA ligand is solvated.

Implications for Catalysis. Poisson methods were used to estimate the impact that solvation of the +2 phosphate group has on its contribution to catalysis. Specifically, we calculated the electrostatic energy for a system with partial charges on only the +2 phosphate group atoms and the sugar of the deoxyuridine, which carries the positive charge in the oxacarbenium cation intermediate. The dielectric boundaries were determined from the protein and DNA molecules; no explicit solvent molecules were considered. Regardless of whether structures from the protonated or neutral end points were used, the interaction between the +2 phosphate group and the sugar is reduced significantly in the HSE/ss simulation with the alternative binding mode (0.7 kcal/mol compared with 2.2−2.6 kcal/mol in Table 3). These data support the hypothesis that recent experiments fail to observe a significant change in catalytic activity upon deletion (*10*) or neutralization (*11*) of the +2 phosphate group due to solvent shielding.

DISCUSSION

In this paper, we use free energy simulations to investigate the basis for an apparent discrepancy between theoretical predictions (*4*) and subsequent experiments (*10*, *11*) concerning the mechanism of the DNA repair protein uracil-DNA glycosylase. The calculations indicate that the terminal (+2) phosphate group of the ssDNA oligomers that were studied is fully solvated. The resulting shielding is sufficient to reduce the catalytic contribution of this group, as indicated by continuum electrostatics calculations using structures from either the protonated or neutral end points of the alchemical simulations. In the dsDNA case, this group is likely to be well-structured and contribute to catalytic activity as suggested previously (*4*).

The calculations support the idea that H148 is doubly protonated over the experimentally studied pH range and thus destabilizes the oxacarbenium cation intermediate and the transition state leading to it, as we suggested previously (*4*). In this regard, it is important to note that the H148L mutant has a specific activity that is only approximately half that of the wild type (*45*). One possible explanation for this observation is that slow dissociation of the apyrimidinic product in the absence of competing protein factors (*46*) masks the effects of introducing the aliphatic side chain [DNA binding is essentially unchanged (*45*)]. Alternatively, H148 is indeed neutral. The stabilization for the protonated tautomer in the HSE/ss simulation derives from the fact that the +1 phosphate group moves into the position occupied by the +2 phosphate group in the starting structure. If this is not the case, and neither the +1 or +2 phosphate group is near H148 (meaning each is likely to be solvated), a neutral tautomer is favored.

That the ssDNA and dsDNA ligands bind in somewhat different manners is consistent with available structural data. In particular, the uracil base in the ssDNA simulations is not as deep in the active site pocket as in the dsDNA structure. Rather, the region that it populates overlaps substantially that of a thymine base in a published X-ray crystallographic structure of UDG in complex with a thymine trinucleotide (*47*). In this structure, the +2 phosphate for that ligand binds in the position occupied by the +1 phosphate in the ssDNA simulations in this study, which indicates that there is a stable phosphate binding position in that location. In the case of UDG bound to a cleaved dUAAp oligomer, no electron density for the abasic trinucleotide is observed, suggesting that it is disordered if present (*48*). These data, combined with the fact that ssDNA oligomers have higher off-rates (*42*, *49*), further support the idea that the ssDNA oligomer samples alternative binding modes.

It is argued by Jiang et al. (footnote 4 of ref *11*) that differences in solvation between the various bound substrates are not likely, based on the fact that deletion and MeP substitution influence the reaction to similar extents. In our simulations, the −2, −1, and +1 phosphates do not move significantly, while the +2 phosphate becomes fully solvated. In comparing the MeP data with the calculations, we must also keep in mind that, although the total charge on these linkers is zero, they are not apolar. In contrast, in the decomposition analyses presented in ref *4*, all the partial charges on a group are removed. Consequently, the changes measured upon MeP substitution represent lower bounds for the contributions of the phosphates to k_{cat}. Of course, both deletion and linker neutralization can destabilize either the ground or transition state more than the other; depending on which state underwent a greater change in free energy, the net change would have a different sign.

We stress that the fact that substrate electrostatic interactions play a very large role in lowering the energy of the rate-limiting transition state for the base excision reaction does not mean that specific protein residues are unimportant. D145, P146, C157, and H268 all make large contributions. In particular, the calculated $\Delta\Delta E$ values (*4*) for these residues are sufficient to account for the hydrolysis of the minimal substrate deoxyuridine (dU) (*10*). Assuming H148 is neutral for the enzyme with dU, as indicated by continuum electrostatics calculations (determination of the pK_a by more rigorous means, such as those employed here, is not warranted in the absence of better structural data for this case), the protein is estimated to contribute 11.5 kcal/mol to cleaving the C1′−N1 bond, which is in good agreement with

the value of 8.3 kcal/mol suggested from differences in k_{cat} for the minimal and optimal substrates for the *E. coli* enzyme.

The charge scaling procedure for treating solvent far from the active site in an efficient but accurate manner (*4*, *16*, *17*) makes the simulations in our study computationally tractable. In this paper, we refined the procedure to account properly for free energy contributions associated with the fact that the electrostatic interactions are truncated only in the molecular dynamics simulations and not the grid-based postprocessing. Given the importance of substrate electrostatic interactions for the catalytic activity of UDG, it is natural to ask about the extent to which the results depend on the choice of protein dielectric. However, the scale factors for the ionizable groups are dominated by the water dielectric ($\epsilon = 80$), so it is unlikely that the conformations that are sampled would differ qualitatively from those in this study.

The results of our study reinforce the conclusions of the earlier QM/MM calculations, which pointed to the large catalytic and anticatalytic roles played by phosphates and H148, respectively, in addition to the "obvious" catalytic residues D145 and H268. With regard to these calculations, it is important to emphasize that, although quantitative accuracy is always desirable, their primary purpose is to provide qualitative insight into the features that give rise to a mechanism. Because they provide information that complements that accessible to experiments, such calculations and the ones described here play an important role in understanding protein behavior as already evidenced by the experiments motivated by our earlier study (*4*, *11*).

ACKNOWLEDGMENT

The simulations were performed in collaboration with Stuart Rice on the Laboratory Computing Resource Center Jazz Cluster at Argonne National Laboratory (Argonne, IL).

SUPPORTING INFORMATION AVAILABLE

Scale factors for ionic groups. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Fersht, A. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman & Co., New York.
2. Wolfenden, R. (1972) Analog approaches to structure of transition-state in enzyme reactions, *Acc. Chem. Res. 5*, 10−18.
3. Garcia-Viloca, M., Gao, J. L., and Karplus, M. (2004) How enzymes work: Analysis by modern rate theory and computer simulations, *Science 303*, 186−195.
4. Dinner, A. R., Blackburn, G. M., and Karplus, M. (2001) Uracil-DNA glycosylase acts by substrate autocatalysis, *Nature 413*, 752−755.
5. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA, *Nature 384*, 87−92.
6. Shroyer, M. J. N., Bennett, S. E., Putnam, C. D., Tainer, J. A., and Mosbaugh, D. W. (1999) Mutation of an active site residue in *Escherichia coli* uracil DNA glycosylase: Effect on DNA binding, uracil inhibition and catalysis, *Biochemistry 38*, 4834−4845.
7. Drohat, A. C., Xiao, G., Tordova, M., Jagadeesh, J., Pankiewicz, K. W., Watanabe, K. A., Gilliland, G. L., and Stivers, J. T. (1999) Heteronuclear NMR and crystallographic studies of wild-type and H187Q *Escherichia coli* uracil DNA glycosylase: Electrophilic catalysis of uracil expulsion by a neutral histidine 187, *Biochemistry 38*, 11876−11886.
8. Stivers, J. T. (2004) Comment on "Uracil DNA glycosylase activity is dispensable for immunoglobulin class switch", *Science 306*, 2042b.
9. Begum, N. A., and Honjo, T. (2004) Response to comment on "Uracil DNA glycosylase activity is dispensable for immunoglobulin class switch", *Science 306*, 2042c.
10. Jiang, Y. L., and Stivers, J. T. (2001) Reconstructing the substrate for uracil DNA glycosylase: Tracking the transmission of binding energy in catalysis, *Biochemistry 40*, 7710−7719.
11. Jiang, Y. L., Ichikawa, Y., Song, F., and Stivers, J. T. (2003) Powering DNA repair through substrate-electrostatic interactions, *Biochemistry 42*, 1922−1929.
12. Parikh, S. S., Walcher, G., Jones, G. D., Slupphaug, G., Krokan, H. E., Blackburn, G. M., and Tainer, J. A. (2000) Uracil-DNA glycosylase-DNA substrate and product structures: Conformational strain promotes catalytic efficiency by coupled stereoelectronic effects, *Proc. Natl. Acad. Sci. U.S.A. 97*, 5083−5088.
13. Simonson, T., Archontis, G., and Karplus, M. (2002) Free energy simulations come of age: Protein-ligand recognition, *Acc. Chem. Res. 35*, 430−437.
14. Kollman, P. (1993) Free-energy calculations: Applications to chemical and biochemical phenomena, *Chem. Rev. 93*, 2395−2417.
15. Gao, J. L., Kuczera, K., Tidor, B., and Karplus, M. (1989) Hidden thermodynamics of mutant proteins: A molecular-dynamics analysis, *Science 244*, 1069−1072.
16. Simonson, T., Archontis, G., and Karplus, M. (1997) Continuum treatment of long-range interactions in free energy calculations. Application to protein-ligand binding, *J. Phys. Chem. B 101*, 8349−8362.
17. Dinner, A. R., Lopez, X., and Karplus, M. (2003) A charge-scaling method to treat solvent in QM/MM simulations, *Theor. Chem. Acc. 109*, 118−124.
18. Merz, K. M., Jr. (1991) Determination of pK_as of ionizable groups in proteins: The pK_a of Glu 7 and 35 in hen egg white lysozyme and Glu 106 in human carbonic anhydrase II, *J. Am. Chem. Soc. 113*, 3572−3575.
19. Dawson, R. M. C., Elliott, D. C., Elliott, W. H., and Jones, K. M. (1969) *Data for biochemical research*, Oxford University Press, Oxford, U.K.
20. Kirkwood, J. G. (1935) Statistical mechanics of fluid mixtures, *J. Chem. Phys. 3*, 300−313.
21. Zwanzig, R. W. (1954) High-temperature equation of state by a perturbation method. 1. Nonpolar gases, *J. Chem. Phys. 22*, 1420−1426.
22. Brooks, C. L., III, Karplus, M., and Pettitt, B. M. (1988) *Proteins*, John Wiley & Sons, New York.
23. Frenkel, D., and Smit, B. (2002) *Understanding Molecular Sumulation*: *From Algorithms to Applications*, Academic Press, London.
24. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem. 4*, 187−217.
25. MacKerell, A. D., Jr., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B 102*, 3586−3616.
26. Simonson, T. (1993) Free energy of particle insertion: An exact analysis of the origin singularity for simple liquids, *Mol. Phys. 80*, 441−447.
27. MacKerell, A. D., Jr., and Banavli, N. (2000) All-atom empirical force field for nucleic acids: 2) Application to molecular dynamics simulations of DNA and RNA in solution, *J. Comput. Chem. 21*, 105−120.
28. Foloppe, N., and MacKerell, A. D., Jr. (2000) All-atom empirical force field for nucleic acids: 1) Parameter optimization based on small molecule and condensed phase macromolecular target data, *J. Comput. Chem. 21*, 86−104.
29. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926−935.

30. Neria, E., Fischer, S., and Karplus, M. (1996) Simulation of activation free energies in molecular systems, *J. Chem. Phys. 105*, 1902−1921.
31. Im, W., Beglov, D., and Roux, B. (1998) Continuum solvation model: Computation of electrostatic forces from numerical solutions to the Poisson-Boltzmann equation, *Comput. Phys. Commun. 111*, 59−75.
32. Brooks, C. L., III, and Karplus, M. (1983) Deformable stochastic boundaries in molecular-dynamics, *J. Chem. Phys. 79*, 6312−6325.
33. Steinbach, P. J., and Brooks, B. R. (1994) New spherical-cutoff methods for long-range forces in macromolecular simulation, *J. Comput. Chem. 15*, 667−683.
34. Brunsteiner, M., and Boresch, S. (2000) Influence of the treatment of electrostatic interactions on the results of free energy calculations of dipolar systems, *J. Chem. Phys. 112*, 6953−6955.
35. Hockney, R. W. (1970) The potential calculation and some applications, *Methods Comput. Phys. 9*, 136−211.
36. Brooks, C. L., III, and Karplus, M. (1989) Solvent effects on protein motion and protein effects on solvent motion: Dynamics of the active site region of lysozyme, *J. Mol. Biol. 208*, 159−181.
37. Tidor, B., and Karplus, M. (1991) Simulation analysis of stability mutant R96H of T4 lysozyme, *Biochemistry 30*, 3217−3228.
38. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes, *J. Comput. Phys. 23*, 327−341.
39. Yang, W., Bitetti-Putzer, R., and Karplus, M. (2004) Free energy simulations: Use of reverse cumulative averaging to determine the equilibrated region and the time required for convergence, *J. Chem. Phys. 120*, 2618−2628.
40. Torrie, G. M., and Valleau, J. P. (1977) Nonphysical sampling distributions in Monte Carlo free energy estimation: Umbrella sampling, *J. Comput. Phys. 23*, 187−199.
41. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A., and Rosenberg, J. M. (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J. Comput. Chem. 13*, 1011−1021.
42. Drohat, A. C., Jagadeesh, J., Ferguson, E., and Stivers, J. T. (1999) Role of electrophilic and general base catalysis in the mechanism of *Escherichia coli* uracil DNA glycosylase, *Biochemistry 38*, 11866−11875.
43. Ma, A., and Dinner, A. R. (2005) Automatic method for identifying reaction coordinates in complex systems, *J. Phys. Chem. B 109*, 6769−6779.
44. Ma, A., Nag, A., and Dinner, A. R. (2006) Dynamic coupling between coordinates in a model for biomolecular isomerization, *J. Chem. Phys. 124*, 144911.
45. Mol, C. D., Arvai, A. S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H. E., and Tainer, J. A. (1995) Crystal structure and mutational analysis of human uracil-DNA glycosylase: Structural basis for specificity and catalysis, *Cell 80*, 869−878.
46. Parikh, S. S., Mol, C. D., Slupphaug, G., Bharati, S., Krokan, H. E., and Tainer, J. A. (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA, *EMBO J. 17*, 5214−5226.
47. Savva, R., McAuley-Hecht, K., Brown, T., and Pearl, L. (1995) The structural basis of specific base excision repair by uracil-DNA glycosylase, *Nature 373*, 487−493.
48. Werner, R. M., Jiang, Y. L., Gordley, R. G., Jagadeesh, G. J., Ladner, J. E., Xiao, G., Tordova, M., Gilliland, G. L., and Stivers, J. T. (2000) Stressing-out DNA? The contribution of serinephosphodiester interactions in catalysis by uracil DNA glycosylase, *Biochemistry 39*, 12585−12594.
49. Slupphaug, G., Eftedal, I., Kavli, B., Bharati, S., Helle, N. M., Haug, T., Levine, D. W., and Krokan, H. E. (1995) Properties of a recombinant human uracil-DNA glycosylase from the *UNG* gene and evidence that *UNG* encodes the major uracil-DNA glycosylase, *Biochemistry 34*, 128−138.

BI061061Y