Published in final edited form as:

J Biomol Struct Dyn. 2014 June ; 32(6): 969–992. doi:10.1080/07391102.2013.799437.

Non-specificity and synergy at the binding site of the carboplatin-induced DNA adduct via molecular dynamics simulations of the MutS α -DNA recognition complex

Lacramioara Negureanu and Freddie R Salsbury Jr*

Department of Physics, Wake Forest University, Winston Salem, NC 27106

Abstract

MutS α is the most abundant mismatch binding factor of human DNA mismatch repair (MMR) proteins. MMR maintain genetic stability by recognizing and repairing DNA defects. Failure to accomplish their function may lead to cancer. In addition, MutS α recognizes at least some types of DNA damage making it a target for anticancer agents. Here, complementing scarce experimental data, we report unique hydrogen bonding motifs associated with the recognition of the carboplatin induced DNA damage by MutS α . These data predict that carboplatin and cisplatin induced damaging DNA adducts are recognized by MutS α in a similar manner. Our simulations also indicate that loss of base pairing at the damage site results in (1) non-specific binding and (2) changes in the atomic flexibility at the lesion site and beyond. To further quantify alterations at MutS α -DNA interface in response to damage recognition non-bonding interactions and salt bridges were investigated. These data indicate (1) possible different packing and (2) disruption of the salt bridges at the MutS α -DNA interface in the damaged complex. These findings (1) underscore the general observation of disruptions at the MutS α -DNA interface and (2) highlight the nature of the anticancer effect of the carboplatin agent. The analysis was carried out from atomistic simulations.

Keywords

MMR proteins; platinum-based anticancer agents; 1,3 platinum DNA adducts; cisplatin

1. Introduction

Carboplatin (*cis*-diammine(cyclobutane-1,1-dicarboxylato)-platinum(II)) is a second generation platinum-based anticancer drug used mainly on ovarian, lung, head and neck cancers (Wheate, Walker, Craig, & Onun, 2010). It was developed in the mid-1980s (Harrap, 1985) as a drug with fewer side-effects than its predecessor, cisplatin (*cis*-diamminechloroplatinum(II)) (Rosenberg, Van Camp, & Krigas, 1965). Clinically, carboplatin main benefits over cisplatin are reduced side-effects and lack of nephrotoxicity (Sharma, Graham, Blagden, & Gabra, 2011) while maintaining similar anti-tumor efficacy

*Corresponding author: salsbufr@wfu.edu, phone: 336.758.4975, fax: 336.758.6142.

Supplementary Material

Supplementary material dealing with changes in DNA flexibility induced by platinum-based adducts and mismatched base, residues sequencing and numbering maps for Msh2 and Msh6, non-specific hydrogen bonding at the 1,3 intra-strand platinum DNA adduct lesion site, C α RMSD and molecular surface representation for the carboplatin-damage recognition complex, hydrogen bonding patterns associated with the recognition of damaged and mismatched DNA substrates by MutS α , a complete list of strong and weak hydrogen bonding type and occupancy in carboplatin and cisplatin-induced MutS α -DNA recognition complexes, simulations conditions and duration, and non-bonding MutS α -DNA interactions at the cisplatin-induced damaged DNA interface is available from the authors directly and can be downloaded free of charge from the author's server at <http://bob.olin.wfu.edu/~web/>.

(Rajeswaran, Trojan, Burnand, & Giannelli, 2008). Thousands of platinum analogues have been synthesized and screened for anticancer effect, but only another such platinum anticancer drug, oxaliplatin, has been recently approved (Wheate et al., 2010).

Platinum-based anticancer agents exert their anticancer effect through formation of adducts with DNA, in which the DNA is severely distorted (Takahara, Rosenzweig, Frederick, & Lippard, 1995; Teuben, Bauer, Wang, & Reedijk, 1999). This damage eventually induces cell death. The exact mechanism by which cell death is induced is not fully understood (Jamieson & Lippard, 1999; Wang & Lippard, 2005; Vasilyeva et al., 2009). One hypothesis is that cell death is induced by the recognition of these platinum-DNA adducts by several specific cellular proteins (Andrews & Howell, 1990; Kartalou & Essigmann, 2001). Recent computational and experimental work supports the hypothesis that these adducts are recognized specifically by the MutS α complex, which is part of the MMR pathway (Salsbury Jr, Clodfelter, Gentry, Hollis, & Scarpinato, 2006; Negureanu & Salsbury Jr, 2012). The cellular processing of these platinum-based adducts by mismatch repair pathway involves the recognition of the DNA damage and the assembly of a multimeric complex of at least 20 polypeptides (Cannavo, Gerrits, Marra, Schlapbach, & Jiricny, 2007). Subsequently, the concerted action of this multimeric complex coordinates repair and cell death events.

Cisplatin and carboplatin each form a variety of intra and inter-strand platinum-DNA adducts (Takahara et al., 1995; Teuben et al., 1999). Although cisplatin and carboplatin structurally share the same *cis*-diammine ligands, the 1,2-d(GG) intra-strand platinum cross-linked DNA is the major adduct formed by cisplatin (Takahara et al., 1995), whereas the 1,3-d(GXG) intra-strand platinum cross-linked adduct is the major adduct formed by carboplatin (Teuben et al., 1999).

Conformational changes induced in DNA by forming the 1,2-d(GG) intra-strand platinum adduct in the absence of the protein have been extensively studied both experimentally, via NMR, X-ray crystallography and AFM studies, and computationally (Elizondo-Riojas & Kozelka, 2001; Spiegel & Magistrato, 2006; Sharma et al., 2007). The NMR study suggests more extreme conformational changes than do the X-ray and AFM studies (Gelasco & Lippard, 1998; Todd & Lippard, 2010; Dutta, 2011). The NMR study found that the 1, 2-d(GG) cisplatin-induced intra-strand adduct bent the DNA helix towards the major groove and unwound it at the site of the platination. This would result in de-stacking of the bases (Gelasco & Lippard, 1998). However, more recent high resolution X-ray (Todd & Lippard, 2010) and AFM (Dutta, 2011) studies indicate more modest conformational changes. A computational study of MutS α (Salsbury Jr et al., 2006) bound to the same adduct agrees with these latter more modest conformational changes found by X-ray and AFM studies. The key interactions found in this paper were indirectly confirmed via mutagenesis analysis combined with cellular assays for repair function and cell death (Salsbury Jr et al., 2006; Vasilyeva et al., 2009).

In contrast, there is limited experimental evidence on the 1,3-d(GXG) adduct, the main adduct formed by carboplatin. A structural NMR (Teuben et al., 1999) study of a 12 base pair DNA fragment with a central G*TG* (* indicates N7 platinated base) 1,3-intra-strand platinum cross-link indicates more substantial changes at the lesion site. The changes are asymmetric in that base pairing is lost for both the 5' G*-C and the central T-A base pair, and the thymine in the middle of the cross-link is flipped out from the minor groove. To accommodate this structural change, the minor groove widens (Teuben et al., 1999). This structural study, however, helps validate our computational investigation. As discussed in the second section of the results we see the same conformational changes in our computational study.

The MutS α protein complex, whose recognition complex with the carboplatin-induced DNA adduct is the focus of our study, is a heterodimer complex (Msh2-Msh6) of DNA mismatch repair (MMR) proteins. The MMR pathway, one of the four cellular pathways responsible for DNA repair (Kartalou & Essigmann, 2001), maintains genetic stability in prokaryotes and eukaryotes by detecting and repairing mismatched bases and insertion/deletions loops mistakenly incorporated during replication (Iyer, Pluciennik, Burdett, & Modrich, 2006). Failure to accomplish this function leads to the accumulation of errors and mutations in a genome, which have been linked to the development of cancer (Iyer et al., 2006). It is not surprising that inherited defects in the MMR pathway, particularly in *MSH2* and *MSH6* genes, are associated with a wide variety of cancers (InSight database: <http://www.insight-group.org>), most notably hereditary non-polyposis colon cancer (HPPC) (Peltomaki, 2003).

In addition to the DNA repair function, MutS α detects and provokes cell death in response to certain types of DNA damage, including the lesions produced by chemotherapeutic agents cisplatin and carboplatin (Duckett et al., 1996; Fink et al., 1996; Mello, Acharya, Fishel, & Essigmann, 1996; Salsbury Jr et al., 2006; Vasilyeva et al., 2009). Many details remain to be elucidated in precisely how cell death is induced, although it has been shown that apoptosis is the ultimate cellular endpoint (Young et al., 2003; Drotschmann, Topping, Clodfelter, & Salsbury Jr, 2004; Salsbury Jr et al., 2006; Vasilyeva et al., 2009). Interest in exploiting this death response has led to work in developing MutS α as a target for drug development (Salsbury Jr et al., 2006; Vasilyeva et al., 2009; Vasilyeva et al., 2010). In this regard, detailed studies have been performed to elucidate the molecular origin of the cell death response to the 1,2d(GG) adduct, i.e., the main adduct formed by cisplatin and other DNA damages (Salsbury Jr et al., 2006; Vasilyeva et al., 2009; Salsbury Jr, 2010; Ghosh, Salsbury Jr, Horita, & Gmeiner, 2011; Negureanu & Salsbury Jr, 2012; Negureanu & Salsbury Jr, 2012).

A natural follow-up to these studies is to examine the protein response to the 1,3d(GXG) intra-strand platinum adduct, i.e., the main adduct formed by carboplatin. Here, we report unique hydrogen bonding motifs associated with the recognition of the carboplatin-induced damage DNA by MutS α via all-atom unbiased simulations of the recognition complex. Consistent with experimental evidence, our simulations also indicate that loss of base pairing at the damage site results in non-specificity and changes in the atomic flexibility at the lesion site and beyond. Further quantifying alterations at the protein-DNA interface in response to damage recognition, predicted non-bonding interactions and salt bridges indicate possible different packing and disruption of the salt bridges at the MutS α -DNA interface in the damaged complex. Finally, in order to encourage further work in this area, especially much needed experimental studies, we predict key residues which would presumably enable MutS α to distinguish between different damaging DNA adducts induced by platinum-based anticancer agents. These findings advance the understanding of the mechanism by which MMR interact with damaged DNA and highlight the nature of the carboplatin cytotoxicity.

2. Methods

2.1 Molecular Dynamics Simulations

Preparing the MutS α -DNA complex—The structural model for our simulations on 1,3-d(GCG) intra-strand platinum-DNA adduct recognition by the MMR machinery was based on the X-ray structure (Warren et al., 2007) of human Msh2-Msh6 heterodimer complex with duplex DNA and two ADP molecules bound in the ATPase sites of the heterodimer (PDB ID 2O8B). Each subunit of MutS α is divided into five domains (Warren et al., 2007): the mismatch binding domain, the connector domain, the lever domain, the clamp domain, and the ATPase domain (Figure 1). At a 2.75 Å resolution and including two ADP molecules (in the context in which the conformational coupling between DNA recognition

and the nucleotide binding site plays a crucial role in MutS α 's function), the PDB ID 2O8B structure is the best currently available structural model for the investigated system.

All titratable residues were left in the dominant protonation state at physiological pH. Coordination of $[cis-Pt(NH_3)_2]^{2+}$ fragment to DNA, in which platinum atom links N7 atoms of two guanines form the same DNA strand, G8 and G10, alters the duplex DNA. The mismatched DNA was used as a template in building the 1,3 intra-strand platinum cross-link. The cross-linked structure was fitted into the binding pocket of the protein to maximize the structural overlap with the mismatched DNA structure, followed by rotations and translations to minimize the energy of the unrelaxed structure using the coordinate manipulation and energy minimization facilities of CHARMM (Brooks et al., 1983). The DNA sequence used was selected to maximize the correspondence with previous experimental and computational studies; (Salsbury Jr et al., 2006; Vasilyeva et al., 2009; Vasilyeva et al., 2010; Negureanu & Salsbury Jr, 2012; Negureanu & Salsbury Jr, 2012), while also allowing for the formation of the 1,3 intra-strand platinum cross-link rather than the 1,2 intra-strand cross-link. However, the precise sequence used to flank the cross-link should not significantly affect the results. N-terminal residues 1-361, not resolved in the crystalline structure, were omitted from the simulations. Note that in our system residue 1 of Msh6 corresponds to residue 362 in the solved structure. Hydrogen atoms were added using the *hbuild* facility of CHARMM. The prepared structure was fully solvated in a rectangular box of TIP3P water (Jorgensen, Chandrasekhar, Madura, Impey, & Klein, 1983) (33,143 molecules of water) using the VMD package (Humphrey, Dalke, & Schulten, 1996), keeping a 10 Å minimum distance between each face of the box and the solute. 57 Na⁺ ions were added to neutralize the total charge of the solvated system using the Autoionize plugin from VMD. There are 855 amino acids in Msh2, 974 amino acids in Msh6, 30 nucleotides in the DNA substrate, and two ADP molecules, a total of 30,046 atoms in the protein-DNA complex and 129,533 atoms in the simulated recognition complex.

Simulations protocol—Simulations were performed using CHARMM27 force field (Brooks et al., 1983) and additional parameters based on pre-existing cisplatin parameters (MacKerell Jr et al., 1998; Scheeff, Briggs, & Hawell, 1999; MacKerell Jr, Banavali, & Foloppe, 2001). CHARMM27 force field has been extensively parameterized for a wide range of biologically important molecules, including amino acids and nucleic acids. Each simulation was the result of a combined CHARMM/NAMD protocol that was derived from earlier protocols (Salsbury Jr, Crowley, & Brooks, 2001; Knaggs, Salsbury Jr, Edgell, & Fetrow, 2007) that have been used in multiple previous studies (Salsbury Jr et al., 2006; Vasilyeva et al., 2009; Vasilyeva et al., 2010; Negureanu & Salsbury Jr, 2012; Negureanu & Salsbury Jr, 2012), and the rationale behind it has been reviewed recently (Salsbury Jr, 2010). The water molecules were briefly minimized for 100 cycles of conjugate gradient minimization with a small (0.25 kcal/(mol·Å²)) harmonic force constant on all protein atoms. The entire system then underwent 250 ps of molecular dynamics simulation to achieve a thermal equilibration using Berendsen pressure regulation with isotropic position scaling (Berendsen, Postma, van Gunsteren, DiNola, & Haak, 1984). The system's temperature was equilibrated by reassigning atom velocities from a Boltzman distribution for a given temperature every 1000 cycles, in a 25 K increment, from an initial temperature of 0 K to a target temperature of 300 K.

Following the equilibration, for each simulation, a 20 ns NPT production simulation at 300K and 1 bar was performed using NAMD package (Kale et al., 1999) with standard parameters and periodic boundary conditions. All bonds involving hydrogen atoms in the protein were constrained using SHAKE constraint algorithm (van Gunsteren & Berendsen, 1977). A 2 fs integration time step was used for pair and bonded interactions, with van der Waals potential switched to zero starting at a radius of 8 Å and ending at 12 Å. Particle Mesh Ewald (PME)

(Darden, York, & Pedersen, 1993) method for full long-range electrostatics was used, with a maximum space between grid points of 1 Å and a local interaction distance of 12 Å, evaluated every two time steps. The short-range non-bonded interactions were calculated at each time step. Constant temperature was achieved using Langevin thermostat, as implemented in NAMD (Kale et al., 1999), with a damping coefficient of 5 ps⁻¹. Berendsen's constant pressure algorithm with a target pressure of 1.01325 bar, a compressibility of 45.7 mbar, a relaxation time of 1 ps, and a pressure frequency of 40 fs provided the pressure control during the simulations.

Four simulations employing the same protocol with different initial velocities and the same coordinates were performed. For the production simulations, initial coordinates, velocities and system dimensions were taken from the final state of the corresponding equilibration simulation.

Cα root-mean-square deviations and total energies are provided in Supplementary Material (SM) Figure S5. These data show that there are two different relaxation timescales, a fast one on the 10s-100s of picosecond timescale, and a slow one on the nanosecond timescale. Data show that most of the relaxation to equilibrium occurs within the first 2 ns, and while there may be additional long-time relaxation, starting the simulations analysis at 8 ns allows for a conservative removal of the majority of the non-equilibration effects. Since our four different simulations started from different initial conditions, it is expected they to show different pathways to equilibration, as well as variation in relaxation. To highlight the differences between the initial structure and the representative structure from molecular dynamics simulations, their superimposition is included in SM, Figure S8.

The structural model and the molecular dynamics simulations protocol for the simulations of the 1,2 intra-strand platinum-DNA MutSα complex are presented elsewhere (Negureanu & Salsbury Jr, 2012; Negureanu & Salsbury Jr, 2012). The structural model, in which platinum atom links N7 atoms of central adjacent guanine bases G8 and G9, was built using the G-T mismatched MutSα-DNA complex (Warren et al., 2007). Employing the same protocol, the 50 ns of all-atom molecular dynamics simulations (Negureanu & Salsbury Jr, 2012; Negureanu & Salsbury Jr, 2012) were extended here to a total of 80 ns. A table describing the conditions and the duration of all the simulations considered in this study is included in SM Table S2.

Sufficient sampling of the configurational portion of phase space is a matter of considerable concern in macromolecular dynamics simulations (Hünenberger, Mark, & van Gunsteren, 1995; Caves, Evanseck, & Karplus, 1998; Negureanu & Salsbury Jr, 2012). Those studies and others (Straub, Rashkin, & Thirumalai, 1994; Auffinger & Westhof, 1997; Horita, Zhang, Smithgall, Gmeiner, & Byrd, 2000) indicate that more effective sampling of the conformational space of proteins is obtained from multiple short trajectories starting from a single conformation with the use of different random distributions for the initial velocities rather than from a single, long trajectory. In general, it has been observed that a single longer trajectory samples only one region of configurational space (Caves et al., 1998; Salsbury Jr, Crowder, Kingsmore, & Huntley, 2009), whereas multiple shorter trajectories starting from the same configuration with different initial velocities sample multiple nearby regions of configurational space (Caves et al., 1998; Horita et al., 2000; Loccisano, Acevedo, DeChancie, Schulze, & Evanseck, 2004; Salsbury Jr et al., 2009). That is, multiple shorter simulations allow for the sampling of different regions of phase space, and so a set of these trajectories is necessary to have reasonable phase-space coverage. Hence, structural and dynamic properties obtained from multiple trajectories are improvements of those obtained from single longer trajectories (Caves et al., 1998; Horita et al., 2000; Loccisano et al., 2004). In this work, fluctuations about the native-state and perturbations due to

carboplatin and cisplatin binding were sampled by multiple molecular dynamics simulations with different initial atomic velocities.

Predictions from similar simulations confirmed by experimental studies—

Employing the same protocol, 2-10 ns of simulation of the *E. Coli* homologue molecular complex (MutS) have been experimentally verified to be sufficient to examine the experimental questions addressed (Drotschmann et al., 2004; Salsbury Jr et al., 2006; Vasilyeva et al., 2010). Responses such as changes in disordered loops associated with the DNA damage response, conformational changes associated with ATP binding/hydrolysis, and key protein-DNA contacts predicted by conformational analysis of MutS complex with cisplatin-damaged DNA (Salsbury Jr et al., 2006; Vasilyeva et al., 2009) were validated by mutational and genetic analysis experiments (Salsbury Jr et al., 2006). Furthermore, the molecular dynamics simulations reported herein, predict conformational and structural DNA changes due to (1,3) cross-linking that are in agreement with NMR structural studies on DNA structural changes due to (1,3) cross-linking, albeit without protein binding (Teuben et al., 1999).

Representative ensemble of structures—The last 12 ns of each of the four independent simulations were considered for analysis, a total of 48 ns of simulations. Snapshots, each containing a record of all atom positions in the protein-DNA complex at a given time, taken every 20 ps intervals from the trajectories of the simulated MutS α -carboplatin-damaged-DNA complex, were collected to obtain a representative ensemble of structures. The explicit water molecules and Na⁺ cations were not included in the snapshots. Similar representative ensembles of structures were generated for the cisplatin-damaged and mismatched DNA MutS α recognition complexes.

2.2 Hydrogen Bonding Analysis

Hydrogen bond analysis was performed using the hydrogen bond analysis tool, HBAT (Tiwari & Panigrahi, 2007), based on (Panigrahi & Desiraju, 2007), as well as the *hbond* tool from CHARMM (Brooks et al., 1983), both based on purely geometrical criteria of search for hydrogen bonding. As a special case of dipole interactions, hydrogen bonding, or rather “hydrogen bridging”, is defined as the attractive force between the hydrogen covalently bound to an electronegative atom of one molecule and an electronegative atom of another molecule or chemical entity. Usually, the electronegative atom is oxygen, nitrogen or fluorine, but hydrogen attached to carbon can also participate in hydrogen bonding when the carbon atom is bound to electronegative atoms (Taylor & Kennard, 1982). Strong hydrogen bonds (O-H...O, N-H...O, O-H...N, N-H...N) in protein-ligand complexes are characterized by small deviations from linearity, while weak hydrogen bonds (C-H...O, C-H...N) have variable geometries (Sarkhel & Desiraju, 2004).

In the strong/weak convention to categorize hydrogen bonding in biomolecules (Jeffrey & Saenger, 1991; Desiraju & Steiner, 1999; Panigrahi & Desiraju, 2007; Panigrahi & Desiraju, 2008; Horowitz & Trievel, 2012), its definition of strong or weak takes into account the electronegativity of the hydrogen's donor and acceptor. Following the above convention, in this study, strong hydrogen bonding refers to strong donor-strong acceptor, while weak hydrogen bonding refers to weak donor-strong acceptor.

Protein-ligand binding is determined by multi-centered interactions and multi-point recognition (Sarkhel & Desiraju, 2004). An acceptor furcation, which is the approach of many donors toward an acceptor, is preferred rather than a donor furcation (Sarkhel & Desiraju, 2004). Because of the prevalence of multi-furcation, the restrictive geometrical criteria set up for hydrogen bonding in small-molecule crystal structures need to be relaxed

in macromolecular structures (Sarkhel & Desiraju, 2004). Consequently, following the recommendations of the above mentioned studies on hydrogen bonding in protein-ligand complexes, the criteria for hydrogen bonding interactions in this study were defined as having a separation distance $d \leq 3.0 \text{ \AA}$ (Sarkhel & Desiraju, 2004) between the hydrogen atom and acceptor (oxygen, nitrogen or sulfur) atom, and the cutoff for the angle θ between the donor-H...acceptor was set at 90° , $90^\circ < \theta \leq 180^\circ$.

2.3 Root Mean Square Deviation and Fluctuation Calculations

To monitor the systems' conformational changes compared with the starting structure, the root-mean-square deviation (RMSD) from the starting structure, averaged over C α atoms of the protein as a function of time was computed in CHARMM (Brooks et al., 1983). Prior to RMSD calculations the protein-DNA complex was aligned to the initial structure by least-square superimpositions performed to remove translational and rotational movements. RMSD is defined as the root-mean-square-average distance between atoms of two superimposed structures. To obtain information on atomic flexibility, the root mean square fluctuations (RMSF), averaged over time as a function of atoms, were calculated around the system's average structure, per nucleotide in the case of DNA substrate or per residue in the case of protein, and averaged over all but hydrogen atoms. RMSFs are proportional to the experimental B-factors (temperature factors), which are frequently used to study local structural flexibility, thermal stability, and heterogeneity in macromolecules (Kuzmanic & Zagrovic, 2010).

2.4 Non-Bonding Interactions

Protein-DNA electrostatic and van der Waals non-bonding interaction energies were calculated from the trajectories by applying CHARMM (Brooks et al., 1983) interaction energy analysis to the collected snapshots. Coulomb's electrostatic interactions and van der Waals interactions based on 6-12 Lennard-Jones potential were shifted to zero using a cut-off value of $22 \text{ \AA}$. Any interactions beyond $22 \text{ \AA}$ were discarded. The spherical cut-off shifting criterion was applied on an atom by atom basis, method which has been proved (Steinbach & Brooks, 1994) to be able to approximate no-cutoff results (computationally unfeasible for this large system) when a cutoff at or beyond $12 \text{ \AA}$ is used. A cut-off radius of $24 \text{ \AA}$ was used to generate the non-bonded pair list, and the non-bonding list was updated at every step. The solvent serves the dual function of solvating each individual charge of the protein-DNA complex and of screening the interaction between charge pairs. To approximate the solvent screening without including explicit water molecules, electrostatic interactions were calculated using a distance-dependent dielectric coefficient (RDIE (Steinbach & Brooks, 1994) implemented in CHARMM (Brooks et al., 1983)), $\epsilon(r)=4r$.

2.5 Salt Bridges Analysis

The salt bridges were evaluated according to the distance between the charged donor atoms, positively charged N ζ of lysine and N ζ , N η 1 and N η 2 of arginine, in the protein, and the charged acceptor atoms, negatively charged oxygen atom in the phosphate group of the DNA nucleotides. The backbone of polynucleotides is highly charged, double stranded DNA having two negative charges per base-pair, corresponding to one unit negative charge per phosphate group. A distance $d \leq 4 \text{ \AA}$, found as a "working definition" for an ion-pair based upon analysis of charged groups in 38 proteins (Barlow & Thornton, 1983), was employed. When the geometry is acceptable, a salt bridge is also counted as a hydrogen bond, and such salt bridge is labeled as hydrogen-bonding salt bridge.

2.6 Clustergrams of the strong hydrogen bonds at the MutS α -DNA interface

To identify the network of hydrogen bonds at the protein-DNA interface with similar correlation patterns during the simulations, clustergrams of the [0,1] time series (0 for lack of hydrogen bonding and 1 for at least one strong hydrogen bonding occupancy) were generated. The method uses hierarchical clustering with Euclidean distance metric and average linkage to generate the hierarchical tree. It clusters first along the columns (the observed strong hydrogen bonds) producing row-clustered data, and then along the rows (the variable 0 or 1 at a given time in the trajectories) producing the dendrogram and the heat map, generically called clustergram (Schonlau, 2002) (Figure 4). In the heat map, red indicates the presence of a hydrogen bond and black indicates its absence.

Structural details in this study were rendered with VMD (Humphrey et al., 1996); the clustergrams (Eisen, Spellman, Brown, & Botstein, 1998; Schonlau, 2002) were created using the *clustergram* tool from Matlab; and time series for hydrogen bonding, RMSD and RMSF data were plotted using Gnuplot (Williams & Kelley, 2011) .

3. Results and Discussion

3.1 Recognition of the 1,3-d(GCG) intra-strand carboplatin-induced platinum DNA adduct by MutS α

It is well-established that biomolecular recognition is mediated by weak, noncovalent chemical interactions, such as hydrogen bonds (Jeffrey & Saenger, 1991; Glusker, 1998; Panigrahi & Desiraju, 2008). Properties such as directionality, cooperativity and reversibility make hydrogen bonding critical in biomolecular recognition and binding. Hydrogen bonds are strong enough to provide specific interactions (Glusker, 1998), but weak enough to be switched on or off in biomolecules at thermal fluctuations energies.

As described in the Methods, in the hydrogen bonding analysis presented in this study we followed the classification of hydrogen bonds into strong and weak hydrogen (Taylor & Kennard, 1982; Jeffrey & Saenger, 1991; Desiraju & Steiner, 1999), (Sarkhel & Desiraju, 2004; Panigrahi & Desiraju, 2008). In this classification scheme, the definition of strong or weak takes into account the electronegativity of the hydrogen's donor and acceptor. Consequently, in this study, strong hydrogen bonding refers to strong donor-strong acceptor interactions, while weak hydrogen bonding refers to weak donor-strong acceptor interactions, so that O-H...O, N-H...O, O-H...N, or N-H...N hydrogen bonds are considered strong and C-H...O, C-H...N, C-H...S are considered weak.

MutS α seems to recognize 1, 2 and 1, 3 intra-strand platinum-DNA adducts in a similar mode—A comprehensive analysis of the protein-DNA hydrogen bonding in the 1,3 intra-strand platinum-DNA MutS α recognition complex is included in the SM Table S1. It indicates 52 possible strong and weak protein-DNA hydrogen bonding interactions, with 75% of them made by residues of the mismatch recognition subunit, Msh6. Among them, 21 hydrogen bonds made by Msh6 and seven hydrogen bonds made by Msh2 were present more than 60% of the simulations time. These predominant hydrogen bonds are presented in Table 1, along with the comparison of the hydrogen bonds present on the cisplatin-induced platinum-DNA adduct recognition complex, as well as on the mismatched-DNA recognition complex. Their specific or non-specific contributions to the platinum-DNA adducts recognition by MutS α will be further discussed below. The occupancy of the strong hydrogen bonds during the simulations is presented in Table 2.

An analysis of the predominate strong (strong donor-strong acceptor) and weak (weak donor-strong acceptor) protein-DNA hydrogen bonds, Table 1, indicates that MutS α

interacts with the lesion sites of the 1,2 and 1,3 intra-strand platinum-DNA adducts in a similar manner. The strong and/or weak hydrogen bonds formed by G69, L88, K92, Y108, V147, and V148 of Msh6 and K528 of Msh2 are indicated to be specific to the 1,2 intra-strand platinum-DNA adduct, whereas the hydrogen bonds formed by Q124, P128 and W609 of Msh6 and K509 of Msh2 are indicated as specific to the 1,3 intra-strand adduct recognition. It is notable that the residues responsible for the specific interactions with the two adducts reside in the mismatched binding or clamp domains of Msh6 and the lever domain of Msh2.

In both adducts, most protein-DNA hydrogen bonds to the damage site are made by residues from a single loop of the Msh6 subunit, comprising residues 68 to 73. We will call this the probing loop because of its interactions with the damage site. While Msh2 makes no hydrogen bonding interactions with the damage site, it does form multiple hydrogen bonds to nucleotides adjacent to the lesions. These interactions are formed through another well-defined loop, residues 545 to 547. This suggests that Msh2 might “assist” Msh6 subunit in either recognition or signaling.

Specifically, the probing loop of Msh6 provides strong and weak hydrogen bonds from residues V68, K70, F61 and E73, Figures 2 and 3. V68 forms specific, weak C-H...N and strong N-H...O hydrogen bonding with the base of the platinated nucleotide G8 (Figure 2B). K70's backbone N atom makes weak C-H...N hydrogen bonding with G8 nucleotide, while its side chain makes specific, strong N-H...O hydrogen bonding with the base of C7 (Figure 2B). In contrast, F71 and E73 are probing both damaged and complementary DNA strands at the damage site. To this end, F71 forms specific, strong N-H...N and N-H...O hydrogen bonds with the base of G8 of the platinated strand and weak C-H...O hydrogen bonding with G9 of the complementary strand (Figure 2C). E73 forms multiple, specific strong N-H...O and weak C-H...N hydrogen bonds with the bases of the G8-C8 pair through its side chain (Figure 3E).

Outside of the probing loop, M91 probes the non-platinated DNA strand at the lesion site by making multiple weak hydrogen bonds with the phosphate group of C8 (C-H...O), the sugar of G7 (C-H...S), and the base of C6 (C-H...O), as depicted in Figure 3D.

Unique for the 1,3 platinum-DNA adduct interaction are the sporadic (Figure 4S, B&E) weak C-H...O(N) and strong N-H...O hydrogen bonds formed by the side chain of H97 with G7 (not depicted) and C8 (Figure 3F) at the lesion site. These interactions occur for 53% of the simulations time (Table 2), suggesting a possible role of H97 in the MutSα' interaction with the 1,3 platinum-DNA adduct.

From the end of the mismatch binding domain of Msh6, the side chain of R150 “anchors” the protein to the 3' side of the 1,3 platinum-DNA adduct damage site (which, as it will be addressed in the following section of this study, is the less damaged site of the adduct) by forming multiple strong N-H...O hydrogen bonds with the phosphate group of G10, Figure 2A, during the entire simulations time (Table 2).

Additionally, the side chains of Q124 (from the mismatch binding domain) and Y484 (from the lever domain) of Msh6 form multiple strong N-H...O and O-H...O hydrogen bonds, respectively, with the phosphate group of C9, the nucleotide extruded from the minor groove at the lesion site, Figure 3, G&F.

Although Msh2 subunit forms no hydrogen bonding interactions with the damage site, the analysis indicates that it might “assist” Msh6 subunit in the recognition of DNA damage. As indicated in Table 1, Msh2 subunit consistently forms contacts with nucleotides adjacent to the damage site. These nucleotides are from both the damaged DNA strand, C5 to C7, and

the complementary DNA strand, nucleotides G5-A4. The former nucleotides make hydrogen bonding interactions with residues Q545, K546, and N547 of the clamp domain, and the latter nucleotides make hydrogen bonding interactions with residues K6, K509, and K512 of the mismatch binding and clamp domains. A depiction of the view of the 1,3 intra-strand platinum-DNA adduct by the protein is presented in Figure 3H. It is noticeable that each subunit establishes most of its damage recognition interactions through a well-defined loop: from Msh6, the mismatch/damage probing loop labeled by G69; and from Msh2, another well-defined loop labeled by K546. Overall, Msh2 subunit forms almost half of its hydrogen bonding interactions with the damaged DNA, particularly with the platinated DNA strand, through this well-defined loop (Table 1).

Given that the sugar-phosphate backbone of the duplex DNA is negatively charged, it would be expected that positively charged residues involved in hydrogen bonding at the MutS α -DNA interface to possibly be involved in close-range electrostatic interactions (salt bridges, see Methods). As summarized in Table 2, our simulations suggest that dominant hydrogen-bonding salt bridges formed by residues R107 and R150 of Msh6 and residues K6 and K546 of Msh2 with the DNA's backbone phosphate groups might contribute to the specificity of the mismatched or damaged DNA substrates recognition by MutS α .

In addition, besides direct hydrogen bonding and salt bridges, water-mediated protein-DNA interactions are recognized to play a role in the specificity of protein-DNA recognition (Janin, 1999). However, the investigation of water-mediated protein-DNA interactions in these MutS α -DNA recognition complexes is beyond the scope of this investigation, but it is under consideration for a follow up study.

Synergism in the platinum-DNA-MutS α recognition complexes—Taking into account only the strong protein-DNA hydrogen bonding present at least 60% of the simulations time (Table 2), it can be concluded that MutS α 's specificity in the recognition of 1,3 intra-strand platinum DNA adduct is accomplished by interactions of residues V68, E73, R107, Q124, R150, Y484, and W609 from Msh6, and residues K6, K512, K546, and N547 from Msh2. Similarly, specificity in the recognition of 1,2 intra-strand platinum-DNA adduct by MutS α is predicted to be carried out by residues E73, K92, R107, V148, R150, and Y484 from Msh6, and residues K6, Q545, K546, and N547 from Msh2. About 60% of these residues responsible for specific recognition of either of the adducts are common, supporting the above prediction that 1,2 and 1,3 intra-strand platinum-DNA adducts are cognized by MutS α in a similar manner.

Time-series of the MutS α -DNA strong hydrogen bonding in the carboplatin and cisplatin-induced damage recognition complexes in a 0 (for lack of hydrogen bonding) and 1 (for at least one hydrogen bond occupancy) representation are presented in SM, Figure S7. In both systems, several patterns of hydrogen bonding networking can be observed. To identify the network of hydrogen bonds with similar time evolution during the simulations, which could collectively enable the recognition of specific damaging DNA adducts by MutS α , a hierarchical clustering of the [0,1] time-series data was performed using the clustergram graph (Schonlau, 2002) from the Bioinformatics Toolbox in Matlab (see Methods).

“Highly significant” for the recognition of 1,2 and 1,3 intra-strand platinum-DNA adducts by MutS α were considered strong protein-DNA hydrogen bonds with high occupancy, the red regions in the heat map, and similar correlation patterns, as indicated by the dendrogram. These highly significant interactions are marked in the clustergrams, Figure 4 A&B.

These results suggest that the collective contribution to the protein-DNA strong hydrogen bonding from residues E73, R107, R150, Y484, and W609 of Msh6, and residues K6, K512,

K546, and N547 of Msh2 may be important in the recognition of 1,3 intra-strand platinum-DNA adduct by MutS α . Similarly, the collective contribution to the protein-DNA strong hydrogen bonding from residues E73, K92, R107, Y108, V148, R150, and Y484 of Msh6, and residues K6, Q545, K546, and N547 of Msh2 may be important in the recognition of 1,2 intra-strand platinum-DNA adduct by MutS α . In the latter recognition complex, in addition to the highly significant population (the region marked in blue in Figure 4B) a subpopulation with similar occupancy but slightly different patterns of correlation is considered (the region marked in magenta in Figure 4B).

Remarkably, there is a significant overlap (about 70%) of the two series of predicted key residues for platinum-DNA adducts recognition. This evidence not only supports the general conclusion of this study, which is that 1,2 and 1,3 intra-strand platinum-DNA adducts are recognized by MutS α in a similar manner, but it also indicates subtle differences. These subtle differences may play a role in the way MutS α distinguish between different damages and/or how different damages are signaled by MutS α . The possible role of these subtle differences in presumably enabling MutS α to distinguish between different types of platinum-based DNA adducts will be explored in the concluding section of this study.

For comparison, key residues associated with the mismatched DNA recognition complex were also predicted in a similar fashion (its [0,1] time-series are presented in SM Figure S7C and its clustergram is presented in Figure 4C). These data indicate that the collective contribution of the strong hydrogen bonding from residues V68, K70, R107, V148, R150 and Y484 of Msh6 and K6, and K546 of Msh2 may be significant for the mismatched DNA recognition complex.

Considering the recognition interactions in all three complexes investigated, five such key residues are common, namely R107, R150, and Y484 from Msh6 and K6 and K546 from Msh2. This indicates a 55% overlap in the carboplatin-mismatched recognition comparison and a 45% overlap in the cisplatin-mismatched recognition comparison, supporting the prediction of a previous study (Negureanu & Salsbury Jr, 2012) which is that mismatched and platinum-based damaged DNA are recognized by MutS in significantly different modes.

Interestingly enough, these common residues, namely R107 and R150 from Msh6 and K6 and K546 from Msh2, are also responsible for the dominant hydrogen-bonding salt bridges at the MutS α interface with either damaged or mismatched DNA substrates (see above), predicting that they contribute to the specific interactions of MutS α with both mismatched and damaged DNA substrates.

3.2 Loss of symmetry and non-specificity at the carboplatin-induced 1,3-d(GCG) intra-strand platinum DNA adduct lesion site

Platinum-based anticancer drugs exert their anticancer effect through the formation of several types of adducts with DNA. In these adducts DNA is chemically modified. Platinum atom covalent bonds to the N7 positions of imidazole ring of purine bases of DNA, primarily guanine and to a less extend adenine, severely distorting the DNA's double-helix. Despite decades of clinical use and substantial research on cisplatin-DNA interactions, the cisplatin-induced structural change in DNA, such as bend angle, bend flexibility and lesion site geometry, are still debated (Wolfe et al., 2012).

Limited experimental evidence on the 3D structure of the 1,3 intra-strand platinum-DNA adduct induced by carboplatin (Teuben et al., 1999) indicates that the different platinum complexes indeed distort DNA differently. The importance of the differences in DNA structural changes induced by these adducts is highlighted by the findings that their interactions with certain cellular proteins are essentially different. In this regard, interactions

of certain cellular proteins with the inter-strand adducts (Deans & West, 2011) are different than with the 1,2-d(GG) intra-strand adducts or with the 1,3-d(GXG) intra-strand adducts (Huang, Zamble, Reardon, Lippard, & Sancar, 1994; Waozniak & Blasiak, 2002; Wang & Lippard, 2005).

Here, we present (1) predicted carboplatin-induced DNA structural and conformational changes which are confirmed by experiment (Teuben et al., 1999), (2) the molecular basis for non-specificity in the protein recognition of the carboplatin-induced damage site, and (3) changes into DNA's atomic flexibility induced by carboplatin and cisplatin-based adducts formation. This work will advance the understanding of the molecular basis of the structural and conformational changes induced by platinum-based anticancer agents, as well as the understanding of how the cellular proteins process these damages.

Predicted carboplatin-induced DNA structural changes confirmed by experiment—In agreement with experimental observations (Teuben et al., 1999), our simulations indicate that the distortion to DNA structure induced by carboplatin-based 1,3-d(GCG) intra-strand platinum adduct spreads out over the entire lesion site, and it is most severe at the 5' site, Figure 5. At the 5' end, G8-C8 and the central C9-G7 base pairs loss their hydrogen bonding during the entire simulations time. The central base C9 is extruded from the minor groove. At the 3' end, G10-C6 base pairing is completely lost about 15% of the simulations time, as depicted in Figure 5B and quantified in Figure 6A.

Loss of symmetry and non-specificity at the carboplatin-induced DNA lesion site—In addition, the time series of G10-C6 hydrogen bonding, Figure 6A, also indicate that its canonical pairing is perturbed more than 94% of the simulations time. Instead, the presence of four G10-C6 strong hydrogen bonds is dominant (present about 59% of the simulations time). On the other hand, G10-C6 complete loss of symmetry does not affect G10's strong hydrogen bonding with the protein, as indicated in Figure 6B. In this regard, one to five strong O-H...O hydrogen bonds can be formed during the simulations between the side chain of R150 from the mismatch recognition subunit Msh6 and the phosphate group of the damaged G10 nucleotide, with two fold hydrogen bonding being dominant (present about 77% of the simulation time). The latter case has been discussed in the previous section and is depicted in Figure 2A. Again, through this study strong hydrogen bonding refers to O-H...O, N-H...O, O-H...N, or N-H...N interactions and details on its definition are included in the Methods section.

The loss of symmetry at the lesion site causes the loss of base pairing and perturbs the base stacking effects. These alterations result in non-specificity and changes in the mechanical flexibility at the lesion site. The base stacking effects are considered to be the origin of the intrinsic rigidity of DNA (Mills & Hagerman, 2004). Stacking perturbation has been reported even for less severe cross-linked structural changes when the Watson-Crick hydrogen bonding is preserved (Huang, Dooley, Harris, Harris, & Stone, 2009) .

Here, the non-specificity at the lesion site is indicated by the lack of well-defined patterns of rigid, strong hydrogen bonding between all but C9 and G10 of the DNA bases at the lesion site with Msh6. Note that Msh2 subunit makes no strong hydrogen bonding with the lesion site, as has been discussed in the above damage recognition section. Our simulations indicate that platinated G8 (Figure 6C) and G7 (Figures 6F and S4 A&B) nucleotides can form up to ten strong hydrogen bonds with varies amino acids from Msh6 (such as V68, F71, E73, H97, and Y484), but none of these rigid interactions is indicated as being present extensively. C6 strong hydrogen bonding with the protein is also indicated as sporadic (Figure 6E), and defined by O-H...O interactions between the side chain of K92 (Msh6) and the nucleotide's phosphate group. C8 is indicated to retain a two-fold strong hydrogen

bonding with the protein, primarily by O-H...N and N-H...N interactions between the nucleotide's base and the backbone of Msh6' V89 (Figure 7G) and G99 (Figure 7H), about 55% of the simulations time (Figure 6D). In certain conformations, the lack of C8-protein strong hydrogen bonding (Figure 6D) is compensated by the presence of C6-protein strong hydrogen bonding (Figure 6E).

Extruded from the DNA's minor groove, nucleotide C9 from the central pair of the lesion site is stabilized during the simulations by extensive, strong hydrogen bonding with residues Q124 (Figure 7I), R150 (Figure 7J) and Y484 (Figure 7K) from Msh6's mismatch binding and lever domains. A surface representation of the 1,3 intra-strand platinum-DNA adduct is included in SI, Figure S1D. It gives additional details on the complete loss of pairing at the G8-C8 and C9-G7 base pairs, the unwinding of the DNA helix, and the extrusion of C9 from the minor groove. All of these predictions are in agreement with the experimental evidence (Teuben et al., 1999).

Calculations of the solvent accessible surface area (SASA) of DNA in the damaged and mismatched complexes further quantify the structural and conformational changes in the carboplatin-damaged DNA adduct. A statistical analysis of the SASA calculations is included in SM, Figure S1E. It indicates that the mean of SASA of DNA in the damaged complex, 4120.25(154.2) Å² (mean(std)), is higher than in the mismatched complex, 3894.12(163.84) Å². It can be speculated that the average difference of 226.13 Å² might count for the loss of base pairing, the unwinding of the DNA helix, and the base extrusion at the carboplatin-induced damage site.

In agreement with the experimental data (Teuben et al., 1999), our simulations confirm that the 1,3-d(GCG)-intra-strand platinum cross-linked adduct induced by the anticancer chemotherapeutic agent carboplatin severely distorts the DNA structure over the entire lesion site, most drastically at the 5' site. These data are consistent with earlier experimental (Marzilli, Saad, Kuklenyik, Keating, & Xu, 2001; Wu et al., 2004) and computational (Sharma et al., 2007) studies on cisplatin and oxaliplatin adducts with DNA, which also report higher distortion at the 5' side of the platinum-GG adducts. G8-C8 and C9-G7 pairs lose their base pairing during the entire simulations time. The non-specificity at the DNA's lesion site is reflected by multiple, sporadic hydrogen bonding made by all, but C9 and G10 of the perturbed nucleotides, with various residues of the protein, particularly from the Msh6's recognition loop. It may be speculated that the non-specific interactions of MMR with the lesion site may result in no immediate cellular response.

Our molecular dynamics simulations also suggest that, "confused" (forms multiple sporadic interactions, as seen above) by the 5' side of the lesion, MutSα "anchors" to the 3' side of the lesion site through residues Q124, R150 and Y484 of the mismatch recognition subunit, Msh6.

These findings highlight the nature of the carboplatin cytotoxicity and shed light into the mechanism by which MMR detects the damaged DNA bases, that is by probing changes in DNA's base pair stacking and flexibility (Yang, 2008), with the latter further discussed subsequently.

Increase of DNA's atomic flexibility in platinum-based adducts—Double helix DNA is a highly dynamic structure that can both bend and twist (Travers, 2004). Functional proteins are not static structures, but rather generally stable soft mechanical constructs that allow certain types of internal motion to enable their biological function. Information on changes in the DNA's flexibility in these particular adducts has implications on the understanding the mechanism of their recognition by MutSα. Based on the crystal structures

and biochemical studies, it has been postulated that MutS α (Yang, 2008) and MutS (its *E. Coli* homologue) (Wang et al., 2003) preferentially bind to the flexible and potentially kinked DNA. The presence of a lesion increases the local flexibility of DNA (Wang et al., 2003; Yang, 2008). The covalent bonding of platinum atom to DNA affects DNA substrate' flexibility directly, as well as indirectly by inducing changes within the canonical hydrogen bonding patterns, as described above. Such changes in hydrogen bonding patterns have been shown to result in altered intrinsic flexibility in biomolecules (Livesay, Huynh, Dallakyan, & Jacobs, 2008).

Our simulations predict that the covalent binding of platinum atom in the 1,2 cisplatin-induced DNA adduct sharply (by about 70%) increases the size of the average thermal fluctuation of the platinated base G8 (SM, Figure S1B) by comparison with the mismatched case (SM, Figure S1C). Remarkably, the 1,2 cisplatin-induced adduct formation also increases the average thermal fluctuations of the nucleotides adjacent to the damage site, namely G9 and G10, by about 73% and 81%, respectively. Overall, the average thermal fluctuations in the 1,2 cisplatin-induced platinum-DNA adduct are slightly higher (2.85(0.91) Å; mean(std)) than in the corresponding mismatched DNA (2.20(0.81) Å). Similar results of increased conformational flexibility at the platinum binding site have been reported in another cisplatin-induced 1,2-GG DNA adduct (Elizondo-Riojas & Kozelka, 2001).

On the other hand, the average thermal fluctuations in the 1,3 intra-strand carboplatin-induced platinum-DNA adduct are much higher (3.56(1.45) Å, SM Figure S1A) by comparison with the 1,2 adduct. Additionally, in the 1,3 adduct, the perturbation in terms of thermal fluctuation as a measure of flexibility extends well beyond the lesion site, presumably making the carboplatin-induced lesion more recognizable by the mismatch repair proteins. Supporting these observations, we also found that the canonical three-fold hydrogen bonding at C11-G5 base pair is replaced by a four-fold hydrogen bonding, present more than 78% of the simulations time (Figure 7L). The Watson-Crick pairing is restored at the T12-A4 site (Figure S4, C&D), which is four base pairs beyond the damage site.

However, by comparison with the 1,2 adduct, in the 1,3 adduct it is not seen a sharp increase in the flexibility of the pairing nucleotides at the lesion site (SM, Figures S1 A&B), suggesting that while more extended, the carboplatin-induced DNA damage may be less toxic.

The analysis of geometrical parameters of DNA and intramolecular interactions of the platinum NH₃ ligands in platinum-DNA adducts has been the subject of several computational investigations (Elizondo-Riojas & Kozelka, 2001; Spiegel & Magistrato, 2006; Sharma et al., 2007). In agreement with these earlier reports, our simulations also indicate that the platinum NH₃ ligands hydrogen bond with the carboplatin or cisplatin damaged bases, presumably stabilizing the adducts, and make no hydrogen bonding interactions with the protein (data not shown).

3.3 Alterations at the MutS α -DNA interface in response to damage recognition

Surface representations of the MutS α -DNA interface in the carboplatin damaged and mismatched complexes are included in Figure 8. In these representations the protein-DNA interface is defined as atoms from the two subunits of the protein within 5 Å of the DNA substrate. These representations indicate again that Msh6 makes extensive contacts with both DNA strands in both damaged and mismatched complexes. By comparison with the damage site, it is noticeable the tight complementarity of the Msh6 surface at the G-T mismatch site, as an indication for stronger interactions. As described above, Msh2 seems to

have an increased role in damage recognition, as suggested by the full surface representation of the carboplatin-damaged DNA-MutS α complex included in SM, Figure S6.

To further investigate differences at the MutS α -DNA interface in response to damage recognition in contrast with mismatch recognition, electrostatic and van der Waals non-bonding interactions were calculated. These data can be used as an indication of the degree of recognition between the protein and DNA. Given the important role played by the close-range electrostatic interactions in protein binding, salt bridges at the protein-DNA interface were also investigated.

Non-bonding interactions at the MutS α -DNA interface—The electrostatic (specific) and van der Waals (non-specific, induced dipole interactions) protein-DNA interactions were calculated in an implicit solvent approximation using a distance-dependent dielectric to mimic the screening effect of the solvent (see Methods). The histograms of the protein-DNA interaction energies in the carboplatin-damaged and mismatched recognition complexes are presented in Figure 9 A-D, and their average values are presented in Table 3. Similar results are seen at the protein-DNA's interface in the MutS α -cisplatin-induced damaged DNA recognition complex, SM Table S3. As expected, the mismatch binding subunit Msh6 associates more strongly than Msh2 subunit with both damaged and mismatched DNA substrates. However, data indicate that the protein-DNA interaction energies differ sharply for the damaged and mismatched complexes.

The electrostatic interactions are dominant at the interface of both subunits with the mismatched DNA, and they are significantly higher than in the damaged complexes. On the contrary, while electrostatic interactions are also dominant at the Msh2 interface with the damaged DNA, the non-specific, van der Waals interactions are dominant at the interface of Msh6 with the damaged DNA. In addition, the average value of the van der Waals interactions energy at the Msh6' interface with the carboplatin-damaged DNA is about three fold stronger than at the Msh6' interface with the mismatched DNA, Table 3.

Overall, a non-bonding interactions energy difference of more than 800 kcal/mol between mismatched and carboplatin-damaged DNA is an indication of the MutS α 's ability to discriminate between the mismatched and damaged DNA substrates. While, these energies are clearly exaggerated, the binding free energy calculations for the protein-DNA complexes formation that would be necessary to more quantitatively and accurately describe the binding differences are far beyond the scope of this work.

These energies differences suggest the possibility of significant conformational changes at the protein-DNA interface in response to damage recognition. To this end, 2D histograms of the electrostatic and van der Waals interactions energies, Figure 10, indicate alterations at the protein-DNA interface in response to carboplatin-damaged DNA recognition. Specifically, there are synchronous (correlation coefficient $r=0.87$) electrostatic and van der Waals interactions at the Msh2-mismatched DNA interface (Figure 10A), which are uncorrelated ($r=0.12$) at the Msh2-carboplatin-damaged DNA interface (Figure 10B). Similarly, there are synchronous ($r=0.80$) electrostatic and van der Waals interactions at the Msh6-mismatched DNA interface (Figure 10C) that are slightly anti-correlated ($r=-0.14$) at the Msh6-damaged DNA interface (Figure 10D).

Possible different packing at the protein-DNA interface in the damage recognition complex—To investigate the specific interactions responsible for the large differences on the non-bonding energies at the MutS α -DNA interface in the carboplatin-damaged and mismatched recognition complexes, the protein-DNA interactions per residue were calculated for the average structures, and their differences are presented in Figure 11.

The differences in protein-DNA electrostatic interactions at the interface of Msh2, Figure 11A, arise mainly from the interactions of the following residues with the mismatched versus carboplatin-damaged DNA substrates: Q545, K546, K509, Q518, N547, D543, and D41. Specifically, Q545-G8 electrostatic interaction is almost twice stronger in the mismatched complex (mismatched G8) than in the carboplatin-damaged complex (platinum-damaged G8), -27.02 kcal/mol versus -15.06 kcal/mol; K546-G8 electrostatic interaction is almost seven times stronger in the mismatched complex than in the carboplatin-damaged complex, -14.13 kcal/mol versus -1.90 kcal/mol; and K509-G8 electrostatic interaction is almost 12 times stronger in the mismatched complex than in the carboplatin-damaged complex, -9.14 kcal/mol versus -0.72 kcal/mol.

The predicted differences in protein-DNA electrostatic interactions at the interface of Msh6, Figure 11B, result mainly from the interactions of the following residues: R150, R613, K70, R146, V68, Y484, F71, V148, and K636. In this regard, R150-G8 electrostatic interaction is more than three times stronger in the mismatched complex than in the carboplatin-damaged complex, -35.20 kcal/mol versus -10.40 kcal/mol; R613-G8 electrostatic interaction is more than eight times stronger in the mismatched complex than in the carboplatin-damaged complex, -25.31 kcal/mol versus -2.98 kcal/mol; and K70-G8 electrostatic interaction is about seven times stronger in the mismatched complex than in the carboplatin-damaged complex, -21.61 kcal/mol versus -3.00 kcal/mol.

The differences in protein-DNA van der Waals interactions at the interface of Msh2, Figure 11C, derive primarily from the interactions of the following residues: K509, N547, Q518, whose net van der Waals interactions with DNA are stronger in the carboplatin-damaged complex than in the mismatched complex, and Q545, K546, and R524, whose net van der Waals interactions with the DNA are indicated to be stronger in the mismatched complex than in the carboplatin-damaged complex. For example, K509-G8 van der Waals interaction in the average structure of the carboplatin-damaged complex is about six times stronger than in the mismatched complex, -4.85 kcal/mol versus -0.74 kcal/mol; and Q518-G8 van der Waals interaction in the carboplatin-damaged complex is about seven times stronger than in the mismatched complex, -1.24 kcal/mol versus -0.18 kcal/mol. On the contrary, the net attractive Q545-G8 van der Waals interaction in the mismatched complex (-1.10 kcal/mol) becomes repulsive in the carboplatin-damaged complex (12.15 kcal/mol). Similarly, the net attractive K546-G8 van der Waals interaction in the mismatched complex (-2.06 kcal/mol) turns repulsive in the carboplatin-damaged complex (6.72 kcal/mol).

The differences in protein-DNA van der Waals interactions at the interface of Msh6, Figure 11D, derive mainly from the interactions of the following residues: M91, P128, V68, whose van der Waals interactions with the carboplatin-damaged DNA are stronger than with the mismatched DNA, and R107, Y484, and R150, whose van der Waals interactions with the mismatched DNA are stronger than with the carboplatin-damaged DNA. In this regard, M91-G8 van der Waals interaction in the carboplatin-damaged complex is four times stronger than in the mismatched complex, -6.80 kcal/mol versus -1.69 kcal/mol; P128-G8 van der Waals interaction in the carboplatin-damaged complex is more than nine times stronger than in the mismatched complex, -4.76 kcal/mol versus -0.50 kcal/mol. On the other hand, the net attractive R107-G10 van der Waals interaction in the mismatched complex (-0.18 kcal/mol) becomes repulsive in the carboplatin-damaged complex (8.73 kcal/mol). Similarly, the net attractive Y484-G8 van der Waals interaction in the mismatched complex (-1.83 kcal/mol) turns repulsive in the carboplatin-damaged complex (1.38 kcal/mol).

Overall, data in Figure 11 indicate that these major differences are a reflection not on how the protein interacts with a given nucleotide, but rather on how the protein interacts with the

entire DNA fragment, suggesting a different packing at the MutS α -DNA interface in the mismatch recognition complex than in the damage recognition complex.

Disruption of salt bridges at the MutS α -damaged DNA interface—Salt bridges or ion pairs are major contributors to the electrostatic interactions in protein binding (Kuman & Nussinov, 2002). Here, the analysis of salt bridges at the MutS α -DNA interface in the mismatched and damaged recognition complexes further highlights differences in response to damage recognition.

The histograms of salt bridges at the Msh2 and Msh6 interfaces with the investigated mismatched and damaged DNA substrates are presented in Figure 9, E&F. While salt bridges at the Msh2-DNA interface display a similar profile in the mismatched (11.54(5.01), mean(sd)), carboplatin-damaged (10.20(4.54)), and cisplatin-damaged (10.34(5.54)) recognition complexes, Figure 9E, a disruption of the salt bridges is indicated at the Msh6 interface with the damaged DNA substrates, Figure 9F. Expressly, in average, about 10 salt bridges from the Msh6 interface with the mismatched DNA fragment (31.33(8.52)) are disrupted at the interface of Msh6 with either carboplatin-damaged (21.00(8.52)) or cisplatin-damaged (20.21(9.04)) DNA substrates.

These findings support the above predictions of stronger electrostatic interactions at the interface of Msh6 with the mismatched DNA than with damaged DNA substrates (more than nine times stronger) and underscore the general observation of disruptions at the MutS α -DNA interface in damaged DNA recognition complexes.

3.4 Putative MMR personalized carboplatin and cisplatin predictions

Since the accidental discovery of its anticancer properties over 40 years ago, cisplatin and its second generation platinum-based anticancer drug (carboplatin) have made a major impact in the chemotherapeutic treatment of testicular and ovarian cancers. These platinum-based drugs are still widely used today against various solid tumors (Kelland, 2007). As a result of the Human Genome Project (http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml) a new set of tools can now be used to understand the complexity of the underlying biology of cancer and its variability. In the emerging field of personalized genomic medicine (Ginsburg & Willard, 2009), the use of genomic information enables us to predict response to therapies with greater precision than has ever been possible.

The clustergrams of the strong MutS α -DNA hydrogen bonding in the investigated damage recognition complexes, Figure 4 A&B, predict that genomic deficiencies in any of the following key residues may lead to MutS α 's inability to recognize and/or stabilize the 1,2 or 1,3 platinum-DNA adducts: E73, R107, R150, and Y484 from Msh6, and K6, K546, and N547 from Msh2. Thus, in these cases, it is predicted that the use of either carboplatin or cisplatin to induce apoptosis in cancerous cells would not be the ideal option. These residues are making prevalent (see the marked regions in Figure 4 clustergrams A&B) hydrogen bonding with the damaged DNA fragments and are clustered together.

Additionally, these data also predict that *MSH6* genes deficient in W609 may recognize the dominant 1,2 adducts induced by cisplatin, but not the dominant 1,3 adducts induced by carboplatin, while *MSH6* genes deficient in K92, Y108, and/or V148 may recognize the dominant 1,3 adducts induced by carboplatin, but not the dominant 1,2 adducts induced by cisplatin. Again, these predicted residues are making extensive (see the marked regions in Figure 4, A&B) hydrogen bonding interactions specific to either of the platinum-based DNA adducts and are clustered together.

And there is evidence highlighting the significance of these predictions. In this regard, residues R107 and V148 of Msh6 and N547 of Msh2 are known to cause cancer when altered (InSight database: <http://www.insight-group.org>). Additionally, predicted by similar computational modeling and validated by mutational and genetic analysis experiments, residue E73 of Msh6 is instrumental in the cellular recognition of cisplatin-DNA adduct and its cytotoxicity, and its mutation was found to increase cisplatin resistance (Drotschmann et al., 2004). Experiments are currently underway to determine the mutational effects of the above predicted key residues R107 of Msh6 and N547 of Msh2.

4. Conclusions

Specificity and synergy of the protein-DNA hydrogen bonding indicate that 1,2 and 1,3 intra-strand platinum-DNA adducts are recognized by MutS α in a similar manner, with subtle differences. While Msh2 subunit makes no hydrogen bonding interactions with the lesion site, the analysis indicates that it “assists” Msh6 subunit in the recognition of the damaged site by making extensive contacts with nucleotides adjacent to it, from both DNA strands. Highlighting the differences, we predict key residues highly relevant to the 1,2 and 1,3 platinum-DNA adducts interactions with MutS α .

Enriching limited experimental data (Teuben et al., 1999), here we predict extensive structural and conformational changes at the lesion site of the 1,3 intra-strand platinum-DNA adduct. Loss of symmetry at the lesion site causes the loss of base pairing resulting in non-specificity and changes in the atomic flexibility at the lesion site and beyond. These findings highlight the nature of the carboplatin anticancer effect and shed light into the mechanism by which MMR detects damaged DNA bases. Such knowledge could provide valuable information for the rational design of more efficient platinum-based drugs targeting the MMR-pathway. One such direction would be the design of drug candidates in which the platinum atom carries into the tumor cell not only the leaving group that would induce 1,2 or 1,3 adducts, but also a molecular fragment that can interfere specifically with MutS α proteins.

Further investigating differences at the protein-DNA interface, non-bonding electrostatic and van der Waals interactions energies indicate a possible different packing at the MutS α -mismatched DNA interface than at the protein interface with the damaged DNA substrates. Furthermore, the analysis of the salt bridges at the MutS α -DNA interface underscores the general observation of disruptions at the protein-DNA interface in the damaged DNA recognition complexes.

Taking into account the role of DNA mismatch repair proteins in maintain human genetic stability and the fact that genetic mutations of MutS α proteins have been associated with a large variety of cancers, we predict key residues highly relevant to the recognition of platinum-based damaged DNA adducts by MutS α . The genomic state, normal or deficient, of these key residues is proposed for further investigation as one of the bases to predict personalized response to carboplatin and cisplatin anticancer therapies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was partially supported by NIH R01CA129373 to FRS. The computational herein were performed on the WFU DEAC cluster; we thank WFU's Provost's office and Information Systems Department for their generous support.

References

- Matlab (7.10.0) Natick, Massachusetts: The MathWorks Inc.
- Andrews P, Howell S. Cellular pharmacology of cisplatin: Perspectives on mechanisms of acquired resistance. *Cancer Cells*. 1990; 2:35–43. [PubMed: 2204382]
- Auffinger P, Westhof E. Three ns of multiple molecular dynamics simulations of solvated tRNA(ASP) anticodon hairpin. *Journal of Molecular Biology*. 1997; 269:326–341. [PubMed: 9199403]
- Barlow DJ, Thornton JM. Ion-pairs in proteins. *Journal of Molecular Biology*. 1983; 168:867–885. [PubMed: 6887253]
- Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics*. 1984; 81:3684–3690.
- Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminatan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*. 1983; 4:187–217.
- Cannavo E, Gerrits B, Marra G, Schlapbach R, Jiricny J. Characterization of the Interactome of the Human MutL Homologues MLH1, PMS1, and PMS2. *Journal of Biological Chemistry*. 2007; 282:2976–2986. [PubMed: 17148452]
- Caves LSD, Evanseck JD, Karplus M. Locally accessible conformations of proteins: Multiple molecular dynamics simulations of crambin. *Protein Science*. 1998; 7:649–666. [PubMed: 9541397]
- Darden T, York D, Pedersen L. Particle Mesh Ewald: an N Log(N) method for Ewald sums in large systems. *Journal of Chemical Physics*. 1993; 98:10089–10092.
- Deans AJ, West SC. DNA interstrand crosslink repair and cancer. *Nature Reviews*. 2011; 11:467–480.
- Desiraju, GR.; Steiner, T. *The Weak Hydrogen Bond In Structural Chemistry and Biology*. Oxford University Press; New York: 1999.
- Drotschmann K, Topping RP, Clodfelter JE, Salsbury FR Jr. Mutations in the nucleotide-binding domain of MutS homologs uncouple cell death from cell survival. *DNA Repair (Amst.)*. 2004; 3:729–742. [PubMed: 15177182]
- Duckett DR, Drummond JT, Alastair IH, Murchie AIH, Reardon JT, Joyce T, Modrich P. Human MutSα recognizes damaged DNA base pairs containing O⁶-methylguanine, O⁴-methylthymine, or the cisplatin-d(GpG) adduct. *Proceedings of the National Academy of Sciences USA*. 1996; 93:6443–6447.
- Dutta, S. Doctoral dissertation. Wake Forest University, USA: 2011. Studying the interaction of cancer and thrombosis therapeutics with protein and DNA.
- Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences USA*. 1998; 95:14863–14868.
- Elizondo-Riojas MA, Kozelka J. Unrestrained 5 ns Molecular Dynamics Simulation of a Cisplatin-DNA 1,2-GG Adduct Provides a Rationale for the NMR Features and Reveals Increased Conformational Flexibility at the Platinum Binding Site. *Journal of Molecular Biology*. 2001; 314:1227–1243. [PubMed: 11743736]
- Fink D, Nebel S, Aebi S, Zheng H, Cenni B, Nehme A, Christen RD, Howell SB. The role of DNA mismatch repair in platinum drug resistance. *Cancer Research*. 1996; 56:4881–4886. [PubMed: 8895738]
- Gelasco A, Lippard SJ. NMR Solution Structure of a DNA Dodecamer Duplex Containing a cis-Diammineplatinum(II) d(GpG) Intrastrand Cross-Link, the Major Adduct of the Anticancer Drug Cisplatin. *Biochemistry*. 1998; 37:9230–9239. [PubMed: 9649303]
- Ghosh S, Salsbury FR Jr, Horita DA, Gmeiner WH. Zn²⁺ selectively stabilizes FdU-substituted DNA through a unique major groove binding motif. *Nucleic Acids Research*. 2011; 39:4490–4498. [PubMed: 21296761]
- Ginsburg GS, Willard HF. Genomic and personalized medicine: foundations and applications. *Translational Research*. 2009; 154:277–287. [PubMed: 19931193]
- Glusker, JP. *Directional aspects of intermolecular interactions*. Springer-Verlag; Berlin: 1998.
- Harrap KR. Preclinical studies identifying carboplatin as a viable cisplatin alternative. *Cancer Treatment Reviews*. 1985; 12:21–33. [PubMed: 3910219]

- Horita DA, Zhang W, Smithgall TE, Gmeiner WH, Byrd RA. Dynamics of the Hck-SH3 domain: comparison of experiment with multiple molecular dynamics simulations. *Protein Science*. 2000; 9:95–103. [PubMed: 10739251]
- Horowitz S, Trievel RC. Carbon-Oxygen Hydrogen Bonding in Biological Structure and Function. *Journal of Biological Chemistry*. 2012; 287:41576–41582. [PubMed: 23048026]
- Huang H, Dooley PA, Harris CM, Harris TM, Stone MP. Differential Base Stacking Interactions Induced by Trimethylene Interstrand DNA Cross-Links in the 5'-CpG-3' and 5'-GpC-3' Sequence Contexts. *Chemical Research in Toxicology*. 2009; 22:1810–1816. [PubMed: 19916525]
- Huang J-C, Zamble DB, Reardon JT, Lippard SJ, Sancar A. HMG-domain proteins specifically inhibit the repair of the major DNA adduct of the anticancer drug cisplatin by human excision nuclease. *Proceedings of the National Academy of Sciences USA*. 1994; 91:10394–10398.
- Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics*. 1996; 14:33–38. [PubMed: 8744570]
- Hunenberger PH, Mark AE, van Gunsteren WF. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *Journal of Molecular Biology*. 1995; 252:492–503. [PubMed: 7563068]
- Iyer RR, Pluciennik A, Burdett V, Modrich PL. DNA Mismatch Repair: Functions and Mechanisms. *Chemical Reviews*. 2006; 106:302–323. [PubMed: 16464007]
- Jamieson ER, Lippard SJ. Structure, Recognition, and Processing of Cisplatin-DNA Adducts. *Chemical Reviews*. 1999; 99:2467–2498. [PubMed: 11749487]
- Janin J. Wet and dry interfaces: the role of solvent in protein-protein and protein-DNA recognition. *Structure*. 1999; 7:R277–R279. [PubMed: 10647173]
- Jeffrey, GA.; Saenger, W. *Hydrogen bonding in biological structures*. Springer Verlag; Heidelberg: 1991.
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulated liquid water. *Journal of Chemical Physics*. 1983; 79:926–935.
- Kale L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, Schulten K. NAMD2: Greater scalability for parallel molecular dynamics. *Journal of Computational Physics*. 1999; 15:1283–1312.
- Kartalou M, Essigmann JM. Recognition of cisplatin adducts by cellular proteins. *Mutation Research*. 2001; 478:1–21. [PubMed: 11406166]
- Kelland L. The resurgence of platinum-based cancer chemotherapy. *Nature*. 2007; 7:573–584.
- Knaggs MH, Salsbury FR Jr, Edgell MH, Fetrow JS. Insights into correlated motions and long-range interactions in CheY derived from molecular dynamics simulations. *Biophysical Journal*. 2007; 15:2062–2079. [PubMed: 17172298]
- Kuman S, Nussinov R. Close-range electrostatic interactions in proteins. *ChemBioChem*. 2002; 3:604–617. [PubMed: 12324994]
- Kuzmanic A, Zagrovic B. Determination of Ensemble-Average Pairwise Root Mean-Square Deviation from Experimental B-Factors. *Biophysical Journal*. 2010; 98:861–871. [PubMed: 20197040]
- Livesay DR, Huynh DH, Dallakyan S, Jacobs DJ. Hydrogen bond networks determine emergent mechanical and thermodynamic properties across a protein family. *Chemistry Central Journal*. 2008; 2 doi:10.1186/1752-1153X-1182-1117.
- Loccisano AE, Acevedo O, DeChancie J, Schulze BG, Evanseck JD. Enhanced sampling by multiple molecular dynamics trajectories: carbonmonoxy myoglobin 10 micros A0-->A(1-3 transition from ten 400 picosecond simulations. *Journal of Molecular Graphics and Modelling*. 2004; 22:369–376. [PubMed: 15099833]
- MacKerell AD Jr, Banavali N, Foloppe N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers*. 2001; 56:257–265. [PubMed: 11754339]
- MacKerell AD Jr, Bashford D, Bellot M, Dunbrack RL, Evanseck JD, Field MJ, Karplus M. All-atom empirical potential for molecular modeling and studies of proteins. *Journal of Physical Chemistry B*. 1998; 102:3586–3616.
- Marzilli GL, Saad JS, Kuklenyik Z, Keating KA, Xu Y. Relationship of Solution and Protein-Bound Structures of DNA Duplexes with the Major Intrastrand Cross-Link Lesions Formed on Cisplatin

- Binding to DNA. *Journal of the American Chemical Society*. 2001; 123:2764–2770. [PubMed: 11456962]
- Mello JA, Acharya S, Fishel R, Essigmann JM. The mismatch-repair protein hMSH2 binds selectively to DNA adducts of the anticancer drug cisplatin. *Chemistry and Biology*. 1996; 3:579–589. [PubMed: 8807890]
- Mills JB, Hagerman PJ. Origin of the intrinsic rigidity of DNA. *Nucleic Acids Research*. 2004; 32:4055–4059. [PubMed: 15289578]
- Negureanu L, Salsbury FR Jr. Insights into Protein-DNA Interactions, Stability and Allosteric Communications: A Computational Study of MutS α -DNA Recognition Complexes. *Journal of Biomolecular Structure and Dynamics*. 2012; 29:757–776. [PubMed: 22208277]
- Negureanu L, Salsbury FR Jr. The molecular origin of the MMR-dependent apoptosis pathway from dynamics analysis of MutS α -DNA complexes. *Journal of Biomolecular Structure and Dynamics*. 2012; 30:347–361. [PubMed: 22712459]
- Panigrahi SK, Desiraju GR. Strong and weak hydrogen bonds in drug-DNA complexes: A statistical analysis. *Journal of Biosciences*. 2007; 32:677–691. [PubMed: 17762141]
- Panigrahi SK, Desiraju GR. Strong and weak hydrogen bonds in the protein-ligand interface. *Proteins*. 2007; 67:128–142. [PubMed: 17206656]
- Panigrahi SK, Desiraju GR. Strong and weak hydrogen bonds in protein-ligand complexes of kinases: a comparative study. *Amino Acids*. 2008; 34:617–633. [PubMed: 18180869]
- Peltomaki P. Role of DNA Mismatch Repair Defects in the Pathogenesis of Human Cancer. *Journal of Clinical Oncology*. 2003; 21:1174–1179. [PubMed: 12637487]
- Rajeswaran A, Trojan A, Burnand B, Giannelli M. Efficacy and side effects of cisplatin- and carboplatin-based doublet chemotherapeutic regimens versus non-platinum-based doublet chemotherapeutic regimens as first line treatment of metastatic non-small cell lung carcinoma: a systematic review of randomized controlled trials. *Lung Cancer*. 2008; 59:1–11. [PubMed: 17720276]
- Rosenberg B, Van Camp L, Krigas T. Inhibition of cell division in *Escherichia coli* by electrolysis products from a platinum electrode. *Nature*. 1965; 205:698–699. [PubMed: 14287410]
- Salsbury FR Jr. Effects of Cisplatin Binding to DNA on the Dynamics of the *E. coli* MutS Dimer. *Protein and Peptide Letters*. 2010; 17:744–750. [PubMed: 19995335]
- Salsbury FR Jr. Molecular Dynamics Simulations of Protein Dynamics and their relevance to drug discovery. *Current Opinion in Pharmacology*. 2010; 10:738–744.
- Salsbury FR Jr, Clodfelter JE, Gentry MB, Hollis T, Scarpinato DK. The molecular mechanism of DNA damage recognition by MutS homologs and its consequences for cell death response. *Nucleic Acids Research*. 2006; 34:2173–2185. [PubMed: 16648361]
- Salsbury FR Jr, Crowder MW, Kingsmore SF, Huntley JJA. Molecular dynamics simulations of the metallo-beta-lactamase from *Bacteroides fragilis* in the presence and absence of a tight-binding inhibitor. *Journal of Molecular Modeling*. 2009; 15:133–145. [PubMed: 19039608]
- Salsbury FR Jr, Crowley MF, Brooks CL. Modeling of the metallo-beta-lactamase from *B. fragilis*: Structural and dynamic effects of inhibitor binding. *Proteins*. 2001; 44:448–459. [PubMed: 11484222]
- Sarkhel S, Desiraju GR. N-H...O, O-H...O, and C-H...O Hydrogen Bonds in Protein-Ligand Complexes: Strong and Weak Interactions in Molecular Recognition. *Proteins*. 2004; 54:247–259. [PubMed: 14696187]
- Scheeff ED, Briggs JM, Hawell SB. Molecular modeling of the intrastrand guanine-guanine DNA adducts produced by cisplatin and oxaliplatin. *Molecular Pharmacology*. 1999; 56:633–643. [PubMed: 10462551]
- Schonlau M. The Clustergram: A graph for visualizing hierarchical and non-hierarchical cluster analysis. *The Stata Journal*. 2002; 3:316–327.
- Sharma R, Graham J, Blagden S, Gabra H. Sustained platelet-sparing effect of weekly low dose paclitaxel allows effective, tolerable delivery of extended dose dense weekly carboplatin in platinum resistant/refractory epithelial ovarian cancer. *BMC Cancer*. 2011; 11:289. [PubMed: 21745358]

- Sharma S, Gong P, Temple B, Bhattacharyya D, Dokholyan NV, Chaney SG. Molecular dynamics simulations of cisplatin- and oxaliplatin-d(GG) intrastrand cross-links reveal differences in their conformational dynamics. *Journal of Molecular Biology*. 2007; 373:1123–1140. [PubMed: 17900616]
- Spiegel K, Magistrato A. Modeling anticancer drug-DNA interactions via mixed QM/MM molecular dynamics simulations. *Organic and Biomolecular Chemistry*. 2006; 4:2507–2517. [PubMed: 16791311]
- Steinbach PJ, Brooks BR. New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulations. *Journal of Computational Chemistry*. 1994; 15:667–683.
- Straub JE, Rashkin AB, Thirumalai D. Dynamics in rugged energy landscapes with applications to the S-peptide and ribonuclease. *Journal of the American Chemical Society*. 1994; 116:2049–2063.
- Takahara PM, Rosenzweig AC, Frederick CA, Lippard SJ. Crystal structure of double-stranded DNA containing the major adduct of the anticancer drug cisplatin. *Nature*. 1995; 377:649–652. [PubMed: 7566180]
- Taylor A, Kennard O. Crystallographic evidence for the existence of C-H...O, C-H...N and C-H...Cl hydrogen bonds. *Journal of the American Chemical Society*. 1982; 104:5063–5070.
- Teuben JM, Bauer C, Wang AHJ, Reedijk J. Solution structure of a DNA duplex containing a cis-diammineplatinum(II) 1,3-d(GTG) intrastrand cross-link, a major adduct in cells treated with the anticancer drug carboplatin. *Biochemistry*. 1999; 38:12305–12312. [PubMed: 10493798]
- Tiwari A, Panigrahi SK. HBAIT: a complete package for analyzing strong and weak hydrogen bonds in macromolecular crystal structures. *In Silico Biology*. 2007; 7:651–661. [PubMed: 18467777]
- Todd RC, Lippard SJ. Structure of duplex DNA containing the cisplatin 1,2-Pt(NH₃)₂2+-d(GpG) cross-link at 1.77 Å resolution. *Journal of Inorganic Biochemistry*. 2010; 104:902–908. [PubMed: 20541266]
- Travers AA. The structural basis for DNA flexibility. *The Philosophical Transactions of the Royal Society London*. 2004; 362:1423–1438.
- van Gunsteren WF, Berendsen HJC. Algorithms for macromolecular dynamics and constraint dynamics. *Molecular Physics*. 1977; 34:1311–1327.
- Vasilyeva A, Clodfelter JE, Gorczynski MJ, Gerardi AR, King SB, Salsbury FR Jr, Scarpinato KD. Parameters of Reserpine Analogs That Induce MSH2/MSH6-Dependent Cytotoxic Response. *Journal of Nucleic Acids*. 2010 Article ID 162018, 13 pages doi:10.4061/2010/162018.
- Vasilyeva A, Clodfelter JE, Reactor B, Hollis T, Scarpinato KD, Salsbury FR Jr. Small molecule induction of MSH2-dependent cell death suggests a vital role of mismatch repair proteins in cell death. *DNA Repair*. 2009; 8:103–113. [PubMed: 18955167]
- Wang D, Lippard SJ. Cellular Processing of Platinum Anticancer Drugs. *Nature*. 2005; 4:307–321.
- Wang H, Yang Y, Schofield MJ, Du C, Fridman Y, Lee SD, Erie DA. DNA bending and unbending by MutS govern mismatch recognition and specificity. *Proceedings of the National Academy of Sciences USA*. 2003; 100:14822–14827.
- Waozniak K, Blasiak J. Recognition and repair of DNA-cisplatin adducts. *Acta Biochimica Polonica*. 2002; 49:583–596. [PubMed: 12422229]
- Warren JJ, Pohlhaus TJ, Changele A, Iyer RR, Modrich PL, Beese LS. Structure of the Human MutS DNA Lesion Recognition Complex. *Molecular Cell*. 2007; 26:579–592. [PubMed: 17531815]
- Wheate NJ, Walker S, Craig GE, Onun R. The status of platinum anticancer drugs in the clinic and in clinical trials. *Dalton Transactions*. 2010; 39:813–827.
- Williams T, Kelley C. 2011 Gnuplot 4.5: an interactive plotting program <http://gnuplot.info>.
- Wolfe KC, Hastings WA, Dutta S, Long A, Shapiro BA, Woolf TB, Guthold M, Chirijian GS. Multiscale Modeling of Double-Helical DNA and RNA: A Unification through Lie Groups. *Journal Physical Chemistry B*. 2012 dx.doi.org/10.1021/jp2126015 |.
- Wu Y, Pradhan P, Havener J, Boysen G, Swenberg JA, Campbell SL, Chaney SG. NMR Solution Structure of an Oxaliplatin 1,2-d(GG) Intrastrand Cross-link in a DNA Dodecamer Duplex. *Journal of Molecular Biology*. 2004; 341:1251–1269. [PubMed: 15321720]
- Yang W. Structure and mechanism for DNA lesion recognition. *Cell Research*. 2008; 18:184–197. [PubMed: 18157156]

Young LC, Peters AC, Maeda T, Edelmann W, Kucherlapati R, Andrew SE, Tron VA. DNA Mismatch Repair Protein Msh6 Is Required for Optimal Levels of Ultraviolet-B-Induced Apoptosis in Primary Mouse Fibroblasts. *Journal of Investigative Dermatology*. 2003; 121:876–880. [PubMed: 14632208]

Figure 1.

MutS α -DNA complex structural model. DNA is shown in light blue (the mismatched bases are depicted in black and brown). The color code for the heterodimer domains is as following: red for the mismatch binding domain, residues 1 to 124 in Msh2 and 1 to 157 in Msh6; yellow for the connector domain, residues 125 to 297 in MSH2 and 158 to 356 in Msh6; green for the lever domain, residues 300 to 456 and 554 to 619 in Msh2, and 357 to 573 and 648 to 714 in Msh6; purple for the clamp domain, residues 457 to 553 in Msh2 and 574 to 647 in Msh6; blue for the ATPase domain, residues 620 to 855 in Msh2 and 715 to 974 in Msh6. Note that in our system, residue 1 of Msh6 corresponds to residue 362 in the solved structure (PDB ID 2O8B); the first 361 residues are unsolved in the X-Ray structure.

Figure 2. MutSa recognition of the carboplatin-induced DNA lesion

A. Strong O-H...O hydrogen bonding between the side chain of R150 and the phosphate group of the platinated G10. **B.** V68 strong N-H...O and weak C-H...N hydrogen bonding with platinated base G8; K70 weak C-H...N hydrogen bonding between its backbone N and G8' sugar, as well as the hydrogen bonding of its side chain with the adjacent base, C7. **C.** F71' strong N-H...N and N-H...O hydrogen bonding with platinated base G8 and weak C-H...O bonding with base G9 from the opposite DNA strand. The weak and strong hydrogen bonding, as defined in the Methods, are indicated by broken black lines. The vdw representation of the platinum atom is depicted in tan.

Figure 3. MutS α recognition of the carboplatin-induced DNA lesion (Continued)

D. Weak C-H...N and C-H...O hydrogen bonding by M91 of the pair of the Pt cross-linked base, C8, and its two adjacent bases in the direction of the perturbation, bases G7 and C6. **E.** Multiple strong N-H...O and weak C-H...N hydrogen bonding by E73 with G8-C8 base pair. **F.** Unique for carboplatin recognition by MutS α , H97 strong N-H...O and weak C-H...O hydrogen bonding with the pair of the one of the platinated bases, C8. Y484 makes specific, strong O-H...O hydrogen bonding with the phosphate group of C9 and the sugar group of G8. **G.** Strong hydrogen bonding made by the side chain of Q124 with the phosphate group of C9 at the damage site. **H.** Global view of the 1,3 platinum-DNA adduct by Msh2 (in silver) and Msh6 (in pink). Msh6 contacts are depicted in red, while Msh2

contacts are depicted in lime. Notice the two well-defined probing loops: one from Msh6 (labeled by G69) and the other one from Msh2 (labeled by K546). The weak and strong hydrogen bonding, as defined in the Methods, are indicated by broken black lines. The vdw representation of the platinum atom is depicted in tan.

Figure 4. Synergism in the carboplatin-, cisplatin-induced, and mismatched MutS α -DNA recognition complexes

A. Carboplatin Complex: hydrogen bonding synergism by dominant strong hydrogen bonds with similar time evolution is provided by residues E73, R107, R150, Y484, W609 of Msh6, and K6, K512, K546, and N547 of Msh2 (the marked region). **B. Cisplatin Complex:** hydrogen bonding synergism by dominant strong hydrogen bonds with similar time evolution is provided by residues E73, K92, R107, Y108, V148, R150, and Y484 of Msh6, and K6, Q545, K546, and N547 of Msh2. In the cisplatin complex case a subpopulation with similar occupancy but slightly different time evolution (region marked in magenta), was also considered. **C. Mismatched Complex:** hydrogen bonding synergism by

dominant strong hydrogen bonds with similar time evolution is provided by residues V68, K70, R107, V148, R150, and Y484 of Msh6, and K6, and K546 of Msh2. The clustergrams (dendrograms and the heat maps) are the result of hierarchical clustering of 0 (for lack of hydrogen bonding) and 1 (for at least one hydrogen bond occupancy) time series. Each row represents a set of 0 or 1 for a given strong hydrogen bond (a column) at a given time during the simulations. The method uses hierarchical clustering with Euclidean distance metric and average linkage to generate the hierarchical tree. It clusters first along the columns (the observed strong hydrogen bonds) producing row-clustered data and then along the rows (the 0 or 1 variables at a given time in the trajectories) producing the dendrogram and the heat map, generically called clustergram. In the heat map, red indicates the presence of a strong hydrogen bond while black indicates its absence.

Figure 5. Structural and conformational changes in the carboplatin-damaged DNA

The perturbation induced by the 1,3 intra-strand covalent bonding of the platinum atom to the N7 atoms from G8 and G10 spreads over the entire lesion site; G8-C8 and C9-C7 lose their pair base hydrogen bonding the entire simulations time (**A&B**); G10-C6 base pairing is less distorted: at least one hydrogen bond is present about 85% of the simulations time (**A**), while a complete loss of pairing (**B**) is indicated for the rest of the simulations time; central base C9 is extruded from the minor groove (**A&B**). In both conformations the vdW representation of the platinum atom is depicted in tan.

Figure 6. Loss of symmetry and non-specificity at the carboplatin-induced DNA lesion site

A. The canonical G10-C6 symmetry is perturbed more than 94% of the simulation time and the presence of four G10-C6 strong hydrogen bonds is prevalent, present 58.64% of the simulations time. **B.** G10-C6 complete loss of symmetry does not affect G10 strong hydrogen bonding with R150 of Msh6; one to five strong O-H...O hydrogen bonds, with two being present more than 77% of the simulations time, are formed between the side chain of R150 and the phosphate group of the G10 nucleotide. **C.** No well-defined (specific) strong hydrogen bonding is indicated for the perturbed G8 with the protein. **D.** The lack of hydrogen bonding of Msh6 to platinated C8 is correlated with the presence of strong

hydrogen bonding to the base pair of the other platinated guanine (G10), namely nucleotide C6 in ⁴⁹panel **E**. No dominant strong hydrogen bonding of C6 (**E**) or G7 (**F**) by Msh6 is indicated.

Figure 7. Loss of symmetry and non-specificity at the carboplatin-induced DNA lesion site (Continued)

Sporadic (non-specific) strong hydrogen bonding of C8 by V89 (G) and G99 (H) from the probing loop of Msh6. Specific hydrogen bonding of C9 by Q124 (I), R150 (J), and Y484 (K) of Msh6. Beyond the platinated site, the canonical base pairing of C11-G5 (L) is replaced by a four-fold hydrogen bonding.

Figure 8. MutSα-DNA interface in the damaged (A) and mismatched (B) recognition complexes
 The protein-DNA interface is defined as atoms within 5 Å of DNA. In silver are depicted atoms from the Msh6 subunit while in pink are depicted atoms from Msh2 subunit. The 1,3 adduct and the G-T mismatch are marked in red on the DNA surface depicted in green. Msh6 makes extensive contact with both DNA strands, in both damaged and mismatched complexes. Msh2 seems to have an increased role in damage recognition. Unlike at the damage site, it is noticeable the tight complementarity of the Msh6 surface at the G-T mismatch site, as an indication for stronger interactions.

Figure 9. Non-bonding interactions at the MutSa-DNA interface in the carboplatin-damaged (A&C) and mismatched (B&D) recognition complexes

Electrostatic interactions at the protein-DNA interface are stronger in the mismatched complex than in the carboplatin-damaged complex: at the Msh2-DNA interface, $-339.99(44.36)$ kcal/mol versus $-64.40(15.69)$ kcal/mol, and at the Msh6-DNA interface $-791.20(66.39)$ kcal/mol versus $-84.08(8.98)$ kcal/mol. On the contrary, van der Waals interactions at the protein-DNA interface are stronger in the carboplatin-damaged complex than in the mismatched complex: at the Msh2-DNA interface $-41.68(8.20)$ kcal/mol versus $-13.27(4.12)$ kcal/mol, and at the Msh6-DNA interface $-117.32(9.38)$ kcal/mol versus $-40.36(4.58)$ kcal/mol. While electrostatic interactions at the Msh6-DNA interface are

dominant in the mismatched complex, van der Waals interactions at the same Msh6-DNA interface are dominant in the carboplatin-damaged recognition complex. Disruption of salt bridges at the MutS α -damaged DNA interface, **histograms E&F**. While the number of salt bridges at the Msh2-DNA interface displays a similar profile in the mismatched (11.54(5.01)), carboplatin-damaged (10.20(4.54)), and cisplatin-damaged (10.34(5.54)) recognition complexes (**E**), a disruption of salt bridges is indicated at the Msh6 interface with the damaged DNA substrates (**F**) (the average number of salt bridges at the protein-DNA interface in the mismatched complex is 31.33(8.52) versus 21.00(8.52) in the carboplatin-damaged and 20.21(9.04) in the cisplatin-damaged complexes). Data presented are mean(standard deviation) for the representative ensemble of structures.

Figure 10. Disruptions at the MutS α -DNA interface in response to damage recognition
 2D histograms of the electrostatic and van der Waals interactions energies for the representative ensemble of structures (24,000 data points) are presented. Synchronous (correlation coefficient $r=0.87$) electrostatic and van der Waals interactions at the Msh2-mismatched DNA interface (**A**) are uncorrelated ($r=0.12$) at the Msh2-carboplatin-damaged DNA interface (**B**). Similarly, synchronous ($r=0.80$) electrostatic and van der Waals interactions at the Msh6-mismatched DNA interface (**C**) are slightly anti-correlated ($r=-0.14$) at the Msh6-carboplatin-damaged DNA interface (**D**).

Figure 11. Major differences in non-bonding interactions at the MutS α -DNA interface in the mismatch versus damage recognition complexes

Electrostatic and van der Waals protein-DNA interactions per residue for the average structures of the mismatched and carboplatin-damaged recognition complexes were considered. **A.** Differences in the Msh2-DNA electrostatic interactions arise mainly from the interactions of the following residues with the mismatched versus carboplatin-damaged DNA substrates: Q545, K546, K509, Q518, N547, D543, and D41. **B.** Differences in the Msh6-DNA electrostatic interactions result mainly from the interactions of the following residues: R150, R613, K70, R146, V68, Y484, F71, V148, and K636. **C.** Differences in the Msh2-DNA van der Waals interactions derive primarily from the interactions of the following residues: K509, N547, Q518, whose non-specific interactions with DNA are stronger in the carboplatin-damaged than in the mismatched complex, and Q545, K546, and R524, whose non-specific interactions with the DNA are indicated to be stronger in the mismatched than in the carboplatin-damaged complex. **D.** Differences in the Msh6-DNA van der Waals interactions derive mainly from the interactions of the following residues: M91, P128, V68, whose van der Waals interactions with the carboplatin-damaged DNA are stronger than with the mismatched DNA, and R107, Y484, and R150, whose van der Waals interactions with the mismatched DNA are stronger than with the carboplatin-damaged DNA.

Table 1
Prevalent weak and strong MutSa-DNA hydrogen bonding in carboplatin and cisplatin induced DNA adducts^a.

Damaged strand	G1	A2	A3	C4	C5	G6	C7	G8/G8	G9/C9	G10	C11	T12	A13	G14	G15
1.T46											P				
2.P47											P				
3.V68								PC	C	P					
4.G69								P	P						
5.K70							PCc	PC							
6.F71							Pc	PCc	C						
7.E73								PpCc	C	P					
8.L88									P						
9.M91							P	Pc	C	c					
10.K92								P	P	P					
11.P101							Pc								
12.R107						Pc	Pc								
13.Y108						P	P								
14.Q124									C						
15.P128						c									
16.V147										P					
17.V148										P					
18.R150									PC	PC					
19.Y484								PC	PC						
20.W609		C	C	C											
21.K639	Pc	Pc													
22.K6*					Pc	Pc									
23.K509*												c	c	C	C
24.K512*											pc	pc			
25.K528*												p	p		
26.Q545*					PC	PC									

Damaged strand	G1	A2	A3	C4	C5	G6	C7	G8/G8	G9/C9	G10	C11	T12	A13	G14	G15
27.K546*						PC	PC								
28. N547*						PC					c	c			
Complementary strand	C15	T14	T13	G12	G11	C10	G9	T8/C8	C7/G7	C6	G5	A4	T3	C2	C1

*Residues that make at least one hydrogen bond of any type more than 60% of the simulations time were considered. Capital letters indicate hydrogen bonding with nucleotides from the DNA upper, damaged strand, while small letters indicate hydrogen interactions with nucleotides of the DNA lower, complementary strand.

* denotes Msh2' s residues. P denotes the 1,2-d(GG) cisplatin-induced platinum cross-linked complex; C denotes the 1,3-d(GCG) carboplatin-induced platinum intra-strand cross-linked complex. In the cisplatin-induced complex C9-G7 base pair is flipped, becoming G9-C7, for allowing the 1,2 intra-strand platinum adduct formation. Nucleotides G8 and G10 are 1,3 intra-strand platinum cross-linked in the carboplatin-induced complex . G8/T8 mismatch corresponds to the cisplatin-induced complex, while G8/C8 pairing corresponds to the carboplatin-induced complex.

Table 2

MutSa-DNA strong hydrogen bonding (hb) and hydrogen bonding salt bridges (sb)^b. Occupancy of at least one hydrogen bond or hydrogen bonding salt bridge of any type during the simulations time is presented.

Residue	Carboplatin		Cisplatin		Mismatched	
	hb	sb	hb	sb	hb	sb
1.T46	20.17		15.76		20.53	
2.G48	11.47		48.01		52.86	
3.V68	59.93		36.55		99.70	
4.G69	0.02		0.00		0.04	
5.K70	38.70	21.72	54.27	18.74	98.39	38.77
6.F71	23.93		0.00		0.01	
7.E73	96.19		71.65		38.47	
8.K92	39.31	19.36	74.76	45.32	34.42	4.48
9.H97	53.47		3.86		0.00	
10.G99	26.64		12.06		1.04	
11.R107	99.74	97.87	99.96	97.87	72.59	70.85
12.Y108	46.20		47.09		45.81	
13.Q124	72.74		57.06		40.40	
14.V148	48.60		70.96		99.44	
15.R150	100	99.61	79.54	79.26	100	99.81
16.Y484	84.11		70.02		81.48	
17.W609	66.53		19.91		47.77	
18.R613	38.74	32.84	30.28	23.74	60.62	53.23
19.K636	26.97	31.91	35.39	34.42	62.56	61.93
20.K639	36.65	27.83	22.37	12.88	24.82	11.75
21.K6	86.00	82.67	84.20	81.42	86.47	83.42
22.K509	51.30	17.00	40.81	24.99	46.14	27.50
23.K512	62.49	58.87	32.37	30.60	61.81	59.82
24.T526	40.07		48.11		62.39	
25.K528	46.98	45.27	56.95	50.93	46.38	42.67

Residue	Carboplatin		Cisplatin		Mismatched	
	hb	sb	hb	sb	hb	sb
26.Q545	56.49		76.16		71.46	
27.K546	99.29	80.48	80.22	66.16	90.60	80.93
28.N547	73.76		65.95		17.42	
29.K550	35.24	32.84	51.89	48.72	28.13	25.55

b In red are residues from Msh2.

Table 3

Non-bonding MutSc-DNA interactions at the carboplatin-damaged and mismatched DNA interfaces; mean(standard deviation), in kcal/mol.

	Carboplatin	Mismatched
Msh2-DNA elect.	−64.40(15.69)	−339.99(44.36)
Msh6-DNA elect.	−84.08(8.98)	−791.20(66.39)
Msh2-DNA vdW	−41.68(8.20)	−13.27(4.12)
Msh6-DNA vdW	−117.32(9.38)	−40.36(4.58)