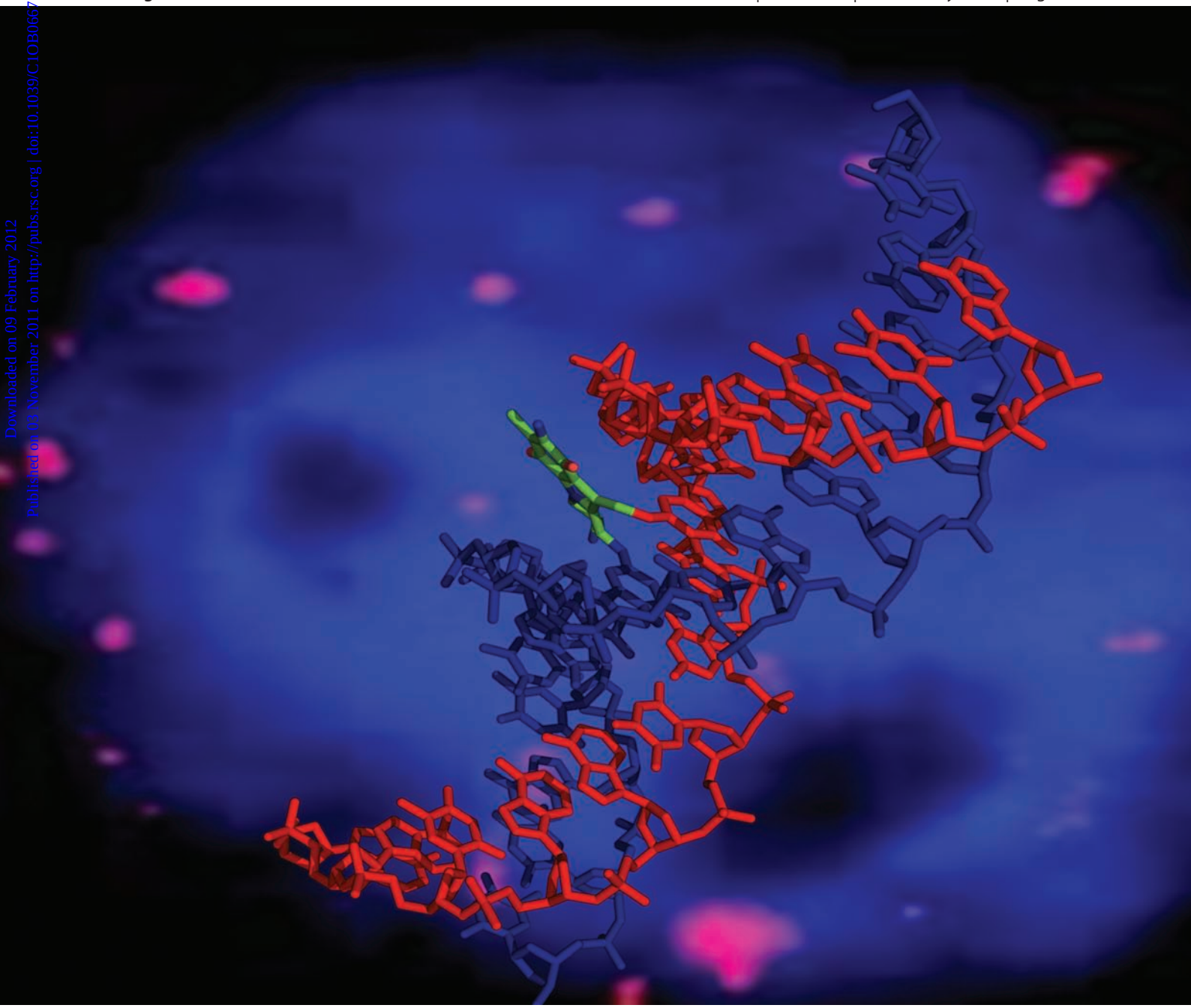
Organic & Biomolecular Chemistry

www.rsc.org/obc

Volume 10 | Number 8 | 28 February 2012 | Pages 1469–1696



ISSN 1477-0520

FULL PAPER

Federico Gago *et al.*

Rationale for the opposite stereochemistry of the major monoadducts and interstrand crosslinks formed by mitomycin C and its decarbamoylated analogue at CpG steps in DNA and the effect of cytosine modification on reactivity

RSC Publishing

Organic & Biomolecular Chemistry

Dynamic Article Links
View Online

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 1543

www.rsc.org/obc

PAPER

Rationale for the opposite stereochemistry of the major monoadducts and interstrand crosslinks formed by mitomycin C and its decarbamoylated analogue at CpG steps in DNA and the effect of cytosine modification on reactivity†

Juan A. Bueren-Calabuig,[a] Ana Negri,‡[a] Antonio Morreale[b] and Federico Gago*[a]

Received 3rd October 2011, Accepted 2nd November 2011
DOI: 10.1039/c1ob06675g

Mitomycin C (MMC) is a potent antitumour agent that forms a covalent bond with the 2-amino group of selected guanines in the minor groove of double-stranded DNA following intracellular reduction of its quinone ring and opening of its aziridine moiety. At some 5′-CG-3′ (CpG) steps the resulting monofunctional adduct can evolve towards a more deleterious bifunctional lesion, which is known as an interstrand crosslink (ICL). MMC reactivity is enhanced when the cytosine bases are methylated (5 MC) and decreased when they are replaced with 5-F-cytosine (5FC) whereas the stereochemical preference of alkylation changes upon decarbamoylation. We have studied three duplex oligonucleotides of general formula d(CGATAAXGCTAACG) in which X stands for C, 5MC or 5FC. Using a combination of molecular dynamics simulations in aqueous solution, quantum mechanics and continuum electrostatics, we have been able to (i) obtain a large series of snapshots that facilitate an understanding in atomic detail of the distinct stereochemistry of monoadduct and ICL formation by MMC and its decarbamoylated analogue, (ii) provide an explanation for the altered reactivity of MMC towards DNA molecules containing 5MC or 5FC, and (iii) show the distinct accommodation in the DNA minor groove of the different covalent modifications, particularly the most cytotoxic C1α and C1β ICLs.

Introduction

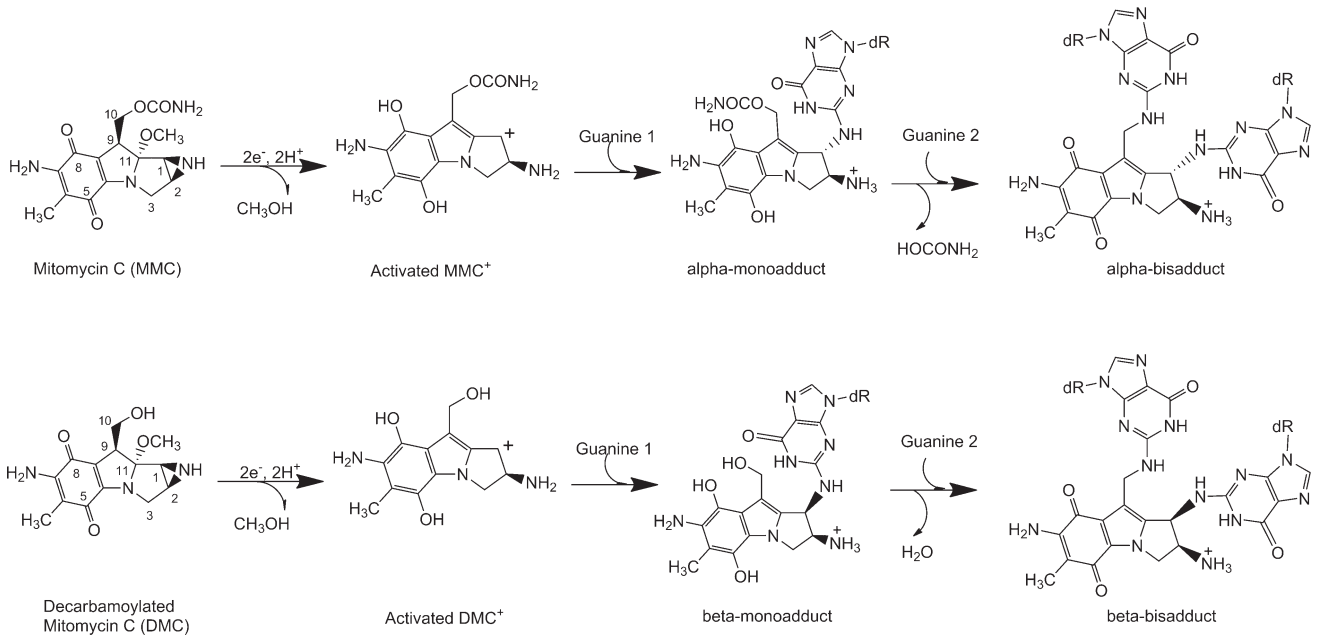
Mitomycin C (MMC) was discovered as a potent anticancer antibiotic produced by *Streptomyces caespitosus*[1] and is biosynthetically obtained nowadays from fermentation cultures of *S. lavendulae* because of the challenge that its total synthesis still represents.[2] MMC's broad spectrum of activity against solid tumours was shown to be primarily due to its potent inhibition of DNA replication mostly because of its ability to crosslink the complementary strands of the double helix[3] specifically at 5′-CG-3′ (CpG) steps[4,5] in hypoxic cells. Nonetheless, the major product upon reaction of MMC with DNA is actually the monofunctional adduct N^2-(2″β,7″-diaminomitosen-l″α-yl)-2′-deoxyguanosine[6] that is formed upon covalent bond formation with the exocyclic amino group of guanine at position 2 of the purine ring in just one strand.[7]

In clinical practice since the 1970's, MMC is intravesically instilled in the chemotherapy of superficial bladder tumours[8] and used as topical adjunctive therapy in various ocular surgeries to inhibit the wound healing response and reduce scarring.[9] MMC is also used parenterally, in combination with other approved drugs, in the therapy of disseminated adenocarcinomas of the stomach or pancreas and as palliative treatment of these malignancies when other modalities have failed.[10] It is noteworthy that hypersensitivity to MMC is a hallmark of cells deficient in one or more of the proteins involved in homologous recombination repair such as Mus81,[11] XPG,[12] Pir51,[13] Brca1,[14] or the Partner and Localizer of Brca2 (PalB2).[15] Interestingly, biallelic inactivating mutations in the gene encoding this latter protein were recently shown to underlie the robust activity of MMC observed in nude mice xenografted with a gemcitabine-resistant human pancreatic cancer as well as the dramatically favourable response attained in the patient from whom the tumour had been resected upon treatment with this drug.[16] Moreover, sensitivity to MMC has been used to classify Fanconi anemia (FA) patients into different complementation groups,[17] can still be used for diagnostic purposes *in vitro*[18]

[a]*Departamento de Farmacología, Universidad de Alcalá, E-28871 Alcalá de Henares, Madrid, Spain. E-mail: federico.gago@uah.es; Fax: +34-918854591; Tel: +34-918854514*
[b]*Unidad de Bioinformática, Centro de Biología Molecular Severo Ochoa, 28049 Madrid, Spain*

†Electronic supplementary information (ESI) available: Additional bonded and nonbonded parameters for drugs and modified bases, four additional figures and a scheme representing the different molecular complexes and techniques employed during the study of the reactions between MMC and DMC with the DNA oligonucleotides. See DOI: 10.1039/c1ob06675g

‡Present address: Department of Structural and Chemical Biology, Mount Sinai School of Medicine, New York, NY 10029-657, USA.



Scheme 1 Reaction pathways leading to the formation of α-MMC–DNA and β-DMC–DNA monoadducts and their corresponding bisadducts or interstrand crosslinks (ICL) of opposite stereochemistry. The DNA-alkylating functionalities of MMC and DMC are unmasked by reductive activation.

and has been instrumental to delineate the so-called FA/BRCA pathway of the DNA damage response.[19]

The deceptively "simple" structure of MMC (Scheme 1) possesses four contiguous stereogenic carbons and economically packs in a constrained architecture (i) a quinone-containing pyrroloindole skeleton, (ii) a methoxy group, (iii) an aziridine ring, and (iv) a carbamate moiety.[2] The aziridine at C1 and the carbamate at C10 are critical to its biological activity because they constitute two masked alkylating functionalities. Indeed, MMC itself is remarkably unreactive toward DNA at pH 7–8[3] but efficient alkylation of this macromolecule is observed following enzymatic, electrochemical or chemical reduction,[20] hence its greater cytotoxicity under hypoxic conditions and the recognition of MMC as the first bioreductive (and bifunctional) alkylating agent.[3,21] Two-electron reduction of the quinone ring of MMC facilitates methoxide elimination from position 9a, formation of a leuco-aziridinomitosene and opening of the aziridine ring to provide the extended quinone methide that is the initial alkylating entity leading to a DNA monoadduct with a guanine (G) in the minor groove.[7,20] At a CpG step, the second alkylating centre of MMC is the iminium ion that forms upon the reverse Michael elimination of carbamic acid at C10 from the monoadduct. It is at this position then that MMC can undergo a second nucleophilic attack either by a water molecule (to yield a C10-decarbamoylated monoadduct)[7] or, exclusively at CpG steps, by the 2-amino group of the guanine in the opposite strand (*i.e.* the G pairing with the cytosine in CpG) thus giving rise to the ICL (Scheme 1). This DNA lesion represents a major threat to cell viability because it is highly effective in blocking replication and transcription forks.[22] The fact that, under standard conditions, alkylation at C10 always follows reaction at C1 explains why ICLs in duplex DNA are observed only at CpG sites. Intrastrand crosslinks at GpG steps,[23] as well as other non-cytotoxic monoadducts that arise from the MMC metabolite 2,7-diaminomitosene,[24] have also been characterized but will not be dealt with in the present work.

The nature of the base on the 3′-side of the guanine that is alkylated by MMC (G) plays a relatively minor role in site selectivity, with reactivity decreasing in the order CGC > CGT > CGG > CGA[5] reportedly due to the possibility of the 3′-pyrimidine assisting in the removal of the proton from the 2-amino group of guanine during the nucleophilic attack.[25] On the other hand, the preference of MMC for different dinucleotide steps, quantified in terms of yields and reaction rates with guanine, follows the order CpG > GpG ≫ TpG ≈ ApG.[5] The presumed involvement of the 2-amino group of a second guanine in the clear enhancement observed at the former two steps, and particularly at the CpG step, through a putative hydrogen bond with the carbamate oxygen attached to C10 (OM) in activated MMC was indirectly supported by the demonstration that replacement of the guanine in the complementary strand with hypoxanthine (as found in deoxyinosine) resulted in much lower reactivity[5] and also by the finding that MMC analogues lacking this OM did not show any selectivity towards CpG steps.[26] Conversely, 5-methylation of the cytosines making up this site was seen to result in a ~2-fold increase in the reactivity of the synthetic oligodeoxynucleotides thus modified. This effect was attributed to either electronic or steric factors, namely, either enhanced nucleophilicity of the 2-amino group of guanine or greater accessibility of the amino group due to a local conformational change.[27] Later work using a 162-bp DNA fragment in which all cytosines had been replaced with either 5-methyl-cytosine (5MC) or 5-F-cytosine (5FC) confirmed this increased reactivity for CpG sites containing 5MC and further showed that the opposite was true when the CpG site contained 5FC in lieu of C. The respective stimulatory and inhibitory effects of 5MC and 5FC were shown to be selective to the G that is paired with the modified pyrimidine. In this case, the explanation put forward to

account for these reactivities, which were said to correlate with the Hammett σ constants, was that the electron-donating effect of the 5-methyl substituent of the cytosine was transmitted to guanine through H-bonding of the 5MC:G base pair.[28] On the basis of density functional calculations on simplified model systems it was later proposed that one of the H-bonding hydrogens from the 2-amino position of the attacking G would be temporarily transferred to the cytosine oxygen with which it pairs to facilitate the reaction[29] but no further support for this hypothesis has been obtained.

It is then clear that MMC is a prodrug unable to bind to DNA unless it has been reduced. But precisely because its reductively activated forms are too reactive and short-lived for binding studies,[7,21] examination of the noncovalent association of this drug or its active metabolite(s) with different double helices has not been feasible. Furthermore, the structural work has been very limited despite the early use of CPK[6] and computer-generated models,[4,30,31] which only in a few instances were built making use of NMR-derived information on the monoalkylated product.[32,33] Thus, to date and to the best of our knowledge, only the solution NMR structure of the d(ICACGTCIT)·d(ACGACGTGC) duplex containing the major MMC monoadduct at the underlined guanine[32] is deposited in the public domain (Nucleic Acid Database[34] entry 199D) and models of an MMC ICL have been reported only for the self-complementary $d(GCATCGATGC)_2$ decamer[4] and $d(TACGTA)_2$ hexamer,[33] which do not represent optimal bonding sites. A common feature of all of these models is the proposal that the N^2-guanine-bonded MMC molecule fits snugly into the minor groove without appreciable perturbation of the DNA structure but none of them provides information about the precovalent complex(es) or the rationale for the stereochemistry of the drug-G covalent bond. In fact, the current paradigm holds that the preference for the CpG site is a consequence of the hydrogen bond between the carbamate oxygen of the MMC activated species and the amino group on the guanine of the nonbonding strand, a proposal that appears to be corroborated by the NMR structural work on the monoadduct.[32] A role for the carbonyl oxygen in this recognition was ruled out in view of the apparently similar bonding patterns obtained for MMC and 10-decarbamoylmitomycin C (DMC, Scheme 1) in early λ exonuclease stop assays[31] and also by the largely decreased sequence selectivity displayed by MMC derivatives bearing a halogen atom in place of the carbamoyl group.[26] Nonetheless, although DMC was initially regarded as a less active monofunctional derivative of MMC[35] later work showed it to produce more adducts and be more cytotoxic than MMC in certain cells, including FA fibroblasts, and to give rise to bisadducts.[36,37] Strikingly, however, the chirality of the mitosene linkage between DMC and the amino group of guanine was shown to be largely C1-β rather than the C1-α that is overwhelmingly observed for MMC, both in the more readily formed monoadducts and in the less frequent ICLs. Therefore the current paradigm cannot account for the opposite stereochemistry of the major DNA adducts produced in mammalian cells by MMC and DMC and for the differential sensitivity of FA cells to these drugs.[37]

To obtain information in atomic detail about the mechanism of CpG recognition and bonding by MMC and DMC, we have focused on a DNA double helix of sequence d(CGATAACGCTAACG) in one strand and the complementary sequence in the opposite strand. Using this 14mer, which embeds a high-affinity CGC site for mono- and bis-alkylation in the middle region, we have simulated in a continuous fashion all the steps of the reaction leading from the initial precovalent complexes to the final (and distinct) ICLs through prior formation of the respective monoadducts. Bond making and breaking during the molecular dynamics (MD) simulations was made possible by coupling a quantum mechanical (QM) function to a classical molecular mechanics (MM) potential, as recently implemented[38] in the popular AMBER suite of programmes (http://ambermd.org/). The main advantage of a hybrid approach involving MD and QM for the solvated drug–DNA complexes over studies that focus on simplified model systems is that changes in atom connectivities and charge redistributions are performed in a dynamic context that realistically allows reorganization of both solutes and water molecules through the whole procedure. Besides, in order to find a rationale to the observation that MMC–DNA monoadduct formation (and therefore subsequent interstrand crosslinking) is altered at CpG sequences bearing cytosine modifications,[27,28] simulations were also performed on similar DNA 14mers in which the cytosines making up the central CpG step contained either a methyl group or a fluorine atom at position 5. For comparison purposes and completeness, the three oligodeoxynucleotides studied as targets for MMC and DMC were also simulated in their free states. The trajectories were analyzed in terms of structural parameters, suitable geometries for nucleophilic attack, molecular electrostatic potentials and solvent-corrected binding energies.

Materials and methods

Geometries, AMBER parameters and charges for the non-standard residues

The X-ray crystal structure of MMC[39] was the starting point for model building this prodrug, as well as DMC and their corresponding reactive intermediates. The geometries of MMC, DMC, activated MMC (both as a hemiquinone methide and as a carbocation, MMC^+), activated DMC (as a carbocation, DMC^+), reduced MMC–G monoadduct, reduced DMC–G monoadduct, activated MMC–G monoadduct, activated DMC–G monoadduct, G–MMC–G and G–DMC–G bisadducts, 5FC, and 5MC were first refined by means of the semiempirical QM program MOPAC2009[40] using the AM1 Hamiltonian[41] and PRECISE stopping criteria, and further optimized using the restricted Hartree–Fock (RHF) method and a 6-31G(d) basis set, as implemented in the *ab initio* quantum chemistry program Gaussian03.[42] The modified nucleic bases incorporated a methyl group on either N^9 (G) or N^1 (C) in place of the deoxyribose ring. For each molecular system the calculated wave function was then used to derive electrostatic potential-derived (ESP) charges employing the RESP methodology.[43] Point charges for deoxyribose and phosphate atoms in the modified nucleotides were restrained to the values these atoms have in the *parmbsc0* AMBER force field,[44] which includes corrections for an improved description of DNA conformations on a multinanosecond time scale.[45] Additional bonded and nonbonded parameters for drugs and modified bases are reported in the ESI.

Construction of the oligodeoxynucleotides

An initial model for the free 14mer of sequence d-(CGATAACGCTAACG)· d(CGTTAGCGTTATCG) was built using the *nucgen* module in AMBER and optimized parameters for B-DNA.[46] Terminal G:C base-pairs on both sides of the A,T-rich regions were used to avoid fraying of the double helix.[47] Replacement of the cytosine base with either 5MC or 5FC in the central CpG step provided initial models for the modified oligonucleotides. The MMC-N^2(G) monoadduct found in NDB entry 199D[32] was used as a template for model building the precovalent intermediates by placing the activated MMC molecule in a similar location and avoiding the steric clash upon bond removal by steepest descent energy minimization.

Molecular dynamics simulations

Each molecular system was immersed in a truncated octahedron of TIP3P water molecules[48] and neutralized by addition of the appropriate number of Na^+ ions[49] at random locations. The standard MD simulations were run using the *pmemd* module in the AMBER 11 suite of programs. Periodic boundary conditions were applied and electrostatic interactions were treated using the smooth particle mesh Ewald method[50] with a grid spacing of 1 Å. The cutoff distance for the non-bonded interactions was 9 Å. The SHAKE algorithm[51] was applied to all bonds involving hydrogens and an integration step of 2.0 fs was used throughout. Solvent molecules and counterions were relaxed by energy minimization and allowed to redistribute around the positionally restrained solute (25 kcal mol^{-1} $Å^{-2}$) during 50 ps of MD at constant temperature (300 K) and pressure (1 atm), essentially as described previously.[52] These initial harmonic restraints were gradually reduced in a series of progressive energy minimizations until they were completely removed. The resulting systems were heated from 100 to 300 K over 20 ps, equilibrated at 300 K for 1 ns in the absence of any restraints and further simulated under the same conditions up to a total time of 10–20 ns during which system coordinates were collected every 20 ps for further analysis.

Hybrid QM/MM calculations

The system was partitioned into two distinct parts: a QM region defined by the *iqmatoms* keyword that consisted of the atoms relevant for the specific bonding reaction being studied and their immediate surroundings, and an MM region with all the remaining atoms, including the solvent molecules and the counterions. The AM1 Hamiltonian[41] and full electrostatic interactions between the QM charge density (expanded in a STO-6G minimal basis set) and the standard RESP/6-31G(d) point charges of the MM atoms were used,[38] as implemented in the *sander* module of AMBER 11. In the case of the precovalent complexes, the QM region included the reactive MMC and DMC derivatives plus the whole CpG step (to ascertain any possible differences involving the cytosines) whereas in the study of ICL formation, the QM region included the drug-G8 monoadduct plus the second guanine (G22) in the opposite strand. Reaction coordinates were specified for covalent bond formation, first between N^2 of G8 and C1 of MMC^+ or DMC^+ and then between N^2 of G22 and C10 in the monoadduct, using the steered MD procedure implemented in *sander*.[53] In brief, the N–C distance between the two atoms to be bonded was shortened linearly from its current value in the precovalent complex to a target value of 1.35 Å over 5 ps using a harmonic restraint of 1000 kcal mol^{-1} $Å^{-2}$ while all other variables were free to change. During the QM/MM part of the MD simulation SHAKE (keyword *qmshake*) was turned off and the integration step was reduced to 0.5 fs. Upon achievement of the final state, the AMBER topology file was updated with the new bond and atom type definitions, and a short energy minimization using the redefined connectivities and point charges allowed this geometry to be optimized in the MM force field prior to continuing the standard MD simulation. A scheme depicting the different MD and QM/MM phases of the whole study is given in the ESI.†

Analysis of the molecular dynamics trajectories and electrostatic energy calculations

Three-dimensional structures and trajectories were visually inspected using the computer graphics program PyMOL.[54] Interatomic distances and angles, as well as root-mean-square deviations (rmsd) from a given structure, were monitored using the *ptraj* module in AMBER. The DNA conformational and helical parameters were analyzed by means of the program Curves+.[55] Molecular electrostatic potentials (MEP) around the low-dielectric solutes ($\varepsilon = 1$) were calculated by treating the surrounding solvent as a continuous high-dielectric medium ($\varepsilon = 78.5$) and solving the non-linear Poisson–Boltzmann (PB) equation by means of a finite difference method as implemented in program DelPhi[56] version 4 and described in more detail elsewhere.[52,57] To highlight the most dissimilar MEP regions between any two DNA molecules, their central CGC triplets were best-fit superimposed and identical grids were calculated for both using the VSDMIP plugin.[58] The energy value at each grid point for one molecule was then subtracted from the value calculated at the same grid point for the other molecule, as reported previously,[59] and the difference map was visualized in PyMOL.

MM–GBSA binding energy calculations

The binding free energy ($\Delta G_{binding}$) between a reactive intermediate of MMC (MMC^+) and a given DNA oligonucleotide (standard, 5MC-DNA or 5FC-DNA) was calculated as:

$$\Delta G_{binding} = \langle G^{DNA-MMC^+}(i) - G^{DNA}(i) - G^{MMC^+}(i) \rangle_i \quad (1)$$

where $\langle\rangle_i$ denotes an ensemble average over i snapshots taken from the MD trajectories of the precovalent complexes. For comparison purposes we took 150 snapshots from 3.0 ns of each simulation during which MMC^+ was similarly oriented in the minor groove of the three complexes studied. The G_i values for each species were calculated using the following energy decomposition scheme, which is implemented in the MM–GBSA method within AMBER 11 (script mm_pbsa.pl):[60]

$$G = E_{gas} + G_{sol} - TS$$

$$E_{gas} = E_{bond} + E_{angle} + E_{torsion} + E_{ele} + E_{vdw}$$

$$G_{sol} = G_{GB} + G_{nonpolar}$$

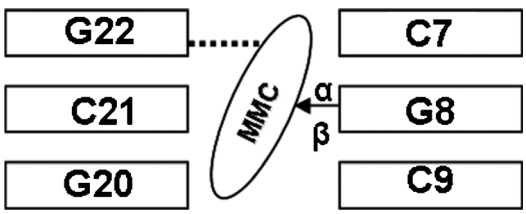
$$G_{nonpolar} = \gamma SASA$$

E_{gas} is the gas phase energy, calculated using the AMBER force field, which encompasses internal energies from bond lengths (E_{bond}), valence angles (E_{angle}) and torsional angles ($E_{torsion}$), as well as coulombic (E_{ele}) and van der Waals (E_{vdw}) non-bonded contributions calculated with no cutoff. In the case of individual trajectories, as in the present case, the contribution of the internal energies to $\Delta G_{binding}$ amounts to zero. The solvation free energy, G_{sol}, was decomposed into polar and nonpolar contributions; the former (G_{GB}) was calculated by solving the generalized Born (GB) equation[61] using dielectric constants of 1 and 78.5 for solute and solvent, respectively, whereas the latter, $G_{nonpolar}$, which is due to due to cavity formation and van der Waals interactions between the solute and the solvent, was estimated from the nonpolar solvent accessible surface area (SASA)[62] using 1.4 Å as the water probe radius and a value of 0.005 kcal mol^{-1} Å^{-2} for the surface tension constant γ. T and S are the temperature and the total solute entropy but this term, which is cumbersome to compute and most likely similar for the three DNA–MMC complexes under study, was obviated in our approximation to the binding free energy differences.[63]

Results and discussion

Drug conformation and orientation in the precovalent complexes

The MD simulation of the initial complex containing the standard oligodeoxynucleotide and an activated MMC molecule (MMC$^+$) in the minor groove of the central region corresponding to the C7:G22/G8:C21 base-pair step (Scheme 2) revealed that the hydrogen bond between the carbamate OM atom and the exocyclic amino group of G22 was not particularly stable in the face of competition with water molecules. This hydrogen bond was reported for the monoadduct structure solved by NMR spectroscopy[32] and presumed to dominate sequence recognition in the precovalent complex as well. However, in addition to the weak character and short life of this interaction (Fig. 1), we also noted that this orientation of the drug in the minor groove was incompatible with a nucleophilic attack by N^2(G8) on the α face of the MMC tetrahydropyrrole ring.



Scheme 2 Simplified diagram representing the alignment of MMC$^+$ in the minor groove of the DNA duplex central region showing the numbering of the base pairs involved in adduct formation.

This unexpected result led us to reconsider the conformation of MMC$^+$ inside the minor groove and explore alternative hydrogen bonding modes. In our lab, previous conformational studies on MMC$^+$ in the gas phase (data not shown) had revealed

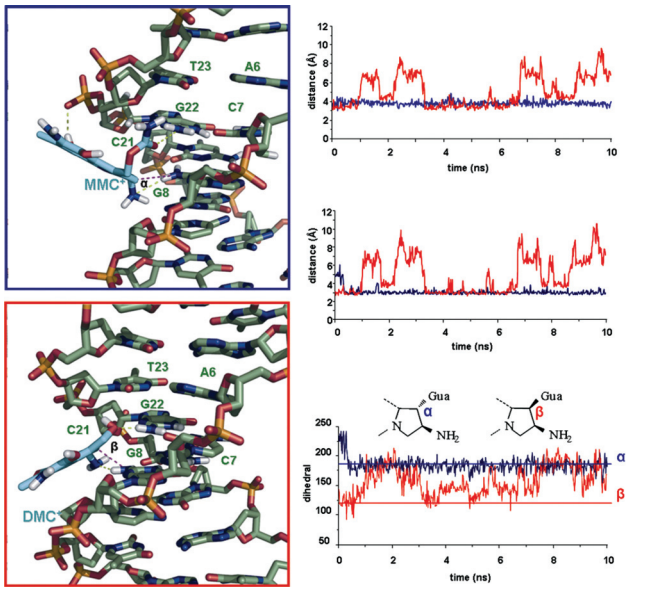


Fig. 1 *Left*: Details of representative MD snapshots from the MD simulations showing the distinct orientations of MMC$^+$ (top) and DMC$^+$ (bottom) in the DNA minor groove prior to nucleophilic attack by the 2-amino group of G8. Drug and DNA carbon atoms are coloured in cyan and green, respectively. *Right*, from top to bottom: Time evolution of N^2(G8)–C1(drug) and N^2(G8)–N2(drug) distances for MMC$^+$ (blue) and DMC$^+$ (red), and the angle for attack.

a preferred geometry that involved formation of an internal hydrogen bond between the phenol group at C8 and oxygen OM, and this conformer was also detected in aqueous solution when the dynamic behaviour of MMC$^+$ was simulated (Supplementary Figure S1†). Besides, additional QM work on a CpG dinucleotide model system had shown that, using this conformation, a reaction coordinate defined to create the covalent bond leading to the α monoadduct could be accomplished. Interestingly, when this particular MMC$^+$ conformer was docked into the oligo's minor groove region comprising C7:G22/G8:C21, we realized that a good hydrogen bond with the amino group of G22 could actually be established by the carbonyl oxygen (OC) instead of OM, as in the previous case. Therefore OC, rather than facing the solvent as in the former simulation, was directed towards the strand opposite to that being alkylated (Fig. 1). The MD trajectory of this new complex allowed us to see that this binding pose remained stable and not only provided a distance between C1 of MMC$^+$ and N^2(G8) that was suitable for nucleophilic attack, *i.e.* ≤4 Å (Fig. 1), but also allowed an approach on the ring face opposite to that of the primary amino group at C2 (Scheme 1). This particular orientation was additionally stabilized by two hydrogen bonds: a strong one between the phenolic O5 and phosphate OP1 oxygens, and another between the sp^2 N^2(G8) and the sp^3 nitrogen of the primary amino group at position 2 of the activated drug. This hydrogen-bonding anchoring, which is maintained for virtually the whole simulation (Supplementary Figure S2†), contributes to placing the mitosene ring system in an asymmetric position within the minor groove so that only one face is in van der Waals contact with the sugar-phosphate backbone of the unmodified strand whereas the other one remains exposed to the aqueous milieu.

In DMC, on the other hand, a hydroxymethyl substituent replaces the methyl carbamate moiety of MMC. Consequently, only one oxygen (OM) is present in the drug's side chain available for hydrogen-bonding to the DNA. When DMC^+ was docked in the minor groove region of the C7:G22/G8:C21 base-pair step similarly to MMC^+ and the resulting complex was simulated using MD, we noticed that OM could indeed accept a hydrogen bond from the amino group of G22 (Fig. 1) but the half-life of this interaction was shorter because of exchange with solvent molecules in the nanosecond timescale (data not shown). Nonetheless, this hydrogen bond largely contributed to placing DMC^+ in an orientation such that attack by N^2(G8) would occur preferentially, although not exclusively, on the β face of the drug's tetrahydropyrrole ring (Fig. 1). In this complex, the activated drug is roughly equidistant from both DNA strands and no direct hydrogen bond with any phosphate group is detected.

Taken together, these results strongly suggest that the very distinct relative occurrence of α and β DNA monoadducts formed in the presence of MMC and DMC[36,37] is largely dictated by the nature of the hydrogen bond involving either the carbamate or the hydroxyl oxygen in the activated drugs and the amino group of the guanine in the unmodified strand that is complementary to the cytosine in the CpG step undergoing G-alkylation. The fact that DMC and MMC show opposite stereochemical preferences supports the roles that we assign to the carbonyl oxygen, OC, which is only present in the former compound, and to OM, which is present in both MMC (as part of the carbamate) and DMC (as a hydroxyl). Altogether, the drug–DNA hydrogen bonds involving these oxygens crucially determine the proposed near-attack conformations and binding modes for MMC and DMC, which are distinct and can be distinguished by monitoring an improper dihedral angle for the incoming 2-amino group relative to the mitosene tetrahydropyrrole ring (Fig. 1).

Precovalent complexes between MMC^+ and DNA containing modified cytosines

Once the putative binding mode for MMC^+ in the minor groove of the central C7:G22/G8:C21 region was identified, we carried out similar MD simulations with the oligodeoxynucleotides in which C7 and C21 had been replaced with either 5MC or 5FC. When we calculated the binding energy ($\Delta G_{binding}$) between MMC^+ and the central CGC triplet using the MM–GBSA method, we found that, relative to the standard oligonucleotide, it was more favourable for the duplex containing 5MC whereas it was significantly decreased in the duplex containing 5FC

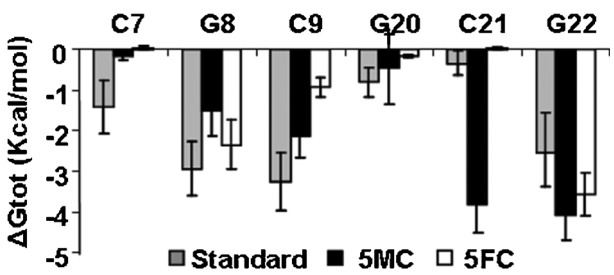


Fig. 2 Calculated binding energies between MMC^+ and individual DNA bases in the precovalent complexes formed with standard DNA, 5MC-DNA and 5FC-DNA.

(−11.2 ± 3.5, −12.1 ± 3.5, and −6.9 ± 1.5 kcal mol^{-1}, respectively). Furthermore, an energy decomposition analysis (Fig. 2) showed that, in both cases, the base providing the largest differences is precisely C21, *i.e.* the cytosine complementary to the guanine undergoing alkylation, which displays an improved interaction with the drug when it is methylated. Our theoretical result that C21, rather than C9, is largely responsible for these differences in $\Delta G_{binding}$ is in very good agreement with experimental evidence showing the prevalence of the 5MC that is paired to the alkylated guanine over that placed 5′ to it in the same strand.[64]

Since the major energy component appeared to be electrostatic we decided to explore the MEP generated in the minor groove of this central region in the three oligonucleotides studied. To facilitate comparisons we also calculated and displayed the MEP differences relative to the duplex containing standard cytosines (Fig. 3). It can be seen that significant differences do arise in the minor groove (and also between the stacked bases), despite the fact that the structural changes are located in the major groove. This type of observation has been used earlier to account for differences in the binding preferences of several bisintercalators (*e.g.* ref. 52 and 59) and can be used here to provide a rationale to the experimental finding that cytosine methylation enhances the rate of DNA alkylation by MMC whereas the opposite is true if C is replaced with 5FC.[27,28] Our interpretation is that the more negative MEP found in the drug-binding region when the cytosines in the CpG step are methylated facilitates and strengthens the electrostatic interaction with the positively charged alkylating species.

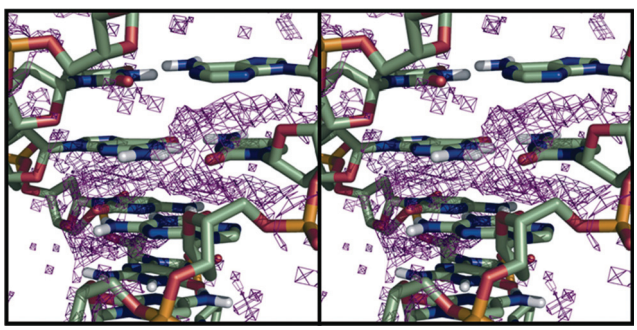


Fig. 3 Close-up view of the CpG region where MMC binds and display of the difference in molecular electrostatic potential (MEP) between 5MC-DNA and standard DNA (contoured at −5 kcal mol^{-1}). Note the more negative character of the MEP around the O^2 of both methylated cytosines.

DNA conformational parameters

To complete the study of the precovalent complexes, we examined their geometrical properties and compared them to those of free DNA, 5MC-DNA and 5FC-DNA. The very similar DNA conformations were well maintained during the course of the MD simulations (data not shown) and only subtle differences involving the critical G8pC9 step were observed. In particular, analysis of the stacking geometry between neighboring G8:C21 and C9:G20 base-pairs revealed a significant displacement towards negative tilt values relative to free DNA (Fig. 4) when MMC^+ is located in the DNA minor groove in a suitable position

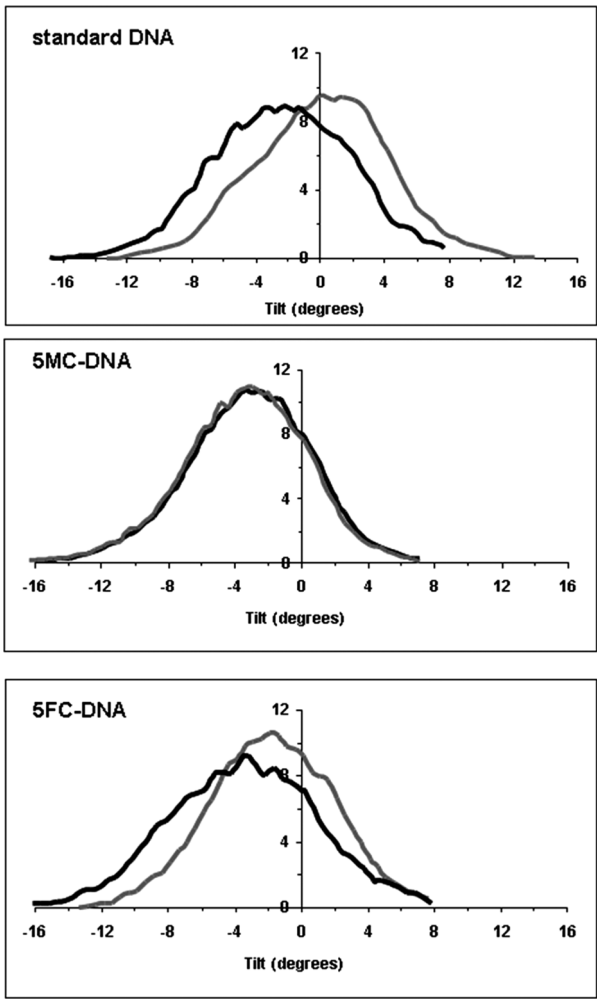


Fig. 4 Distribution of tilt values at the G8pC9 step of the central CGC triplet over the MD simulations of free (grey) standard DNA (top), 5MC-DNA (middle) and 5FC-DNA (bottom) and their respective precovalent complexes with MMC^+ (black). No significant differences were observed at the C7pG8 step.

for nucleophilic attack by N^2(G8). This means that the angle formed between these base-pairs needs to be opened towards the unmodified DNA strand in order for the reaction to proceed. Remarkably, this slightly distorted geometry is naturally present, in the absence of bound drug, when C9 and C21 are methylated and also, to a lesser extent, when C is replaced with 5FC. It is therefore likely that this structural preorganization for MMC^+ binding in the DNA containing the modified cytosines acts in concert with the more negative MEP in the minor groove to enhance alkylation by MMC when the 5 substituent is an electron-donating methyl group. On the other hand, in the case of the DNA containing 5FC in the central region, it seems that the less favourable electrostatic interaction with MMC^+ prevails over the more preorganized geometry to decrease drug binding to this site.

Formation and description of the MMC and DMC monoadducts

We used different snapshots from the MD trajectories of the MMC^+–DNA and DMC^+–DNA precovalent complexes reported above to start simulating the QM reaction coordinate that leads to covalent bond formation between N^2 of the attacking guanine (G8) and C1 of the drugs by systematically reducing this distance using the hybrid QM/MM method. Interestingly, a consistent pattern was found that revealed that the opposite α or β stereochemistry at C1 of the resulting adduct (Schemes 1 and 2) was entirely dependent on the conformation-driven positioning of the activated drug in the DNA minor groove. Thus, in the simulation involving MMC^+, hydrogen-bonding of the drug's OC to N^2(G22) determined formation of the C1α-monoadduct whereas the C1β-stereoisomer was systematically obtained when this hydrogen bond was lost. An identical result was obtained when the same procedure was used on the precovalent (5MC) DNA–MMC^+ and (5FC)DNA–MMC^+ complexes. The free energy profiles for this step of the reaction, on the other hand, showed no significant differences among the three complexes (Supplementary Fig. 2A†).

In the resulting MMC–DNA covalent intermediate N^2(G8) still bears two hydrogens, one of which (H^{21}) can still be used to pair with the complementary O^2(C21) following relaxation of the complex. But the extra H^{22} must be transferred to the amino group of the bonded drug (Scheme 1) and, given the geometry of the resulting adduct, this proton shuttle could only be effected with the concourse of the nucleobase 3′ to the alkylated guanine, in our case O^2(C9). Nonetheless, simulation of these reactions using the QM/MM method turned out to be unfeasible. We then realized that in the precovalent complexes reported above for both MMC^+ and DMC^+ the sp^2-hybridized N^2(G8) was close not only to C1 of the drugs but even more so to the sp^3-hybridized amino group at C2, with which it could actually establish a good hydrogen bond. In fact, both sets of interatomic distances were found to be strictly correlated during the MD simulations (Fig. 1), which suggests that transfer of H^{22} to N2(MMC^+) or N2(DMC^+) could be concomitant with the formation of the C1–N^2(G8) bond, but this alternative reaction pathway was not explored further.

The structures of the resulting MMC–DNA and DMC–DNA complexes containing, respectively, the α- and β-monoadducts were relaxed upon relocation of the extra H^{22} on N^2(G8) to the amino group of the bonded drug and equilibrated for 10 ns of unrestrained MD simulations. The MMC–DNA complex faithfully reproduced the major features of the models originally refined on the basis of NOE intensities and deposited in the NDB with code 199D,[32] despite differences in base composition and sequence on both sides of the central CpG step and the fact that the NMR solution structure considered a benzoquinone ring in MMC rather than the hydroquinone present in our model and necessary for evolution towards a bisadduct. Most noteworthy are (i) the absence of significant deviations in helical parameters with respect to free DNA (data not shown), (ii) a slight widening of the minor groove (Fig. 5) where the covalently bonded drug remains in an off-centre location despite the loss of the O5–OP1 hydrogen bond, and (iii) alternative orientations for the carbamate side chain of MMC.

In the case of the DMC β-adduct, the drug is lodged in a slightly compressed minor groove (Fig. 5), where it interacts simultaneously with the sugar-phosphate backbone of both strands, and the hydroxyl group at C10 keeps its hydrogen-bonding interaction with O^2(T23).

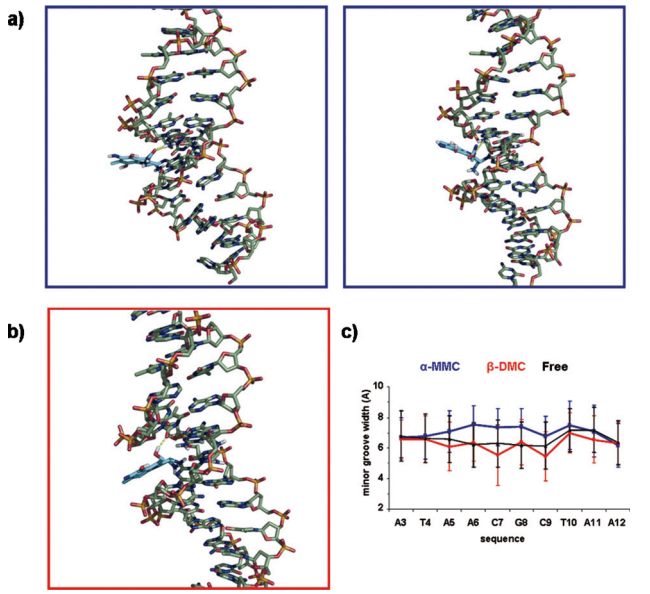


Fig. 5 a) Representative MMC–DNA α-monoadducts showing the two most populated orientations of the drug in the minor groove depending on whether the carbonyl oxygen of the drug's carbamate group is hydrogen bonded to N^2(G22) (left) or not (right). b) Representative DMC–DNA β-monoadduct showing the hydrogen bond between the drug's hydroxyl and O^2(T23). c) Widths of the minor groove in free DNA and in the two types of adducts.

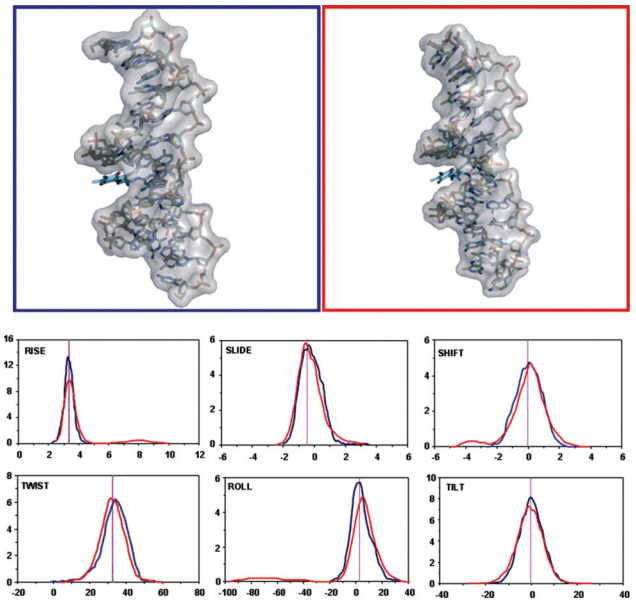


Fig. 6 *Top*: representative snapshots of the duplexes containing the C1α (left, blue frame) and C1β (right, red frame) bis-adducts that crosslink the two DNA strands. The DNA atoms are surrounded by a semi-transparent van der Waals surface. *Bottom*: Distributions of base-pair helical parameters (excluding both ends) in the duplexes containing the C1α (blue) and C1β (red) ICL. Twist, tilt and roll are in degrees. Shift, slide and rise are in Å. The vertical line represents the average for the simulated drug-free DNA.

From the above, it must be noted that a DMC–DNA complex resulting from water-mediated decarbamoylation of DNA-bonded MMC,[36] which appears to be more abundant in cells than the original adduct,[37] is stereochemically identical (C1α) to only one of the two monoadducts that readily originate from direct attack of DMC to DNA (Scheme 1), although in this case the C1β is obtained in a significantly larger proportion.[37]

Interstrand crosslinking by DNA-bonded MMC and DMC at the CpG step

Once the structures of the duplexes containing the α- and β-monoadducts were equilibrated, we proceeded to remove the carbamate (in the case of MMC) and the hydroxyl (in the case of DMC) so as to generate the second alkylating species. After 10 ns of equilibration of these new complexes using standard MD, we could observe that the separation between the exocyclic N^2(G22) and C10 of the drug fluctuated less in the case of DMC and remained close to 4 Å (Supplementary Fig. 3†), a suitable distance for nucleophilic attack and similar to that reported above for monoadduct formation. By switching to the QM/MM method we shortened the N^2–C10 distance so as to create the second covalent bond and give rise to the corresponding bis-adducts, which were equilibrated thereafter for 20 ns of unrestrained MD simulation.

Our results show that the DNA is able to accommodate this dual covalent modification (on G8 in one strand and G22 in the opposite strand) undergoing minimal distortion only in the case of the C1α-ICL, in agreement with early work that employed MM exclusively.[4] In this bis-adduct, the crosslinking mitosene unit is settled in the minor groove of the CpG step in close

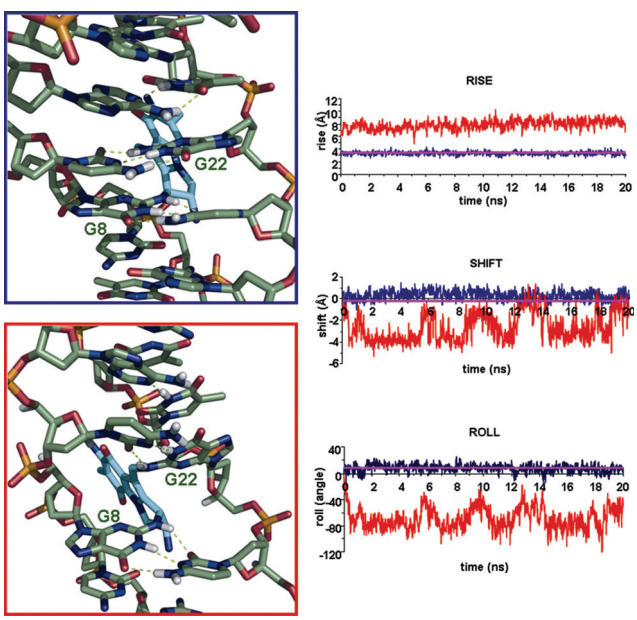


Fig. 7 *Left*: detail of the central region of the duplexes containing the C1α (top, blue frame) and C1β (bottom, red frame) bis-adducts that crosslink the two DNA strands, as seen from the major groove. Note the loss of intrastrand stacking interactions at the central CpG step in the C1β ICL. *Right*: Time evolution of rise (top), shift (middle) and roll (bottom) parameters at the C7pG8 step during the MD simulations of the bis-adducts. The pink horizontal line in each plot represents the average value for drug-free DNA.

contact with the strand containing the second alkylated guanine; the global structural parameters of the DNA molecule hardly

change with respect to the drug-free duplex (Fig. 6). On the contrary, in the C1β-ICL, whose structure has not been reported yet as far as we know, the mitosene unit is more equidistant to both DNA strands but the rise, shift and roll parameters at the C7pG8 step significantly deviate from the values in free DNA and this causes a marked change in the distribution of these values (Fig. 6). A more detailed view from the major groove is shown in Fig. 7 where the loss of stacking interactions at this step can be visually appreciated and also deduced from the increased rise values that are monitored along the MD trajectory. The negative shift is a consequence of the displacement of the drug-bonded guanines into the minor groove whereas the correlated negative roll indicates bending of the helix toward the minor groove, a situation contrary to what is observed in DNA duplexes containing an adduct with the monofunctional antitumour tetrahydroisoquinolines trabectedin (Yondelis®) or PM01183, which have been shown to bend the helix toward the major groove by introducing positive roll[65,66] and can functionally mimic an ICL.[67]

Conclusions

The oligodeoxynucleotide used in the present study contains a highly preferred CGC triplet site in the middle region for MMC and DMC binding. Furthermore the initial monoadducts formed with G8 can evolve towards an ICL by forming a second covalent bond with G22 in the opposite strand. By simulating each step of the reaction using a combination of MD and QM/MM calculations, we have been able to obtain information in atomic detail about the recognition events prior to monoadduct formation as well as about the structural and electronic requirements that possibly account for the increased reactivity of MMC towards CpG steps containing 5MC in place of cytosine. Moreover, we provide a possible explanation to the origin of the formation of either C1α or C1β monoadducts depending on the drug–DNA hydrogen bonds that are established in the precovalent DNA–MMC$^+$ and DNA–DMC$^+$ complexes. According to our findings, formation of a hydrogen bond between the carbamate oxygen of activated MMC and the guanine in the strand opposite to that containing the G that undergoes the first alkylation reaction appears to favour formation of the C1α-monoadduct. A change in the positioning of its decarbamoylated analogue DMC due to a distinct hydrogen-bonding pattern leads to a preferred attack on the opposite face of the pyrrolidine ring and creation of the C1β-monoadduct.

The DNA helical structure at the target CpG step appears to be more preorganized for binding the positively charged activated MMC, and also to possess a more negative MEP in the minor groove, when the cytosines are methylated. Both factors acting in concert may well explain the enhanced reactivity of MMC towards this 5MC-DNA that has been detected experimentally.

Our unrestrained MD simulations in aqueous solution faithfully reproduce the drug–DNA interactions previously described for the NMR solution structure of the major MMC–DNA C1α-monoadduct and provide further insight into (i) the earlier precovalent complex, (ii) the reaction of the activated monoadduct with a second guanine in the opposite strand at a CpG step, and (iii) the final ICL, which is embedded in a minimally distorted DNA helix that does not show any appreciable curvature, in good agreement with the experimental evidence.[33] We then studied in a similar way the formation of the major DMC–DNA monoadduct with opposite stereochemistry at C1 and its stepwise evolution to an ICL. The three-dimensional structures of these C1β-adducts have not been reported before, to the best of our knowledge, and they were found to display some distinct features. Most notably, the drug-bonded CpG step is severely distorted in the β-ICL and the helix is bent towards the minor groove.

The differences we have detected in the conformations of these adducts, which appear in different proportions in the DNA of cells that have been exposed to either MMC or DMC, are likely to pose dissimilar challenges to the DNA repair machineries. This may account for the experimental findings that FA cells are hypersensitive to MMC but have normal sensitivity to DMC whereas in other cell types DMC is more cytotoxic than MMC.[37]

References

1 T. Hata, T. Hoshi, K. Kanamori, A. Matsumae, Y. Sano, T. Shima and R. Sugawara, *J. Antibiot. (Tokyo)*, 1956, **9**, 141–146.
2 J.-C. Andrez, *Beilstein J. Org. Chem.*, 2009, **5**, 1–36.
3 V. N. Iyer and W. Szybalski, *Proc. Natl. Acad. Sci. U. S. A.*, 1963, **50**, 355–362.
4 M. Tomasz, R. Lipman, D. Chowdary, J. Pawlak, G. L. Verdine and K. Nakanishi, *Science*, 1987, **235**, 1204–1208.
5 S. Kumar, R. Lipman and M. Tomasz, *Biochemistry*, 1992, **31**, 1399–1407.
6 M. Tomasz, D. Chowdary, R. Lipman, S. Shimotakahara, D. Veiro, V. Walker and G. L. Verdine, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 6702–6706.
7 M. Tomasz, *Chem. Biol.*, 1995, **2**, 575–579.
8 M. D. Shelley, M. D. Mason and H. Kynaston, *Cancer Treat. Rev.*, 2010, **36**, 195–205.
9 L. M. Abraham, D. Selva, R. Casson and I. Leibovitch, *Drugs*, 2006, **66**, 321–340.
10 R.-D. Hofheinz, U. Beyer, S.-E. Al-Batran and J. T. Hartmann, *Onkologie*, 2008, **31**, 271–281.
11 J. P. McPherson, B. Lemmers, R. Chahwan, A. Pamidi, E. Migon, E. Matysiak-Zablocki, M. E. Moynahan, J. Essers, K. Hanada, A. Poonepalli, O. Sanchez-Sweatman, R. Khokha, R. Kanaar, M. Jasin, M. P. Hande and R. Hakem, *Science*, 2004, **304**, 1822–1826.
12 Y.-J. Lee, S.-J. Park, S. L. M. Ciccone, C.-R. Kim and S.-H. Lee, *Carcinogenesis*, 2006, **27**, 446–453.
13 S. E. Henson, S.-C. Tsai, C. S. Malone, S. V. Soghomonian, Y. Ouyang, R. Wall, Y. Marahrens and M. A. Teitell, *Mutat. Res., Fundam. Mol. Mech. Mutagen.*, 2006, **601**, 113–124.
14 M. E. Moynahan, T. Y. Cui and M. Jasin, *Cancer Res.*, 2001, **61**, 4842–4850.
15 F. Zhang, J. Ma, J. Wu, L. Ye, H. Cai, B. Xia and X. Yu, *Curr. Biol.*, 2009, **19**, 524–529.
16 M. C. Villarroel, N. V. Rajeshkumar, I. Garrido-Laguna, A. De Jesus-Acosta, S. Jones, A. Maitra, R. H. Hruban, J. R. Eshleman, A. Klein, D. Laheru, R. Donehower and M. Hidalgo, *Mol. Cancer Ther.*, 2011, **10**, 3–8.
17 C. A. Strathdee, A. M. Duncan and M. Buchwald, *Nat. Genet.*, 1992, **1**, 196–198.
18 J. Cervenka, D. Arthur and C. Yasis, *Pediatrics*, 1981, **67**, 119–127.
19 K. D. Mirchandani and A. D. D'Andrea, *Exp. Cell Res.*, 2006, **312**, 2647–2653.
20 S. E. Wolkenberg and D. L. Boger, *Chem. Rev.*, 2002, **102**, 2477–2495.
21 M. Tomasz and Y. Palom, *Pharmacol. Ther.*, 1997, **76**, 73–87.
22 D. M. Noll, T. M. Mason and P. S. Miller, *Chem. Rev.*, 2006, **106**, 277–301.
23 R. Bizanek, B. F. McGuinness, K. Nakanishi and M. Tomasz, *Biochemistry*, 1992, **31**, 3084–3091.

24 Y. Palom, M. F. Belcourt, S. M. Musser, A. C. Sartorelli, S. Rockwell and M. Tomasz, *Chem. Res. Toxicol.*, 2000, **13**, 479–488.
25 H. Borowy-Borowski, R. Lipman and M. Tomasz, *Biochemistry*, 1990, **29**, 2999–3006.
26 V.-S. Li, D. Choi, Z. Wang, L. S. Jimenez, M.-s. Tang and H. Kohn, *J. Am. Chem. Soc.*, 1996, **118**, 2326–2331.
27 J. T. Millard and T. M. Beachy, *Biochemistry*, 1993, **32**, 12850–12856.
28 A. Das, K. S. Tang, S. Gopalakrishnan, M. J. Waring and M. Tomasz, *Chem. Biol.*, 1999, **6**, 461–471.
29 J. J. Dannenberg and M. Tomasz, *J. Am. Chem. Soc.*, 2000, **122**, 2062–2068.
30 S. N. Rao, U. C. Singh and P. A. Kollman, *J. Am. Chem. Soc.*, 1986, **108**, 2058–2068.
31 V. S. Li and H. Kohn, *J. Am. Chem. Soc.*, 1991, **113**, 275–283.
32 M. Sastry, R. Fiala, R. Lipman, M. Tomasz and D. J. Patel, *J. Mol. Biol.*, 1995, **247**, 338–359.
33 D. Norman, D. Live, M. Sastry, R. Lipman, B. E. Hingerty, M. Tomasz, S. Broyde and D. J. Patel, *Biochemistry*, 1990, **29**, 2861–2875.
34 H. M. Berman, W. K. Olson, D. L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S. H. Hsieh, A. R. Srinivasan and B. Schneider, *Biophys. J.*, 1992, **63**, 751–759.
35 S. Y. Kim and S. Rockwell, *Oncol. Res.*, 1995, **7**, 39–47.
36 Y. Palom, G. Suresh Kumar, L.-Q. Tang, M. M. Paz, S. M. Musser, S. Rockwell and M. Tomasz, *Chem. Res. Toxicol.*, 2002, **15**, 1398–1406.
37 M. M. Paz, S. Ladwa, E. Champeil, Y. Liu, S. Rockwell, E. K. Boamah, J. Bargonetti, J. Callahan, J. Roach and M. Tomasz, *Chem. Res. Toxicol.*, 2008, **21**, 2370–2378.
38 R. C. Walker, M. F. Crowley and D. A. Case, *J. Comput. Chem.*, 2008, **29**, 1019–1031.
39 K. Shirahata and N. Hirayama, *J. Am. Chem. Soc.*, 1983, **105**, 7199–7200.
40 J. J. P. Steward, *MOPAC2009*, (2009), Colorado Springs, CO 80921, USA.
41 M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902–3909.
42 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *GAUSSIAN 03 (Revision B.04)*, Gaussian, Inc., Wallingford, CT, 2004.
43 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.
44 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.
45 A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham Iii, C. A. Laughton and M. Orozco, *Biophys. J.*, 2007, **92**, 3817–3829.
46 S. Arnott and D. W. Hukins, *Biochem. Biophys. Res. Commun.*, 1972, **47**, 1504–1509.
47 J. A. Bueren-Calabuig, C. Giraudon, C. M. Galmarini, J. M. Egly and F. Gago, *Nucleic Acids Res.*, 2011, **39**, 8248–8257.
48 W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey and M. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
49 J. Åqvist, *J. Phys. Chem.*, 1990, **94**, 8021–8024.
50 T. A. Darden, D. York and L. G. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
51 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
52 E. Marco, A. Negri, F. J. Luque and F. Gago, *Nucleic Acids Res.*, 2005, **33**, 6214–6224.
53 V. Babin, *J. Chem. Phys.*, 2008, **128**, 134101.
54 W. D. DeLano, *PyMOL*, 2006.
55 R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute and K. Zakrzewska, *Nucleic Acids Res.*, 2009, **37**, 5917–5929.
56 M. K. Gilson, K. A. Sharp and B. H. Honig, *J. Comput. Chem.*, 1988, **9**, 327–335.
57 E. Marco, W. Laine, C. Tardy, A. Lansiaux, M. Iwao, F. Ishibashi, C. Bailly and F. Gago, *J. Med. Chem.*, 2005, **48**, 3796–3807.
58 A. C. Cabrera, R. Gil-Redondo, A. Perona, F. Gago and A. Morreale, *J. Comput.-Aided Mol. Des.*, 2011, **25**, 813–824.
59 C. Bailly, S. Echepare, F. Gago and M. J. Waring, *Anticancer Drug Des.*, 1999, **14**, 291–303.
60 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, *Acc. Chem. Res.*, 2000, **33**, 889–897.
61 A. Onufriev, D. Bashford and D. A. Case, *J. Phys. Chem. B*, 2000, **104**, 3712–3720.
62 D. Sitkoff, K. A. Sharp and B. Honig, *J. Phys. Chem.*, 1994, **98**, 1978–1988.
63 A. Negri, D. Rodríguez-Larrea, E. Marco, A. Jiménez-Ruiz, J. M. Sánchez-Ruiz and F. Gago, *Proteins: Struct., Funct., Bioinf.*, 2010, **78**, 36–51.
64 W. S. Johnson, Q. Y. He and M. Tomasz, *Bioorg. Med. Chem.*, 1995, **3**, 851–860.
65 R. García-Nieto, I. Manzanares, C. Cuevas and F. Gago, *J. Am. Chem. Soc.*, 2000, **122**, 7172–7182.
66 J. F. Leal, M. Martinez-Diez, V. Garcia-Hernandez, V. Moneo, A. Domingo, J. A. Bueren-Calabuig, A. Negri, F. Gago, M. J. Guillen-Navarro, P. Aviles, C. Cuevas, L. F. Garcia-Fernandez and C. M. Galmarini, *Br. J. Pharmacol.*, 2010, **161**, 1099–1110.
67 S. Feuerhahn, C. Giraudon, M. Martinez-Diez, J. A. Bueren-Calabuig, C. M. Galmarini, F. Gago and J. M. Egly, *Chem. Biol.*, 2011, **18**, 988–999.

