

Unusual DNA Structure and DNA Damage Recognition: Structure and Dynamic Markers

Markus W. Germann*, Christopher N. Johnson, and Alexander M. Spring

Abstract: Nucleic acids play a central role in many biological processes, including information storage, gene expression, serving as messengers or structural components and even catalysis. Their diverse roles have made them targets of interest to diagnose and treat an array of human disorders such as infections, degenerative diseases and cancer. Nature has evolved proteins and ligands that recognize specific nucleic acid sequences or structures and control their function, demonstrating that this can be efficiently accomplished. This has led to the development of wide variety of synthetic molecules that selectively bind to nucleic acids. In turn, this has precipitated numerous studies which showed that nucleic acid structures and their dynamic properties must be understood in order to efficiently target specific sequences or structures.

Keywords: Alpha anomeric · DNA structure · Parallel stranded · Repair



Markus W. Germann was born in 1959 in Moutier, Switzerland. Following an apprenticeship with Ciba-Geigy, he earned his M.S. in chemical engineering from Winterthur Polytechnic. He obtained his PhD in 1989 from the University of Calgary, Canada. Subsequently he worked as an NMR application specialist with Bruker-Spectrospin. In 1993 he joined the faculty at Thomas Jefferson University in Philadelphia as an Assistant Professor. He was promoted to Associate Professor in 2001 and in the same year he moved to Georgia State University as a Georgia State Distinguished Cancer Scientist where he was promoted to Professor in 2006. His research interests include DNA damage and repair, macromolecular structures and dynamics and design of antiviral agents.

He obtained his PhD in 1989 from the University of Calgary, Canada. Subsequently he worked as an NMR application specialist with Bruker-Spectrospin. In 1993 he joined the faculty at Thomas Jefferson University in Philadelphia as an Assistant Professor. He was promoted to Associate Professor in 2001 and in the same year he moved to Georgia State University as a Georgia State Distinguished Cancer Scientist where he was promoted to Professor in 2006. His research interests include DNA damage and repair, macromolecular structures and dynamics and design of antiviral agents.

Nucleic acids play a central role in many biological processes. Their functions include information storage, controlling gene expression, serving as messengers or structural components and even catalysis. Given their diverse roles, they are targets of interest to diagnose and treat an array of human disorders such as infections, degenerative diseases and cancer. Nature has produced an assortment of protein and non-protein ligands that are able to recognize specific nucleic acid sequences or structures. The development of molecules that selectively bind to nucleic acids such as metal complexes, peptides, oligonucleotides and a wide array of organic compounds has provided many details and showed that nucleic acid structures must be understood in order to target them.

In the following, we focus on the involvement of our laboratories in the discovery of unusual DNA structures and DNA damage recognition. It has been known for some time that DNA possesses conformational variability far beyond the original B-form double helix proposed by Watson and Crick in 1953.^[1] Common properties in the three major families A, B and Z-DNA duplexes are the antiparallel orientation of the complementary strands, Watson-Crick (WC) type base pairing and the presence of 5'-3' phosphodiester linkages. The existence of parallel stranded structures was reported as early as 1961 for poly(A) and later for oligodeoxy- and polydeoxyribonucleotides at low pH.^[2] In addition to low pH, parallel stranded duplex structures are promoted by chemical modifications and bulky substituents that block WC base pairing as well as ligand nucleic acids interaction.^[3,4]

1. Parallel-stranded Duplex DNA

In 1986 Pattabiraman suggested that the homooligomeric AT tracts (dA₆:dT₆) could form a parallel-stranded right-handed double helical structure.^[5] In this model each individual DNA strand essentially maintains a B-like conformation with 2' endo sugar puckering and anti orientation of the bases. To test this prediction artificial 3'-3' or 5'-5' phosphodiester linkages were employed to enforce an intramolecular parallel-stranded orientation in the sequence dA₁₀:dT₁₀.^[6] These and similar sequences were shown to form double helical PS DNA with reverse Watson-Crick base pairing under physiological conditions (Fig. 1).^[7]

A characteristic feature of PS DNA is the equal size of the grooves (Fig. 2). This is

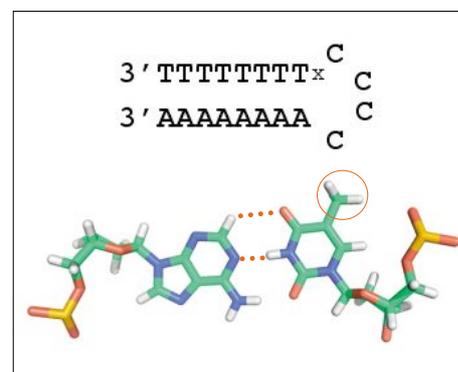


Fig. 1. Top: 3'-d(T₈x C₄A₈)-3' DNA hairpin. X denotes a 5'-5' linkage. Note polarity reversal caused by the 5'-5' linkage. Bottom: Reverse Watson Crick base pair in a parallel stranded DNA hairpin.^[7] Hydrogen bonds and location of T Methyl group are marked.

*Correspondence: Prof. M.W. Germann
Departments of Chemistry and Biology
Georgia State University
Atlanta, Georgia 30303, USA
Tel.: +1 404 651 1576
E-mail: mwg@gsu.edu

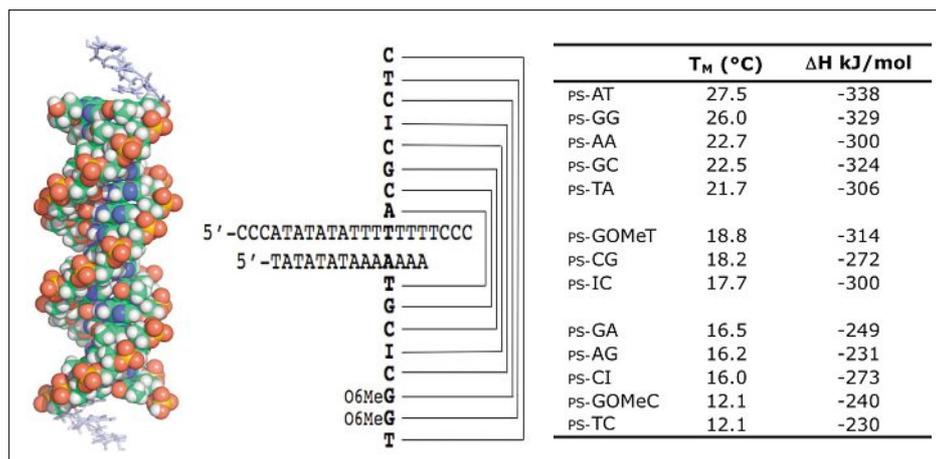


Fig. 2. Left: Model of an intramolecularly formed duplex PS DNA with homooligomeric and alternating AT base pairs.^[8] Middle: Duplex and substitution pattern to test alternate base combination. I: Inosine, O6MeG is G where the O6 position is methylated and the N1 functionality is changed from a H-bond donor to acceptor. Right: Melting temperatures (T_m) and enthalpies of PS duplexes (10 μ M) containing different base combination in 15 mM MgCl₂, 50 mM Tris, pH 7.2, measured at 260 nm. The first letter in the name refers to the substitution in the 15mer.

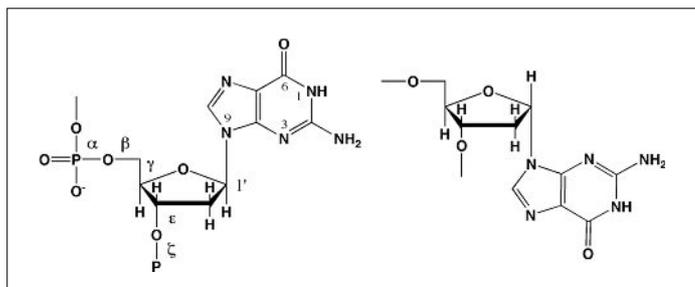


Fig. 3. Natural β - and α -anomeric dG nucleotide. The nucleotides differ at the stereochemistry at the C(1') position. Also indicated are numbering and torsion angles for the β -dG nucleotide.

in stark contrast to B-DNA which features distinct major and minor grooves. For PS AT sequences, the same functional groups are present in the former minor groove, which is enlarged and now contains the thymine methyl groups (Fig. 1). Hoechst 33258, an AT specific minor groove binding drug is therefore expected to bind to PS DNA but with altered affinity and spectral properties. This was confirmed by the observation that PS-AT hairpin has a 41-fold lower affinity for the drug than an antiparallel control while removal of the methyl groups (PS-AU hairpin) increases the binding constant 8-fold compared to the PS-AT sequence.^[9]

Because of the same hydrogen bonding potential, alternating AT sequences (*i.e.* d(AT)_n) can also form PS DNA.^[9-11] The stability of alternating AT base pairs is however distinctly lower than homooligomeric AT pairs, largely because of a smaller helix growth parameter.^[9] Parallel-stranded DNA is not restricted to systems containing unusual linkages but occurs in unconstrained and unmodified DNA with the appropriate sequence composition (Fig. 2).^[8,12] Subsequently, it was shown that PS DNA can also tolerate the presence of GC base pairs albeit with compromised stability compared to an AT base pair.^[13] The GC

pair is predicted to form with a single hydrogen bond, although a wobble pair with two hydrogen bonds remains a possibility. The dimer system shown in Fig. 2, served as a test bed to evaluate the stability of other base combinations and revealed that GG and AA but not GA or TC base combinations are tolerated. The observation that the GO6Me-C is not stable while the GO6Me-T combination implicates O(6) and N(1) of G in base pair formation (Fig. 3).

It has been known for some time that oligonucleotides containing (GA)_n repeats are inefficient agents for either triplex formation or antisense approaches, partly due to the inherent tendency of these sequences to self-associate.^[14] This provided an impetus for the study of (GA)_n sequences that has resulted in several proposed structures.^[15,16] Based on chemical modification and model building, Rippe *et al.* proposed that (GA)_n sequences form parallel-stranded duplexes with a dinucleotide repeat (with G-syn and A-anti glycosidic torsion) and GG respectively AA base pairs.^[15] An NMR-based study at low pH on a related sequence favors an anti conformation of all bases and different base pairing scheme for G.^[17] Our work on 3'-d(GA)₃-5'-T₄(AG)₃-3' and 3'-d(GA)₄-5'-5'-T₄(AG)₄-3' DNA hairpins confirms the formation of a

parallel-stranded structure; no evidence for triple-stranded structures could be observed. Further support for GG and AA pairing was obtained from a dimer duplex where either GG and AA or alternatively, GA pairing was enforced by sequence constraints. Only duplexes with GG and AA base pairs were found to be stable. This prompted us to examine additional sequences for parallel DNA formation. In the mammalian genome (GA)_n and (GT)_n tracts occur frequently and may potentially also form PS structures *in vitro*.^[18] This was demonstrated with the sequence 3'-d(GT)₃-5'-5'-T₄(AG)₃-3' which formed a parallel-stranded hairpin with remarkable stability, similar to the (GA)_n sequences. Using the appropriate sequence design we could show that GG and AT parallel base pairs form.^[19] A parallel strand disposition is also obtained in duplexes formed between complementary α - and β -anomeric sequences^[20] (Fig. 3).

Considering this, we hypothesized that introduction of an alpha anomeric component in a GA duplex will again invert the polarity and therefore result in a novel antiparallel duplex. This is indeed the case, 5'-d(GA)₄T₄ α [(AG)₄T-3', forms an intramolecular antiparallel duplex with G- α G and A- α A base pairs.^[21]

2. DNA•DNA and DNA•RNA Duplexes with Localized Parallel-stranded Segments

The promise of controlling gene expression at the translational level in a specific and efficient manner has spurred a large research effort into antisense oligodeoxynucleotide (ODN) therapy.^[22] The basic principle is to interfere with the production of an undesired protein by inactivating or degrading the corresponding mRNA. Desired properties of potential antisense ODNs include: high nuclease resistance to preserve integrity in an organism, high affinity and selectivity towards the intended target mRNA, cellular uptake, synthetic feasibility and RNase H cleavage of the mRNA in complex with the ODN. The latter takes advantage of a cellular enzyme and enables an ODN to potentially mediate the destruction of a large number of mRNA molecules. To meet these different challenges a flurry of modifications to bases, sugars and backbones were explored.^[20,23] With very few exceptions however, most modifications result in ODNs that do not support the desired RNase H activity. This has led to the design of chimeric ODNs containing nuclease resistant components and segments that elicit RNase H activity.^[24] Our strategy and that of others was to employ α -anomeric nucleotides in conjunction with 3'-3' and 5'-5' linkages to

generate a new class of ODNs.^[25] The 3'-3' and 5'-5' linkages allow the local inversion of the strand polarity, enabling the parallel stranded α -anomeric tracts to form Watson-Crick base pairs with the RNA target. This results in ODNs that permit RNase H to destroy the RNA component of an ODN•RNA heteroduplex^[26–28] and selectively inhibit growth in two cervical carcinoma cell lines as well as in tumor-bearing mice.^[29] These ODNs bind more strongly to the target RNA than phosphorothioates due to the enantiomeric purity of α -anomeric nucleotides and the inherent stability of α -ODN/ β -RNA hybrids.^[20]

2.1 Alpha-winged ODN/RNA Duplexes Are Susceptible to Cleavage by RNase H

The α/β -ODNs containing polarity reversals that we are using in our proof-of-principle studies target *erbB-2* mRNA, which is overexpressed in >30% of early and advanced mammary tumors.^[30] This target has been tested with six different overlapping antisense sequences and several control sequences and cancer cell lines.^[31] In order to compare the efficacy of the α/β -gapmer ODNs, control sequences comprised of all α -anomeric nucleotides (alpha), all phosphorothioate backbone, and gapmers featuring four or five anomeric residues flanked by 2'-O-methyl stretches (2'-O-Me4 and 2'-O-Me5) were also studied.

A β -anomeric window of three nucleotides (two natural phosphodiester linkages) is sufficient to allow limited *E. coli* RNase H activity (Fig. 4). Increasing the window size markedly improves the substrate quality as measured from the digestion of the labeled 15-mer RNA target. Clearly, the α -winged ODNs are better substrates than the corresponding 2'-O-Me sequences. For mammalian RNase H a four β -anomeric window (three natural phosphodiester linkages) is sufficient. Again we observe that the α -winged oligonucleotides outperform the 2'-O-Me sequences. The preferred site of RNA hydrolysis for the mammalian enzyme is at the center of the β -anomeric window while *E. coli* RNase H cuts the RNA across the 5'-5' linkage. All gapmer constructs performed considerably better than the PS ODN for which only 49% of the RNA was hydrolyzed. Our work also showed that a substoichiometric amount of ODN was able to digest an excess (8-fold) of RNA, demonstrating that the ODN can mediate the destruction of multiple RNA targets, which is of biological interest.^[28]

2.2 NMR Structural Studies

The previous section demonstrates the need to understand the structural and dynamic details of α -anomeric residues combined with linkage reversals in nucle-

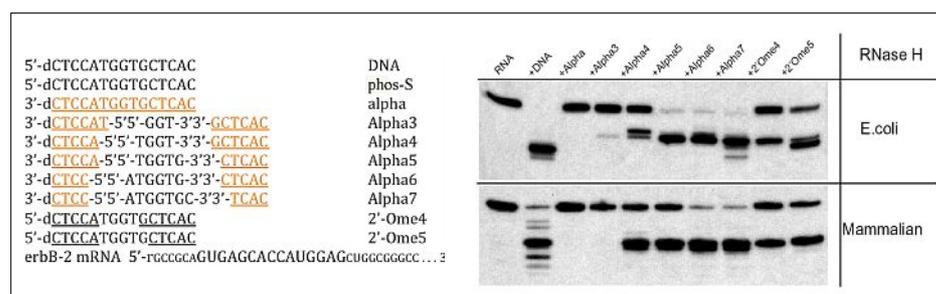


Fig. 4. Left: 15-mer ODN sequences targeting *erbB-2* mRNA. Nuclease resistant α -anomeric nucleotides shown in red and are underlined, while 2'-O methyl nucleotides are underlined only. The 2'-O methyl group substitutes the ribose 2'-OH group, in phosphorothioates phos-S one of the non-bridging oxygens in the phosphodiester linkage is replaced by a sulphur. *erbB-2* mRNA: RNA target; the 15-mer used in the RNase H reactions is shown in larger lettering. Right: Denaturing gel electrophoresis of RNase H substrates. Reactions contain ³²P- RNA and RNase H (*E. coli*, or rabbit reticulocyte lysate).

ic acids. Inserting an α T residue in an antiparallel orientation (*i.e.* in absence of the linkage) in a DNA sequence was found to severely destabilize duplex formation.^[27] Using 3'-3' and 5'-5' linkages we designed DNA•DNA and DNA•RNA duplexes containing single α -anomeric residues (α T, α A, α G and α C) (Fig. 4). NMR solution structures were determined using standard techniques. Briefly, the MARDIGRAS (Matrix Analysis of Relaxation for Discerning the Geometry of an Aqueous Structure) program was used to generate proton–proton distance restraints from integrated NOESY (Nuclear Overhauser Spectroscopy) peaks in the presence of spin diffusion and torsion angles defining the deoxyribose ring systems were calculated from three-bond proton–proton scalar couplings.^[32]

The DNA•RNA nonamer hybrid corresponds to the core of the *erbB-2* RNA target presented in the previous section. Thermodynamic and spectroscopic (UV and circular dichroism) studies established that the presence of a single 3'-3'- α N-5'-5' modification per strand results in a relatively small decrease in thermostability (T_m and ΔH° : control > α T \approx α A \approx α G > α C). Both DNA and hybrid duplexes were fully Watson-Crick base paired and the parallel α -anomeric residues were stacked into the helix (Fig. 5). The DNA duplex adopts over all a B-type structure that is highly superimposable with the unmodified control. However, a number of local perturbations are evident within each region encompassing the modifications. Beside its own intrinsic high S deoxyribose, α T also resulted in an altered sugar ring conformation (shift to S conformation) of the nucleotide following the 5'-5' linkage and unusual backbone torsion angles at the 3'-3' and 5'-5' linkages (Fig. 5).

In the case of the DNA•RNA hybrid, the polarity-reversed α T fits snugly into an overall A-like duplex that still retains some

B-type characteristics. Analysis of the deoxyribose proton couplings^[33] revealed that despite the largely unperturbed structure the sugar pucker and hence the dynamics of the nucleic acids is drastically altered. A more accurate description of the conformational dynamics was achieved using time-averaged restraints (MDtar) where the NMR distance restraints are not required to be fulfilled all the time but only over a specified time frame which permits dynamic behavior to emerge. This resulted in a markedly improved agreement of the obviously dynamic duplexes with the experimental data.^[34] In both DNA•DNA and DNA•RNA the 5'-5' linkage introduces a dynamic hotspot which is evident from both the sugar pucker and backbone behavior (ϵ and ζ torsion angles). The 3'-3' linkage is in contrast much more static and in case of the DNA•RNA duplex shifts the two β -anomeric sugars to a pure S-type conformation (Fig. 5). For the DNA duplex this effect cannot be observed since these residues are already in a 2' endo conformation. Together with the α -anomeric residue this results in an unusual stretch of high S sugars in the DNA strand of the hybrid with a concomitant reduction of the minor groove width. These observations demonstrate that we can locally control the sugar pucker and minor groove topology, which has ramifications for the substrate quality of the hybrid duplex for RNase H.

3. Alpha Anomeric Damage and Recognition

We previously demonstrated that single α -anomeric nucleotides flanked by 3'-3' and 5'-5' linkages fit well into helices, suggesting that 'normal' 3'-5' α -anomeric residues would be destabilizing and could possibly be extrahelical.^[27] This notion was supported by previous work since a single 3'-5' linked α -anomeric thymidine

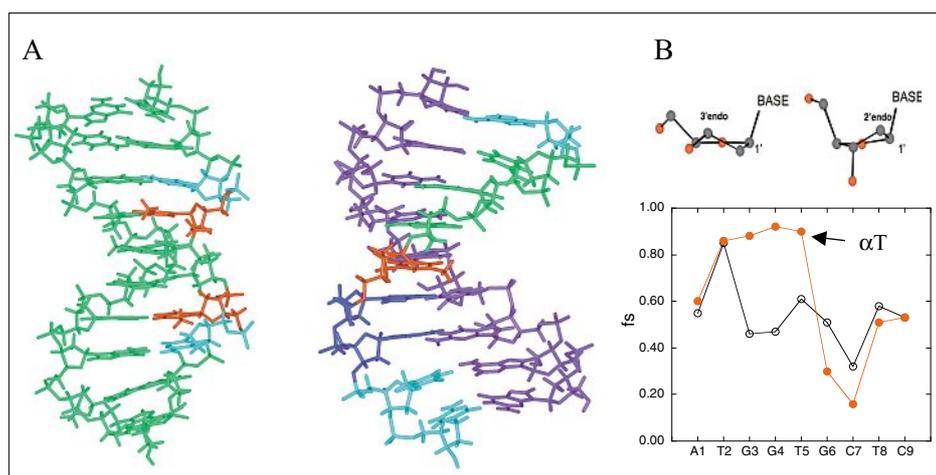


Fig. 5 A) NMR structures of *EcoRI* DNA•DNA (left) and *erbB-2* DNA•RNA (right) calculated using AMBER 4.1; PDB accession numbers 1BX5 and 1C2Q. Color scheme: α T, red; high S- β -deoxyriboses ($fs > 0.67$), green; intermediate N-/S- β -deoxyriboses ($0.33 < fs < 0.67$), cyan; high N- β -deoxyriboses ($fs < 0.33$), blue; RNA strand, purple. Unperturbed torsion angles α , β , γ , ϵ , ζ shift to α^+ , β^+ , γ^+ , ϵ^+ , ζ^+ at the α T residue, where the superscripts +, -, t indicate *gauche* +, - or *trans* angles. B) Top: Deoxyribose structures found in DNA; 3' endo belongs to the N(orthern) conformational range while 2' endo is a member of the S(southern) region. The puckered sugar ring is generally described as an equilibrium between 3' endo (N) and 2' endo (S) conformation. In B-type DNA exhibits $fs > 0.70$ while a low fs or predominant 3' endo conformation is found in A-DNA helices. Bottom: Sugar pucker in the DNA strands of the α T DNA•RNA hybrid (red) and the unmodified control hybrid (black) expressed as fraction of S (2' endo) conformation.

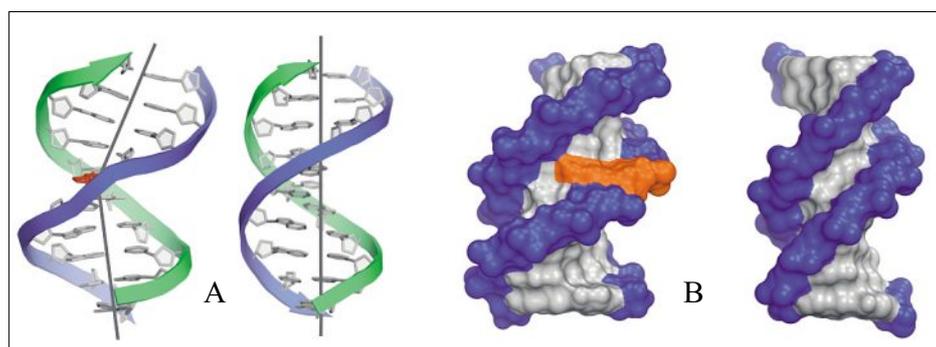


Fig. 6. A) Ribbon representation of the kinked α -A decamer duplex (left) and Arnott B-DNA (right). The sugar-phosphate backbones are shown in blue and green, α -A is shown in red. The helical axes were obtained by CURVES fits of the base-paired segments above and below the lesion. B) View into the minor groove, backbone in indicated in blue. Left: the α -A decamer with the lesion site in red. Right: Arnott B-DNA.

embedded in normal duplex severely destabilized the structure.^[27] In DNA, naturally occurring α -anomeric lesions are produced by inversion of the chirality at C(1') position following the abstraction of the H(1') proton by a hydroxyl radical. α -adenosine (α -A) is the major lesion produced by γ -irradiation under anoxic conditions.^[35] The lesion is mutagenic since it directs the incorporation of dC, dA and dT *in vitro* and generates deletions in an *in vivo* system.^[36] Consequently such damage must be repaired to maintain genetic integrity. The base excision repair (BER) pathway is one of the most prevalent DNA repair pathways and proceeds in two steps. In the damage specific stage, DNA glycosylases detect specific

damaged bases and cleave the N-C(1') glycosidic bond. This results in a toxic apurine/apyrimidine (AP) site which must be rapidly removed by a 'damage general' AP endonuclease such as Endo IV that cleaves the DNA backbone at the AP site and sets the stage for the rest of the repair process.^[37] Alpha-anomeric lesions have a unique repair pathway.^[35,36,38] Remarkably, *E. coli* endonuclease IV (Endo IV) is able to both detect and process α -anomeric damage without requiring a damage specific glycosylase.^[35] Model building of Endo IV revealed a cavity that capable of accepting an α - but not β -anomeric nucleotide.^[39] This activity is not restricted to bacteria, since the human analog of exonuclease III, APE-1, has

also been shown to recognize and process α -anomeric lesions.^[38]

To understand signaling and recruitment of the repair enzyme Endo IV, we determined the structural consequences of a single α -anomeric adenosine in a DNA duplex substrate for Endo IV.^[40] Interestingly, unlike the inclusion of an α -anomeric thymidine, incorporation of an α -A residue revealed only a slightly lower stability compared to the unmodified control. This is further supported by the NMR solution structure of the DNA decamer, which reveals that the lesion is intrahelical and stacked within the duplex in a reverse Watson-Crick fashion. Stacking the α -A within the duplex can be achieved by altering only one backbone torsion angle, ζ . This results in a concomitant increase in the minor groove width (3' downstream of the lesion), a significant increase in roll and decrease in twist for the base-pair steps preceding and following the lesion, respectively. Together these structural features result in an overall 18° kink of the helix, which is supported by numerous local restraints at the site of the lesion (Fig. 6). It is interesting to note that even though the lesion and the perturbations caused by it are local, they change the global appearance of the DNA. This suggests that a local static change can give rise to a structural marker for repair enzyme recognition. In this case, the kinked helical axis and rather enlarged minor groove around the lesion in the α A structure have significant ramifications for its recognition by Endo IV. Both of these features are also present in the crystal structure of an apurinic DNA substrate with Endo IV (1QUM) albeit to a much larger extent.^[39,40]

Our results suggest that the α -A lesion facilitates enzymatic access by causing the initial distortion. In addition, because of the altered stacking pattern, the energetic cost of driving the nucleic acid into its final conformation in the complex with the repair enzyme is reduced. Moreover, the energetics of the distortion is also expected to be modulated by flanking sequences. Enzymatic studies tie the modulation of repair efficiency to changes in the flanking sequence surrounding the α A lesion site. The structural and dynamic properties of these flanking sequence changes are currently under investigation.

4. Mismatch Recognition by MSH2-MSH6 Dynamic Aspects

In contrast to a lesion, a DNA mismatch is not a foreign component and is taken care of by the mismatch repair (MMR) pathway. In bacteria, MutS is responsible for the repair of small insertion/deletion loops and mismatches. This dimeric enzyme recog-

nizes mismatches by probing the DNA substrate with a phenylalanine. Upon binding, phenylalanine intercalates at the mismatch resulting in a kink in the DNA along with conformational changes in the protein to form a clamp around the substrate.^[41] The hetero dimeric enzyme hMSH2-hMSH6 is the human homolog to bacterial MutS. Defective hMSH2-hMSH6 is associated with hereditary nonpolyposis colorectal cancer, underlying its importance.^[42] In addition to a wide range of mismatches, hMSH2-hMSH6 also recognizes single nucleotide small insertion/deletion, loops and damaged nucleotides such as 8-oxo-G and O⁶-methylguanine.^[42,43] Upon recognition of a mismatch, hMSH2-hMSH6 hydrolyses ATP and forms a clamp around the DNA substrate causing recruitment of other enzymes to continue the repair process.^[42–44] The ATPase activity (k_{cat}) is a convenient measure of the recognition/repair efficiency. The wide array of chemically disparate substrates suggests that features other than specific groups serve as the signpost for recognition. In addition, base pairs flanking the mismatch have a significant effect on mismatch recognition and repair.

For mismatches symmetrically flanked by 3' purines (*e.g.* CgA where g stands for the guanine involved in the mismatch) a k_{cat} of 21.7 min⁻¹ and 17.8 min⁻¹ was observed for G/T and G/A substrates.^[44] When the flanking sequence was altered to 3' symmetric pyrimidines (*i.e.* AgC) k_{cat} was reduced to 14.5 min⁻¹ and 11.6 min⁻¹, respectively. This clearly demonstrates that from the perspective of the enzyme it is the flanking sequence that determines the 'reparability' of the mismatch. Thermodynamic properties (ΔH , ΔG , T_m) of mismatched substrates fail to provide a rational for the modulation of the substrate quality. For example, the T_m of G/T mismatched duplexes are nearly identical with AgC 31.3 °C and CgA 31.7 °C while for the G/A mismatch the AgC duplex was significantly less stable than the CgA combination. This prompted us to carry out a concerted investigation of the structure and enzymatic properties of several DNA substrates (Fig. 7). In a regular duplex the base–base stacking results in characteristic close distances between base hydrogens and hence in observable NOE's between bases and sugar protons. All duplexes were found to be base paired including the G/T and G/A mismatches, regardless of the flanking sequence. As anticipated the G/T duplexes behave essentially like a regular duplex and exhibit normal base–base stacking and base-H(1') distances but with local perturbations. For the G/A duplexes strong base–base contacts were only observed for the top strand of the G/A-AgC duplex. All other strands showed several breaks in both base–base as well as base-H(1') connectivity. In ad-

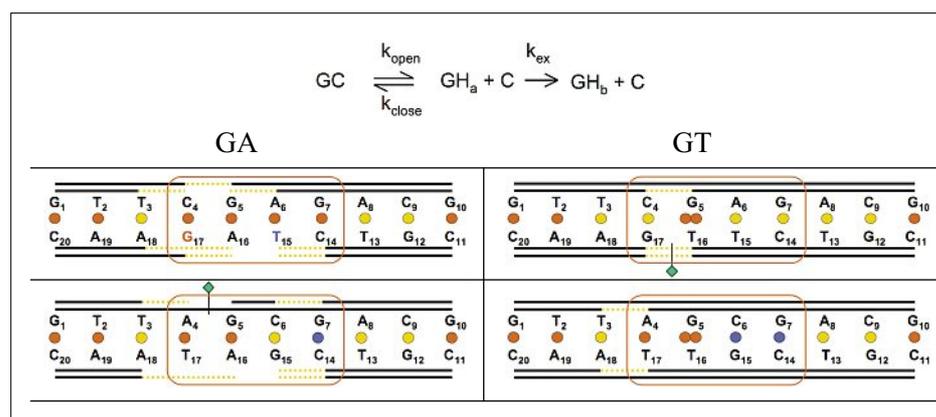


Fig. 7. Top: Kinetic pathways governing imino proton exchange. For an exchange to take place the base pair must assume an open conformation (k_{open}) from which exchange with the solvent (k_{ex}) is possible. [45]. Bottom: DNA substrates. Left: G/A mismatches. Right: G/T mismatches. Base pair life times, ● < 5ms, ○ 6ms – 30ms, ◆ 31ms – 65ms, ◆ ³¹P backbone perturbation, base-base NOE contacts shown by inner lines, and Base-H(1') NOE contacts shown by outer lines. Strong NOE contacts are shown by a solid black line, medium/weak contacts by an orange dashed line, and weak or missing contacts are blank. Opening rates marked by the red box were measured at pH 8.8, 288 K, all others were measured at pH 8.4, 276 K. Multiple sample conditions were required to capture the lifetimes of all base pairs.

dition, DNA backbone perturbations result in unique phosphorus chemical shifts. A single resonance was shifted downfield for one of the G/T and one of the G/A duplexes indicating perturbations in ϵ and ζ (Fig. 7). However this is unlikely to serve as a recognition feature because it occurs in both poorly and well-recognized substrates.^[44] The lack of NOE data signifies that the dynamic properties are not confined to the mismatched pair and that multiple conformations must be considered. Considering the widely different properties of the G/T and G/A duplexes it is not readily evident why the CgA duplexes for G/T and G/A mismatches should be recognized similarly. In order to probe the dynamics of these duplexes in the ms range we focused on base pair opening. Base pair lifetimes are extrapolated to infinite catalyst concentrations when each opening of the base pair results in the exchange of the NMR visible T or G imino protons.

Typical AT base pairs have lifetimes of 2–5 ms, while GC base pairs range from 40–60 ms depending on temperature, sequence and solvent.^[45] As expected, base pairs located near the ends in the duplexes have much shorter life times than internal positions, and base pair lifetimes involved in either mismatch were estimated to be <1 ms. The life times of the base pairs upstream of the mismatch (6–15 and 7–14) are highly dependent on the flanking sequence. When the mismatch site is surrounded by purines in the 3' position (AgC) these base pairs have a much longer lifetime for both G/T and G/A mismatches compared to pyrimidines. These differences are solely determined by the flanking sequences which provides a rational for the sequence dependence of the recognition by hMSH2-

hMSH6. In the MutS complex with a GT mismatch substrate a phenylalanine is inserted on the 3' side of the mismatched G-15.^[41,46] Intercalation of the conserved phenylalanine residue, which is also present in hMSH2-hMSH6, is considered one of the key initiation features of mismatch recognition.^[41,47]

We therefore conclude that base pairs with long life times hamper the probing of the DNA and subsequent insertion of the phenylalanine thereby impeding recognition and consequently also repair. We are currently extending our work using additional substrates to test our hypothesis dynamics of a GG mismatched duplex on an extended time scale to more fully establish the role of local dynamics on the recognition/repair process.

Acknowledgements

We enjoyed collaborations with numerous people. Special thanks to Hans van de Sande, Hans Vogel, Bernd Kalisch, Richard Pon, James Aramini, Anwer Mujeeb, Richard Cunningham, Richard Fishel and Antony Mazurek. Our work was supported by NIH Grants GM55404, AI/GM47459, the Georgia Cancer Coalition and Molecular Basis of Disease and Brain and Behavior Fellowships to A.M.S. and C.N.J

Received: August 14, 2009

- [1] J. D. Watson, F. H. Crick, *Nature* **1953**, *171*, 737.
- [2] A. Rich, D. R. Davies, F. H. Crick, J. D. Watson, *J. Mol. Biol.* **1961**, *3*, 71; D. M. Brown, D. M. Grey, M. H. Patrick, R. L. Ratliff, *Biochem.* **1985**, *24*, 1676; M. H. Sarma, G. Gupta, *FEBS Lett.* **1986**, *205*, 223; M. W. Germann, N. Zhou, J. H. van de Sande, J. H. Vogel, *Methods Enzymol.* **1995**, *261*, 207; K. Rippe, T. M. Jovin, *Methods Enzymol.* **1992**, *211*, 199.

- [3] L. H. Koole, M. H. P. van Genderen, H. M. Buck, *J. Am. Chem. Soc.* **1987**, *109*, 3916; N. T. Thuong, U. Asseline, V. Roig, M. Takasugi, C. Hélène, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 5129; T. Hakoshima, T. Fukui, M. Ikehara, K. I. Tomita, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 7309.
- [4] E. Westhof, M. Sundaralingam, *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 1852.
- [5] N. Pattabiraman, *Biopolymers* **1986**, *25*, 1603.
- [6] J. H. van de Sande, N. B. Ramsing, M. W. Germann, W. Elhorst, B. W. Kalisch, E. V. Kitzing, R. T. Pon, R. C. Clegg, T. M. Jovin, *Science* **1988**, *241*, 551.
- [7] M. W. Germann, J. H. Vogel, R. T. Pon, J. H. van de Sande, *Biochemistry* **1989**, *28*, 6220; N. Zhou, M. W. Germann, J. H. van de Sande, N. Pattabiraman, J. H. Vogel, *Biochemistry* **1993**, *32*, 646.
- [8] M. W. Germann, B. W. Kalisch, J. H. van de Sande, *Biochemistry* **1988**, *27*, 8302.
- [9] M. W. Germann, B. W. Kalisch, R. T. Pon, J. H. van de Sande, *J. Biomol. Struct. Dyn.* **1996**, *13*, 953.
- [10] M. W. Germann, B. W. Kalisch, R. T. Pon, J. H. van de Sande, *Biochemistry* **1990**, *29*, 9426.
- [11] A. K. Shchyolkina, Y. P. Lysov, I. A. Il'ichova, A. A. Chernyi, Y. B. Golova, B. K. Ckernov, B. P. Gottikh, V. L. Florentiev, *FEBS Lett.* **1989**, *244*, 39.
- [12] N. B. Ramsing, T. M. Jovin, *Nucleic Acids Res.* **1988**, *14*, 6659; H. Fritzsche, A. Akhebat, E. Taillandier, K. Rippe, T. M. Jovin, *Nucleic Acid Res.* **1993**, *21*, 5085.
- [13] K. Rippe, N. B. Ramsing, R. Klement, T. M. Jovin, *J. Biomol. Struct. Dyn.* **1990**, *7*, 1199; D. Rentzperis, K. Rippe, T. M. Jovin, L. A. Marky, *J. Am. Chem. Soc.* **1992**, *114*, 5926; N. A. Tchurikov, B. K. Chernov, Y. B. Golova, Y. D. Nechipurenko, *FEBS Lett.* **1989**, *257*, 415; O. F. Borisova, A. K. Shchyolkina, B. K. Ckernov, N. A. Tchurikov, *FEBS Lett.* **1993**, *322*, 304.
- [14] S. B. Noonberg, J. C. François, T. Garestier, C. Hélène, *Nucleic Acids Res.* **1995**, *23*, 1956.
- [15] K. Rippe, V. Fritsch, E. Westhof, T. M. Jovin, *EMBO J.* **1992**, *11*, 3777.
- [16] J. M. Casasnovas, D. Huertas, M. Ortiz-Lombardia, J. Kypr, F. Azorin, *J. Mol. Biol.* **1993**, *233*, 671; N. G. Dolinnaya, A. Ulku, J. R. Fresco, *Nucleic Acids Res.* **1997**, *25*, 1100; N. G. Dolinnaya, J. R. Fresco, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 9242; P. Rajagopalan, J. Feigon, *Biochemistry* **1998**, *28*, 7859.
- [17] H. Robinson, A. H. Wang, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 5224.
- [18] H. Manor, R. B. Sridhara, R. G. Martin, *J. Mol. Evol.* **1988**, *27*, 96.
- [19] M. W. Germann, B. W. Kalisch, J. H. van de Sande, *Biochemistry*, **1998**, *37*, 12962.
- [20] J. Paoletti, D. Bazile, F. Morvan, J. L. Imbach, C. Paoletti, *Nucleic Acids Res.* **1989**, *17*, 2693.
- [21] B. W. Kalisch, M. W. Germann, J. H. van de Sande, *FEBS Lett.* **1998**, *427*, 301.
- [22] B. F. Baker, B. P. Monia, *Biochim. Biophys. Acta* **1999**, *1489*, 3; S. T. Crooke, *Methods Enzymol.* **2000**, 313.
- [23] A. De Mesmaeker, R. Häner, P. Martin, H. E. Moser, *Acc. Chem. Res.* **1995**, *28*, 366; A. Eschenmoser, *Science*, **1999**, *284*, 2118.
- [24] B. P. Monia, E. A. Lesnik, C. Gonzalez, W. F. Lima, D. McGee, C. J. Guinasso, A. M. Kawasaki, P. D. Cook, S. M. Freier, *J. Biol. Chem.* **1993**, *268*, 14514; B. Larrouy, C. Boiziau, B. Sproat, J. J. Toulmé, *Nucleic Acids Res.* **1995**, *23*, 3434; S. Agrawal, Z. Jiang, Q. Zhao, D. Shaw, Q. Cai, A. Roskey, L. Channavajjala, C. Saxinger, R. Zhang, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 2620.
- [25] F. Debart, G. Tosquellas, B. Rayner, J. L. Imbach, *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1041; M. Koga, A. Wilk, M. F. Moore, C. L. Scremin, L. Zhou, S. L. Beaucage, *J. Org. Chem.* **1995**, *60*, 1520.
- [26] C. Boiziau, F. Debart, B. Rayner, J. L. Imbach, J. J. Toulmé, *FEBS Lett.* **1995**, *361*, 41.
- [27] J. M. Aramini, J. H. van de Sande, M. W. Germann, *ACS Symp. Ser.* **1998**, *682*, 92.
- [28] M. W. Germann, J. M. Aramini, B. W. Kalisch, R. T. Pon, J. H. van de Sande, *Nucleosides Nucleotides Nucleic Acids*, **1997**, *16*, 1481; M. W. Germann, J. M. Aramini, B. W. Kalisch, J. H. van de Sande, *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 4.
- [29] T. M. C. Tan, B. W. Kalisch, J. H. van de Sande, R. C. Y. Ting, Y. H. Tan, *Antisense Nucleic Acid Drug Dev.* **1998**, *8*, 95.
- [30] D. J. Slamon, G. M. Clark, S. G. Wong, W. J. Levin, A. Ullrich, W. L. McGuire, *Science* **1987**, *235*, 177; D. J. Slamon, W. Godolphin, L. A. Jones, J. A. Holt, S. G. Wong, D. E. Keith, W. J. Levin, S. G. Stuart, J. Udove, A. Ullrich, M. F. Press, *Science* **1989**, *244*, 707.
- [31] W. Brysch, E. Magal, J. C. Louis, M. Kunst, I. Klinger, R. Schlingensiepen, K. H. Schlingensiepen, *Cancer Gene Ther.* **1994**, *1*, 99; J. P. Vaughn, J. D. Iglehart, S. Demirdji, P. Davis, L. E. Babiss, M. H. Caruthers, J. R. Marks, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8338.
- [32] J. M. Aramini, B. W. Kalisch, R. T. Pon, J. H. van de Sande, M. W. Germann, *Biochemistry* **1996**, *35*, 9355; J. M. Aramini, J. H. van de Sande, M. W. Germann, *Biochemistry* **1997**, *36*, 9715; J. M. Aramini, A. Mujeeb, M. W. Germann, *Nucleic Acids Res.* **1998**, *26*, 5644; J. M. Aramini, M. W. Germann, *Biochemistry* **1999**, *38*, 15448.
- [33] L. J. Rinkel, C. Altona, *J. Biomol. Struct. Dyn.* **1987**, *4*, 621.
- [34] J. M. Aramini, A. Mujeeb, N. B. Ulyanov, M. W. Germann, *J. Biomol. NMR* **2000**, *18*, 287.
- [35] H. Ide, K. Tedzuka, H. Shimizu, Y. Kimura, A. A. Purmal, S. S. Wallace, Kow, *Biochemistry*, **1994**, *33*, 7842.
- [36] H. Shimizu, R. Yagi, Y. Kimura, K. Makino, H. Terato, Y. Ohyama, H. Ide, *Nucleic Acids Res.* **1997**, *25*, 597.
- [37] R. P. Cunningham, *Mutation Research*, **1997**, *383*, 189; P. W. Doetsch, R. P. Cunningham, *Mutation Research*, **1990**, *236*, 173.
- [38] A. A. Ischenko, M. K. Saparbaev, *Nature* **2002**, *415*, 183; A. Ischenko, E. Deprez, A. Maksimenko, J. C. Brochon, P. Tauc, M. Saparbaev, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2564.
- [39] D. Hosfield, Y. Guan, B. J. Hass, R. P. Cunningham, A. J. Tainer, *Cell*, **1999**, *98*, 397.
- [40] J. Aramini, S. Cleaver, R. Pon, R. Cunningham, M. W. Germann, *J. Mol. Biol.* **2004**, *338*, 77.
- [41] M. H. Lamers, A. Perrakis, J. H. Enzlin, H. H. K. Winterwerp, N. D. Wind, T. K. Sixma, *Nature* **2000**, *407*, 711.
- [42] Z. Yu, J. Chen, B. N. Ford, M. E. Brackley, B. W. Glickman, *Environ. Mol. Mutagen.* **1999**, *33*, 3.
- [43] R. D. Kolodner, G. T. Marsischky, *Curr. Opin. Genet. Dev.* **1999**, *9*, 89; A. Mazurek, M. Berardini, R. Fishel, *J. Biol. Chem.* **2002**, *277*, 8260.
- [44] A. Mazurek, C. N. Johnson, M. W. Germann, R. Fishel, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 4177.
- [45] M. Guéron, J. L. Leroy, *Methods Enzymol.* **1995**, *261*, 383; J. G. Moe, I. M. Russu, *Biochemistry* **1992**, *31*, 8421.
- [46] J. H. G. Lebbink, G. Dubravka, G. Natrajan, A. Fish, H. H. K. Winterwerp, T. K. Sixma, N. D. Wind, *EMBO J.* **2006**, *25*, 409.
- [47] F. R. Salsbury, J. E. Clodfelter, M. B. Gentry, T. Hollis, K. D. Scarpinato, *Nucleic Acids Res.* **2006**, *34*, 2173; G. Obmolova, C. Ban, P. Hsieh, W. Yang, *Nature* **2000**, *407*, 703.