J Mol Model (2005) 11: 141–159
DOI 10.1007/s00894-005-0238-9

ORIGINAL PAPER

Tsvetan G. Gantchev · Sylvain Cecchini
Darel J. Hunting

Dynamic conformational states of DNA containing T·T or BrdU·T mispaired bases: wobble H-bond pairing versus cross-strand inter-atomic contacts

Received: 27 August 2004 / Accepted: 1 December 2004 / Published online: 18 February 2005

Abstract The dynamic structure of 11-mer DNA duplexes of different sequences with or without homopyrimidine (T·T, or BrdU·T) mismatches was studied by molecular dynamics (MD) simulations on a time scale from 200 ps to 1 ns. The conformational analysis suggests that in mismatched duplexes the formation of classical T·T wobble H-bonding pairing is nearest-neighbor sequence-dependent and, in most cases, three-centered H-bonds and numerous alternative close cross-strand interatomic contacts exist. Thus, in duplex *W1*, where the central triplet is 5′d(C**T**A)·d(**TT**G), two wobble conformations W$\uparrow$ ($\alpha\beta$) and W$\downarrow$ ($\beta\alpha$) are formed and exchange rapidly at 300 K. In contrast, when the central triplet is 5′d(**TTT**)·d(A**T**A) (*W2* duplex) wobble conformations are rarely observed at 300 K, and the T·T mispair most often adopts a ''twisted'' conformation with one largely persistent normal H-bond, plus a stable cross-strand contact involving a T flanking base. However, at elevated temperature (400 K) the same *W2* duplex shows frequent exchange between the two classical wobble conformations ($\alpha\beta\leftrightarrow\beta\alpha$), as is in the case when the central triplet is 5′d(T**BrdU**T)·d(A**T**A) (*W3* duplex at 300 K). It is suggested that in the *W2* sequence, restrictions due to thymine-methyl/π interactions prevent the formation of wobble pairing and thermal activation energy, and/or the chemical replacement of T by BrdU are required in order for the T(BrdU)·T mismatch to adopt and exchange between wobble conformations. The specific short and/or long-lived (double/triple) cross-strand dynamic interactions in *W1*, *W2* and *W3* duplexes are throughout characterized. These frequent atomic encounters exemplify possible inter-strand charge transfer pathways in the studied DNA molecules.

T. G. Gantchev (✉) · S. Cecchini · D. J. Hunting
Department of Nuclear Medicine and Radiobiology,
Faculty of Medicine, Université de Sherbrooke,
Sherbrooke, QC, J1H 5N4, Canada
E-mail: Tsvetan.Gantchev@USherbrooke.ca
Tel.: +1-819-5645601
Fax: +1-819-5645442

Keywords DNA conformation · molecular dynamics · H-bond · cross-strand interactions

Introduction

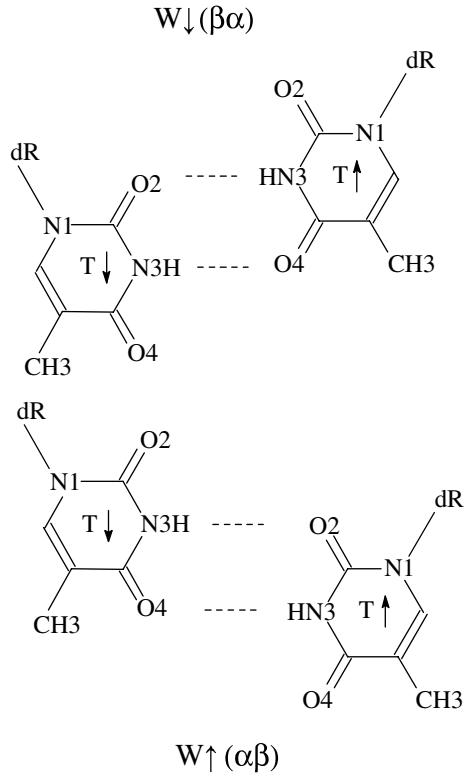
Base-pair mismatches are formed through erroneous incorporation of deoxyribonucleotides by DNA polymerases during replicative or repair synthesis, or via heteroduplex formation after homologous recombination. Some mismatches arise independently of replication, repair and recombination after deamination of 5-methyl cytosine. In the majority of cases, the misincorporated base will be identified and excised by specific enzyme systems. Mismatches that escape the proofreading activity of DNA polymerases are typically recognized and repaired by the post-replicative mismatch repair pathway (Mut-S, L, Y proteins) and the base excision repair pathway (glycosylase enzyme families). The structural features of mismatched DNA that enable the repair enzymes to discriminate between specific lesions are of great interest and have been the subject of numerous X-ray diffraction, [1–3] NMR [4–7] and computational studies. [8, 9] On the other hand, base modification, or insertion of base analogues into DNA, e.g. halogenated thymine analogs such as BrdU, IdU and FdU, is widely used in cancer chemotherapy and radiation therapy (radiosensitization) [10]. In all instances the local structure and internal dynamics of modified duplex DNA raises a number of issues concerning its functionality.

Experimental studies have shown that a single mismatched base-pair inserted into double stranded DNA is usually accommodated in the double helix with little distortion of the global helical conformation [1, 2]. However, mismatched DNA is clearly destabilized relative to normal DNA, and the degree of local perturbation is a function of the type of the mismatch and the sequence context in which it occurs [5]. The

destabilization is due to frequent dynamic reorientations demonstrated by alternative H-bonding contacts within the mismatched base pair, disruption of base stacking, formation of twisted conformations which give rise to bifurcated H-bonds and other short-lived cross-strand contacts involving the bases flanking the mismatch bases. In thermodynamic terms, mismatches lead to differences of free energy of about 1–3 kcal mol^{-1}, and affect the double helix stability by loss of enthalpy (interruption of base-stacking) and gain of entropy (higher degree of reorientation freedom), both manifested macroscopically by lower DNA melting temperatures [11, 12]. Molecular motion in DNA is on the time scale from several ns to μs and can be analyzed by magnetic resonance (EPR and NMR) [13]. However, more rapid internal oscillations are difficult to probe experimentally. Pico to nanosecond-scale molecular dynamics (MD) studies provide valuable information about the structure and dynamic behavior of DNA by complementing experimental approaches [8, 9, 14, 15].

Three principal schemes of hydrogen bonding between non-complementary bases have been proposed: "wobble", rare tautomer and ionized [1, 2, 16]. The solution dynamic structure of homopyridine mismatch T·T has been a subject of several NMR/NOESY and molecular dynamics investigations [17, 18, 19]. It has been shown that under physiological conditions most, but not all, of the studied oligonucleotides incorporate the mismatch by assuming "wobble" conformations, where both thymidines are in the *anti* orientation, forming two complementary hydrogen bonds between the imino N3H and O4 and between N3H and O2 of opposite bases (Scheme1).

Due to the homocharacter of the T·T mismatch, pairing is achieved in two conformations, W↑ and W↓ (denoted here as "canonical wobble", $\alpha\beta$ and $\beta\alpha$ conformations). The W↓ ($\beta\alpha$) conformation can be obtained from W↑ ($\alpha\beta$) by 180° rotation about the pseudidyadic axis (and vice versa). The MD simulations on the 5′d(GCCACTAGCTC)·d(GAGCTTGTGGC) duplex have further shown that the $\alpha\beta \leftrightarrow \beta\alpha$ interchange occurs with a frequency of $\sim 5\times10^{10}$ Hz (picosecond time-scale) [19]. However, the dynamic properties of the DNA segments containing the T·T mismatch are more complex and depend on the sequence context, with the short-range interactions playing the dominant role, i.e. nearest-neighbor induced steric hindrance of base movement [8]. Formation of "twisted" conformations is usually accompanied by the formation of cross-strand interatomic contacts. The pattern of cross-strand H-bonding has been introduced by Yoon et al. [20] when analyzing crystal structures of AT-rich DNA helices. Several NMR observations, which include detection of high field chemical shifts and broadening of aromatic and methyl protons of adjacent bases, are consistent with the existence of more than one conformation in the T.T mismatched site and structural distortion of the bases flanking the mismatch; however, NMR studies failed to identify unequivocally dynamic structures involving cross-strand contacts [17, 18, 21]. The main reason is that cross-strand interactions likely occur in a sub-ns time domain, and the ensuing atypical conformations are too rapid for the NMR time-scale. Conversely, analysis of such dynamic interactions (intramolecular encounter complexes) is of primary importance in processes which involve electron migration in DNA, e.g. during radiation-induced DNA damage.



Sch. 1 The two canonical wobble conformations for the T·T mispair

In this study, we report results from MD simulations performed on DNA duplexes of different sequences with (*W1* and *W2*), or without (*N1* and *N2*) a T.T mismatch at the central position. We also present results of the dynamic structure of the *W3* duplex, derived from *W2* by replacing one of the T bases by BrdU, thus introducing a new mismatch, BrdU.T. Apart from the important implications in radiation-induced DNA damage and radiosensitization [22], the later modification allowed us to reveal, by comparison, the sequence-related effects of steric hindrance on base motion in T-enriched B-DNA tracts.

DNA molecules and simulation details

Molecules

The studied DNA molecules are 11-mer duplexes with GC clamps at their ends. The three normal DNA duplexes are:

5′d(GCCACTAGCTC)·d(GAGCTAGTGGC) (*N1*), 5′d(ACGATTTACGA)·d(TCGTAAATCGT) (*N2*) and 5′d(ACGATBrdUTACGA)·d(TCGTAAATCGT) (*N3*). Although the central base pair triplets of these sequences are significantly different, the terminal GC enrichment is known to help the formation of stable Watson-Crick double helices (B-type) in solution, even when molecules contain a central base-pair mismatch, such as the wobble base pair, T·T [23]. The thymidine–thymidine mismatched sequences were derived from the normal ones (*N1* and *N2*) after mutation of the central **A**17 of the second strand by **T**, thus yielding the wobble pair T6·T17 in the following DNA molecules 5′d(GCCACTAGCTC)·d(GAGCTTGTGGC) (*W1*) and 5′d(ACGATTTACGA)·d(TCGTATATCGT) (*W2*); The DNA containing the BrdU6·T17 mismatch was constructed either by mutation of **A**17 by **T** in *N3*, or by replacement of **T**6 by **BrdU**6 in *W2*, to give the 11-mer duplex 5′d(ACGATBrdUTACGA)·d(TCGTATATCGT) (*W3*). The solution structure of sequence (*W1*) has been previously analyzed by NMR and MD [19]. In the present study, it was selected both as a reference structure for the validation of our molecular dynamics approach as well as to underscore nearest-neighbor sequence-dependent conformational differences within one and the same mismatched (T·T) base pair. Sequences *N2*, *N3*, *W2* and *W3* are part of a truncated 25-mer oligonucleotides presently used by us in radiosensitization experiments to study the formation of single and double strand DNA breaks [22].

Initial model building and molecular mechanics

All structure simulations, molecular mechanics and dynamics calculations were performed using Advanced Computation, Biopolymer and Dynamics modules as implemented in the Sybyl (Version 6.7 and 6.8) molecular modeling package (Tripos Inc.) running on Octane SGI platform (250 MHz, MIPS R10000 IP30). The two normal DNA 11-mer sequences *N1* and *N2* were built either using the average B-DNA parameters [24], or using the backbone coordinates of the reported solution structure of an 11-mer oligonucleotide containing a central C·T mismatch (PDB entry 1 fky [25]), followed by base replacement to achieve the desired sequences. The later approach was preferred in cases of mismatched structures, and therefore, all MD results reported here were obtained using molecules initially built by this technique. Atoms were assigned AMBER force field all atom types and charges. The BrdU 3D-structure with Mulliken electron population analysis was derived by semi-empirical AM1-Hamiltonian computations and before adding it to the DNA AMBER-95 Sybyl dictionary, charges were scaled to match Kollman force field all atom charges on deoxyribose and phosphate backbone. The latter was achieved by normalizing the AM1 partial charges to C1′ charge (Kollman atomic charge of 0.068). It is interesting that this procedure gives equal net charge (0.004) on the CH3-group (T) and the Br atom (BrdU). The most important atomic charge difference is on C5 (0.003 vs. −0.033, in T and BrdU, respectively; complete data are available as supplementary material). Molecules were submitted to molecular mechanics unconstrained minimization using the AMBER-95 force field, 8-Å cutoff distance, steepest descent (Powell energy gradient) for 20 iterations, followed by a conjugated gradient minimization. Solvation effects were mimicked using a linear distance-dependent dielectric constant, $D_{ij} = Cd_{ij}$ ($C = 4$) [26]. This simplification is justified by many previous studies and demonstrated excellent conformity with the NMR data. [5, 6, 19, 25] Moreover, an earlier study [8] did not indicate any differences between wobble dynamic states during MD in vacuum and in the presence of explicit water molecules. It is to be noted that with $C = 4$ we observed a slight widening of the DNA major groove and deepening of the minor groove, as compared to structures obtained when C was set to 1 or 2. All initially built structures were of B-form DNA, with total energy ~ -500 kcal mol^{-1}, bases in predominantly *anti* conformation and deoxyribose in the C2′ *endo* conformation. Typical 11-mer DNA structures, before and after 200 ps MD are shown in Fig.1.

Molecular dynamics

The approach applied by us is similar to the NMR DNA structure assignment by implementation of inter-proton distance restraints coupled with a restricted molecular dynamics protocol. Following the molecular mechanics structure refinement, different constraints were defined similarly to those deduced from NOESY NMR spectra for mismatched base pair DNA molecules extensively studied by Fazakerley and co-workers. [5, 19, 25, 27, 28]. These include 44 H8(H6)-H1′/H2″ intra-proton (including the mismatched pair), and 33 H8(H6)-H1′/H2″ inter-proton distances (excluding the mismatched pair and nearest neighbors). Equilibrium distances were set equal to those of the optimized structures with an initial force constant, $k = 5$ kcal mol^{-1} Å^2. This constant was gradually reduced to zero in separate sets of subsequent molecular dynamics runs. The Watson–Crick H-bonding distances of the two edge base pairs were defined as range constraints, 1.8–2.8 Å. To avoid ‘‘frying’’ when molecules were heated, the four end bases were also set as aggregates. A preferential C2′-endo deoxyribose conformation has been indicated by NMR [19], and therefore torsional angles, δ, were also weakly restrained to 140° ($k = 2$ kcal mol^{-1} rad^2). However, this constraint was not crucial and could be omitted. No additional angle or torsion constraints were applied. To enlarge the explored conformational space, the restrained energy minimized DNA structures were first subjected to 10 cycles of Simulated Annealing (SA): 1 ps step heating phase up to 500 K and an exponential 5 ps annealing phase down to 50 K. The ten end-cycle po-

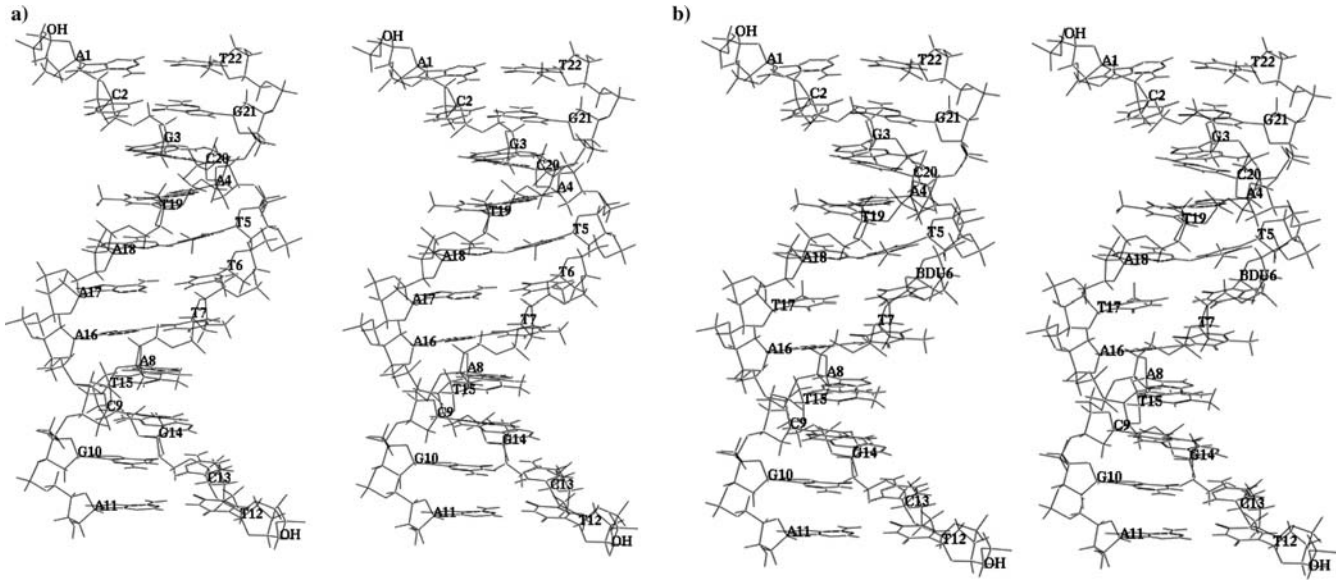


Fig. 1 Stereo views of 11-mer DNA structures: *N2* duplex after simulated annealing and minimization (**a**); *W3* duplex averaged over a 200 ps MD run (**b**)

tential energy-minimum molecules were selected for further molecular mechanics refinement as above (constraint force constant, $k = 0$ kcal mol^{-1} Å^2). In the case of mismatch base-pair molecules, special attention was given to the generation of conformers with alternative wobble ($\alpha\beta$ and $\beta\alpha$) base-pairing. In each case, at least two different canonical wobble ($\alpha\beta$; $\beta\alpha$) conformers of lowest potential energy were selected for MD runs. Previous NMR-resticted MD studies [5, 19, 25, 27, 28] have shown that molecular dynamics simulations for 100–250 ps usually adequately represent the dynamic behavior of mismatched DNA duplexes. Therefore, and for conformity with these studies, we first performed 200 ps MD simulations. In special cases, e.g. when wobble conformations were rarely observed, as for the *W2* duplex at 300 K, MD runs were extended up to 1 ns.

Preparation for molecular dynamics

Preparation of DNA molecules for MD simulations was performed following the published protocol [26–28] with some modifications. The preparation procedure comprised four phases: heating, randomization, equilibration and relaxation (if required). Random thermal velocities were assigned corresponding to a Gaussian distribution with mean temperature at 50 K (thermal coupling, 50 fs) for a period of 1 ps, followed by heating the molecule from 50 to 300 K (or higher) in steps of 50 K, 2 ps each (coupling, 100 fs). The next stage, the randomization phase (total prolongation 5 ps, coupling 2 fs), was performed in which the heated molecule was repeatedly reassigned random atomic velocities during 25 time intervals (200 fs each). Further equilibrium took place over the following 4 ps (coupling, 200 fs). In a separate set of simulation experiments, molecules were subjected to relaxation (minimum length of 10 ps), during which the initially imposed weak harmonic constraints of 5 kcal mol^{-1} Å^2 were gradually reduced in steps of 1 kcal mol^{-1} down to 0 kcal mol^{-1}. The MD runs with constrained and unconstrained DNA molecules showed negligible conformational differences (mainly more frequent oscillations in the later case). For conformity, the presented MD results with different DNA molecules are from the MD runs where all force constants of the distance constraints were reduced to 2 kcal mol^{-1} Å^2. In trial runs, bond lengths and angles between heavy atoms and hydrogens were optimized using the SHAKE algorithm. This optimization, however, was omitted in the final MD simulations. When simulations were performed at 300 K, the average temperature during the 200 ps period was in the range of 298 ± 6 K.

Table 1 200 ps MD of normal DNA duplex (*N2*): typical hydrogen bonding and C1′–C1′ distances

Atom pair	Mean distance (Å)	SD (± Å)	High (Å)	Low (Å)	Sweep (Å)
T6N3H-N1A17	2.1	0.15	2.7	1.7	1.0
T6O4-H6N6A17	2.2	0.28	3.4	1.6	1.8
C9O2-H22N2G14	2.0	0.14	2.7	1.6	1.1
C9N3-H1N1G14	2.1	0.11	2.6	1.8	0.8
C9N4H41-O6G14	2.1	0.18	2.9	1.7	1.2
T6C1′-C1′A17	11.0	0.25	11.7	10.2	1.5
C9C1′-C1′G14	10.9	0.18	11.4	10.3	1.2

Analysis of molecular dynamics results

The outputs from MD simulations were analyzed using Sybyl Dynamics Table/Spreadsheet and Graphics tools. These readily allow convenient visual presentation of trajectory plots of atomic movement and population

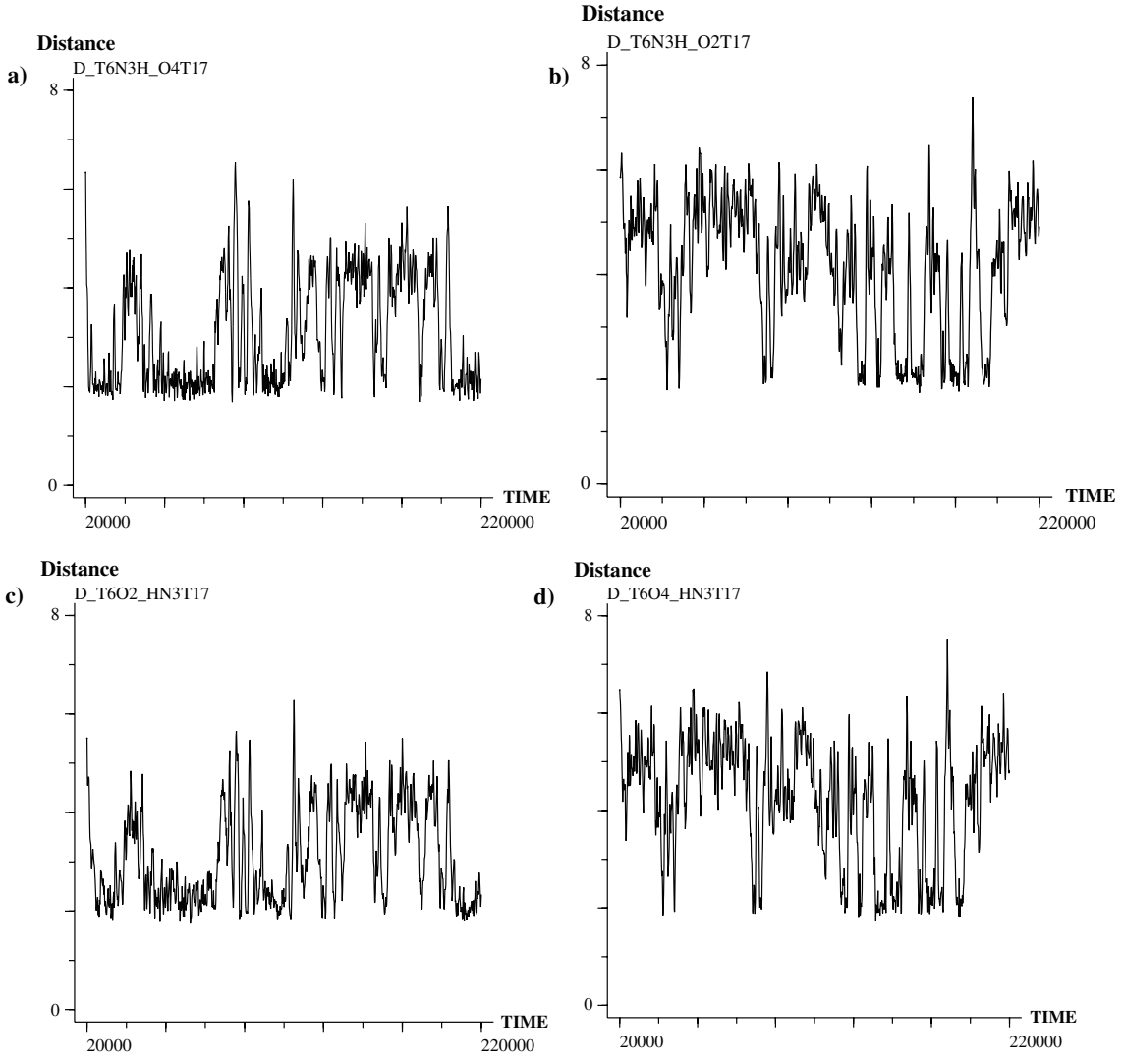


Fig. 2 *W1* duplex; 200 ps MD trajectory plots (*fms*) of H-bond donor-acceptor distances (Å) within T6·T17 mismatch: T6N3H-O4T17 (**a**), T6N3H-O2T17 (**b**), T6O2-HN3T17 (**c**) and T6O4-HN3T17 (**d**). Trajectory plots of glycosidic bond/C1′-C1′ vector angle, λ at T17 (**e**) and inter-atomic distance C5NH41-O4T17 (**f**). Population histograms over 200 ps MD for the inter-atomic distances: T6NH3-O4T17 (**g**) and T6NH3-O2T17 (**h**)

density histograms, complete topological analysis and 3D-presentation of every conformation taken at 20 fs snapshot intervals. In a typical 200 ps dynamics run, 10,000 or more conformers (Dynamics Table rows) were generated. Sybyl tools further assist selections (and cross-section selection) of conformers (table rows) which fulfil given geometrical/conformational criteria, such as range of inter-atomic distances, angles, torsions, centre of mass, etc. After extraction from the original Dynamics table, these conformers can be analyzed separately, an option most often applied in this study. However, in the Sybyl environment local and global DNA parameters are not pre-defined (e.g. as in ''Curves'' algorithm [29]), and an automatic analysis of DNA curvature is not available. Therefore, and in view of our interest focused on structure-related pathways of inter-base charge transfer, the conformational analysis presented here is limited to description of dynamic DNA conformers (e.g. within the central base-pair triplet) defined by the formation of close normal and cross-strand inter-atomic contacts (electrostatic or H-bonding). The presence of canonical Watson-Crick, and/or W$\uparrow$ and W$\downarrow$ wobble H-bonding (denoted here as $\alpha\beta$ or $\beta\alpha$ dual inter-atomic contacts) was the primary criteria for description of conformer dynamic population, followed by the presence of untypical cross-strand close inter-atomic contacts (usually within a distance range of 1.8–2.9 Å). Depending on the lifetime tracked in trajectory plots, conformers can be divided in two main groups: long-lived (≥5–10 ps), and short-lived (<1 ps). Molecular conformations in mismatches are labelled as follows: the canonical wobble (two H-bonds) by two Greek symbols, e.g. $\alpha\beta$ (W$\uparrow$); the conformations with only one H-bond as α' or β', when the N3H group of the strand-I mismatched base (T6, or BrdU6) is in contact with O4 or O2, respectively of the partner (T17) base of the strand-II; similarly, α'' and β'' denote the ''half-

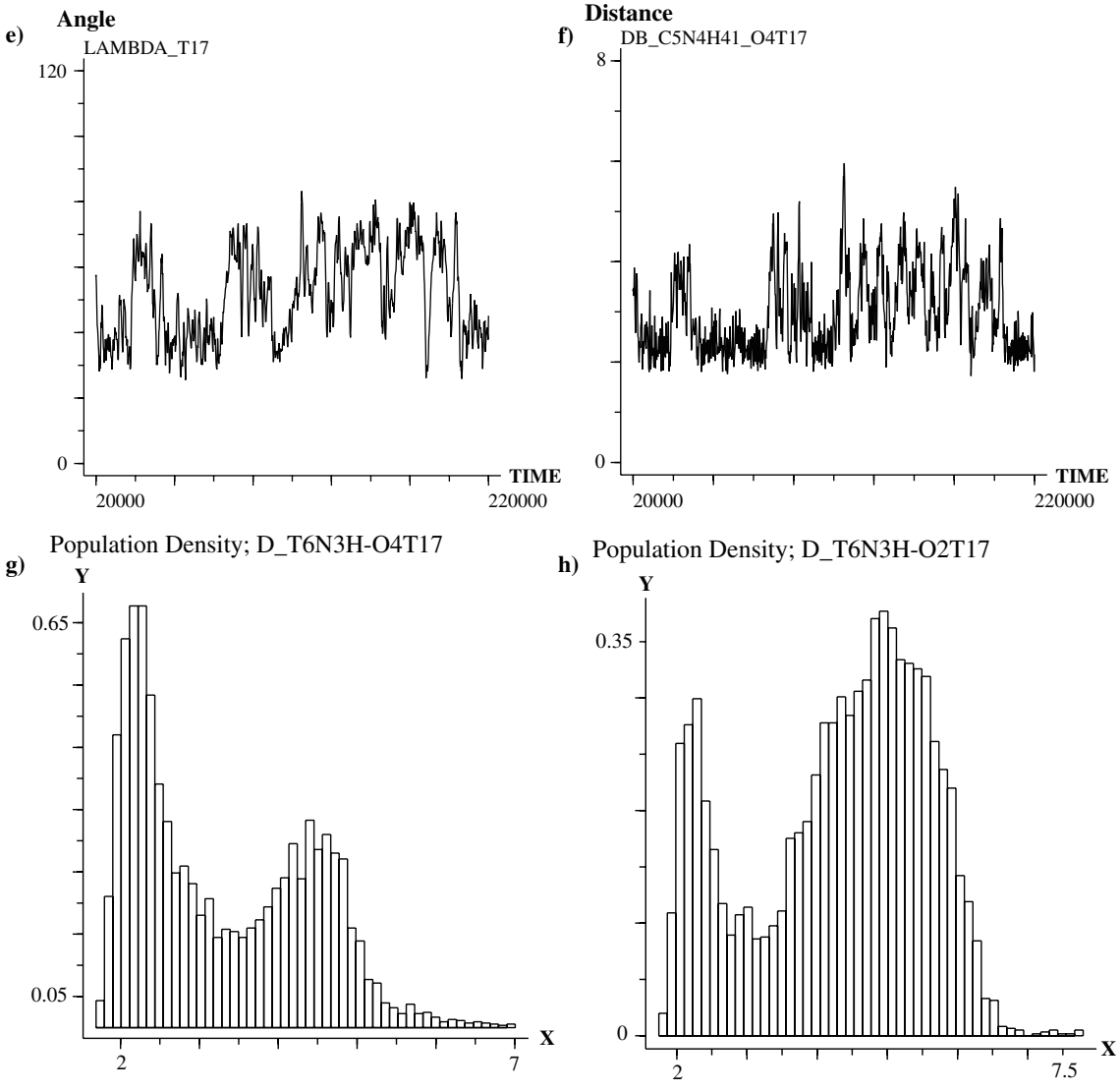


Fig. 2 (Contd.)

wobble" conformation when the N3H group of the mismatched base on the second strand (T17) is H-bonded to O4, or O2 atoms of strand-I base. Other single-letter or double-letter Greek symbols denote conformers (usually short-lived) when single, or double cross-strand (non-canonical) contacts exist between potential H-bonding partners, or more generally, between atoms that undergo frequent dynamic encounters and show strong Coulombic interactions. In the Results section, we describe distinctive structural features that characterize classical wobble and unusual base pairing in the studied mismatched DNA oligonucleotides. The results are presented graphically as 2D structures, trajectory plots and population histograms. For clarity, conformer properties of individual molecules are summarized in separate tables, which are abatements from the corresponding multi-column/row Dynamics Tables, generated in Sybyl.

Results

Normal (A.T) and (A.BrdU) duplexes at 300 K

The MD simulations applied to the three normal duplexes: 5′d(GCCAC**T**AGCTC)·d(GAGCT**A**GTGGC) (*N1*), 5′d(ACGAT**T**TACGA)·d(TCGTA**A**ATCGT) (*N2*) 5′d(ACGA T**BrdU**TACGA)·d(TCGTA**A**ATCGT) (*N3*) at 300 K following the protocol described above always resulted in B-type DNA conformation and sugar pucker pseudorotation, P steady within the 145–152° [30]. The only noted deviation from a classical B-DNA model was the narrowing of the minor groove within the central AT enriched domain (groove opening as measured by phosphate–phosphate distance). The minor groove narrowing was even more pronounced in *N2* and *N3* nucleotides and mismatched sequences, where the shortest 5′ $P(n)$–3′ $P(n-4)$ distances across the groove reached values as low as 8.2 Å (B-DNA average is 11.6 Å). The above is known

as one of the most fundamental properties of AT-enriched regions of B-DNA, with a potential for the formation of cross-chain (bifurcated) A-T hydrogen bonds [20]. With the studied normal oligonucleotides, however, bifurcated H-bonding, accompanied by high propeller twist, was scarcely observed. In general, canonical Watson–Crick hydrogen bonding was apparent in >99% of the analyzed 20 fs snapshot intervals (~10,000 snapshots). Table1 lists H-bonding and C1′–C1′ interatomic distances in selected AT and CG pairs, averaged over entire 200 ps MD runs. The lowest and highest values and the difference between them (Sweep) represent the extent of fluctuations and are given for a comparison with the dynamic fluctuations in the mismatched sequences (below). Not surprisingly, the interatomic fluctuations in AT pairs are somewhat higher than in CG pairs. The MD results were essentially the same with, or without applying distance and torsion constraints and the only noticeable differences were the slightly larger fluctuation amplitudes observed in the plotted trajectories in the later case (data available as supplementary material).

Dynamic conformations of the T·T mispair (*W1* duplex, at 300 K)

Distance restrained and unrestrained MD calculations with the 5′d(GCCAC**T**AGCTC)· d(GAGCT**T**GTGGC) oligonucleotide (*W1*) were performed with energy-refined structures (after Simulated Annealing), starting from either of the two energy minimized wobble conformations involving the T·T mispair W↑ ($\alpha\beta$) or W↓ ($\beta\alpha$), Scheme 1. An average duplex structure from 200 ps MD is presented in Fig.1. During every 200 ps MD run an exchange between $\alpha\beta$ and $\beta\alpha$ conformations (Fig.2) was observed which is clearly seen in the trajectory plots of the corresponding H-bonding distances (Fig.2a,c vs. b,d). This H-bonding (conformational) exchange was previously shown to correspond to the change of only one global DNA duplex parameter, shear [19]. The same study has also shown that only one angular parameter specifically represents the two canonical wobble conformations and was defined as an angle, λ formed between the C1′–C1′ vector and the glycosidic bond C1′–N1 (in T·T), Fig.2e. In accordance, we did not observe any other concerted topographical rearrangements, e.g. involving bond, and/or torsion angles. The sugar pucker was constant (average value of ca. 159°), and sugar moiety was invariably in C2′ *endo* conformation and the base stacking is mostly preserved in canonical wobble H-bonding. Conformations $\alpha\beta$ and $\beta\alpha$ when formed are relatively long-lived (>10 ps), however the central region containing the mispaired bases also adopts long-lived and short-lived (<1 ps) intermediate conformations, often lacking T·T double H-bonding. Using Sybyl-Table tools we analyzed the families of recurrently adopted conformations within the central 5′d(C5**T6**A7)·d(T16**T17**G18) DNA triplet (Table2). The most frequently formed conformation is the one for which the T6N3H-O4T17 distance is in the H-bonding range of 1.8–2.9 Å (50% of total, labeled as α'). When cross-selected with T6O2-HN3T17 = 1.8–2.9 Å, thus extracting conformers in the $\alpha\beta$-wobble conformation, the resultant population covers 47% of the total conformational space. The β' conformation, identified by T6N3H-O2T17 = 1.8–2.9 Å represents 20%, while the complimentary wobble double H-bonding, $\beta\alpha$ fills 18% of the total conformational space. The T6C1′-C1′T17 distance in the mismatch (8.8–10.3 Å) is notably shorter than in normal B-DNA (~ 10.9 Å), and is the shortest (8.8 Å) in $\beta\alpha$ conformation. In the $\alpha\beta$ conformation the

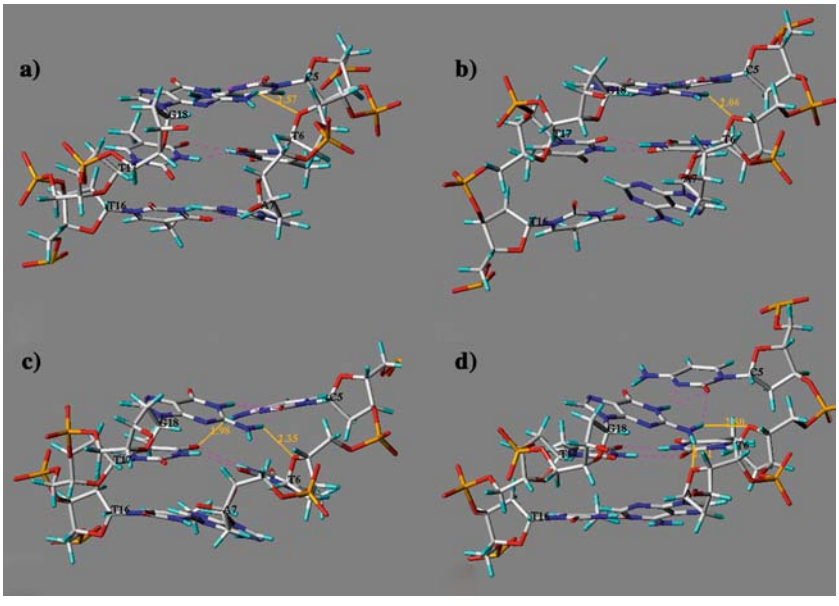


Fig. 3 *W1* duplex central triplet 3D structure (*snapshots*) of the most often encountered conformers during 200 ps MD: $\alpha\beta + \sigma'$ (**a**); $\beta\alpha + \sigma'$ (**b**); $\alpha\beta + \sigma' + \omega'$ (**c**); $\alpha\beta + \sigma' + \tau'$ (**d**). *Dashed magenta lines* indicate H-bonds. *Solid orange lines* represent close cross-strand contacts (Å)

Table 2 DNA duplex *W1* at 300 K. Extractions from dynamics table: selected distance range (1.8–2.9 Å) and cross-selected ranges (*x*) represent wobble H-bonding and cross-strand inter-atomic contacts (conformers)

Selection	None	T6N3H O4T17	T6N3H-O4T17×T6O2-HN3T17	T6N3H-O2T17	T6N3H-O2T17×T6O4-HN3T17	T6O4′-H2N2G18	A7O4′-H1N2G18	T6N3H-O4T16	C5N4H4-O4T17
Conf.	Aver. 200 ps	α'	$\alpha\beta$	β'	$\beta\alpha$	σ'	τ'	μ'	ω'
Population (% of total)	100	50	47	20	18	36	7	2	23
Average inter-atomic distances, Å									
T6N3H-O4T17	3.3	2.3	2.3	4.6	4.6	3.2	3.4	3.3	2.5
T6N3H-O2T17	4.3	5.1	5.1	2.3	2.3	4.4	4.8	4.8	5.2
T6O2-HN3T17	3.3	2.4	2.4	4.5	4.5	3.1	3.4	2.9	3.3
T6O4-HN3T17	4.4	5.2	5.2	2.4	2.4	4.6	4.9	5	4.4
C5N4H1-O4T17	3.2	2.7	2.7	4.2	4.2	3.1	3.2	2.8	2.3
T6N3H-O4T16	3.8	3.7	3.6	4.2	4.2	3.6	3.8	2.4	3.5
T6O4′-H2N2G18	2.7	2.6	2.6	2.7	2.7	2.3	2.7	2.5	2.6
A7O4′-H1N2G18	3.6	3.6	3.6	3.9	4	3.4	2.3	3.4	3.6
C5C1′-C1′G18	10.6	10.6	10.6	10.6	10.6	10.6	10.7	10.6	9.3
T6C1′-C1′T17	9.2	9.3	9.2	8.8	8.8	9.1	9.6	9.2	10.3
A7C1′-C1′T18	10.4	10.3	10.3	10.5	10.5	10.4	10.5	10.6	9.3

flanking the mismatch base pairs retain normal Watson–Crick H-bonding, C5.G18 (three bonds) and A7.T16 (two bonds), but often the amino hydrogen in C5 is involved in a bifurcated (cross-strand) H-bonding with the O4 of T17 (Fig.3). As shown in Table2 (see also Scheme2, Discussion), additional close inter-atomic contacts are also frequent (labeled: σ', τ', μ' and ω'), implying cross-strand atom encounters, usually accompanied by base stacking disruption, especially on the 5′(T16**T**17G18) side. Snapshots of 3D structures, representative of the most important conformers are shown in Fig.3. The most prominent cross-strand contacts are T6O4′-H22N2G18 (σ', average distance of 2.7 Å over 200 ps) and C5N4H41-O4T17 (ω', average distance of 2.5 Å). Occasionally, infrequent close encounter contacts between atoms; A7O4′-H21N2G18 (τ') and T6N3H-O4T16 (μ') are also formed. The average proximity of the T6O4′- H22N2G18 atomic pair is a persistent property of the central *W1* triplet containing the mismatch. The O4′ (T6) and H22N2 (G18) atoms are in a close range in a large number of conformers (36% of the total 200 ps assemble). In most cases, the α' conformation (strong T6N3H-O4T17 H-bond) is preserved, while in general the σ' conformation is concomitant with canonical $\alpha\beta$ and $\beta\alpha$ conformations, albeit at weaker H-bonding. Thus, when the T6O4′-H22N2G18 distance is small (2.1–2.4 Å), the mismatched T6·T17 pair adopts a conformation resembling $\alpha\beta$ conformation but all H-bonding partners are at least $\sim$ 3.2 Å apart and base stacking is disrupted. In an extreme situation when the distance T6O4′-H22N2G18 $\leq$ 2.1 Å, the T6 and T17 bases are pushed away from the helix axis towards the major groove and canonical wobble conformations are abolished. In conformation ω' (close cross-strand contact between atoms C5N4H41-O4T17), the average interatomic distance is 2.5 Å, which implies possible H-bonding. The 200 ps trajectory of this distance follows closely the ones of α' and β'' (Fig.2f). The cross-section of ω' with α' (overlapping states and bifurcated H-bonding) is high, $\alpha'\times\omega' = 71\%$, for both distances ranging from 1.8 to 2.9 Å, but is 0%, in the range of 1.8–2.1 Å. Further analysis of trajectory plots shows that there is a rapid exchange in the extremes between α' and ω' conformations (fluctuations of $\sim$1.2 Å sweep), i.e. when C5N4H41-O4T17 $\approx$ 2 Å, T6N3H-O4T17 $\approx$ 3.2 Å, and vice verso. The ω' and $\beta\alpha$ conformations, however are mutually exclusive, i.e. $(\beta\alpha)\times\omega' \sim 0$.

Dynamic conformations of the T·T mispair (*W2* duplex, 300 K)

The duplex oligonucleotide (*W2*) of the sequence 5′d(ACGAT**T**TACGA)·d(TCGTA**T**ATCGT) was first studied by performing MD simulations at 300 K. Under these conditions, as demonstrated by the 200 ps trajectory diagrams (Fig.4) and the selected interatomic distances (Table3), the mismatched T6.T17 base pair forms one steady H-bond (T6N3H-O4T17, referred to as α' conformation) and seldom adopts the W$\uparrow$ ($\alpha\beta$) conformation. During the 200 ps MD run the $\alpha\beta$ (W$\uparrow$) conformation is assumed only once (population density = 1.4% from total), while the alternative wobble H-bonding, β' and/or $\beta\alpha$ (W$\downarrow$) conformation are totally missing. To verify that the 200 ps dynamics simulation adequately represents the behavior of the *W2* duplex, and does not adopt any unusual conformations at longer time intervals, we extended the 200 ps dynamics run up to 1 ns. No new conformational states, different from those observed during the 200 ps run were formed. Examples of H-bonding distance trajectories of T6N3H-O4T17 (α') and T6O2-HN3T17 (β'') during the 1 ns dynamics simulation are presented in Fig.4e and f. The 1 ns run also confirmed that the $\alpha\beta$ (W$\uparrow$) conformation is rarely formed: on average once per 200 ps, the population density is $\sim$ 1.5% of the total, and the

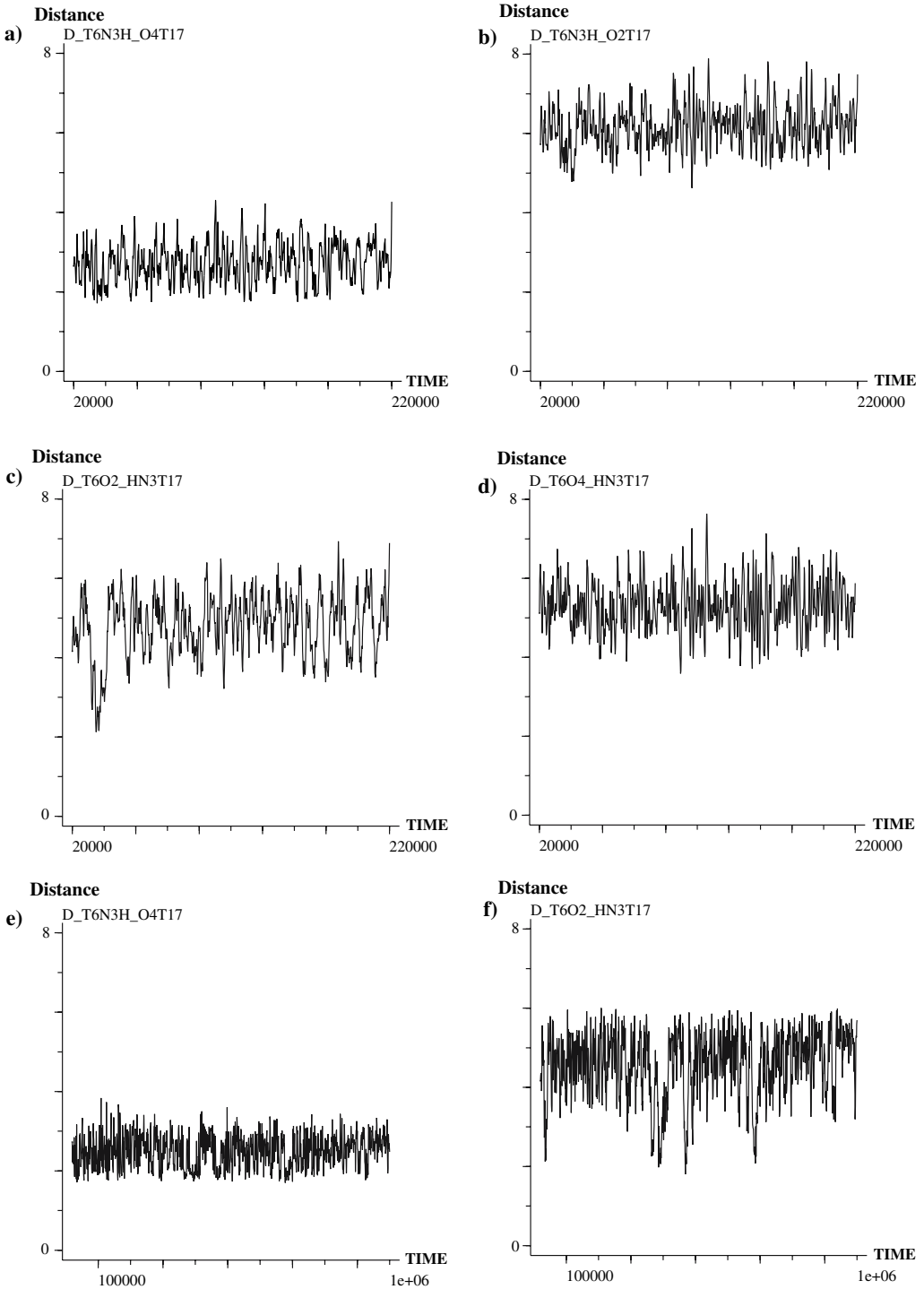


Fig. 4 *W2* duplex at 300 K: 200 ps MD trajectories (*fms*) of inter-atomic distances (Å): T6N3H-O4T17 (**a**), T6N3H-O2T17 (**b**), T6O2-HN3T17 (**c**) and T6O4-HN3T17 (**d**). 1 ns MD trajectories of H-bonding distances: T6N3H-O4T17 (**e**) and T6O2-HN3T17 (**f**); the "canonical" $\alpha\beta$ ($W\uparrow$) wobble conformation exists when in addition to the short T6N3H-O4T17 (α') distance, the conjugated T6O2-HN3T17 (β'') distance falls within the range of 1.8–2.9 Å (once during 200 ps MD, and five times during 1 ns MD). However, no $\beta\alpha$ ($W\downarrow$) wobble conformation is formed, i.e. the distances T6N3H-O2T17 (β') and T6O4-HN3T17 (α'') are always long (shown only for the 200 ps MD run)

lifetime of this wobble conformational state ranges from 5.3 to 8.4 ps. Both during the 200 ps and the 1 ns MD simulations, we observed a number of peculiar orientations (conformers) involving all six nucleotides in the central triplet which are manifested by the formation of bifurcated H-bonds and close cross-strand contacts. Table3, which presents selected distance cross-sections from the 200 ps MD table lists the most important ones. The most frequently observed (96%)

cross-strand interatomic contact is between O2(T7) and HN3 (T17) atoms, often as low as 2 Å (Table3). The later implies not only a possible H-bonding between these atoms, but also a strong van der Waals overlap. This contact (conformers denoted as γ') is persistent and does not exist only when the central triplet adopts briefly the canonical wobble $\alpha\beta$ conformation (Fig.5b). In the ''twisted'', γ' conformation the T6N3H-O4T17 H-bond (α') is always preserved (long-lived). Another group of frequently observed conformers (δ'), characterized by a close T6N3H-N1A18 cross-strand contact, occupy 42% of the total conformational space (Fig.5c). A group of ρ' conformers showing a close inter-atomic contact, T6O4-H62N6A16 which represents ~30% of the total conformational space is also present, in addition to two minor sets of conformers, labeled as η', and ϵ' (Table3). The ϵ' conformation co-exists with the $\alpha\beta$ wobble state (Fig.5). The C1′–C1′ distances within the central triplet of *W2* remained relatively high (as compared with the *W1*- oligonucleotide), as well as there were no cross-strand contacts involving sugar backbone atoms (e.g. with O4′ as in *W1*). Interestingly, when most of the described cross-strand interactions are present, the H-bonding in the 3′-flanking the mismatch, T7·A16 pair is preserved up to 93% (T7O4-H66N6A16) and 100% (T7N3H-N1A18) during the entire 200 ps MD run. In contrast, the H-bonding in the 5′-flanking T5·A18 base pair is more frequently disrupted by cross-strand and bifurcated H-bonding. The T5N3H-N1A18 H-bond is present only in 68% of total conformers, and the Watson–Crick double H-bonding occurs in 66% of the total. However, the complimentary H-bond between T5O4-H62N6A18 is mostly preserved (98%). The more frequent disruption of the T5N3H-N1A18 H-bonding is accompanied by high propeller twist and/or buckle (measured as a higher than average base plane angle).

The most frequently encountered conformers γ', δ', ρ' and η', show interesting interrelationships. The crossed-conformational space $\gamma'\times\delta'$ (two cross-strand close contacts) is present in 40% of total conformations and δ' can be assumed to be sub-conformation of γ' (both γ' and δ' are exclusive in respect to the $\alpha\beta$ conformation). In the $\gamma\delta$ conformation (Table3), the T6C1′-C1′T17 distances are larger than those in the canonical wobble conformation and at least one thymine base (usually T6) is largely exposed in the major groove. Within conformers, ρ' (T6O4-H62N6A16 distance between 1.8 and 2.9 Å) exists a group of sub-conformers when this contact is co-present with the $\alpha\beta$ conformation, and thus it does not preclude the formation of the later. When the above distance is within H-bonding range (1.8–2.9 Å) the majority of conformers also show a short T7O2-HN3T17 distance (γ'), and the cross-section $\rho'\times\gamma' = 90\%$. A population of ρ' conformers is also present, when both, γ' and δ' cross-strand contacts exist simultaneously, i.e. $\rho'\times(\gamma'\times\delta') = 20\%$ of the ρ' conformational space. The conformers, η' (T7N3H-O4T17 = 1.8–2.9 Å) are present in about 8% of the total and it appears they constitute a complete set of γ' sub-conformation (all η' are also γ', as $\eta'\times\gamma' = 8\%$), and partially overlap with δ' (i.e. not all η' conformers are in δ', since $\eta'\times\delta' = 5\%$).

Dynamic conformations of the T·T mispair (*W2* duplex, 400 K)

To gain further insight in the dynamic properties of the T·T mismatch inserted in the central T-enriched 5′d(ACGAT**T**TACGA)·d(TCGTA**T**ATCGT) DNA

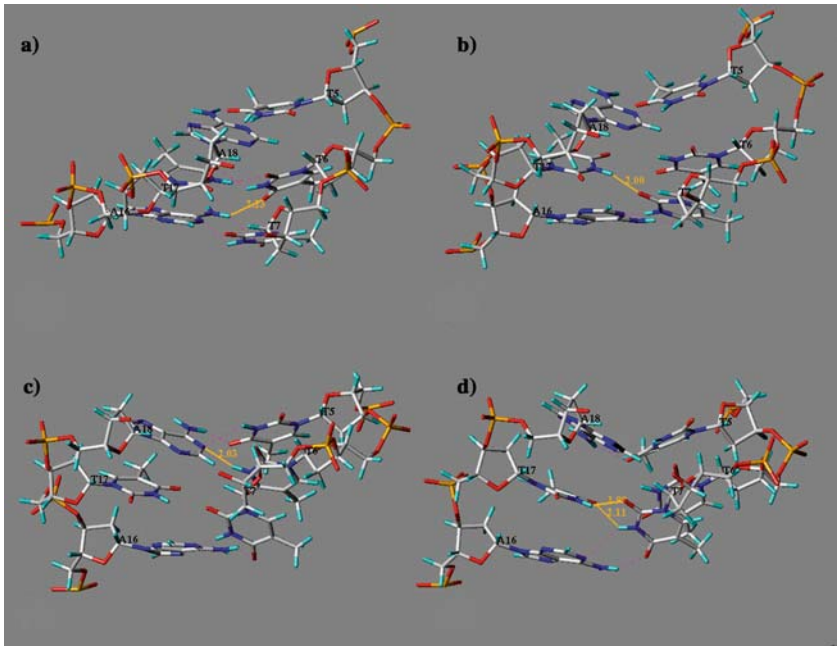


Fig. 5 3D structure of conformers observed within the *W2* central triplet during 200 ps MD at 300 K: $\alpha\beta + \rho'$ (**a**); γ' (**b**); δ' (**c**); $\gamma' + \eta'$ (**d**). H-bonds (*magenta*) and close cross-strand contacts, Å (*orange*)

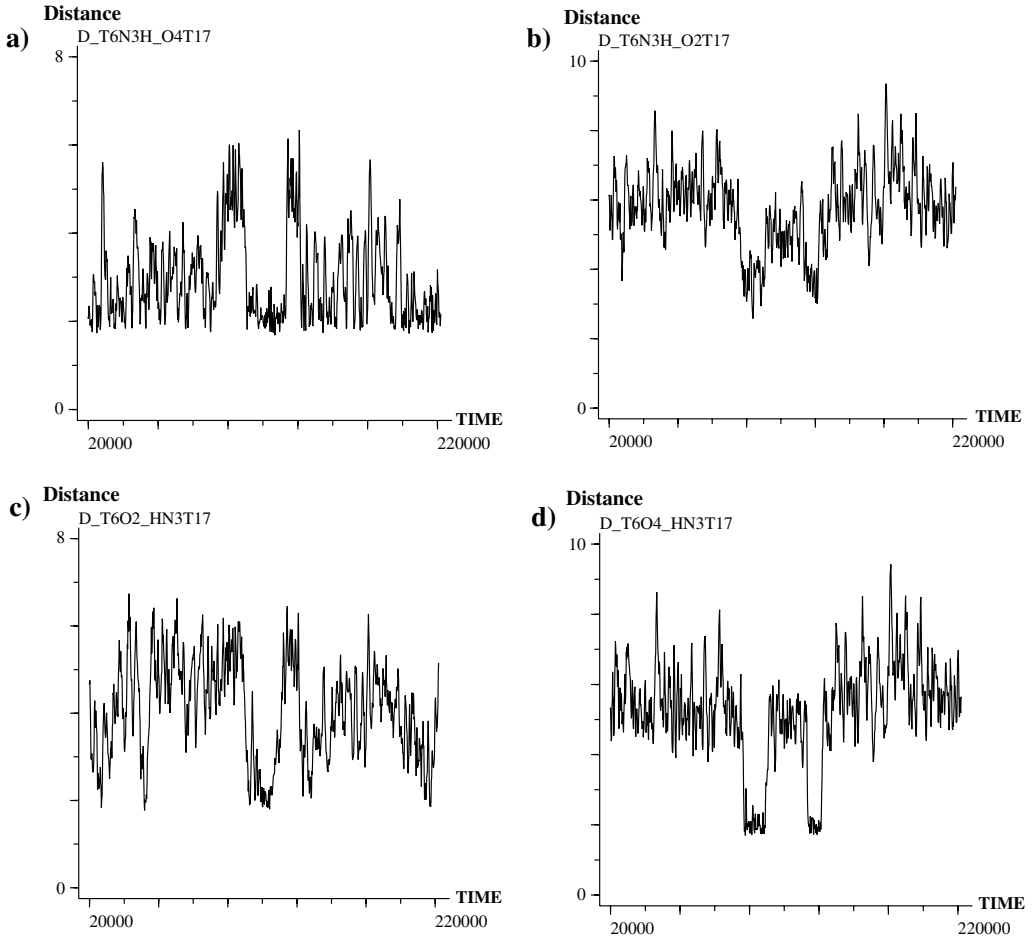


Fig. 6 *W2* duplex at 400 K: 200 ps MD trajectory plots (*fms*) of H-bond donor-acceptor distances (Å) within T6·T17 mismatch: T6N3H-O4T17 (**a**), T6N3H-O2T17 (**b**), T6O2-HN3T17 (**c**) and T6O4-HN3T17 (**d**). Trajectories of the cross-strand inter-atomic distance T7O2-HN3T17, representing conformers, γ' (**e**), and angle λ at T6 (**f**). Population histograms over 200 ps MD for the inter-atomic distances: T6NH3-O4T17 (**g**) and T6O4-HN3T17 (**h**)

duplex we performed MD simulations at elevated temperature (350 and 400 K). MD simulations at 350 K did not result in the formation of new conformations that were different from those described in the previous section. In contrast, new dynamic conformational states were formed when the temperature was set to 400 K. Examination of 200 ps trajectory plots of various distances (Fig.6, Table4) shows that fluctuations are more intensive (due to the higher kinetic energy of the atoms) and that the T6N3H-O4T17 H-bond is again preserved in the majority of conformers (52%). The $\alpha\beta$ wobble conformation is more frequent (12 vs. 1.4% at 300 K for the 200 ps MD), and is present in the time regions when, in addition to the largely persistent T6N3H-O4T17 H-bonding, the complimentary H-bonding distance, T6O2-HN3T17, is between 2.0 and 2.5 Å, (Fig.7a). The T6O4-HN3T17 (α'') and T6N3H-O2T17 (β') H-bonds, which both feature the $\beta\alpha$ wobble conformation, are also formed (Fig.6b, d). As in *W2* at 300 K the γ', δ', η' and ρ' conformational states characterized by the corresponding cross-strand contacts are frequently formed. Additional cross-strand contacts, which do not form in *W2* at 300 K, are represented by κ' and π' (low population) conformers. The conformation, denoted as κ' and characterized by a close interatomic contact T6N3H -N3A18 is accompanied by breakage of the one of the H-bonds in the 5′-flanking pair (T5O4-H62N6A18), due to the high reorientation angle of A18. Conformers in the κ' dynamic state show high buckle and propeller twist angle of the mismatched pair T6·T17, too (Fig.7c). Conversely, when the 3′-end flanking base pair is perturbed (e.g. T7·A16 adopting high inclination), the cross-strand contact T7N3H-O4T17 (η') is formed. Table4 presents the characteristic distances for the most frequently observed dynamic states (conformers) during the 400 K, 200 ps MD simulations. In summary, it is seen that at 400 K the $\beta\alpha$ wobble conformation is formed, albeit at low yield (0.4%). Similarly to the 300 K MD of *W2*, the γ' conformers, characterized by the close T7O2-HN3T17 contact are also present, however, the population is diminished to 65%, as is the case with δ' conformation (T6N3H-N1A18 = 1.8–2.9 Å), now present in only 8% from total. The most important average interatomic distances, when double cross-strand contacts exist (overlapping conformational states) are

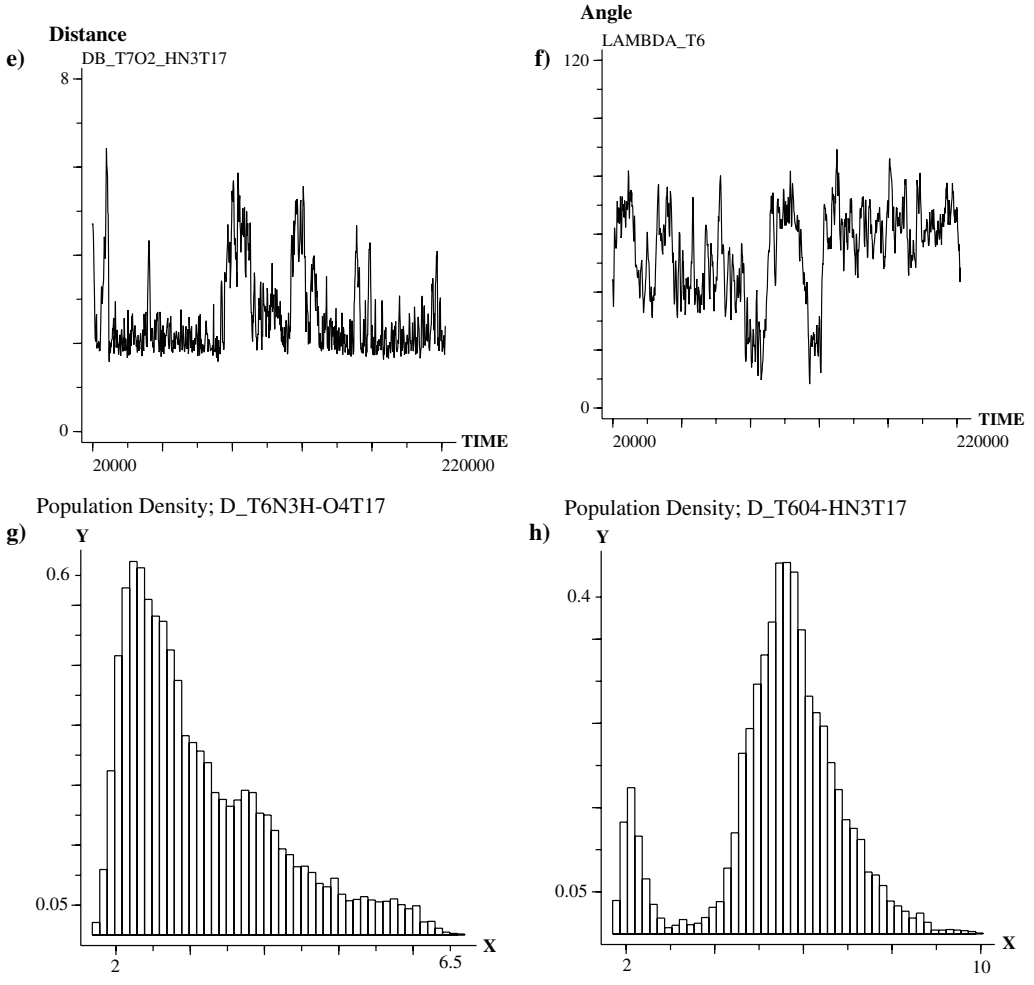


Fig. 6 (Contd.)

also listed in Table4. It can be seen that $\gamma\delta$ is less frequent as compared to MD simulations at 300 K (7 vs. 40% at 300 K), and as before it is exclusive to $\alpha\beta$, $\beta\alpha$. The ρ' state (T6O4-H62N6A16 = 1.8–2.9 Å), however shows nearly unchanged population density (32%). In contrast, the ϵ' population (T5N3H-O4T17 cross-strand contact) is increased to 11%. Several double cross-strand contacts form simultaneously. Similarly to MD simulations at 300 K, the ρ' conformational state (long-lived) partially overlaps with γ' ($\gamma' \times \rho' = 18\%$ from total), and with ϵ' ($\epsilon' \times \rho' = 5\%$ from total), while about 50% of ϵ' conformers are also in the ρ' state. At the same time, the cross section, $\rho' \times \alpha'' = 12\%$, and all $\beta\alpha$ conformers are also in the ρ' conformational state ($(\beta\alpha) \times \rho' = 100\%$) showing an average cross-strand distance, T6O4-H62N6A16 = 2.7 Å.

Dynamic conformations of the BrdU·T mispair (*W3* duplex, at 300 K)

The *W3* duplex was subjected to MD simulations at 300 K for 200 ps with subsequent extension up to 1 ns. No new conformational states were generated during the 1 ns run, as compared to the 200 ps simulation period. For comparison with other duplexes we present the analysis from a typical 200 ps MD run. As can be seen from the 200 ps trajectory plots of H-bonding distances within the mismatch (Fig.8a–d), substitution of T by BrdU in the oligonucleotide *W2* to yield the sequence *W3*, 5′d(ACGAT**BrdU**TACGA)·d(TCGTA**T**ATCGT) changes remarkably the preferred dynamic orientation of the bases in the central triplet containing the mismatched pair, as compared to *W2* at 300 K. The mismatched pair adopts frequently pure "canonical" wobble W↑ ($\alpha\beta$) and W↓ ($\beta\alpha$) conformations, similar to the *W1* oligonucleotide. Wobble $\alpha\beta$ and $\beta\alpha$ dynamic states are relatively long-lived and again the $\alpha\beta$ (12%) is favored over $\beta\alpha$ (6%) (Table5). Likewise, α' and α'' conformations are more often present (47 and 38% of total, respectively), as compared with the conjugated H-bonded conformers, β' and β'' (Table5). When the characteristic distances BrdU6O2-HN3T17 (β'') or BrdU6N3H-O2T17 (β') are short, $\alpha\beta$ and $\beta\alpha$ conformations are definitively present (Fig.8). However, the existence of a single H-bond within the BrdU·T pair (as in conformations α' or α'') is not always a prerequisite for the adoption of the double H-bonded $\alpha\beta$, or $\beta\alpha$

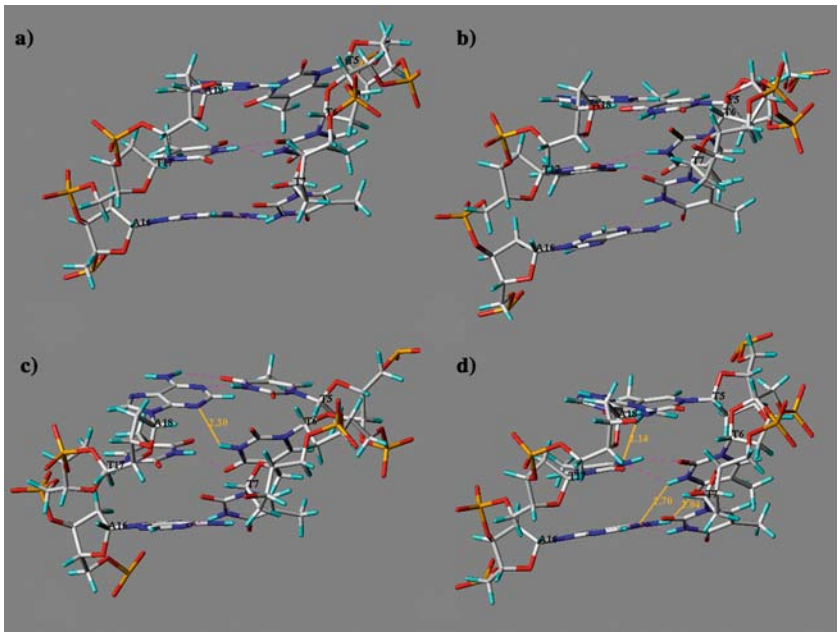


Fig. 7 3D structure snapshots of repeatedly formed conformers within the *W2* central triplet during 200 ps MD at 400 K: $\alpha\beta$ (**a**); $\beta\alpha$ (**b**); $\beta' + \kappa'$ (**c**); $\alpha\beta + \epsilon' + \rho' + \pi'$ (**d**). H-bonds (*magenta*) and close cross-strand contacts, Å (*orange*)

orientation. From the trajectory plots and structural analysis (Fig.9) it was possible to trace the nearest-neighbor topology when canonical wobble conformations are not formed. These include several situations when: (i) at least one of any of the four H-bonding distances within the T·BrdU mismatch is too long (> 4.0 Å), e.g. bases are displaced far in the major groove; (ii) intermediate states, when all H-bonding partners are within the range of ~ 3.0–4.0 Å (corresponding to transitions: $\alpha\beta \leftrightarrow \beta\alpha$); (iii) when a single very strong H-bonding contact is formed between BrdU6 and T17 (e.g. the distance, BrdU6N3H-O4T17 ≤ 2 Å), a concomitant cross-strand contact T7O2-HN3T17 (γ') exists, thus precluding canonical wobble H-bonding. The trajectory plot of the later distance (γ') follows closely those of BrU6N3H-O4T17 (α') and BrdU6O2-HN3T17 (β'') distances and is long-lived (Fig.8e). All close cross-strand contacts (interatomic distances: 1.8–2.9 Å) are essentially the same as for *W2* at 400 K, but show a different population density, ordered as: $\rho' > \gamma' > \kappa' > \epsilon' > \delta' > \eta' > \pi'$ (Table5). Conformers ρ', represented by a short BrdU6O4-H62N6A16 distance fill 62% of the total 200 ps conformational space. Although, conformers ρ' partially overlap with $\alpha\beta$ and $\beta\alpha$, i.e. $(\alpha\beta)\times\rho' = 56\%$ and $(\beta\alpha)\times\rho' = 61\%$, and the population cross-sections with α' and α'' are 50 and 57%, respectively, the characteristic interatomic distance undergoes extensive oscillations, and the trajectory does not change synchronously with the $\alpha\beta \leftrightarrow \beta\alpha$ transitions. This conformation is usually accompanied by a higher angular reorientation of A16 and displacement of T17 towards the major groove, while BrdU6 is positioned closer to the helix axis. The conformer set, denoted as γ' (T7O2-HN3T17 = 1.8–2.9 Å) is one of the most frequent cross-strand contact in the *W2* and *W3* duplexes. These conformers largely overlap with α' (53%), β'' (29%), and both, $(\alpha\beta)\times\gamma' = 29\%$), but are mutually exclusive with $\beta\alpha$, i.e. $(\beta\alpha)\times\gamma' = 0\%$. The conformational state when the distance BrdU6N3H-N3A18 becomes low (labeled as κ') overlaps strongly with $\beta\alpha$ (63%), but is exclusive to $\alpha\beta$. Topologically, it appears that the conformation δ' (T6N3H-N1A18 = 1.8–2.9 Å) is close to $\alpha\beta$, but the H-bonding partners are apart and the bonds are weak. From the positions on the trajectory plots it can be deduced that these conformers, similar to the γ' conformation, are likely a transition state during $\alpha\beta \leftrightarrow \beta\alpha$ interchange. At very short (< 2.1 Å) T7O2-HN3T17 (γ'), or BrdU6O4-H62N6A16 (ρ') distances (Fig.9c, d) the formation of the canonical wobble states is suppressed due to larger perturbations within the central base triplet. The other minor conformations are: η', which may coexist with α' (short BrdU6N3H-O4T17 distance), but is completely exclusive to $\alpha\beta$ and $\beta\alpha$; π', (BrdU6NH3−N1A16 = 1.8–2.9 Å) which is likely a state representing the $\alpha\beta \rightarrow \beta\alpha$ transition; and conformation, ϵ' (close T5N3H-O4T17 cross-strand contact), where the distance BrU6N3H-O4T17 is always short, i.e. ϵ' is a sub-conformation of α', but not of $\alpha\beta$. In general, when cross-strand contacts are formed, the 3D-structures show high plane rotation of T5 and T7, while BrdU6 is positioned into the major groove. In these conformers the, 5′ flanking base pair forms only one H-bond, T5O4-H62N6A18, but the 3′ base pair shows two normal bonds, plus a bifurcated one, T7O2-N3HT17. The rare conformation, η', appears to be overlapping with γ' (formation of a double cross-strand H-bonding contact), but is exclusive towards κ' ($\kappa'\times\eta' = 0$), and may be considered as a sub-conformation of γ' and δ' conformation. The conformer ϵ' can co-exist in some cases with κ' ($\kappa\epsilon$ cross-population of 1.1%), and with γ' ($\gamma\epsilon$ cross-population of 4.4%). At the same time, κ' and π' are strictly mutually exclusive ($\kappa'\times\pi' = 0$). Further anal-

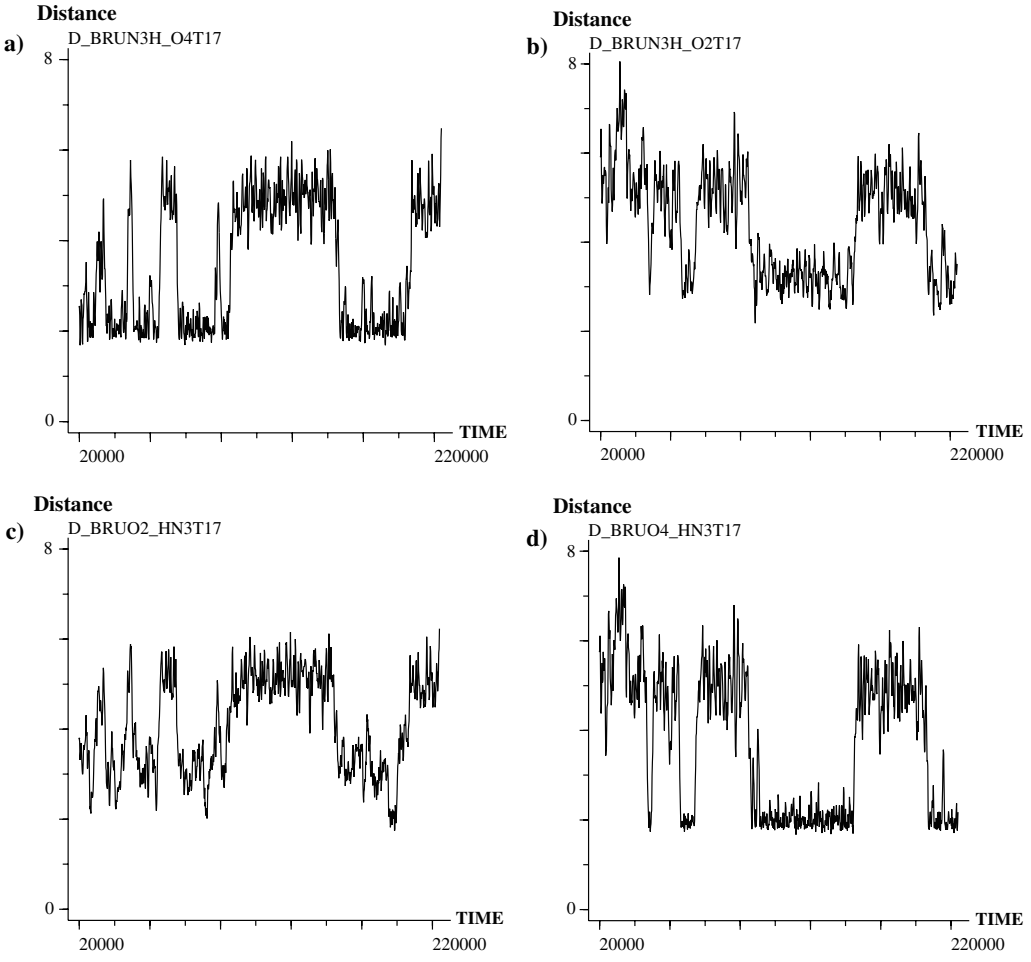


Fig. 8 *W3* duplex: 200 ps MD trajectory plots (*fms*) of H-bond donor-acceptor distances (Å) within BrdU6·T17 mismatch: BrdU6N3H-O4T17 (**a**), BrdU6N3H-O2T17 (**b**), BrdU6O2-HN3T17 (**c**) and BrdU6O4-HN3T17 (**d**). Trajectories of the cross-strand inter-atomic distance T7O2-HN3T17, representing conformers, γ' (**e**), and angle λ at BrdU6 (**f**). Population histograms over 200 ps MD for the inter-atomic distances: BrdU6O4-HN3T17 (**g**) and BrdU6O2-HN3T17 (**h**)

ysis of the interrelations between the frequently formed conformational states (overlaps) during 200 ps MD reveals that: $\gamma'\times\delta' = 16.3\%$; $\gamma'\times\rho' = 12.0\%$; $\gamma'\times\epsilon' = 1.5\%$; $\rho'\times\delta' = 0.8\%$; $\gamma'\times\eta' = 0.6\%$, but $\eta'\times\gamma' = 98\%$, while $\eta'\times\delta' = 0\%$ (i.e. almost all η'conformers are also in γ', but none in δ'). At the same time, η' is exclusive to κ' and ϵ', since $\eta'\times\kappa' = \eta'\times\epsilon' = 0$. Likewise, $\kappa\times\gamma' = \epsilon'\times\kappa' = 0$, and are mutually exclusive pairs of conformational states, as well.

Discussion

We have carried out restrained and unrestrained MD simulations of 11-mer DNA duplexes of different sequences with, or without a single T·T, or BrdU·T mismatches. Most of the simulations were performed for a 200 ps period. In the case of *W2* and *W3* duplexes at 300 K the MD runs were extended up to 1 ns. However, the results show that 200 ps MD conformational space sufficiently covers the variations of the dynamic states of interest. Due to the unavailability of X-ray diffraction or NMR data, all structures have been built "in silico" using the tools described under Simulation Details. Throughout the present study we used AMBER force field applied with different Tripos algorithms; MAXIMIN2, Simulated annealing and Molecular dynamics. It is widely accepted that AMBER force field gives the most accurate B-DNA structure conversion [15]. The reliability of our computational approach is confirmed by the highly stable distance trajectory plots and temperatures during restrained and unrestrained MD simulation of the three normal DNA (*N1*, *N2 and N3*) sequences. In addition, all normal DNA structures retained characteristic B-type overall parameters, including low RMSD, when compared with experimental B-DNA structures; typical average backbone torsion angles, H-bonding distances, groove widths and acceptably low propeller twist, ca. −13–20°. As well, our MD simulations with the *W1* DNA duplex, 5′d(GCCAC**T**AGCTC)·d(GAGCT**T**GTGGC), show an excellent conformity, e.g. the dynamics of $\alpha\beta\leftrightarrow\beta\alpha$ "canonical" wobble conformation exchange, when compared with the results reported earlier and obtained

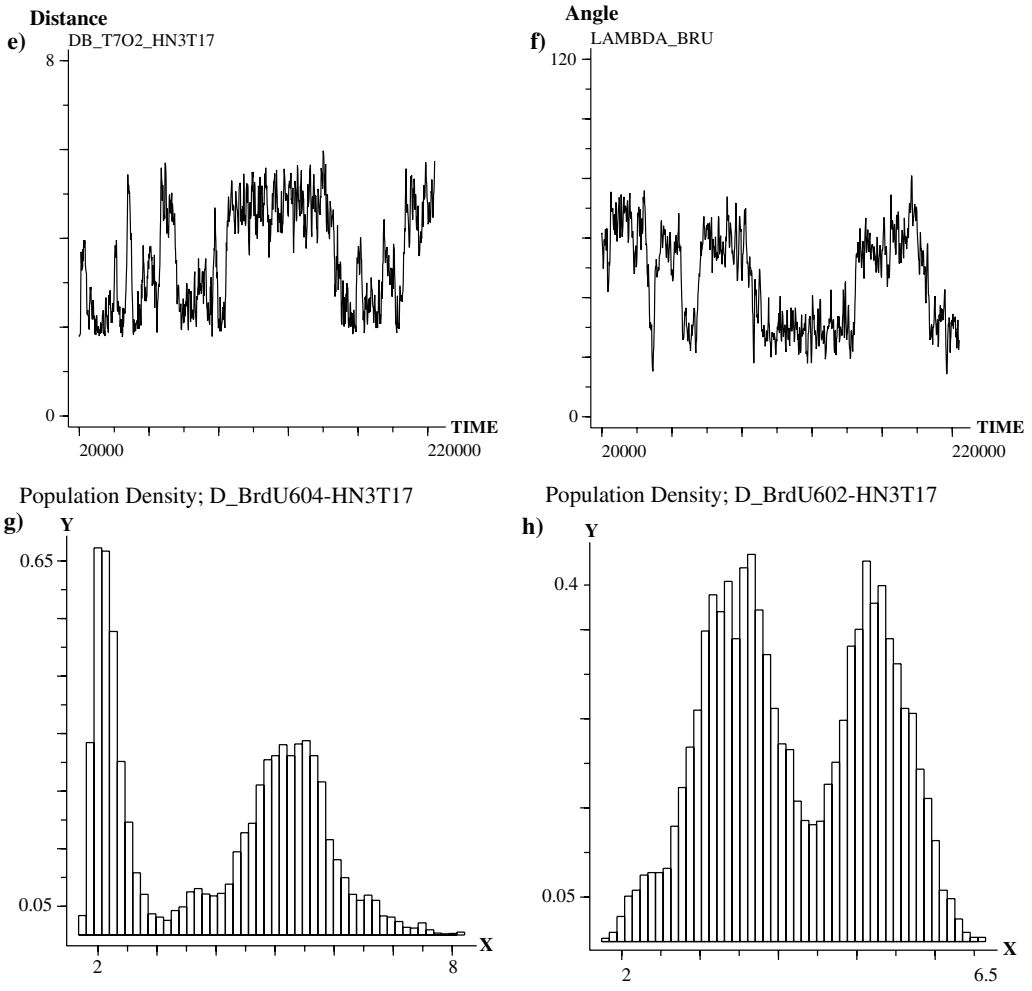


Fig. 8 (Contd.)

by restricted MD using distance constraints, directly derived by NMR [19].

The conformational states of the T·T mismatch are sequence context dependent and depend on the nature of nearest neighbors. Thus, when the mismatch is incorporated into the *W1* duplex (central triplet, 5′d(C**T**A)·d(T**T**G)) it frequently adopts "canonical" wobble conformations at 300 K, accompanied by a rapid exchange between the two states, $\alpha\beta \leftrightarrow \beta\alpha$ (Table2). This is in contrast with the MD results with the *W2* duplex at 300 K (central triplet, 5′d(T**T**T)·d(A**T**A)), where the $\alpha\beta$ wobble conformation is rarely adopted, and $\beta\alpha$ is totally missing, while the T·T mismatch most frequently assumes a propeller-twisted conformation with one persistent T6N3H-O4T17 "wobble" H-bond (α'), and one steady cross-strand contact, T7O2-HN3T17 (γ') with a conformational overlap, $\alpha' \times \gamma' \sim$ 90% (Table3). This behavior, although unusual, was anticipated in view of the restricted motion due to the steric hindrance of the central triplet-T in *W2*. Previous NMR [17, 18] and MD [8] studies have pointed out that when the T·T mismatch is inserted in sequences which contain only a doublet-T, such as 5′d(A**T**T)·d(T**T**A), the motion is highly restrained, and hence $\alpha\beta$ (W$\uparrow$)and $\beta\alpha$ (W$\downarrow$) wobble pairing is infrequent and/or absent. In the later case, the inhibition of "canonical" wobble formation was independent if the MD simulations were performed in vacuum, or in the presence of explicit water as solvent [8]. Venable et al. [8] also addressed the question of whether the predominance of the observed propeller-twisted conformation is an artifact of the force field used in the simulation. They found no evidence of force field artifacts, but proposed that the phenomenon is related to the inherent AT-regional flexibility. The latter is in line with the results reported by [20], depicting higher twisting angles in A-tract B-DNA. A more recent study [31] tackled the deformability of DNA containing AT-steps and A-tract DNA by emphasizing on the importance of thymine-methyl/π interactions (both, electrostatic and steric), which are postulated to take place between the thymine CH3 group and the aromatic ring of 5′-adenine flanking base. Although, analysis of the nature of inter-base interactions was beyond the scope of the present work, our simulation results support the above theory by showing that (i) in order to adopt "canonical" wobble conformations the bases in the central triplet *W2* duplex, 5′d(ACGAT**T**TACGA)·d(TCGTA**T**ATCGT) must overcome a thermody-

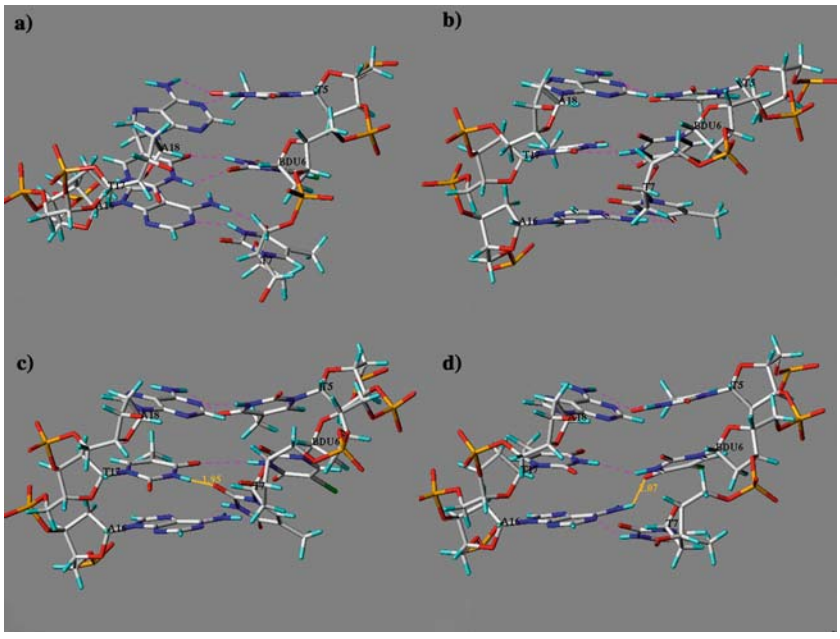


Fig. 9 3D structure snapshots of wobble and frequent overlapping conformers formed within the *W3* central triplet during 200 ps MD: $\alpha\beta$ (**a**); $\beta\alpha$ (**b**); $\alpha' + \gamma'$ (**c**); $\alpha'' + \rho'$ (**d**). H-bonds (*magenta*) and close cross-strand contacts, Å (*orange*)

namic activation barrier due interactions stabilizing the T.T propeller-twisted conformation. This is evident from the comparison of the results obtained at 300 and 400 K (Tables3 vs. 4); (ii) the results from 200 ps MD simulations with the *W3* duplex, 5′d(AC-GAT**Brd**UTACGA)·d(TCGT A**T**ATCGT), which differs from *W2* by the substitution of T6 with BrdU6, indicate that at 300 K the mismatch in *W3* (as with *W1*) frequently adopts "canonical" wobble states, $\alpha\beta$ and $\beta\alpha$, and there is a rapid exchange between these states. Therefore, it is obvious that the replacement of the 5-CH3 group by Br eliminates intra-strand interactions which would normally prevent the formation of wobble conformations. Moreover, the dynamic states of the central triplet in *W2* at 400 K and *W3* at 300 K share a very common pattern with respect to the observed set of cross-strand interatomic contacts (Tables4 vs. 5).

Apart from the evaluation of sequence dependent differences concerning the T·T "canonical" wobble pairing dynamic features, the scope of the present study is to bring deeper insight in the interatomic contacts characteristic for various dynamic states assumed by the central triplet containing the mismatched bases. The H-bonding distances within the central triplet show the

Table 3 DNA duplex *W2* at 300 K. Extractions from dynamics table: selected distance range (1.8–2.9 Å) and cross-selected ranges (*x*) represent wobble H-bonding, cross-strand inter-atomic contacts and illustrate properties of overlapping conformers ($\gamma\delta$)

Selection	None	T6NH-O4T17	T6N3H-O4T17× T6O2-HN3T17	T6N3H-O2T17	T6N3H-O2T17×T 6O4-HN3T17	T7O2-HN3T17	T6N3H-N1A18	T7N3H-O4T17	T5N3H-O4T17	T6O4-H2N6A16	T6N3H-N1A18× T7O2-HN3T17
Conf.	Aver. 200 ps	α'	$\alpha\beta$	β'	$\beta\alpha$	γ'	δ'	η'	ϵ'	ρ'	$\gamma\delta$
Population (% of total)	100	50	1.4	0	0	96	42	8	0.4	30	40
Average inter-atomic distances, Å											
T6N3H-O4T17	2.9	2.5	2.2			2.9	3	3	2.3	2.9	3
T6N3H-O2T17	6.4	6.1	5.6			6.4	6.4	6.8	5.6	6.3	6.4
T6O2-HN3T17	5	4.6	2.6			5	5.6	5.6	2.7	4.8	5.6
T6O4-HN3T17	5.5	5.5	5.6			5.5	5.2	5.8	5.6	5.6	5.2
T5N3H-N1A18	2.7	2.6	2.2			2.7	2	2.9	2.2	2.6	3
T5N3H-O4T17	4.6	4.4	3.3			4.6	5.1	5.1	2.7	4.5	5
T7O2-HN3T17	2.2	2.2	2.9			2.2	2.3	2.1	2.7	2.2	2.3
T7N3H-O4T17	3.6	3.6	4.3			3.5	3.4	2.7	4	3.6	3.4
T6N3H-N1A18	3.1	3.3	4.5			3.1	2.5	2.8	4.4	3.4	2.5
T6N3H-N3A18	4.7	4.8	5.8			4.7	4.2	4.5	5.6	5	4.2
T6O4-H2N6A16	3.2	3.1	2.8			3.2	3.3	3.4	2.8	2.6	2.5
T5C1′-C1′A18	11.8	11.7	10.6			11.8	12.1	12	10.4	11.7	12
T6C1′-C1′T17	11.8	11.5	9.8			11.9	12.1	12.2	9.8	11.7	12.2
T7C1′-C1′A16	10.6	10.6	10.1			10.6	10.7	10.8	10.3	10.6	10.7

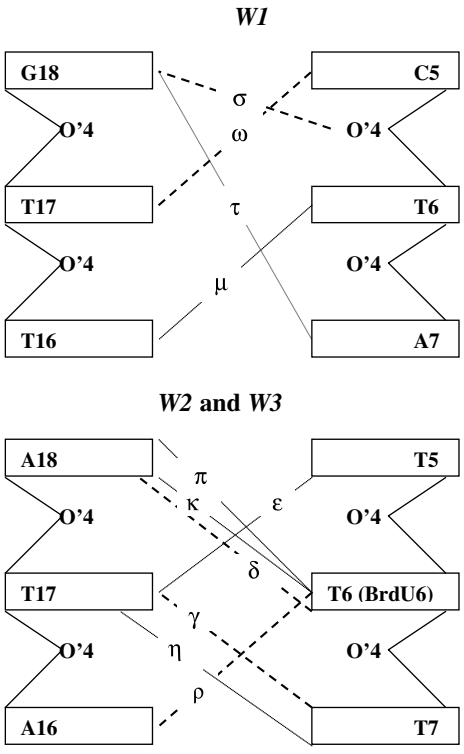
Table 4 DNA duplex *W2* at 400 K. Extractions from dynamics table: selected distance range (1.8–2.9 Å) and cross-selected ranges (x) represent wobble H-bonding, cross-strand inter-atomic contacts (conformers). The last column illustrates properties of overlapping conformers ($\gamma\delta$)

Selection	No	T6N3H-O4T17	T6N3H-O4T17×T6O2-HN3T17	T6N3H-O2T17	T6O4-HN3T17	T6N3H-O2T17×T6O4-HN3T17	T7O2-HN3T17	T6N3H-N1A18	T7N3H-O4T17	T5N3H-O4T17	T6N3H-N3A18	T6O4-H2N6A16	T7O2-HN3T17×T6N3H-N1A18
Conf.	Aver. 200 ps	α'	$\alpha\beta$	β'	α''	$\beta\alpha$	γ'	δ'	η'	ϵ'	κ'	ρ'	$\gamma\delta$
Population (% of total)	100	52	12	2.0	9.5	0.4	65	7.7	4	10.8	8	32	7
Average inter-atomic distances, Å													
T6N3H-O4T17	3.1	2.4	2.3	4.6	5.2	4.6	2.9	3.2	3.2	2.9	5.2	3.1	3.2
T6N3H-O2T17	6.0	5.8	5.3	2.8	4.0	2.8	6.4	6.5	7.0	5.4	4.0	5.6	6.5
T6O2-HN3T17	4.3	3.6	2.5	4.7	5.6	4.7	4.4	5.7	5..3	3..3	5.6	4.0	5.8
T6O4-HN3T17	5.5	5.6	5.4	2.2	2.2	2.2	6.0	5.2	6.4	5.1	2.2	5.1	5.2
T6N3H-N1A18	4.0	4.0	4.4	3.7	4.0	3.7	4.0	2.5	3.5	4.4	3.9	4.1	2.5
T6O4′-HC2A18	5.1	4.7	3.7	4.0	5.4	4.0	5.3	6.4	5.8	4.4	5.4	4.9	6.5
T5N3H-O4T17	3.9	3.6	3.2	3.2	3.4	3.2	4.0	4.8	4.7	2.6	3.4	3.6	4.9
T7O2-HN3T17	2.9	2.7	3.2	5.5	4.8	5.5	2.3	2.5	2.1	3.2	4.8	3.0	2.4
T7N3H-O4T17	4.2	4.0	4.5	6.2	5.8	6.2	3.7	3.5	2.7	4.6	5.8	4.4	3.4
T6O4-H2N6A16	3.3	3.2	3.0	2.7	3.0	2.7	3.4	3.4	3.7	3.0	3.1	2.6	3.4
T6N3-N3A18	5.1	5.5	5.6	3.0	2.7	3.0	5.4	4.1	5.2	5.2	2.6	5.0	4.1
T5C1′-C1′A18	11.2	11.1	10.6	10.2	10.4	10.2	11.4	12.0	11.7	10.6	10.4	11.0	12.1
T6C1′-C1′T17	11.0	10.6	9.4	9.9	11.0	9.9	11.3	12.3	12.0	10.1	11.1	10.7	12.3
T7C1′-C1′A1	10.8	10.7	10.4	11.9	12.8	11.9	10.6	10.8	11.0	10.8	12.0	10.7	10.7

Table 5 DNA duplex *W3* at 300 K. Extractions from dynamics table: selected distance range (1.8–2.9 Å) and cross-selected ranges (x) represent wobble H-bonding and close cross-strand inter-atomic contacts (conformers)

Selection	None	BrdU6N3H-O4T1	BrdU6O2-HN3T17	T6N3H-O4T17× BrdU6O2-HN3T17	BrdU6O4-HN3T17	BrdUN3H-O2T17	BrdU6O4-HN3T17× BrdU6N3H-O2T17	T7O2-HN3T17	BrdU6N3H-N1A18	T5N3H-O4T17	BrdU6N3H-N3A18	BrdU6O4-HN6A16	T7O2-HN3T17 × BrdU6O4-HN6A16
Conf.	Aver. 200 ps	α'	β'	$\alpha\beta$	α''	β'	$\beta\alpha$	γ'	δ'	ϵ'	κ'	ρ'	$\gamma\delta$
Population (% of total)	100	47	12	12	38	6.3	6	32	1.1	3	24	62	12
Average inter-atomic distances, Å													
BrdU6N3H-O4T17	3.6	2.3	2.2	2.2	5.2	5.1	5.1	2.4	3.4	2.4	5.2	3.6	2.3
BrdU6N3H-O2T17	4.6	5.4	5.2	5.2	3.4	2.8	2.8	5.7	4.8	5.4	3.4	4.4	5.6
BrdU6O2-HN3T17	4.2	3.3	2.6	2.6	5.3	5.1	5.1	3.6	4.4	3.0	5.4	4.2	3.5
BrdU6O4-HN3T17	4.0	5.3	5.2	5.2	2.2	2.1	2.1	5.6	4.0	5.3	2.2	3.8	5.4
BrdU6N3H-N3A18	4.0	4.8	5.0	5.0	2.9	2.9	2.9	4.9	3.5	5.0	2.4	3.9	4.7
T5N3H-O4T17	4.0	3.6	3.4	3.4	4.3	4.5	4.5	3.7	4.3	2.9	4.3	3.9	3.6
T7O2-HN3T17	3.8	3.0	3.2	3.2	5.0	5.2	5.2	2.5	3.6	3.1	5.0	4.0	2.5
BrdU6O4-HN6A1	2.9	2.9	2.9	2.9	2.9	2.8	2.8	3.1	2.9	2.8	2.9	2.7	2.6
BrdU6N3H-N1A18	3.8	3.8	4.0	4.0	3.9	3.8	3.8	3.8	2.9	4.0	3.8	4.1	3.8
T5C1′-C1′A18	10.8	10.8	10.6	10.6	10.6	11.4	10.4	11.0	11.0	10.7	10.6	10.7	11.0
BrdU6C1′–C1′T17	10.6	10.4	9.7	9.7	10.6	10.0	10.0	10.6	10.9	10.1	10.6	10.5	10.6
T7C1′-C1′A16	10.9	10.7	10.5	10.4	11.2	11.1	11.0	10.6	11.0	10.6	11.2	10.9	10.6

The last column illustrates properties of overlapping conformers, $\gamma\rho$



Sch. 2 Schematic presentation of the most frequent cross-strand inter-base contacts

presence of a variety of conformers, encompassing states with strong H-bonds (~ 2 Å), weak (~ 3 Å), or absent (> 4 Å). Conformational states that feature close cross-strand interatomic contacts and in some instances three-center (bifurcated) hydrogen bonds were repeatedly formed (Figs.3, 5, 7, 9, Tables2, 3, 4, 5). Among the inter-base parameters, the exchange between the wobble T·T conformations is synergistically accompanied by only one intra-base pair parameter, shear [19]. Most of the states represented by unusual cross-strand contacts show random changes of intra-base pair parameters (propeller twist, buckle and opening), angular axis-base pair (inclination and tip), and inter-base pair (tilt, roll, twist, etc.) within the central base pair triplet.

The formation of inter-base (cross-strand) encounter complexes (both long and short lived, transient conformers) is the most remarkable property of the mismatched duplexes. A simplified presentation of the frequent cross-strand contacts is given in Scheme2, where the more frequent contacts are represented by bold dashed lines. The properties of the most prominent unusual states for the three studied mismatched DNA molecules; *W1*, *W2* and *W3* are summarized in the Results section (see also Tables3, 4, 5). It is important to underline that in all mismatched duplexes, and particularly in the *W3* duplex, wobble pairing H-bonding coexists with some close cross-strand contacts: e.g. $(\alpha\beta)\times\gamma' = 29\%$; $(\beta\alpha)\times\kappa' = 63\%$; $(\alpha\beta)\times\rho' = 56\%$ and $(\beta\alpha)\times\rho' = 61\%$, etc. Although rarely, triple close cross-strand contacts are formed in all studied mismatched duplexes. These findings may be of special importance for understanding the pathways of electron-transfer reactions in DNA subjected to ionizing radiation. In order to extend our work to all aspects of the radiation-induced DNA degradation processes, it will be necessary to carry out additional simulations. First of all, it will be essential to perform MD simulations including explicit solvent (water) molecules, instead of the used averaging of electrostatic interactions by a distance dependent dielectric function. Solvent water molecules, apart from eventual stabilization of wobble, and/or other unusual conformations, can play a crucial role in the primary and secondary radiation processes in DNA. Second, in order to distinguish between many primary electron-transfer reactions, it will be important to study MD conformational states of DNA structures containing oxidized and/or reduced base free radicals, together with a proper assignment of the corresponding electrochemical potentials.

Conclusions

Our results show that the dynamic conformational states of the T·T mismatch depend on the DNA sequence context in which they occur. The T·T mismatch may or may not adopt "canonical" wobble conformations. These findings will lead to a better understanding of how certain DNA sequences may confound the recognition and excision processes performed by DNA repair systems. In all cases, the bases in the central triplet containing the mismatch form specific (sequence dependent) cross-strand contacts. Some of these contacts are short lived but frequent, thus implying numerous atomic collisions. Often the measured distances are < 2.0 Å, which suggests strong atomic overlap. Frequent atomic collisions are a prerequisite for charge transfer interactions, as described in the Marcus theory of electron transfer reactions [32]. The specified cross-strand inter-atomic contacts in *W1*, *W2* and *W3* duplexes present an intricate network of numerous co-existing (double, triple and bifurcated) H-bonding, electrostatic and/or van der Waals interactions. This enables further analysis encompassing specific base electron donor-acceptor properties for the evaluation of inter-strand electron (charge) transfer pathways in these particular DNA molecules. The results may be especially important in the case when radiation-induced dissociation of the Br atom (*W3*) takes place and the resultant oxidized uracil radical is generated, serving as an initiator of a cascade of intra-strand and inter-strand electron-transfer processes most often resulting in DNA backbone scission.

Acknowledgements This work was supported by the NCI of Canada. Supplementary material: BrdU AM1 structure, MM optimized and MD DNA 3D-structures (mol2 and pdb formats), dynamics history files, etc. are available on request.

References

1. Brown T, Kennard O, Kneale G, Rabinovich D (1985) Nature 315:604–606
2. Brown T, Kneale G, Hunter WN, Kennard O (1986) Nucl Acids Res 14:1801–1809
3. Privé GG, Heinemann U, Chandrasegaran S, Kan L-S, Kopka ML, Dickerson RE (1987) Science 238:498–504
4. Maskos K, Gunn MB, LeBlanc DA, Morden KM (1993) Biochemistry 32:3583–3595
5. Boulard Y, Cognet JAH, Gabarro-Arpa J, Le Bret M, Carbonnaux C, Fazakerley GV (1995) J Mol Biol 246:194–208
6. Boulard Y, Fazakerley GV, Sowers LC (2002) Nucl Acids Res 30:1371–1378
7. Isaacs RJ, Rayens WS, Spielmann HP (2002) J Mol Biol 319:191–207
8. Venable RM, Widmalm G, Brooks BR, Egan W, Pastor RW (1992) Biopolymers 32:783–794
9. Špačková N, Berger I, Šponer J (2000) J Am Chem Soc 122:7564–7572
10. Limoli CL, Ward JF (1993) Radiat Res 134:160–169
11. Gralla J, Crothers DM (1973) J Mol Biol 78:301–319
12. Peyret N, Seneviratne PA, Allawi HT, SantaLucia Jr J (1999) Biochemistry 38:3468–3477
13. Robinson BH, Mailer C, Drobny G (1997) Annu Rev Biophys Biomol Struct 26:629–658
14. Swaminathan S, Ravishanker G, Beveridge DL (1991) J Am Chem Soc 113:5027–5040
15. McAteer K, Kennedy MA (2003) J Biomol Struct Dynam 20:487–506
16. Sowers LC, Goodman MF, Eritja R, Kaplan B, Fazakerley GV (1989) J Mol Biol 205:437–447
17. Arnold FH, Wolk S, Cruz P, Tinoco Jr I (1987) Biochemistry 26:4068–4075
18. Kouchakdjian M, Li BFL, Swann PF, Patel DJ (1988) J Mol Biol 202:139–155
19. Gervais V, Cognet JAH, Le Bret M, Sowers LC, Fazakerley GV (1995) Eur J Biochem 228:279–290
20. Yoon C, Privé GG, Goodsell DS, Dickerson RE (1988) Proc Natl Acad Sci USA 85:6332–6336
21. Sommers MF, Byrd RA, Gallo KA, Samson CJ, Zon G, Egan W (1985) Nucl Acids Res 13:6375–6386
22. Cecchini S, Girouard S, Huels MA, Sanche L, Hunting DJ (2004) Radiat Res 162:604–615
23. Myers RM, Fischer SG, Maniatis T, Lerman LS (1985) Nucl Acids Res 13:3111–3129
24. Arnott S, Hukins DWL (1972) Biochem Biophys Res Comm 47:1504–1509
25. Boulard Y, Cognet JAH, Fazakerley GV (1997) J Mol Biol 268:331–347
26. Cognet JAH, Gabarro-Arpa, Cuniasse Ph, Fazakeley GV, Le Bret M (1990) J Biomol Struct Dynam 7:1095–1115
27. Boulard Y, Cognet JAH, Gabarro-Arpa J, Le Bret M, Sowers LC, Fazakerley GV (1992) Nucl Acids Res 20:1933–1941
28. Cognet JAH, Boulard Y, Fazakerley GV (1995) J Mol Biol 246:209–226
29. Lavery R, Sklenar H (1989) J Biomol Struct Dynam 6:655–667
30. Real AN, Greenall RJ (2000) J Mol Model 6:654–658
31. Umezawa Y, Nishio M (2002) Nucl Acids Res 30:2183–2192
32. Marcus RA, Sutin N (1985) Biochim Biophys Acta 811:265–322