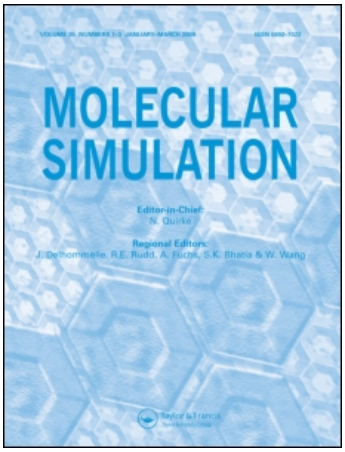
This article was downloaded by:
On: *14 January 2011*
Access details: *Access Details: Free Access*
Publisher *Taylor & Francis*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Molecular Dynamics Simulations of Novel Hoogsteen-Like Bases That Recognize the T-A Base Pair by DNA Triplex Formation

Jeffrey H. Rothman[a]; W. Graham Richards[a]
[a] Physical and Theoretical Chemistry Laboratory, Oxford University, Oxford, UK

To cite this Article Rothman, Jeffrey H. and Richards, W. Graham(1996) 'Molecular Dynamics Simulations of Novel Hoogsteen-Like Bases That Recognize the T-A Base Pair by DNA Triplex Formation', Molecular Simulation, 18: 1, 13 — 42
To link to this Article: DOI: 10.1080/08927029608022352
URL: http://dx.doi.org/10.1080/08927029608022352

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Molecular Simulation, 1996, Vol. 18, pp. 13–42
Reprints available directly from the publisher
Photocopying permitted by license only

© 1996 OPA (Overseas Publishers Association)
Amsterdam B.V. Published in The Nethrelands under
license by Gordon and Breach Science Publishers SA
Printed in Malaysia

MOLECULAR DYNAMICS SIMULATIONS OF NOVEL HOOGSTEEN-LIKE BASES THAT RECOGNIZE THE T-A BASE PAIR BY DNA TRIPLEX FORMATION

JEFFREY H. ROTHMAN and W. GRAHAM RICHARDS

Physical and Theoretical Chemistry Laboratory, Oxford University, South Parks Road, Oxford, UK OX1 3QZ

(Received February 1996; accepted March 1996)

Downloaded At: 19:15 14 January 2011

Effective sequence-specific recognition of duplex DNA is possible by triplex formation with natural oligonucleotides via Hoogsteen H-bonding. However, triplex formation is in practice limited to pyrimidine oligonucleotides binding duplex A-T or G-C base pair DNA sequences specifically at homopurine sites in the major groove as T·A-T and C^+·G-C triplets. Here we report the successful molecular dynamics modelling of novel unnatural nucleosides that recognize the T-A DNA base pair by Hoogsteen interaction. These X·T-A triplets were placed within $(T \cdot A\text{-}T)_{11}$ triplexes of A-type or B-type starting configuration. These triplex structures underwent energy minimization and 300 ps molecular dynamics simulation with counter-ions and solvent. The same procedures were applied to unmodified $(T \cdot A\text{-}T)_{11}$ triplexes and their trajectories compared to those of the central triplet modified triplexes of similar starting configuration. The phosphodiester backbone dihedrals, χ-dihedrals, and H-bonding distances of the central and adjacent triplets of the modified and unmodified triplexes show similar average values and deviations from starting configurations. This indicates that these X·T-A triplets do not induce any significant perturbations for the duration of the simulation.

Keywords: DNA triplex; Hoogsteen; novel nucleosides; molecular dynamics.

BACKGROUND

The ever increasing knowledge of gene sequences has made DNA a suitable drug target. It has become worthwhile to consider sequence selective ligands such as DNA triple helix forming oligonucleotides (TFOs), as one of the more promising routes. Effective sequence-specific recognition of duplex DNA by triplex formation is facilitated via major groove Hoogsteen

H-bonding to the duplex Watson-Crick base pairs. Oligodeoxynucleotide-implemented triple helix formation furnishes one of the most versatile methods for sequence-specific recognition of double helical DNA[1,2].The ability to target a broad scope of DNA sequences, the high stabilities of the triplex structure, and the sensitivity to single-base mismatches make this a powerful method for binding exclusive sites within large segments of duplex DNA. Since, statistically, the base sequence of a 17-mer oligonucleotide occurs just once in the sequence of the human genome, extremely selective intervention ought to be possible [3]. However, that approach is severely limited if we restrict attention to natural nucleotides since triplex formation is limited to pyrimidine TFOs binding duplex A-T or G-C base pair DNA sequences specifically at homopurine sites in the major groove parallel to the homopurine strand as T·A-T or C^+·G-C triplets. Helix-coil transition melting temperature, UV mixing curve, and ^1H NMR experiments all give credence to these homopolymeric structures [4–7]. However, the construction of homopolymeric triplex structures structurally and configurationally analogous to T·A-T and C^+·G-C via TFOs binding in the major groove of a T-A or C-G duplex parallel to the homopyrimidine strand have yet to be experimentally confirmed.

Downloaded At: 19:15 14 January 2011

Application in DNA Duplex Recognition

Relying solely upon the recognition inherent in the T·A-T and C^+·G-C triplets, the utilization of natural TFOs has been successful in mimicing repressors and the construction of artificial restriction enzymes. [8,9,10] For example, TFOs have been successful in achieving single or double site specific cleavage of yeast and human chromosomal DNA. [11] Sequence-specific inhibition of DNA binding proteins such as prokaryotic modifying enzymes and eukaryotic transcription factor have been successful at micromolar TFO concentrations. [12] TFOs have also been shown to be useful as competitors for DNA-binding proteins and as site-specific DNA damage or cleavage reagents. [9,10,13] Suppression of gene expression via triplex formation has potential as demonstrated by suppression of human *c-myc* gene transcription with nanomolar TFO concentrations [14]. Successful suppression of transcription has also been implemented by blocking the promoter region thereby inhibiting the binding of the eukaryotic transcription factor [15]. Also depending upon natural nucleotide triplet scheme, alternate strategies of TFOs binding in the major groove have attempted to recognize T-A and C-G base pairs. Modelling of CGG triplexes is consistent with experimental results indicating antiparallel TFO

orientation with respect to the purine [15,16,17]. Simultaneous binding of the pyrimidine TFO to purines on alternate strands of the duplex has been engineered by connecting TFOs in 3'-3' fashion via a crossover link [18]. In this manner the TFO pyrimidine is expected to bind the purine strand in both parts of the major groove. However, this approach lacks specificity due to the binding deficit in the crossover region. In light of the adaptability portrayed in these examples, this method demonstrates potential to be a universal solution for DNA recognition.

Triplex Conformations from ^{1}HNMR and Molecular Dynamics Results

Design studies such as these rely heavily upon structural data from experiment. The first DNA triplex models were derived from the Arnott fibre diffraction data [19] where the triplex configuration is considered to be similar to A-DNA with C3′-endo sugar puckers. At present, there are still no reports of crystal diffraction studies of DNA triplex. In hope of a more complete understanding of the triplex structure and energetics, molecular dynamics studies have been performed [20,21]. These studies have found specific time-averaged behavior of the triplex systems to be consistent with ^{1}H NMR NOE solution structure data [22]. However, ^{1}H NMR solution studies [23,24,25] and molecular dynamics [20] both report structural deviations from the fiber diffraction model [19]. Recent solution IR spectroscopy studies [26,27] of T·A-T triplex and solution NMR studies [28] of various triplex oligonucleotide systems suggest that many of the nucleotide residues have sugar puckers nearer to a C2′ -endo configuration which would be in better agreement with a B-DNA type conformation. Consequently, in light of these new findings a more appropriate model T·A-T triplex has been devised [29,30] in which all three strands have the same phosphodiester geometry and C2′-endo sugar pucker that characterizes B-type DNA geometry. Recent molecular dynamics simulations [31] of T·A-T DNA triplexes from both the A-type and B-type starting configurations show similar trajectories which converge to a structure that is structurally equidistant from, but not very similar to either of the initial A-type or B-type structures. Although it may be difficult to produce an accurate representation of a time-averaged helical unit from these dynamics simulations, the resultant convergence to similar structure suggests available configurational pathways between A-type and B-type DNA triplex conformations. Unfortunately, overwhelming proof or rejection of the preference of one type of conformation over the other is not evident.

Downloaded At: 19:15 14 January 2011

Downloaded At: 19:15 14 January 2011

Nonnatural Base Design

Much effort has gone into the design of nonnatural nucleotide bases for TFOs, especially those that aim to bind specifically to T-A or C-G base pairs via Hoogsteen H-bonding with the same parallel orientational geometry as the known T·A-T and C^+·G-C natural triplexes. [32–36] Unnatural bases have also been devised for a triplex RNA starting structure as analysed by molecular mechanics [37]. Due to the significant conformational freedom of DNA, automated ligand design routines are not well suited to binding sites that need not remain close to their starting geometry for optimal binding interactions. Nonnatural bases may bind with comparable energetics with respect to known stable triplets, but are allowed conformational latitude without energetic penalty to the host triplex. These changes within the target structure may easily eclipse or extend van der Waals contact limits for the ligand design search parameters.

In this paper we report the modelling of X·T-A triplets where X is a Hoogsteen-binding unnatural nucleotide centrally placed within a host T·A-T 11mer triplex. Different Hoogsteen bases were devised for X·T-A triplets in the A-type or B-type triplex configuration. As a first step, triplets composed of novel Hoogsteen nucleosides designed to bind in the major groove of T-A or C-G base pairs should show structural stability enclosed within a known stable triplex structure. A successful base should demonstrate favorable stacking and Hoogsteen interaction energy and a comparable energy minimized phosphodiester backbone and nucleotide geometry to that of the known natural triplets.

Previous energy minimization results of unnatural Hoogsteen bases X2, X3, X4, and X5 ((*E*) 4-(1-(2-deoxy-β-D-ribofuranosyl)-1-buten-3-yne) imidazol-2-one, (*E*) 4-(1-(2-deoxy-β-D-ribofuranosyl)-1-buten-3-yne) thiazol-2-one, (*E*) 4-(1-(2-deoxy-β-D-ribofuranosyl)-1-buten-3-yne) 1-*N*-methyl, 5-methyl imidazol-2-one, and (*E*) 4-(1-(2-deoxy-β-D-ribofuranosyl)-1-ethene) 1-*N*-methyl, 5-methyl imidazol-2-one (Figures 1a, 1b, 2a, 2b respectively) in a centrally placed X·T-A triplet have shown successful triplex conformational binding preferences [38]. The unnatural Hoogsteen X-bases, X2 and X3, which incorporate a 1-buten-3-yne linker show more favorable binding energies within an A-type X·T-A triplet configuration than a B-type whereas the unnatural Hoogsteen base, X5, which incorporates an ethenyl linker shows a slight binding preference in the B-type triplet configuration. The structural viability of an X·T-A triplet was initially assessed by energy minimization of the entire triplex structure. The resulting energies were subsequently compared to those of a centrally placed T·A-T triplet within

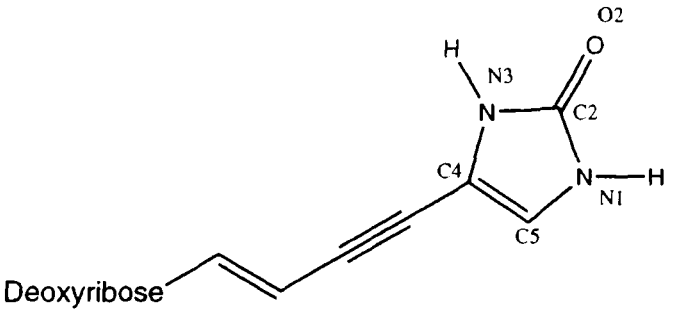


FIGURE 1a Hoogsteen Base X2.

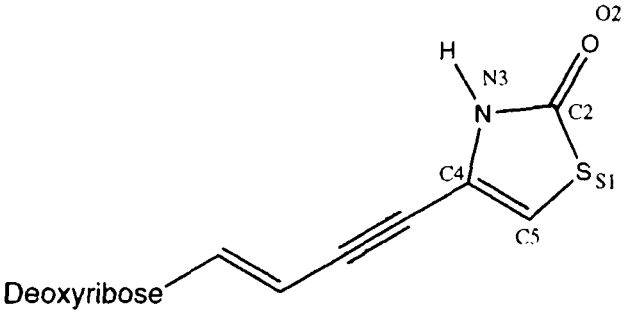


FIGURE 1b Hoogsteen Base X3.

Downloaded At: 19:15 14 January 2011

the same $(T \cdot A\text{-}T)_{11}$ triplex. The configuration and energetics of the energy minimized structures were then assessed with respect to any significant structural or energetic perturbations imposed upon adjacent nucleotides due to replacement of the central T·A-T triplet by the X·T-A triplet. Further testing includes comparison of time dependent structure of the control $(T \cdot A\text{-}T)_{11}$ triplex to that of the $(T \cdot A\text{-}T)_5(X \cdot T\text{-}A)(T \cdot A\text{-}T)_5$ triplex through molecular dynamics. Assessing the triplet's stability and effects upon the rest of the triplex should be facilitated by exposing the generation of any deviations. In order to reduce unchecked χ-dihedral angle rotation during molecular dynamics simulations, methyl groups are incorporated into the azalone ring structure of X2 thus creating X4 (Fig. 2a), and with the ethenyl linker, X5 (Fig. 2b).

Structural perturbations caused by the test triplet to the entire triplex structure with respect to the control triplex should be more easily exposed through dynamic simulation. By demonstrating that the test triplet does not induce any or minimal perturbation into the rest of the triplex structure

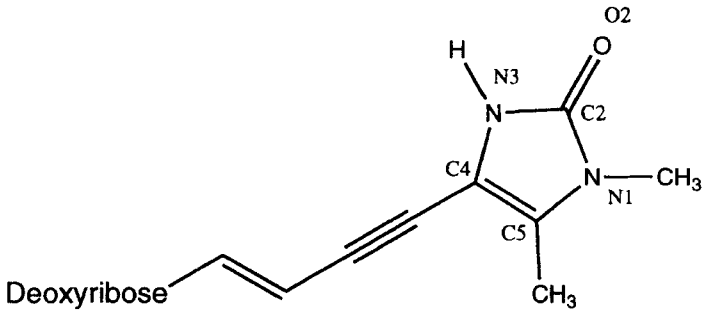


FIGURE 2a Hoogsteen Base X4.

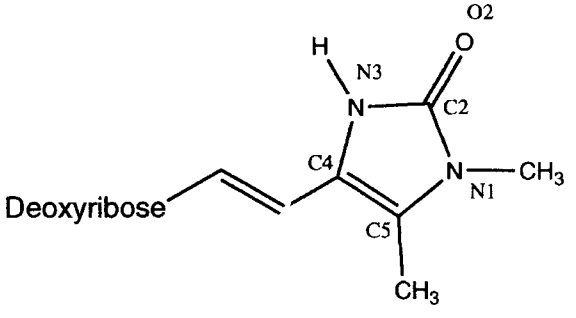


FIGURE 2b Hoogsteen Base X5.

Downloaded At: 19:15 14 January 2011

with respect to the triplex without this triplet, there is greater probability that it will be successful experimentally. Here we report the successful molecular dynamics modelling of novel unnatural nucleosides, X4, (*E*) 4-(1-(2-deoxy-β-D-ribofuranosyl)-1-buten-3-yne) 1-*N*-methyl, 5-methyl imidazol-2-one (Fig. 2a), and X5, (*E*) 4-(1-(2-deoxy-β-D-ribofuranosyl)-1-ethene) 1-*N*-methyl, 5-methyl imidazol-2-one (Fig. 2b) that bind the major groove of a T-A base pair in the center of a T·A-T 11mer triplex for the A-type and B-type starting configurations, respectively.

METHODS

General

All minimizations and dynamics calculations were performed on a Silicon Graphics Power Challenge XM with two processors. The AMBER 4.1 [39] suite of programs were used for the molecular minimizations, dynamics, and

analysis. Visualization and interactive modelling were performed using QUANTA [40].

Construction of System

The initial starting structure of the A-type configuration $(T \cdot A\text{-}T)_{11}$ triple helix was generated from the Arnott fiber diffraction model [19] where the triplet step height of 3.26 Å and a turn angle of 30.0° were used. This is related to placement of a Hoogsteen binding pyrimidine strand into the major groove of an A-DNA duplex parallel to the purine strand. The initial starting structure for the B-type configuration $(T \cdot A\text{-}T)_{11}$ triplex was obtained from the proposed molecular mechanics model of Sasisekharan [29–30] based on IR [26,27] and ^1H NMR studies. [23–25,28] All three strands have the same ribophosphodiester geometry, nevertheless the step height, turn angle, and strand orientation are equivalent to those of the A-type configuration.

Counterion placement was achieved through non-linear Poisson-Boltzmann methods of DELPHI [41]. In this manner sodium ions were placed iteratively at the position of most negative electrostatic potential at a 3.5 Å distance from the surface of the triplex. This process was repeated by taking the previously placed ions into consideration until net neutrality was achieved. The triplex and counterions were then immersed in a box of Monte Carlo waters so that their placement was a minimum of 1.7 Å distance between the solute and solvent waters. Boundaries for solvation were limited to 9.0 Å between the box edge and any solute atoms along the helix axis and 8.0 Å along the other axes. The starting system assembled contains 1083 solute atoms including 30 sodium counterions. For the A-type $(T \cdot A\text{-}T)_{11}$ triplex simulation the starting configuration includes 3501 water molecules in a box of dimension 59.4 Å × 44.3 Å × 44.1 Å whereas the starting configuration of the B-type $(T \cdot A\text{-}T)_{11}$ triplex simulation includes 3271 water molecules in a box of dimension 57.1 Å × 44.3 Å × 43.5 Å.

Construction of test triplet and placement in triplex

The center triplet used for testing novel nucleotide bases was constructed by switching the thymine and adenine bases to the opposite Watson-Crick strands while retaining the χ-dihedral angles initially associated with each strand. Initial design and visual screening of proposed Hoogsteen bases were performed on QUANTA [40]. For purposes of calculating charges, the geometries of the proposed bases were determined with MOPAC [42] using

Downloaded At: 19:15 14 January 2011

the AM1 hamiltonian and MMOK force field constraint. The charges of the MOPAC determined structure were then procured from GAUSSIAN90 [43] with an RHF/STO-3G basis set run with PRECISE and EF parameters. These charges were then fitted using the CHELPG [44] method and scaled to the AMBER force field for natural nucleotides.

Construction and initial data of triplexes with X·T-A central triplets

The placement of the unnatural Hoogsteen bases into the central triplet of the triplex follows the same χ-dihedral angle and geometry as specified for the particular configurational type of the triplex. The X4 Hoogsteen base is installed within an A-type configuration triplex whereas the X5 Hoogsteen base is designated for the B-type. The solvation of the 11mer T·A-T triplexes with X·T-A central triplets follows the same procedures for counterion placement and water addition as were applied to the $(T \cdot A\text{-}T)_{11}$ triplex. The assembled $(T \cdot A\text{-}T)_5$X4.T-A$(T \cdot A\text{-}T)_5$ triplex system contains 1090 solute atoms including 30 sodium counterions. Its starting cofiguration includes 3518 water molecules in a box of dimension 60.7 Å × 43.7 Å × 44.0 Å. The assembled $(T \cdot A\text{-}T)_5$ X5·T-A$(T \cdot A\text{-}T)_5$ triplex system contains 1088 solute atoms including 30 sodium counterions. Its starting configuration includes 3269 water molecules in a box of dimension 57.1 Å × 44.3 Å × 43.5 Å.

Minimization and Dynamics Procedures

Initially to test the viability of the triplet energy and structure, minimizations were performed upon the triplex with distance dependent dielectric and implicit solvation. Energy minimizations comprising 100 steps deepest descent and 1000 steps of conjugate gradient method were performed to ensure a tolerance of structural RMS deviations less than 0.05 Å. A nonbonded cutoff distance of 8.0 Å was used, and a linear distance dependent dielectric of 4 modelled the implicit solvation.

Energy minimizations and dynamics of the explicitly solvated triplex with counter ions were performed using an 8.0 Å nonbonded cutoff and a pairlist that was updated every 10 steps. Distance and dihedral endconstraints were applied to H-bonds in the terminal and its adjacent triplets. Initial conditioning of the entire system consisted of 500 cycles of energy minimization followed by 25 ps of dynamics on the water and ions alone starting at 10° K to 298° K under constant pressure. This was then followed by energy minimization of the entire system for 500 cycles. Molecular dynamics then

Downloaded At: 19:15 14 January 2011

proceeded on the entire system with starting velocities assigned at the 10° K initial temperature. The system was then heated to 298° K by coupling to a heat bath and run under constant pressure with periodic boundary conditions. The timestep was 0.02 ps using the SHAKE algorithm [45] and structures were saved every 1.0 ps for six serial 50 ps trajectories.

RESULTS

Reason for Design

In this paper we probe and analyze the conformational and energy profile of the Hoogsteen binding region of a T-A base pair with conformationally adjustable Hoogsteen bases, X4 for the A-type and X5 for the B-type starting triplex configuration. As a starting point, initial energy minimizations revealed minimal backbone geometry perturbation [38]. However, this alone is merely an indicator. By performing molecular dynamics simulations and assessing the conformational preferences of the adjustable base probe the further design of more purpose built Hoogsteen bases can be facilitated.

In our particular study we are interested in finding the optimal placement of a rigid azalone structure capable of forming one H-bond with each Watson-Crick base and the extent to which these are maintained during the dynamics simulation. By connecting this H-bonding structure via a 1-buten-3-yne linker for the A-type and an ethene linker for the B-type, some rotation around the azalone is allowed in a crankshaft-type motion thus permiting it to maintain more easily coplanarity with the bases above and below it while adjusting its distance to the Watson-Crick bases in its triplet. There are many variables that cannot be addressed by just one test structure such as varying the innate structural van der Waals and electrostatic potentials, and iteratively exploring them would be prohibitively time-consuming. However, the aim is to have optimized as many structural variables as best as possible by screening through iterative minimization trials of different test structures before probing conformational binding profiles through molecular dynamics.

Analysis Outline

In order to assess the overall comparative conformational integrity, certain distances and angles must be chosen which best represent the characteristics

Downloaded At: 19:15 14 January 2011

under scrutiny. Since little is known about the solution structure of the conformation of DNA triplexes except for NOE data [46,47,22] describing the furan conformation, the initial configurational analysis of the X·T-A triplets are compared to those of the proposed A-type and B-type canonical configuration.

Predicting the binding efficacity of an unnatural Hoogsteen nucleotide within a triplet can be assessed by direct comparison of its conformational similarities to those of known triplets. Using the backbone geometry of phosphates, furans, and χ-dihedral angles of the A-type and B-type T·A-T triplex starting configuration, comparison of the energy minimized triplet structures within a stable triplex is an efficient means for screening viable triplets. A molecular dynamics simulation is then an efficient means for assessing the response of the triplet's stability and effects upon the rest of the triplex by exposing any conformational perturbations in the process of sampling more conformational space.

The unnatural base under scrutiny must be as conformationally stable as the known (control) Hoogsteen base as part of a triplet within a timescale where the entire Hoogsteen oligonucleotide strand remains bound to the duplex. Starting with an energy minimized structure with ions and waters, the deviations of the center and its adjacent triplet nucleotide dihedral angles are monitored during the molecular dynamics simulation. In order to increase the simulation time in which the Hoogsteen strand remains bound to the duplex in solution, the fraying at the ends are checked by distance and dihedral angle constraints on the H-bonds of the outer two triplets at each end. These constraints are meant mechanically to approximate the effects of (computationally expensive) longer stretches of DNA oligonucleotides.

Downloaded At: 19:15 14 January 2011

Previous Energy Minimization Studies for A-type Triplex

Of the initial bases subjected to minimization trials the azalone, X2, and thioazolone, X3, heterocycles attached to an 1-buten-3-yne link fulfilled the energetic and structural criteria (Fig. 3a,b) with respect to mimicing a natural A-type DNA T·A-T triplet (Fig. 3c). The central placement of these X·T-A triplets within an A-type T·A-T 11 mer produces a similar Hoogsteen interaction energy profile to that of the unmodified T·A-T 11 mer [38,48]. The Hoogsteen base plane stacking energies and Watson-Crick base interactions also remain virtually unperturbed in comparison to those of the unmodified T·A-T 11 mer [38,48]. Additionally, there are no significant structural perturbations evident upon comparison of the energy

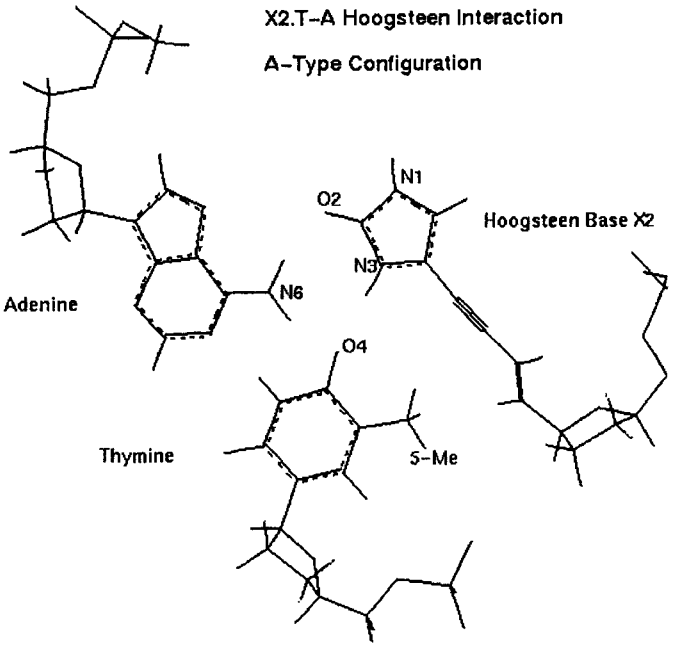


FIGURE 3a Energy minimized X2. T-A triplet configuration within the $(T.A\text{-}T)_5$- (X2.T-A)-$(T.A\text{-}T)_5$ triplex from the A-type configuration.

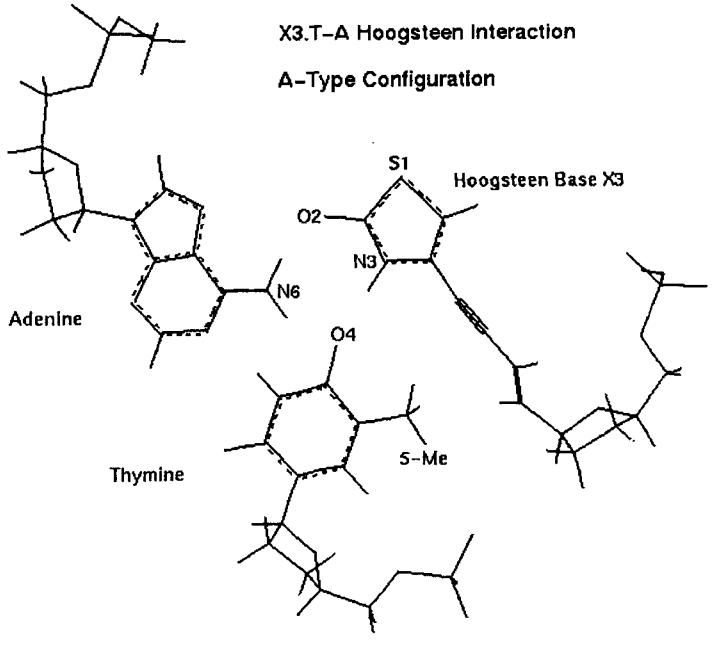


FIGURE 3b Energy minimized X3.T-A triplet configuration within the $(T.A\text{-}T)_5$-(X3.T-A)-$(T.A\text{-}T)_5$ triplex from the A-type configuration.

Downloaded At: 19:15 14 January 2011

minimized T·A-T 11 mer and the X·T-A centrally inserted T·A-T 11 mer, with regard to phosphodiester backbone dihedrals and χ-dihedrals of the base-furan bond. Upon initial molecular dynamics simulation studies of the X2·T-A azalone inclusive triplex, the standard deviation of the Hoogsteen H-bonding distances in the X2·T-A triplet was found to be much greater than that of the analogous Hoogsteen bond of T·A-T center triplet. This is due to the 1-NH amide forming H-bonds with the adjacent triplets and consequently tipping the azalone out of the triplet base plane for significant time durations of the simulation. Competition with the Hoogsteen bonding configuration would be minimized by preventing such unwanted interactions as in the design of X3.

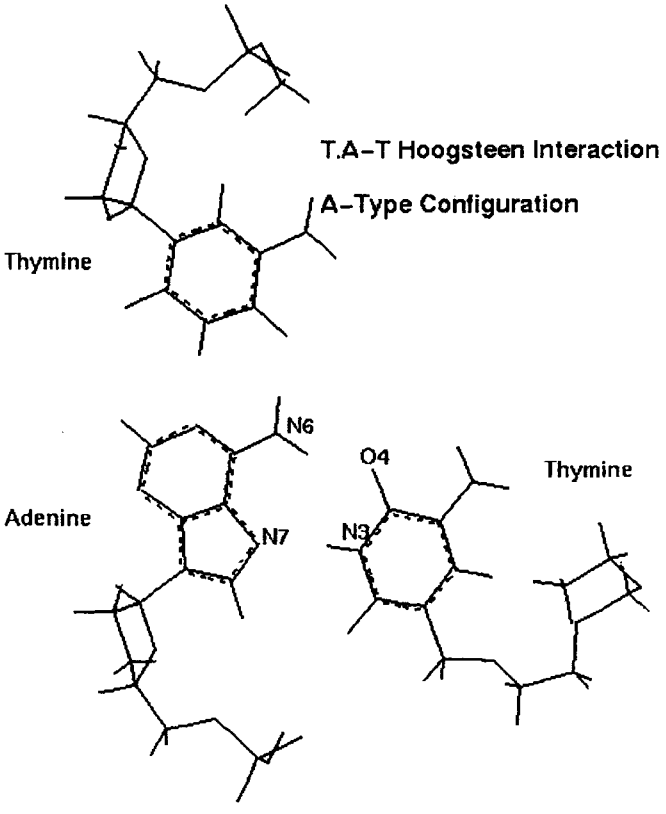


FIGURE 3c Energy minimized T.A-T triplet configuration within the $(T.A\text{-}T)_{11}$ triplex from the A-type configuration.

Downloaded At: 19:15 14 January 2011

Downloaded At: 19:15 14 January 2011

Construction of X4 for A-type Triplet

Although the thiazolone base X3 does not have undelegated H-bonding potential like X2, its Hoogsteen bonding orientation is not as optimal as X2. This is evident in comparing the base χ-dihedrals of the associated energy minimized triplex structures. In order to avoid potentially disorienting H-bond interactions with adjacent triplets and maintain the optimal Hoogsteen bonding orientation by extending the base plane, a 1-*N*-methyl, and 5-methyl modification to the azalone X2 were considered next to construct X4. Minimization trials show that the X4-Hoogsteen intra-triplet interaction energy is 108% that of the analogous interaction energy of an A-type central T·A-T triplet, however its stacking interaction is only 70% due to its poorer overlap. These values are similar to those of an X2·T-A triplet, as well as a Hoogsteen base total interaction energy (which includes the intra-triplet and adjacent triplet interactions) of 84% to that of an A-type central T·A-T triplet (Fig. 4a). Most importantly there are minimal structural differences upon comparison of the central and adjacent triplets of the unmodified A-type $(T \cdot A\text{-}T)_{11}$ triplex and its X4·T-A centrally modified version (Fig. 4b).

Construction of X5 for B-type Triplet

Initially the construction of a Hoogsteen base for an X·T-A triplet in the B-type configuration, X5, was based upon properly orienting the 1-N-methyl

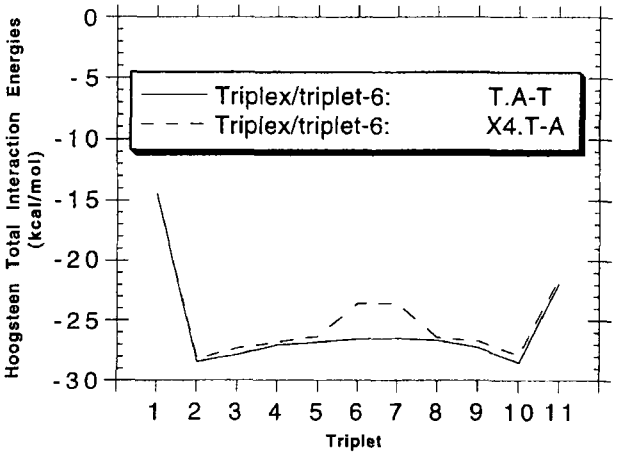


FIGURE 4a Interaction energies for each Hoogsteen nucleoside with the rest of the triplex with central triplets: T.A-T and X4.T-A for the A-type configuration.

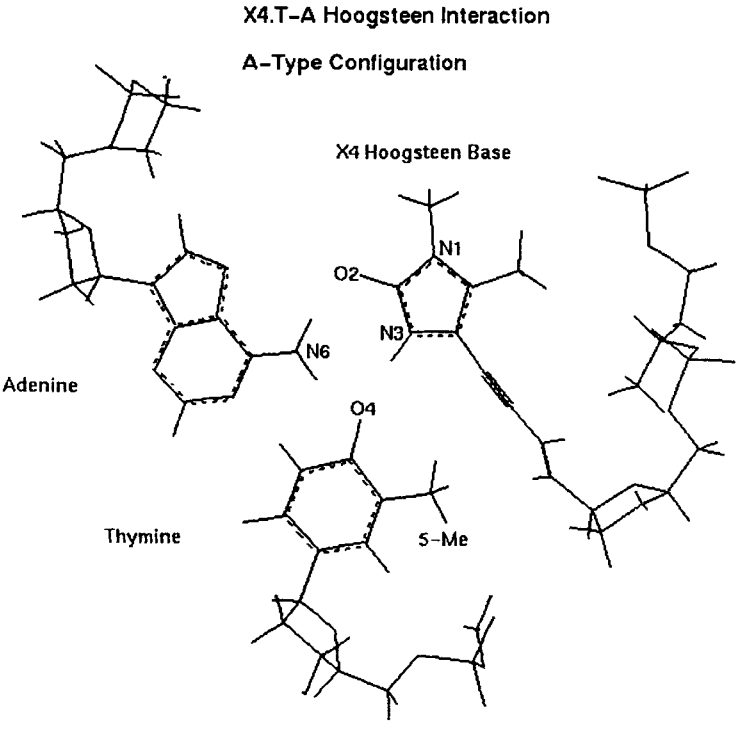


FIGURE 4b Energy minimized X4.T-A triplet configuration within the (T. A-T)$_5$- (X4.T-A)-(T.A-T)$_5$ triplex from the A-type configuration.

Downloaded At: 19:15 14 January 2011

substituted azalone portion of X4 within the B-type triplet to allow suitable Hoogsteen H-bonding distances (Fig. 5a) with respect to mimicing a natural B-type DNA T·A-T (Fig. 5b). This was accomplished by using an ethenyl linker to attach the substituted azalone to the C1′ of the furan in a *trans* orientation. As in the comparison of the X4·T-A triplet to an A-type central T·A-T triplet, data from the energy minimized central X5·T-A inclusive triplex indicates a similar structural geometry and a Hoogsteen base total interaction energy of 84% to that of a B-type central T·A-T triplet (Fig. 5c). However, the X5-Hoogsteen intra-triplet interaction energy is only 65% that of the analogous interaction energy of a central T·A-T triplex. Due to better π-overlap than X4 its stacking interaction is 88% that of the central T·A-T triplet stacking interaction of the unmodified B-type (T·A-T)$_{11}$ triplex.

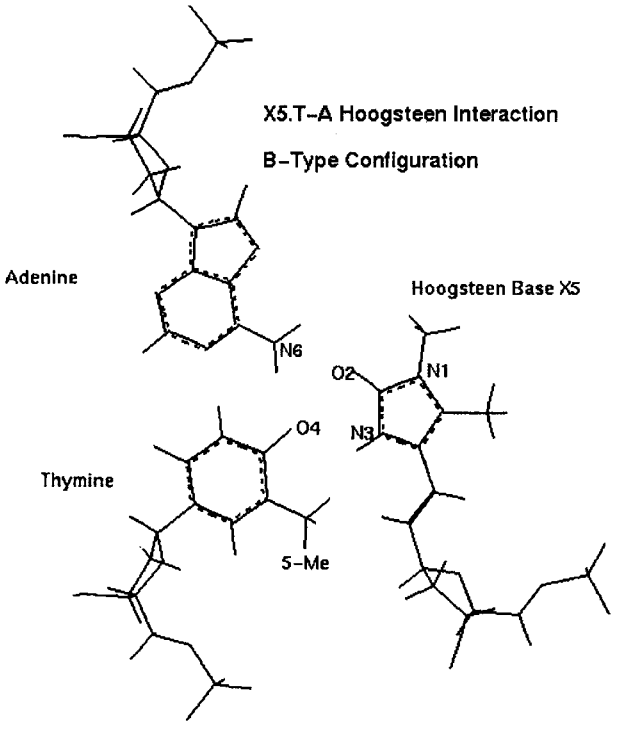


FIGURE 5a Energy minimized X5.T-A triplet configuration within the $(T.A\text{-}T)_5$- (X5.T-A)-$(T.A\text{-}T)_5$ triplex from the B-type configuration.

Downloaded At: 19:15 14 January 2011

Measurement of Molecular Dynamics Simulation Results

Measurement of the stability of the proposed X·T-A central triplets and any possible perturbations upon its adjacent triplets can be best determined from the more energy dependent structural features. Stabilities in DNA binding recognition can be directly associated with the effectiveness of Watson-Crick base pair and Hoogsteen base H-bonding which can be followed by monitoring H-bonding distance data. In the same manner the π-stacking interactions can be followed by monitoring base-ribose χ-dihedrals. Additionally, phosphodiester backbone dihedrals are used to monitor overall configurational stabilities. Comparison of serial structures through internal coordinates allows representations of gross structural rms deviations without the necessity for superimposing centers of mass.

During molecular dynamics simulations structurally and conformationally homogeneous starting structures such as these triplexes become

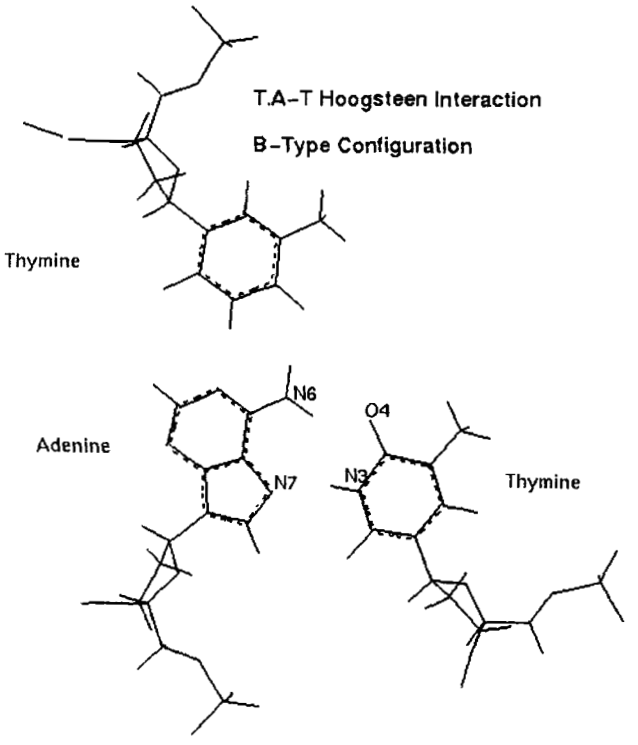


FIGURE 5b Energy minimized T.A-T triplet configuration within the $(T.A\text{-}T)_{11}$ triplex from the B-type configuration.

conformationally unhomogeneous. [31] It is therefore more purposeful to use empirical parameters that describe changes in motion over specified portions of the structure for comparative purposes. For this reason the changes in rms deviations for a specific group of dihedrals during the simulation are used as a basis to compare dynamic similarities between the unmodified $(T \cdot A\text{-}T)_{11}$ triplex and the related centrally substituted triplex of the same starting configuration. The initial structure of a simulation serves as the reference in determining the dihedral deviations whereas correlation of movements between related structures with respect to simulation time would be an unsatisfactory indicator in that it is sensitive to harmonic phase displacement. The initial simulation structures of the compared triplexes are chosen as references due to their similar configuration after energy minimization.

Downloaded At: 19:15 14 January 2011

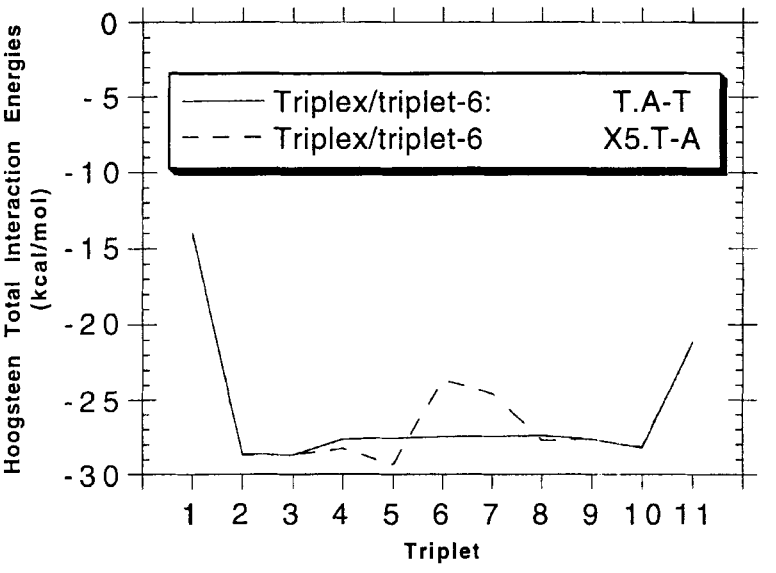


FIGURE 5c Interaction energies for each Hoogsteen nucleoside with the rest of the triplex with central triplets: T.A-T and X5.T-A for the B-type configuration.

Downloaded At: 19:15 14 January 2011

Phosphodiester backbone

Average phosphodiester backbone rms deviations measured over the three central triplets were plotted with respect to simulation time in order to follow gross structural changes in overall configuration. The rms deviation of the backbone dihedrals of the A-type (Fig. 6a) and B-type (Fig. 6b) starting configurations of the $(T \cdot A\text{-}T)_{11}$ triplex show a slower initial increase than that of their analogous X4·T-A (Fig. 6c) and X5·T-A (Fig. 6d) centrally substituted triplexes. The difference is not surprising in that the X · T-A triplets have not been as minimized to the same extent as the T · A-T triplet unit in the unmodified $(T \cdot A\text{-}T)_{11}$ triplex. It is important to note that the calculated rms deviations of the dihedrals for each strand have similar profiles and values indicating that conformational strand dihedral deviations are occuring in unison. This is also consistent with significant interstrand interactions resulting in correlated strand motion. However, these conclusions are meaningless unless supported by evidence indicating that other interstrand interactions, Watson-Crick and Hoogsteen H-bonds, and π-stacking interactions remain intact throughout most of the simulation.

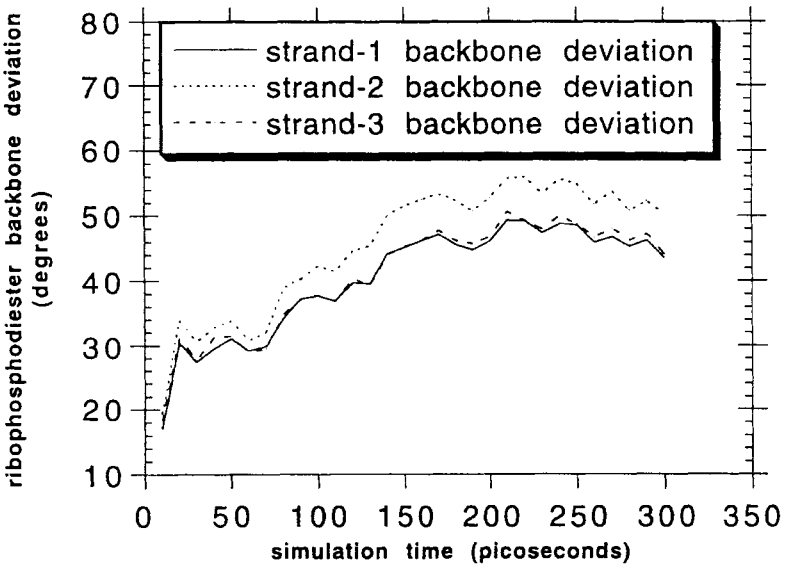


FIGURE 6a Ribophosphodiester backbone dihedral deviation from starting configuration; deviations are averaged over the three cental triplets of A-type $(T.AT)_{11}$ in blocks of 10 ps.

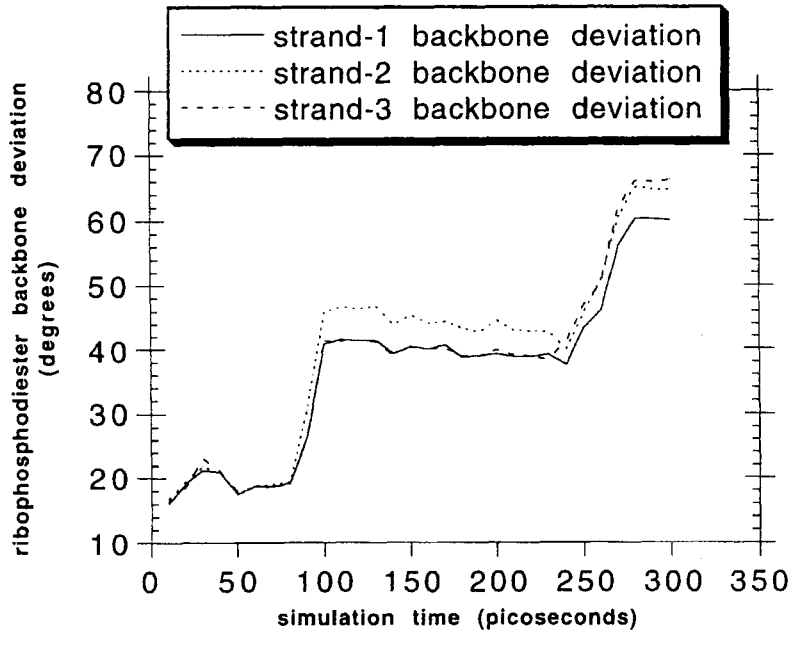


FIGURE 6b Ribophosphodiester backbone dihedral deviation from starting configuration; deviations are averaged over the three central triplets of B-type $(T.AT)_{11}$ in blocks of 10 ps.

Downloaded At: 19:15 14 January 2011

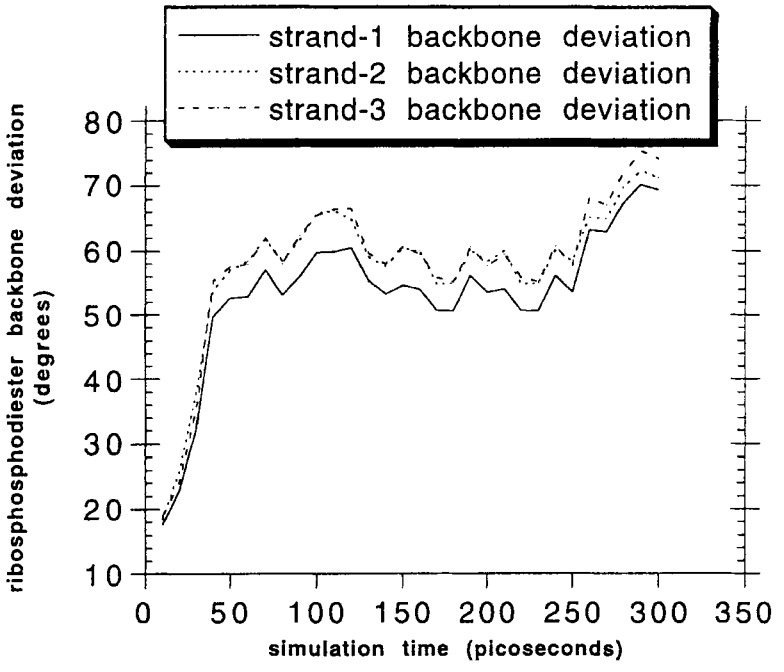


FIGURE 6c Ribophosphodiester backbone dihedral deviation from starting configuration; deviations are averaged over the three central triplets of A-type $(T.A\text{-}T)_5$-(X4.T-A)-$(T.A\text{-}T)_5$ in blocks of 10 ps.

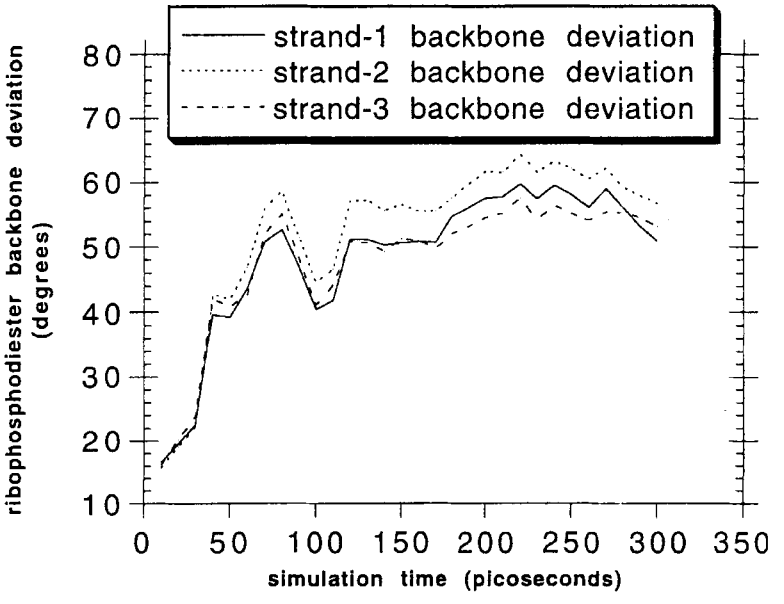


FIGURE 6d Ribophosphodiester backbone dihedral deviation from starting configuration; deviations are averaged over the three central triplets of B-type $(T.A\text{-}T)_5$-(X5.T-A)-$(T.A\text{-}T)_5$ in blocks of 10 ps.

Downloaded At: 19:15 14 January 2011

Downloaded At: 19:15 14 January 2011

Base-ribose χ-dihedral

The π-stacking orientation of nucleotide bases are examined through the χ-dihedral measured along the heterocycle attachment and its Cl's attachment on the furan. The rms deviations of these dihedrals with respect to their starting configurations measured over the three central triplets appear to approach similar values at the end of the simulations for each configuration. Comparison of the A-type triplex configurations reveals that there is initially more χ-dihedral deviation associated with the X4·T-A inclusive triplex than that of the unmodified A-type T·A-T 11 mer (Fig. 7c,d) during the first 150 ps of the simulation, but the former approaches the latter χ-dihedral deviation by 5°–10° during the remaining portion of the simulation. This type of behavior is most consistent with local restructuring during the initial phase of the simulation followed by slow equilibration towards the rms deviation values associated with the unmodified $(T \cdot A\text{-}T)_{11}$ structure. However, comparison of the B-type configurations reveals that the χ-dihedral rms deviation trajectories associated with X5·T-A inclusive $(T \cdot A\text{-}T)_{11}$ mer are within 5° rms deviation of each other throughout the entire simulation (Fig. 7a,b). This result is consistent with that either the X5·T-A triplet causes minimal perturbation or the χ-dihedrals of the B-type

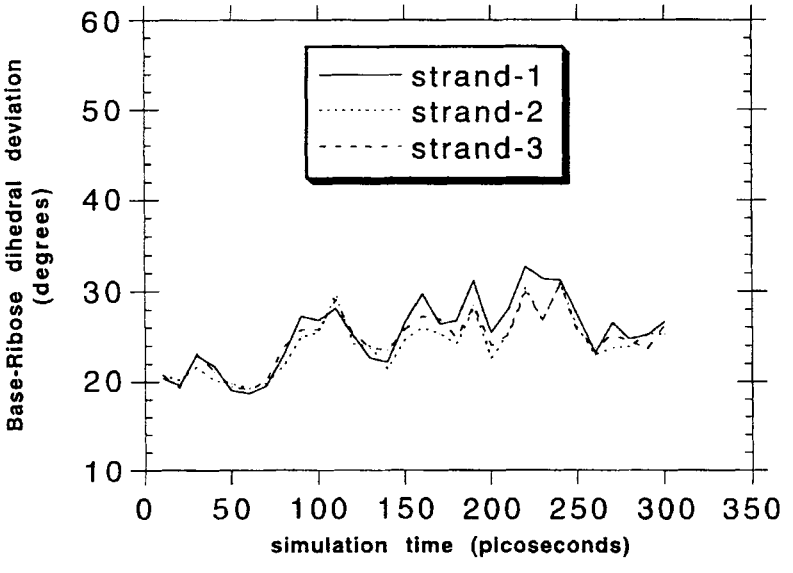


FIGURE 7a Base-Ribose χ-dihedral deviation from starting configuration; deviations are averaged over the three central triplets of B-type $(T.A\text{-}T)_5$-(X5.T-A)-$(T.A\text{-}T)_5$ in blocks of 10 ps.

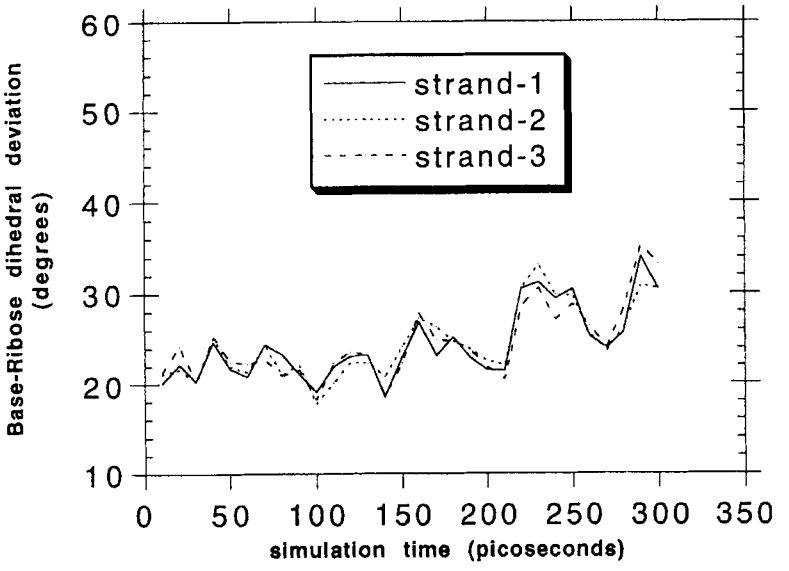


FIGURE 7b Base-Ribose χ-dihedral deviation from starting configuration; deviations are averaged over the three central triplets of B-type $(T.A\text{-}T)_{11}$ in the blocks of 10 ps.

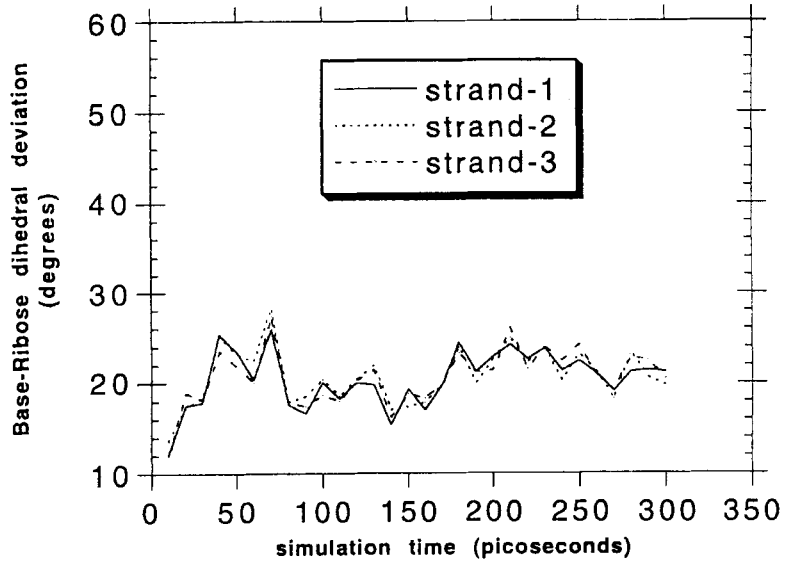


FIGURE 7c Base-Ribose χ-dihedral deviation from starting configuration; deviations are averaged over the three central triplets of A-type $(T.A\text{-}T)_5$-(X4.T-A)-$(T.A\text{-}T)_5$ in blocks of 10 ps.

Downloaded At: 19:15 14 January 2011

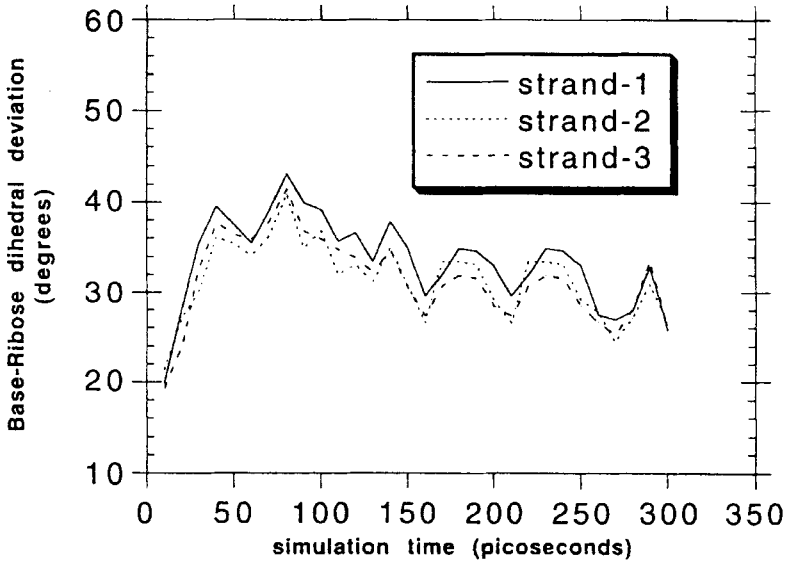


FIGURE 7d Base-Ribose χ-dihedral deviation from starting configuration; deviations are averaged over the three central triplets of A-type $(T.A\text{-}T)_{11}$ in blocks of 10 ps.

Downloaded At: 19:15 14 January 2011

configuration are not particularily as sensitive with respect to those of the A-type configuration. For all triplex structures the deviation plots (Fig. 7a–d) show that the central three triplets of each strand of a given structure closely follow a similar course in rms χ-dihedral deviation over the entire course of their simulations. This evidence is consistent with concerted χ-dihedral fluctuations within a given triplet.

Comparison of the individual time-averaged χ-dihedrals between the unmodified $(T \cdot A\text{-}T)_{11}$ triplex and their central triplet modified versions more directly show the structural similarities. As shown in Table IIa the three central average base-ribose dihedrals on all three strands of the A-type $(T \cdot A\text{-}T)_{11}$ triplex are well within standard deviation of the corresponding (analogous) base-ribose dihedrals of its central triplet modified version. The same can be stated for the comparison of these dihedrals between the B-type $(T \cdot A\text{-}T)_{11}$ triplex and its central triplet modified version as seen in Table IIb. Although the difference in the means of all analogous χ-dihedrals are statistically significant by t-test, the dihedral population averages of the compared dihedrals do significantly overlap with respect to their standard deviations.

TABLE IIa Simulation averages and standard deviations for specific χ-angles for the A-type(T.A-T)$_{11}$ and (T.A-T)$_5$-(X4.T-A)-(T.A-T)$_5$ triplexes

Strand-I	*Average χ-Dihedral Angle*		*Standard Deviation*	
Specified central triplet:	T.A-T	X4.T-A	T.A-T	X5.T-A
Triplet-5	30.03	33.79	12.03	14.03
Triplet-6	33.60	23.67	12.95	20.29
Triplet-7	30.35	32.08	14.51	17.80
Strand-II	*Average χ-Dihedral Angle*		*Standard Deviation*	
Specified central triplet:	T.A-T	X5.T-A	T.A-T	X5.T-A
Triplet-5	12.48	15.88	24.06	19.01
Triplet-6	36.08	38.44	17.18	16.53
Triplet-7	37.53	28.06	20.57	18.01
Strand-III	*Average χ-Dihedral Angle*		*Standard Deviation*	
Specified central triplet:	T.A-T	X4.T-A	T.A-T	X4.T-A
Triplet-5	43.61	49.46	14.25	21.56
Triplet-6	24.02	53.61	12.28	20.56
Triplet-7	38.77	34.38	15.71	13.10

TABLE IIb Simulation averages and standard deviations for specific χ-angles for the B-type(T.A-T)$_{11}$ and (T.A-T)$_5$-(X5.T-A)-(T.A-T)$_5$ triplexes

Strand-I	*Average χ-Dihedral Angle*		*Standard Deviation*	
Specified central triplet:	T.A-T	X4.T-A	T.A-T	X4.T-A
Triplet-5	31.21	33.51	16.32	15.75
Triplet-6	12.03	19.63	12.49	16.63
Triplet-7	29.79	37.48	12.30	14.30
Strand-II	*Average χ-Dihedral Angle*		*Standard Deviation*	
Specified central triplet:	T.A-T	X4.T-A	T.A-T	X4.T-A
Triplet-5	30.28	7.71	16.26	22.74
Triplet-6	28.06	43.25	17.61	13.06
Triplet-7	35.14	40.43	15.41	24.56
Strand-III	*Average χ-Dihedral Angle*		*Standard Deviation*	
Specified central triplet:	T.A-T	X4.T-A	T.A-T	X4.T-A
Triplet-5	28.07	49.84	14.15	19.37
Triplet-6	28.98	25.92	13.62	16.61
Triplet-7	35.51	26.44	16.14	14.21

Downloaded At: 19:15 14 January 2011

Hydrogen bonding

The H-bonding distances of the Watson-Crick and Hoogsteen interactions are the most important factor in determining the viability of the tested unnatural base in its triplet and its interactions with adjacent triplets. Recognition of the Watson-Crick base pair is not only dependent upon but also can be evaluated by the distance of the Hoogsteen H-bonds. The average distances between the H-bonding heteroatoms of the three central triplets of the triplexes were calculated from the simulation. Of primary importance it is necessary to show that the unmodified $(T \cdot A\text{-}T)_{11}$ control structure maintains viable H-bonding distances throughout simulation in the A-type and B-type configuration. As shown in Tables IIIa,b it is evident that the average Watson-Crick and Hoogsteen H-bonding heteroatom distances for both the A-type and B-type starting configurations of the unmodified $(T \cdot A\text{-}T)_{11}$ triplex are all within 0.1 Å of the analogous distances in their respective starting models. The standard deviations of the distances from the simulations follow these results by showing the narrow 0.1–0.2 Å range around the optimal bonding distance.

As shown in Table IIIa the Watson-Crick and Hoogsteen H-bonds in the simulations of A-type $(T \cdot A\text{-}T)_{11}$ starting triplex show similar average

TABLE IIIa Simulation averages and standard deviations for the H-bonds of the central three triplets of the A-type $(T.A\text{-}T)_{11}$

Triplet-5: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Thy-O4–Ade-N6	2.95	0.16
Watson-Crick/I–II	Thy-N3–Ade-N1	2.91	0.10
Hoogsteen/III–II	Thy-O4–Ade-N6	2.96	0.19
Hoogsteen/III–II	Thy-N3–Ade-N7	2.88	0.11
Triplet-6: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Thy-O4–Ade-N6	2.97	0.14
Watson-Crick/I–II	Thy-N3–Ade-N1	2.87	0.09
Hoogsteen/III–II	Thy-O4–Ade-N6	2.96	0.19
Hoogsteen/III–II	Thy-N3–Ade-N7	2.88	0.14
Triplet-7: Bond Type Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Thy-O4 Ade-N6	2.95	0.14
Watson-Crick/I–II	Thy-N3–Ade-N1	2.90	0.10
Hoogsteen/III–II	Thy-O4–Ade-N6	3.02	0.22
Hoogsteen/III–II	Thy-N3–Ade-N7	2.84	0.12

Downloaded At: 19:15 14 January 2011

TABLE IIIb Simulation averages and standard deviations of the H-bonds of the central three triplets of the B-type $(T.A\text{-}T)_{11}$

Triplet-5: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Thy-O4–Ade-N6	2.90	0.13
Watson-Crick/I–II	Thy-N3–Ade-N1	2.94	0.12
Hoogsteen/III–II	Thy-O4–Ade-N6	2.95	0.17
Hoogsteen/III–II	Thy-N3–Ade-N7	2.89	0.14
Triplet-6: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Thy-O4–Ade-N6	2.95	0.14
Watson-Crick/I–II	Thy-N3–Ade-N1	2.89	0.10
Hoogsteen/III–II	Thy-O4–Ade-N6	2.90	0.13
Hoogsteen/III–II	Thy-N3–Ade-N7	2.88	0.12
Triplet-7: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Thy-O4–Ade-N6	2.94	0.15
Watson-Crick/I–II	Thy-N3–Ade-N1	2.91	0.12
Hoogsteen/III–II	Thy-O4–Ade-N6	2.92	0.14
Hoogsteen/III–II	Thy-N3–Ade-N7	2.91	0.12

Downloaded At: 19:15 14 January 2011

distances and deviations to those of its X4·T-A centrally inserted triplex version (Tab. IIIc). Accordingly, all of the corresponding H-bonding distances are all well within a standard deviation of each other. Of importance to notice is the extent of the success in which the X4 Hoogsteen base H-bonds to both Watson Crick bases whereas the Hoogsteen base in the central T·A-T triplet manages to only H-bond the adenine Watson Crick base. The X4 base achieves a 2.94 Å H-bonding distance from its 3-NH to the thymine 4-CO with standard deviation of 0.23 Å and a 2.95 Å H-bonding distance from its 2-CO to the adenine 6-NH_2 with a standard deviation of 0.27 Å while the Hoogsteen thymine of a T·A-T triplet achieves a 2.96 Å H-bonding distance from its 4-CO to the adenine 6-NH_2 and a 2.88 Å distance from 3-NH to the adenine 7-N with 0.19 Å and 0.14 Å standard deviations, respectively. The adjacent triplets of these two central triplets are also in close accordance with their H-bonding data from their simulations as seen in Table IIIc.

The H-bonding distance data from the simulations of the B-type starting configuration for the X5·T-A centrally inserted triplex version are shown in Table IIId. As in simulations for the triplex pair of A-type starting configuration, the B-type unmodified $(T\cdot A\text{-}T)_{11}$ triplex and its centrally X5·T-A

TABLE IIIc Simulation averages and standard deviations for the H-bonds of the central three triplets of the A-type $(T.A\text{-}T)_5$-(X4.T-A)-$(T.A\text{-}T)_5$ triplex

Triplet-5: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Thy-O4–Ade-N6	2.90	0.13
Watson-Crick/I–II	Thy-N3–Ade-N1	2.91	0.13
Hoogsteen/III–II	Thy-O4–Ade-N6	3.10	0.31
Hoogsteen/III–II	Thy-N3–Ade-N7	2.93	0.20
Triplet-6: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Ade-N6–Thy-O4	2.89	0.10
Watson-Crick/I–II	Ade-N1–Thy-N3	2.97	0.18
Hoogsteen/III–II	X4-O2–Ade-N6	2.94	0.23
Hoogsteen/III–II	X4-N3–Thy-O4	2.95	0.27
Triplet-7: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Thy-O4–Ade-N6	3.03	0.22
Watson-Crick/I–II	Thy-N3–Ade-N1	2.94	0.15
Hoogsteen/III–II	Thy-O4–Ade-N6	2.94	0.16
Hoogsteen/III–II	Thy-N3–Ade-N7	2.90	0.12

TABLE IIId Simulation averages and standard deviations of the H-bonds of the central three triplets of the B-type $(T.A\text{-}T)_5$-(X5.T-A)-$(T.A\text{-}T)_5$ triplex

Triplet-5: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Thy-O4–Ade-N6	2.96	0.17
Watson-Crick/I–II	Thy-N3–Ade-N1	2.91	0.13
Hoogsteen/III–II	Thy-O4–Ade-N6	3.02	0.27
Hoogsteen/III–II	Thy-N3–Ade-N7	3.01	0.23
Triplet-6: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Ade-N6–Thy-O4	2.87	0.10
Watson-Crick/I–II	Ade-N1–Thy-N3	2.98	0.16
Hoogsteen/III–II	X5-O2–Ade-N6	3.31	0.35
Hoogsteen/III–II	X5-N3–Thy-O4	3.15	0.39
Triplet-7: Bond Type/ Strand Type	*H-bonding Interaction*	*Average H-bonding distance (Å)*	*Standard Deviation (Å)*
Watson-Crick/I–II	Thy-O4–Ade-N6	2.93	0.13
Watson-Crick/I–II	Thy-N3–Ade-N1	3.00	0.26
Hoogsteen/III–II	Thy-O4–Ade-N6	2.94	0.16
Hoogsteen/III–II	Thy-N3–Ade-N7	2.86	0.10

Downloaded At: 19:15 14 January 2011

modified version shown similar Watson-Crick base pair average H-bond distances for the central and adjacent triplets. However, the average distances of Hoogsteen H-bonds within the X5·T-A triplet are not as near to those of the central T·A-T triplet in the B-type triplex starting configuration simulation as the X4·T-A triplet Hoogsteen bond distances are to those of the central T·A-T triplet in the A-type triplex starting configuration simulation. From the preliminary energy minimization results, this result should not be too surprising owing to the weaker Hoogsteen interaction energy for X5 in the X5·T-A triplet in comparison to that in the central T·A-T triplet of the unmodified control triplex. However, the Hoogsteen H-bonded interactions of the adjacent triplets are unaltered with respect to those of the unmodified $(T \cdot A\text{-}T)_{11}$ triplex as indicated by average H-bond distances and deviations in Table IIId.

DISCUSSION

Although the X4 Hoogsteen base appears to be more successful as a Hoogsteen base designed to fit an A-type starting configuration than X5 is to fit a B-type starting configuration, these simulations just represent local configurational explorations of two configurations for a structure that in reality inevitably samples both. Whether DNA triplex in solution samples a significant portion of conformational space that is either A-type or B-type or mostly states between these two energetic minima [31] can best be answered by experimental studies (solution ^{1}HNMR NOESY, UV-T_m) of triplexes containing these X·T-A triplets. By incorporating Hoogsteen base probes such as these into triplex structures that barely perturb adjacent triplets, comparative studies of triplex stabilities eg. UV-T_m melting experiments would prove more definitive in detecting conformational preferences. Conformational analyses such as described in this paper are used to screen the configurational response of triplets in localized conformational space. Future studies should include sufficiently prolonged simulation times to allow both starting configurations to converge into a similar conformational range. In this manner not only may a more complete conformational assessment be achieved, but also a more accurate knowledge of the conformational requirements needed for designing unnatural Hoogsteen bases can be obtained. Due to the large time investment in running long simulations these preliminary studies just utilize the basic homogeneous A-type and B-type configurations. The studies performed here are a reference point to guide the further design of unnatural Hoogsteen bases for recognition of the

Downloaded At: 19:15 14 January 2011

T-A base pair. However, the final test must be synthesis and thermodynamic binding studies. These simulations should provide encouragement for experimental work.

Acknowledgement

J. H. R. is supported by a Hitchings-Elion postdoctoral fellowship from the Burroughs-Wellcome Fund.

References

[1] Strobel, S. A., Moser, H. E. and Dervan, P. B. (1988) Double-Strand cleavage of genomic DNA at a single site by triple helix formation", *J. Am. Chem. Soc.*, **110**, 7929–9.
[2] Povsic, T. J. and Dervan, P. B. (1989) Triple helix formation by oligonucleotides on DNA extended to the physiological pH range", *J. Am. Chem. Soc.*, **111**, 3059–61.
[3] Uhlmann, E. and Peyman, A. (1990) Antisense oligonucleotides: A new therapeutic principle", *Chem. Rev.*, **90**, 544–579.
[4] Pilch, D. S., Brousseau, R. and Shafer, R. H. (1990) Thermodynamics of triple helix formation", *Nucl. Acids Res.*, **18**, 5743–50.
[5] Pilch, D. S., Levenson, C. and Shafer, R. H. (1990) Structural analysis of the $(dA)_{10}$ $2(dT)_{10}$ triple helix", *Proc. Natl. Acad. Sci. USA*, **87**, 1942–6.
[6] Pilch, D. S., Levenson, C. and Shafer, R. H. (1991) Structure, stability, and thermodynamics of a short intermolecular purine-purine-pyrimidine triple helix", *Biochemistry*, **30**, 6081–7.
[7] Plum, E. G., Park, Y. W., Singleton, S. F., Dervan, P. B. and Breslauer, K. J. (1990) Thermodynamic characteristic of the stability and the melting behavior of a DNA triplex", *Proc. Natl. Acad. Sci. USA*, **87**, 9436–40.
[8] Moser, H. E. and Dervan, P. B. (1987) "Sequence-Specific cleavage of double helical DNA by triple helix formation", *Science*, **238**, 645–650.
[9] Strobel, S. A., Moser, H. E. and Dervan, P. B. (1988) Double-strand cleavage of genomic DNA at a single site by triple-helix formation", *J. Am. Chem. Soc.*, **110**, 7927–29.
[10] Le Doan, T., Perroualt, L., Praseuth, D. and Helene, C. (1987) Sequence Specific Recognition, photocrosslinking and cleavage of the DNA double helix by oligo [a] thymidylate covalently linked to an azidoproflavine derivative", *Nucleic Acid Res.*, **15**, 1749–60.
[11] Strobel, S. A. and Dervan, P. B. (1990) Site-specific cleavage of a yeast chromosome by oligonucleotide-directed triple-helix formation", *Science*, **249**, 73–5.
[12] Maher, L. J., Wold, B. and Dervan, P. B. (1989) Inhibition of DNA binding proteins by oligonucleotide-directed triple helix formation, *Science*, **245**, 725–30.
[13] Povsic, T. J. and Dervan, P. B. (1990) Sequence-specific alkylation of double-helical DNA by oligonucleotide-directed triple-helix formation", *J. Am. Chem. Soc.*, **112**, 9428–30.
[14] Cooney, M., Czernuszewicz, G., Postal, E. H., Flint, S. J. and Hogan, M. E. (1988) Site-specific oligonucleotide binding represses transcription of the human c-myc gene in vitro",*Science*, **241**, 456–9.
[15] Durland, R. H., Kessler, D. J., Gunnell, S., Duvic, M., Pettitt, B. M. and Hogan, M. E. (1991) Binding of triple helix forming oligonucleotides to sites in gene promoters",*Biochemistry*, **30**, 9246–55.
[16] Beal, P. A. and Dervan, P. B. (1992) "The influence of single base triplet changes on the stability of a pur.pur.pyr triple helix determined by affinity cleaving", *Nucleic Acids Res.*, **20**, 2773–6.
[17] Cheng, Y. K. and Pettitt, B. M. (1992) "Hoogsteen versus reversed-Hoogsteen base pairing: DNA triple helices", *J. Am. Chem. Soc.*, **114**, 4465–74.

Downloaded At: 19:15 14 January 2011

[18] Horne, D. A. and Dervan, P. B. (1990) "Recognition of mixed sequence duplex DNA by alternate -strand triple-helix formation", *J. Am. Chem. Soc.*,**112**, 2435–7.
[19] Arnott, S., Bond, P. J., Selsing, E. and Smith, P. J. C. (1976) "Models of triple-stranded polynucleotides with optimized stereochemistry", *Nucleic Acids Res.*, **3**, 10, 2459.
[20] Laughton, C. A. and Niedle, S. (1992) "Molecular dynamics simulation of the DNA triplex $d(TC)_5.h(GA)_5.d(C^+T)_5$", *J.Mol. Biol.*, **223**, 519.
[21] Laughton, C. A. and Niedle, S. (1991) "DNA triple helices. A molecular dynamics study ", *J. Chim. Phys.*, **88**, 2597–2603.
[22] Umemoto, K., Sarma, M. H., Gupta, G., Luo, J. and Sarma, R. H. (1992) "Structure and stability of a DNA triple helix in solution - NMR studies on $d(T)_6.d(A)_6.d(T)_6$ and its complex with a minor groove binding drug", *J. Am. Chem. Soc.*, **112**, 4539–4545.
[23] Radhakrishnan, I. and Patel, D. J. (1994) "DNA triplexes: Solution structures, hydrations sites, energetics, interactions, and function", *Biochemistry*, **33**, 1405–16.
[24] Radhakrishnan, I. and Patel, D. J. (1994) "Solution structure of a pyrimidine.purine.pyrimidine DNA triplex containing T. AT, C+GC and G.TA triplets", *Structure*, **2**, 17–32.
[25] Radhakrishnan, I. and Patel, D. J. (1994) "Solution structure and hydration patterns of a pyrimidine.purine.pyrimidine DNA triplex containing a novel T.CG base triple ", *J. Mol. Biol.*, **241**, 600–619.
[26] Thomas, G. A. and Peticolas, W. L. (1991) "A Raman study of low frequency intrahelical modes in A-DNA, B-DNA, and C-DNA", *J. Biomol. Struct. Dyn.*, **9**, 437–445.
[27] Howard, F. B., Miles, H. T., Liu, K., Frazier, J., Raghunathan, G. and Sasisekharan, V. (1992) "Structure of d(T)n.d(A)n.d(T)n the DNA triple helix has B-form geometry with C2'-endo sugar pucker", *Biochemistry*, **31**, 10671–7.
[28] Macaya, R. F., Schultze, P. and Feigon, J. (1992) "Sugar conformations in intramolecular DNA triplexes determined by coupling constants obtained by automated simulation of COSY cross peaks", *J. Am. Chem. Soc.*, **114**, 781–3.
[29] Raghunathan, G., Miles, H. T. and Sasisekharan, V. (1993) "Symmetry and molecular structure of a DNA triple helix $d(T)_n.d(A)_n.d(T)_n$", *Biochemistry*, **32**, 455–62.
[30] Liu, K., Miles, H. T., Parris, K. D. and Sasisekharan, V. (1994) "Fiber-type X-ray diffraction patterns from single crystals of triple helical DNA", *Nat. Struct. Biol.*, **1**, 11–12.
[31] Laughton, C. A. (1995) "Moleculalr dynamics simulations of DNA triple-helices. Does the triplex $d(A)_{10}.d(T)_{10}.d(T)_{10}$ have A-form or B-form geometry?", *Molecular Simulation*, **14**, 275–89.
[32] Griffin, L. C., Kiessling, L. L., Beal, P. A., Gillespie, P. and Dervan, P. B. (1992) "Recognition of All Four Base Pairs of Double-Helical DNA by Triple-Helix Formation: Design of Nonnatural Deoxyribonucleosides for Pyrimidine. Purine Base Pair Binding", *J. Am. Chem. Soc.*, **114**, 7976–82.
[33] Griffin, L. C. and Dervan, P. B. (1989) "Recognition of Thymine. Adenine Base Pairs by Guanine in a Pyrimidine Triple Helix Motif", *Science*, **245**, 967–71.
[34] Koshlap, K. M., Gilliespie, P., Dervan, P. B. and Feigon, J. "Nonnatural Deosyribonucleotide D3 Incorporated in an Intramolecular DNA Triplex Binds Sequence-Specifically by Intercalation", *J. Am. Chem. Soc.*, **115**, 7908–9.
[35] Ono, A., Ts'o P. O. and Kan, L. (1991) "Triplex Formation of Oligonucleotides Conataining 2'-O-Methylpseudoisocytidine in Substitution for 2'-Deoxycytidine", *J.Am. Chem. Soc.*, **113**, 4032–3.
[36] Koh, J. S. and Dervan, P. B. (1992) "Design of a nonnatural deoxyribosucleoside for recognition of GC base pairs by oligonucleotide-directed triple helix formation", *J. Am. Chem. Soc.*, **114**, 1470–8.
[37] Mohan, V., Cheng, Y. K., Marlow, G. E. and Pettitt, B. M. (1993) "Molecular recognition of Watson-Crick base-pair reversals in triple-helix formation: use of nonnatural oligonucleotide bases", *Biopolymers*, **33**, 1317–25.
[38] Rothman, J. H. and Richards, W. G. "Molecular dynamics simulations of novel Hoogsteen-like bases that recognize the T-A base pair by DNA triplex formation", *Biopolymers* (in press).
[39] Pearlman,D. A., Case, D. A., Caldwell, J.C., Siebel, G. L., Singh, U. C., Wiener, P. A. and Kollman, P. A. (1991) AMBER 4.1, University of California, San Francisco.
[40] QUANTA 4.0 (1994) Molecular Simulations, Inc., Burlington, MA 01803–5297

Downloaded At: 19:15 14 January 2011

[41] Gilson, M. K., Sharp, K. A. and Honig, B. (1987) "Calculating the electrostatic potentail of molecules in solution-method and error assessment", *J. Comp. Chem.*, **12**, 435.
[42] Stewart, J. J. P. (1990) "MOPAC - A semiemperical molecular-orbital program", *J. Comp. Aided Mol. Design*, **4**, 1.
[43] Frisch, M. J., Head-Gordon, M., Trucks, G. W., Foresman, J. B., Schlegel, H. B., Ragavachari, K., Robb, M., Binkley, J. S., Gonzalez, C., Defrees, D. J., Fox, D. J., Whiteside, R. A., Seeger, R., Melius, C. F., Baker, J., Martin, L. R., Kahn, L. R., Kahn, L. R., Stewart, J. J. P., Topiol, S. and Pople, J. A. GAUSSIAN 90, 1990, Revision J. Gaussian Inc., Pittsburgh, PA.
[44] Chirlian, L. E. and Francl, M. M. (1987) "Atomic charges derived from electrostatic potentials-a detailed study", *J. Comp. Chem.*, **8**, 894.
[45] Ryckaert, J. P., Ciccotti, G. and Berendsen, H. J. C. (1977) "Numerical intergration of the cartesion equations of motion of a system with constraints: molecular dynamics of n-alkanes", *J. Comput. Phys.*, **23**, 327.
[46] Radhakrishnan, I., Gao, X., de los Santos, C., Live, D. and Patel, D. J. (1991) "NMR structural studies of intramolecular $(Y^+)_n.(R^+)_n.(Y^-)_n$ DNA triplexes in solution-imino and amino proton and nitrogen markers of G·TA base triple formation", *Biochemistry*, **30**, 9022–9030.
[47] Radhakrishnan, I., Patel, D. J. and Gao, X. (1992) "3-dimensional homonuclear NOESY-TOCSY of an intramolecular pyrimidine.purine.pyrimidine DNA triplex containing a central G. TA triplex-nonexchangeable proton assignments and structural implications", *Biochemistry*, **31**, 2514–2523.
[48] Rothman, J. H. and Richards, W. G. (1995) "Novel nucleotide bases for DNA duplex recognition by triple helix formation", *J. Chem. Soc. Chem. Comm.*, 1589–90.

Downloaded At: 19:15 14 January 2011