A comparison of the solution structures of an LNA:DNA duplex and the unmodified DNA:DNA duplex

Gitte A. Jensen,[a] Sanjay K. Singh,[b] Ravindra Kumar,[b] Jesper Wengel[b] and Jens Peter Jacobsen*[a]

[a] *Department of Chemistry, University of Southern Denmark, Odense University, DK-5230 Odense M, Denmark. E-mail: jpj@chem.sdu.dk*
[b] *Center for Synthetic Bioorganic Chemistry, Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark*

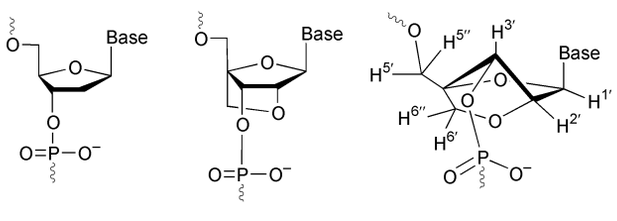
Received (in Cambridge, UK) 19th October 2000, Accepted 30th April 2001
First published as an Advance Article on the web 31st May 2001

Modified oligonucleotides, containing restricted nucleotides with a 2′-*O*,4′-*C*-methylene bridge (LNA), hybridized toward either DNA or RNA display an unprecedented increase in melting temperatures. In order to understand the structural basis for this high stability we have used ^{1}H NMR spectroscopy to determine the high resolution solution structures of an LNA-modified oligonucleotide, as well as the structure of the corresponding unmodified duplex. The modified duplex is an LNA:DNA duplex containing three thymidine LNA modifications ($\mathbf{T^L}$), d($C_1\mathbf{T^L}_2G_3A_4\mathbf{T^L}_5A_6\mathbf{T^L}_7G_8C_9$): d($G_{10}C_{11}A_{12}T_{13}A_{14}T_{15}C_{16}A_{17}G_{18}$). A full relaxation matrix approach by the program RANDMARDI was used to obtain interproton distance bounds from NOESY cross peak intensities. These distance bounds were used as restraints in molecular dynamics (rMD) calculations. Both duplexes have right-handed helix conformations with all bases in the *anti* conformation forming normal Watson–Crick base pairs. The LNA strand in the modified duplex has predominantly N-type sugar conformations compared to the S-type conformations of the complementary strand. The unmodified DNA:DNA strand has almost exclusively S-type sugar conformations. The structural strain introduced by the conformational changes of the ribose rings in the LNA:DNA duplex is released by unwinding the helix and widening the minor groove, but as a whole the structure of the duplex is surprisingly unaffected by introducing the modified LNA nucleotides.

Introduction

Modified oligonucleotides have for the last decade been acknowledged as potentially very useful in developing antisense agents and a vast number of chemically modified nucleotides have been synthesized.[1] The requirements for being a potent and nontoxic antisense oligonucleotide include stability towards cellular nucleases, effective delivery to cells *in vivo*, low toxicity, high affinity and sequence selectivity towards cognate RNA, and the formation of duplexes with the RNA that are amenable to RNase H mediated degradation.[2]

Recently, it was demonstrated[3] that a conformationally restricted DNA analog called "Locked Nucleic Acid" (LNA;[4,5] containing one or more 2′-*O*,4′-*C*-methylene linked ribofuranosyl nucleotide monomers, Scheme 1) actually fulfils most

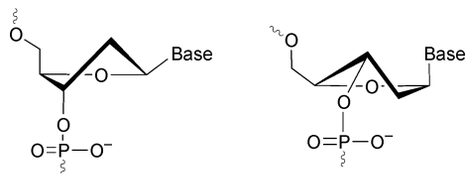


Scheme 1 Structures of a DNA and an LNA monomer. The locked conformation of an LNA monomer together with the numbering scheme is shown to the right.

of these requirements. Importantly, oligomerization of 3′-*O*-phosphoamidite LNA monomers proceeds efficiently on an automated DNA synthesizer using standard procedures.[4] Thermal denaturation studies of duplexes formed between LNAs and complementary DNA and RNA show that LNA recognizes both with remarkable affinities and selectivities. Thus, duplexes involving LNAs (hybridized towards either DNA or RNA) display an extremely large increase in melting temperatures of between +4.0 to +9.3 °C per modification compared to the corresponding unmodified reference duplexes.[4] Incorporation of a given number of LNA monomers into an oligonucleotide therefore appears to be a very convenient and predictable way of improving the stability of duplexes toward complementary DNA or RNA. In living MCF-7 breast cancer cells cellular uptake of a fully modified LNA 15-mer oligonucleotide has been demonstrated.[3] Unlike unmodified oligonucleotides, LNAs are not degraded in blood serum and cell extracts, but partly modified LNA-containing oligonucleotides were able to activate RNase H when hybridized to RNA.[3] LNA oligonucleotides exhibited potent antisense activity on assay systems as different as a G-protein-coupled receptor in living rat brain and an *E. coli* reporter gene.[3] At present, LNA is probably the most promising antisense candidate among modified oligonucleotides.

LNAs contain one or more 2′-*O*,4′-*C*-methylene linked bicyclic ribofuranosyl nucleotides locked in a C3′-*endo* (N-type) conformation (Scheme 1). In general, the conformations of the flexible deoxyribose rings influence the overall structure of a (deoxy)ribonucleic acid duplex. Duplexes in the A-type conformation contain nucleotides with an N-type (C3′-*endo*, Scheme 2) sugar conformation while B-type duplexes contain nucleotides with an S-type (C2′-*endo*, Scheme 2) sugar conformation.[6] In a recent publication,[7] we have used 2D NMR spectroscopy to study the sugar conformations of a single stranded LNA (ssLNA) oligomer and the corresponding unmodified single stranded DNA (ssDNA) as well as the sugar

 DOI: 10.1039/b008431j



Scheme 2 S- and N-type nucleotide sugar conformations.

conformation of three LNA:DNA duplexes with different compositions and a different number of modifications. The results obtained for this partly modified LNA indicate that the ssLNA is preorganized and suggest that the increased stability of the LNA:DNA duplex may, at least partly, be explained by conformational changes from C2′-*endo* to C3′-*endo* of the LNA-nucleotide.

In an attempt to further unravel the structural changes at the atomic level in a double stranded DNA (dsDNA) duplex, when introducing LNA nucleotide monomers, we have used 2D ^{1}H NMR spectroscopy to determine the solution structure of both the unmodified DNA:DNA duplex as well as the LNA:DNA duplex containing three thymidine LNA modifications ($\mathbf{T^L}$), d(**CT^L**GA**T^L**A**T^L**GC):d(GCATATCAG) duplex (called the "LNA:DNA duplex" in the following, Scheme 3). A total

1 2 3 4 5 6 7 8 9
5′ - C *T^L* G A *T^L* A *T^L* G C - 3′
3′ - G A C T A T A C G - 5′
18 17 16 15 14 13 12 11 10

Scheme 3 The numbering scheme used for the LNA:DNA duplex (and similarly for the DNA:DNA).

relaxation matrix approach was used to obtain interproton distance bounds from NOESY cross peak intensities.[8] These distance bounds were used as restraints in molecular dynamics (rMD) calculations. Since many NOE contacts were observed, the resulting structures have high resolution and allowed the determination of the changes in the local conformational features when introducing the LNA modifications in a dsDNA duplex.

Results

Spectral analysis

The 1D ^{1}H NMR spectra of the samples of both the LNA:DNA duplex and the DNA:DNA duplex consist of sharp lines. The NOESY spectra of the duplexes exhibit the characteristic features of dsDNA sequential connectivities. Parts of the NOESY spectrum of the two duplexes with 200 ms mixing time are shown in Fig. 1. The assignments of the nonexchangeable protons in the modified duplex were performed using standard methods.[9] Aromatic (H6, H5, H8, and H2) and sugar protons (H1′, H2′, H2″, H3′, H4′, and H5′, H5″) were assigned. The chemical shift values and the NOE connectivity pattern of the imino protons were observed to be in accordance with normal Watson–Crick base pairing in both duplexes. The NOESY spectrum with short mixing time (50 ms) allowed unambiguous assignments of the H2′ and H2″ resonances. The assignment of the exchangeable protons was obtained from the NOESY spectrum in H_2O.[10] The chemical shift values are given in Table 1 and compared to the corresponding values of the unmodified duplex.

Backbone and sugar conformation

The *J*-scaled ^{1}H–^{31}P HMBC spectra were used to measure the three-bond coupling constants between H3′(*i*) and P5′(*i* + 1). In the case of the LNA:DNA duplex the values obtained were in the interval between 7 and 8 Hz except for the end-nucleotides C9 and G18 which had values between 4 and 6 Hz. However, no cross peaks from phosphorus nuclei to the H2″ region of the spectrum of the LNA:DNA duplex were observed, so the values of the dihedral angles, ε, in the phosphate backbone were assumed in this duplex to be in the interval from 150 to 210° (the *trans* domain). Due to spectral overlap no reliable coupling constants at all were obtained for the DNA:DNA duplex.

The sugar conformations of the LNA:DNA duplex were determined earlier.[7] The sugar conformations of the DNA:DNA duplex were determined in this work in a similar way. The nucleotides C1(71%), G3(75%), T7(84%), G8(74%), C9(85%), G10(90%), **T**15(87%), C16(82%), A17(83%) and G18(88%) in the DNA:DNA duplex were shown to have predominantly the S-type sugar conformation with the fractions of the S-type conformation given in parentheses. Due to spectral overlap it was not possible to determine the precise sugar conformations of the remaining nucleotides in the DNA:DNA duplex, but qualitatively the cross-peak patterns in the DQF COSY spectra of these nucleotides were also in agreement with a predominant S-type conformation.

Structure calculations

More than 800 NOE cross peaks were observed in the NOESY spectrum obtained with a mixing time of 200 ms in both duplexes. The accuracy of the integration of some cross peaks was hampered by spectral overlap. These cross peaks were therefore not included in the RANDMARDI calculations. Integrations of the cross peaks in the four NOESY spectra of each duplex were done separately for each side of the diagonal. Some of the cross peaks used in the structure calculations resulted predominantly from spin diffusion. This is taken into account during the statistical analysis of the results generated by the RANDMARDI procedure.[8a] Cross peak integrals that corresponded to fixed distances in the dsDNA were used for internal calibrations in MARDIGRAS, and therefore not converted into distance restraints for use in the rMD simulations. The MARDIGRAS calculations were performed by using a value of the correlation time of $\tau_c = 2$ ns. This value was justified from earlier work.[11,12] Furthermore, it turns out that the results of the calculations are rather insensitive to the exact values of τ_c used. Calculations with various values of τ_c in the interval from 1.5 to 7.0 ns yielded exactly the same structure but with slightly larger RMSD deviations. This insensitivity of the calculations towards values of τ_c is similar to what has been found earlier.[11–13]

The RANDMARDI calculations returned 389 interproton distances in the case of the LNA:DNA duplex and 347 interproton distances in the case of the DNA:DNA duplex. The distance bounds used in rMD simulations were determined by combining the results from all of the individual MARDIGRAS calculations performed during the RANDMARDI procedure into one set. They were calculated individually for each proton pair corresponding to a NOESY cross peak included in the RANDMARDI calculations. Each individual restraint was generated from the MARDIGRAS calculations. Fig. 2 gives a summary of the distribution of the distance restraints that were actually used in the rMD calculations. The distance bounds involving the methyl groups were omitted for both duplexes and the experimental distance bounds for the end nucleotides in the DNA:DNA duplex were substituted by torsional restraints involving the α, β, γ and ζ dihedral angles in order to avoid the influence of dynamic motions on the structure. The range of differences between the upper and lower bounds for the NOE derived restraint bounds varied, but 50% were smaller than 0.6 Å in the case of the LNA:DNA duplex and 50% were smaller than 0.5 Å in the case of the DNA:DNA duplex.

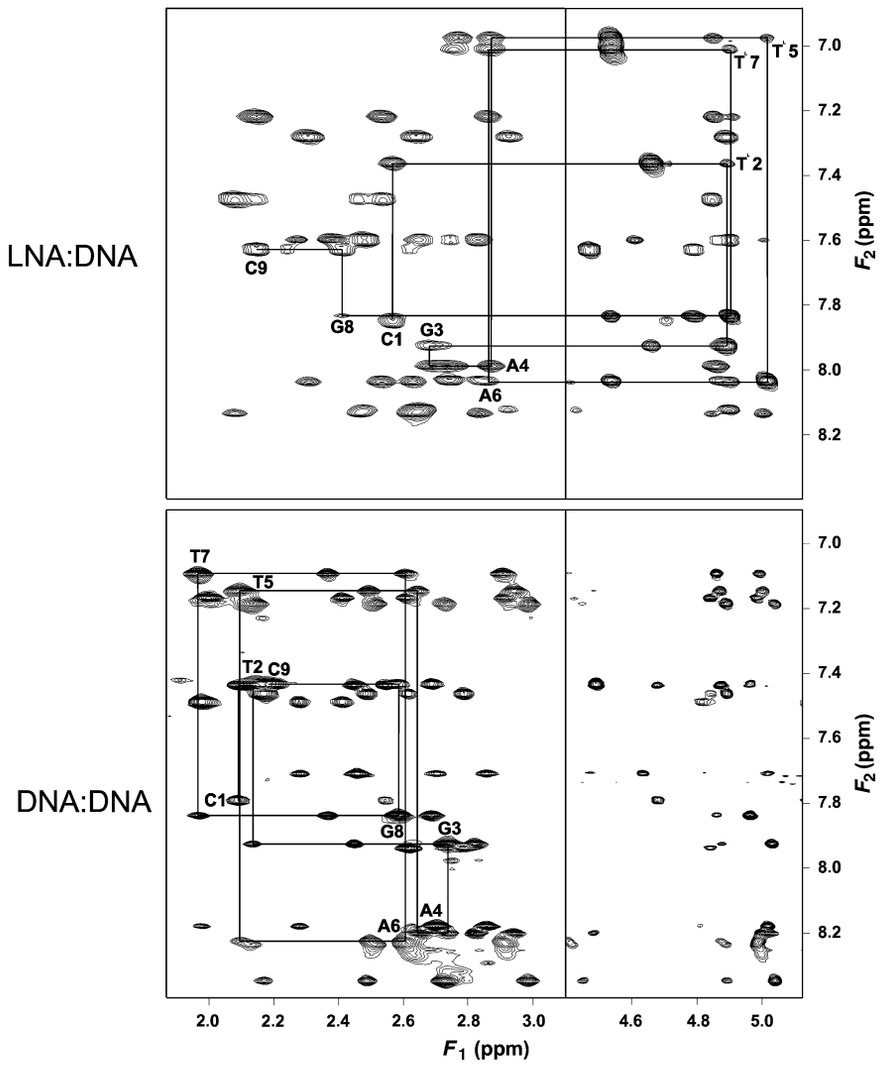


Fig. 1 Parts of the 200 ms NOESY spectrum of the d(C**T**^{L}G**A**^{L}**A**^{L}**T**^{L}GC):d(GCATATCAG) duplex (*top*) and the unmodified duplex (*bottom*). The sequential connectivity pattern H6/8 and H2′ on the modified strand is indicated with a full line in the top spectrum. The corresponding connectivity pattern for the unmodified duplex is shown below. Some of the strong cross peaks in the top spectrum are between H3′ and H6/8, indicative of C3′-*endo* conformations.

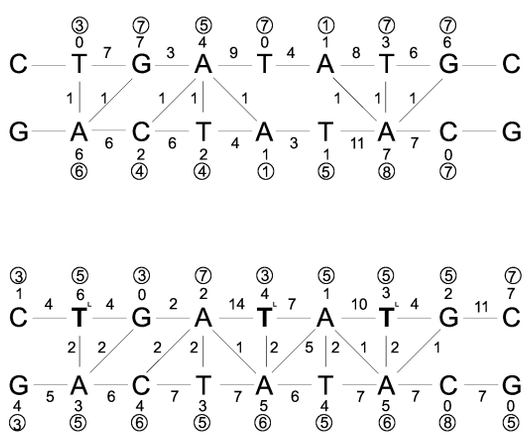


Fig. 2 Number of distance restraints obtained from RANDMARDI calculations and used in the rMD calculations of the d(C**T**^{L}G**A**^{L}-**A**^{L}**T**^{L}GC):d(GCATATCAG) duplex (*bottom*) and the unmodified duplex (*top*). The distances are grouped as intra- and inter-nucleotide distances. The intra-nucleotide distances are further grouped as sugar–sugar distances and sugar–base distances (marked with a circle).

Evaluation of the spectra obtained in H_2O of both duplexes shows normal Watson–Crick hydrogen bonding, justifying the inclusion of 22 hydrogen bond distance restraints. Three hydrogen bonds were included for each of the six GC base pairs and two hydrogen bonds were included for each of the four AT base pairs with upper and lower bounds of 1.74 Å and 2.10 Å, respectively. Additional restraints with loose bounds involving labile protons were derived using the isolated spin pair approximation on the NOESY spectra recorded of the sample in H_2O.

The sugar conformations of the LNA:DNA duplex were reported earlier[7] as a mixture of N-type and S-type. The DNA:DNA duplex has in this work been shown to have predominantly the S-type sugar conformation but with non-negligible N-type contributions. This mixed situation forced us not to include any dihedral angle restraints of the sugar ring in the rMD calculations. Doing that, it would have been necessary to include a description of the dynamic ring puckering motions of the sugars. Such a description is not available. However, we did include restraints for the dihedral angle ε in the phosphate backbone according to the observed values of the corresponding coupling constants.

Ten final structures for each of the two starting structures were generated by rMD calculations using the simulated annealing protocols described in EXPERIMENTAL (*vide infra*). All the structures converged to one family of conformations. The average values of all the pair-wise root-mean-square deviations (RMSD) between all the structures obtained including all atoms are given in Table 2. Obviously, the main contribution to the RMSD arises from the end base pairs. Taking only the five inner base pairs into account, both duplexes converged to an RMSD of approximately 0.6 Å for both starting models.

Table 1 Chemical shift values (in ppm) in the d(CT^{L}GAT^{L}AT^{L}GC):d(GCATATCAG) duplex [a]

	H6/H8	H5/H2/CH_3	H1′	H2′	H2″	H3′	H4′	H5′	H5″	H1/H3	*H*4	H4
C1	7.85	5.92	5.97	2.57	2.57	4.71	4.13	3.80	3.80		7.08	8.06
	(7.79)	(5.91)	(5.91)	(2.09)	(2.54)	(4.69)	(4.10)	(3.76)	3.76		(7.05)	(7.86)
T^{L}2	7.36	1.69	5.41	4.89	—	4.66	—	4.31	4.35	13.68		
	(7.44)	(1.69)	(5.70)	(2.14)	(2.45)	(4.87)	(4.15)	(4.06)	(4.04)	(13.82)		
G3	7.92		6.01	2.68	2.68	4.89	4.39	4.31	4.17	12.09		
	(7.93)		(5.64)	(2.73)	(2.82)	(5.03)	(4.37)	(4.26)	(4.06)	(12.55)		
A4	7.99	7.69	6.18	2.87	2.77	4.85	4.38	4.25	4.21			
	(8.20)	(7.72)	(6.23)	(2.64)	(2.94)	(5.00)	(4.48)	(4.26)	(4.14)			
T^{L}5	6.97	1.40	5.32	5.01	—	4.54	—	4.35	4.35	13.17		
	(7.15)	(1.42)	(5.68)	(2.10)	(2.49)	(4.87)	(4.16)	(4.26)	(4.15)	(13.19)		
A6	8.04	6.98	6.19	2.87	2.76	4.54	4.29	—	—			
	(8.22)	(7.21)	(6.22)	(2.60)	(2.91)	(4.99)	(4.41)	(4.15)	(4.20)			
T^{L}7	7.01	1.26	5.26	4.90	—	4.54	—	4.16	4.27	13.55		
	(7.10)	(1.35)	(5.76)	(1.98)	(2.37)	(4.86)	(4.17)	(4.16)	(4.27)	(13.53)		
G8	7.83		6.06	2.41	2.62	4.79	4.22	4.11	4.30	12.65		
	(7.85)		(5.92)	(2.60)	(2.70)	(4.97)	(4.35)	(4.25)	(4.09)	(12.74)		
C9	7.63	5.41	6.18	2.15	2.25	4.47	4.03	4.16	4.02		6.88	8.22
	(7.43)	(5.39)	(6.17)	(2.19)	(2.19)	(4.49)	(4.06)	(4.23)	(4.23)		(6.62)	(8.11)
G10	8.03		6.05	2.75	2.83	4.86	4.38	3.80	3.80	—		
	(7.94)		(5.97)	(2.62)	(2.80)	(4.85)	(4.24)	(3.72)	(3.72)	(—)		
C11	7.60	5.45	6.06	2.49	2.65	4.90	4.34	4.24	4.24		6.77	8.63
	(7.47)	(5.42)	(5.70)	(2.18)	(2.49)	(4.90)	(4.22)	(4.14)	(4.14)		(6.51)	(8.38)
A12	8.13	7.44	6.28	2.65	2.93	4.90	4.43	4.28	4.28			
	(8.35)	(7.66)	(6.30)	(2.73)	(2.98)	(5.04)	(4.45)	(4.12)	(4.21)			
T13	7.28	1.35	5.92	2.31	2.63	4.89	4.27	4.28	4.21	13.06		
	(7.19)	(1.51)	(5.69)	(2.13)	(2.51)	(4.89)	(4.21)	(4.29)	(4.18)	(13.16)		
A14	8.04	7.13	6.21	2.54	2.86	4.91	4.42	—	—			
	(8.23)	(7.19)	(6.22)	(2.61)	(2.91)	(4.99)	(4.41)	(4.28)	(4.41)			
T15	7.22	1.17	5.96	2.15	2.54	4.85	4.24	4.26	4.16	13.41		
	(7.17)	(1.32)	(5.91)	(2.00)	(2.40)	(4.85)	(4.18)	(4.14)	(4.29)	(13.53)		
C16	7.48	5.54	5.84	2.08	2.47	4.85	4.17	4.16	4.16		6.69	8.40
	(7.49)	(5.63)	(5.41)	(1.97)	(2.27)	(4.81)	(4.07)	(4.07)	(4.07)		(6.72)	(8.46)
A17	8.14	7.40	6.05	2.65	2.84	5.00	4.39	4.17	4.10			
	(8.18)	(7.74)	(6.02)	(2.70)	(2.85)	(5.02)	(4.38)	(4.01)	(4.11)			
G18	7.60		6.05	2.38	2.26	4.61	4.19	4.22	4.13	—		
	(7.72)		(6.02)	(2.46)	(2.27)	(4.63)	(4.17)	(4.24)	(4.12)	(—)		

[a] The values for the unmodified duplex are given in parentheses. The values are given at 25 °C relative to DSS. The protons in the C2′,C4′ linker have the following chemical shift values: **T^{L}**2 H6′/H6″: 3.99/4.14; **T^{L}**5 H6′/H6″: 4.10/4.15; **T^{L}**7 H6′/H6″: 4.00/4.09.

Table 2 The root mean square deviation (RMSD in Å) obtained in the rMD calculations of the LNA:DNA and the DNA:DNA duplexes by using different starting models. All calculations converged for each of the two duplexes to the same family of structures

Starting model	Number of base pairs [a]	LNA:DNA	DNA:DNA
	9	1.02	0.80
	7	0.84	0.68
A-model	5	0.68	0.60
	9	0.83	1.05
	7	0.69	0.78
B-model	5	0.60	0.61

[a] Number of base pairs used in the fitting to calculate the RMSD. The outer base pairs are the ones that have been omitted in some of the calculations.

The convergence of the structure was further improved by including backbone restraints on the dihedral angles α, β, γ and ζ that are in accordance with both A-type and B-type models as described by Aramini.[14] A superposition of the structures of the d(CT^{L}GAT^{L}AT^{L}GC):d(GCATATCAG) duplex (*in dark gray*) and the unmodified duplex (*in light gray*), obtained without backbone restraints, is shown in Fig. 3, including only the bases of the (5′-GAT^{L}AT^{L}-3′):(5′-ATATC-3′) part for clarity.

The reliability of the structure calculations was carefully evaluated by changing the values of the force constants in the pseudo-energy terms and removing particular distance restraints.

Helix parameters

Helical parameters for the 20 final structures were analysed with the program CURVES 5.1.[15] Plots of some of the global helical parameters for the LNA duplex are shown Fig. 4.

Discussion

Spectral data

The chemical shift values given in Table 1 demonstrate close agreement between the modified and unmodified duplex except for the protons near the modification sites. As expected from the locked configurations, the chemical shifts of the protons in the LNA nucleotides differ from those in the unmodified nucleotides. The protons on the other nucleotides show only small differences in chemical shift values between the modified and the unmodified duplex, indicating that only minor differences between the two structures are to be expected. In particular, the small differences between the chemical shift values of the adenine H2 indicate that no major changes in the stacking of the bases occur. This is in contrast to what was observed for the d(CCGCT^{L}AGCG):d(CGCTAGCGG)[12b] and the d(CT^{L}-GCT^{L}T^{L}CT^{L}GC):d(GCAGAAGCAG)[12a] duplexes for which considerable changes of these chemical shift values indicated that a change of the stacking had occurred upon modification.

A particularly important observation is related to the chemical shift values of the imino protons. These are, within ±0.15 ppm, identical in the two duplexes, which implies that no major changes in hydrogen bonding and stacking have occurred. This is also in contrast to observations made for the

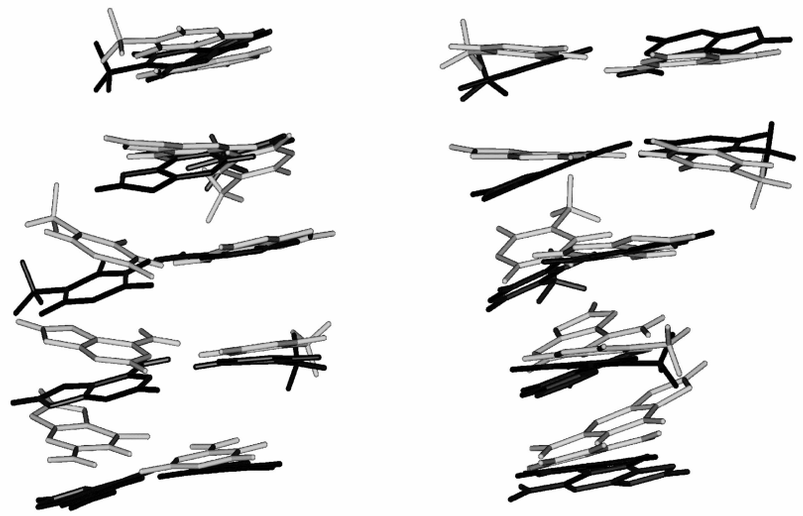


Fig. 3 A superposition of the bases of the (5′-GA**T**^{L}A**T**^{L}-3′):(5′-ATATC-3′) part of the d(C**T**^{L}GA**T**^{L}A**T**^{L}GC):d(GCATATCAG) duplex (*in dark gray*) and the unmodified duplex (*in light gray*). The overlaid structures were both randomly chosen from the 20 calculated structures for each of the two duplexes, and the Cartesian coordinates of the atoms in five bases A12–C16 were fitted to each other using a root-mean-square deviation protocol. The base pairs T7–A12 and **T**^{L}7–A12, respectively, are seen at the top while the base pairs G3–C16 from both duplexes are seen at the bottom. On the left the perspective is directly into the minor groove defined by the three lower base pairs. The drawing on the right is rotated approximately 90° clockwise around the helix axis compared to the drawing on the left, so the perspective of this drawing is directly into the minor groove defined by the two upper base pairs.

d(CCGC**T**^{L}AGCG):d(CGCTAGCGG)[12b] and the d(C**T**^{L}GC-**T**^{L}**T**^{L}C**T**^{L}GC):d(GCAGAAGCAG)[12a] duplexes.

Description of the structures

Qualitatively, the NOESY spectra indicate that both duplexes adopt right-handed helix conformations. All bases are in the *anti* conformation and they all form normal Watson–Crick base pairs. The sugar conformations in the DNA:DNA duplex are, as expected, predominantly S-type for all the nucleotides. The LNA nucleotides are locked in a C3′-*endo* (N-type) conformation. The conformations of the other nucleotides in the LNA:DNA duplex are predominantly S-type, except for the nucleotides A4, A6 and G8 which have a substantial contribution from the N-type conformation.[7] The nucleotide A6 is in an almost completely N-type conformation, whereas the other two nucleotides have mixed conformations. The differences in the sugar conformations of the LNA:DNA and the unmodified DNA:DNA duplex indicate that a change toward A-type duplex structure occurs when introducing LNA modifications. This change is further illustrated by comparing the width of the minor groove of the two structures as shown in Fig. 5. In canonical B-type the minor groove width is 5.7 Å (the shortest P–P distance is 11.5 Å), while it is 11.0 Å (the shortest P–P distance is 16.8 Å) in canonical A-type DNA. Clearly, the unmodified DNA:DNA duplex is typical B-type dsDNA whereas the LNA:DNA duplex is much more A-type DNA.

The distance between the phosphorus atoms on adjacent nucleotides on the same strand is 5.9 Å in canonical A-type dsDNA and 7.0 Å in canonical B-type DNA. The values obtained in the two duplexes are shown in Fig. 6. It is obvious that the DNA:DNA duplex with respect to this distance resembles a B-type dsDNA whereas a change toward A-type dsDNA has occurred in the LNA:DNA duplex. It is also obvious that this change is local and not uniform since the **T**^{L}2 modification does not seem to induce a major difference.

Experimentally derived NMR restraints are distributed anisotropically and are short range in nature. This largely establishes local structures with some helical parameters better defined by the experimental parameters than others. We have found a number of NMR-derived inter-strand restraints that improved the ability to define the structure. The helix parameters calculated by CURVES largely establish that the two structures are irregular duplexes. Except for *Twist*, *Inclination* and *Buckle*, no other single helix parameter deviates significantly between the two structures. Consequently, most of the helix parameters are not useful in the attempt to describe the structural changes that occur when introducing modified LNA duplexes. It is especially surprising that the values of *Rise* (the distance between adjacent base pairs) were calculated to be in between the values of A-type end B-type dsDNA (2.7 and 3.4 Å) for the base pairs in both duplexes. This is in accordance with the finding that there is no significant change in chemical shift values of the aromatic protons between the two duplexes, but is in contradiction to what was found for the d(CCGC-**T**^{L}AGCG):d(CGCTA-GCGG)[12b] and the d(C**T**^{L}GC**T**^{L}**T**^{L}-C**T**^{L}GC):d(GCAGAAGCAG)[12a] duplexes. In particular, the latter LNA duplex containing four modifications exhibits low values of the *Rise* parameter in agreement with a more A-type like DNA structure.

The average values of *Twist* are 34.2 ± 1.7° in the case of the DNA:DNA duplex and 30.7 ± 1.5° in the case of the LNA:DNA duplex. This establishes that an unwinding of approximately 40° on a complete helix turn is the result of introducing the modifications. The average twist in the d(CCGC**T**^{L}AGCG):d(CGCTAGCGG)[12b] duplex was found to be 35.9 ± 0.3°, close to canonical B-DNA, and 32 ± 1° in the d(C**T**^{L}GC**T**^{L}**T**^{L}C**T**^{L}GC):d(GCAGAAGCAG)[12a] duplex. Although the latter LNA:DNA duplex contains four modifications the unwinding is smaller than in the LNA:DNA duplex studied in this work, but in both cases the average twist is close to the values in canonical A-DNA of 32.7°.

A comparison between the structures of the LNA duplexes examined shows that a change of the sugar conformations of the modified strand towards the N-type conformation is a common trend. The structural strain that these conformational changes introduce is partly released by unwinding the helix and widening the minor groove. Furthermore, there seems to be a sequence specific possibility of further relaxing the structure by either decreasing the distance between adjacent base pairs or adjusting the *Tip, Inclination* and *Buckle*. However, the duplex structures as a whole are surprisingly unaffected by introducing the modified LNA nucleotide.

The stability of the LNA:DNA duplex

In the d(G**T**^{L}GA**T**^{L}A**T**^{L}GC):d(GCATATCAC) duplex the

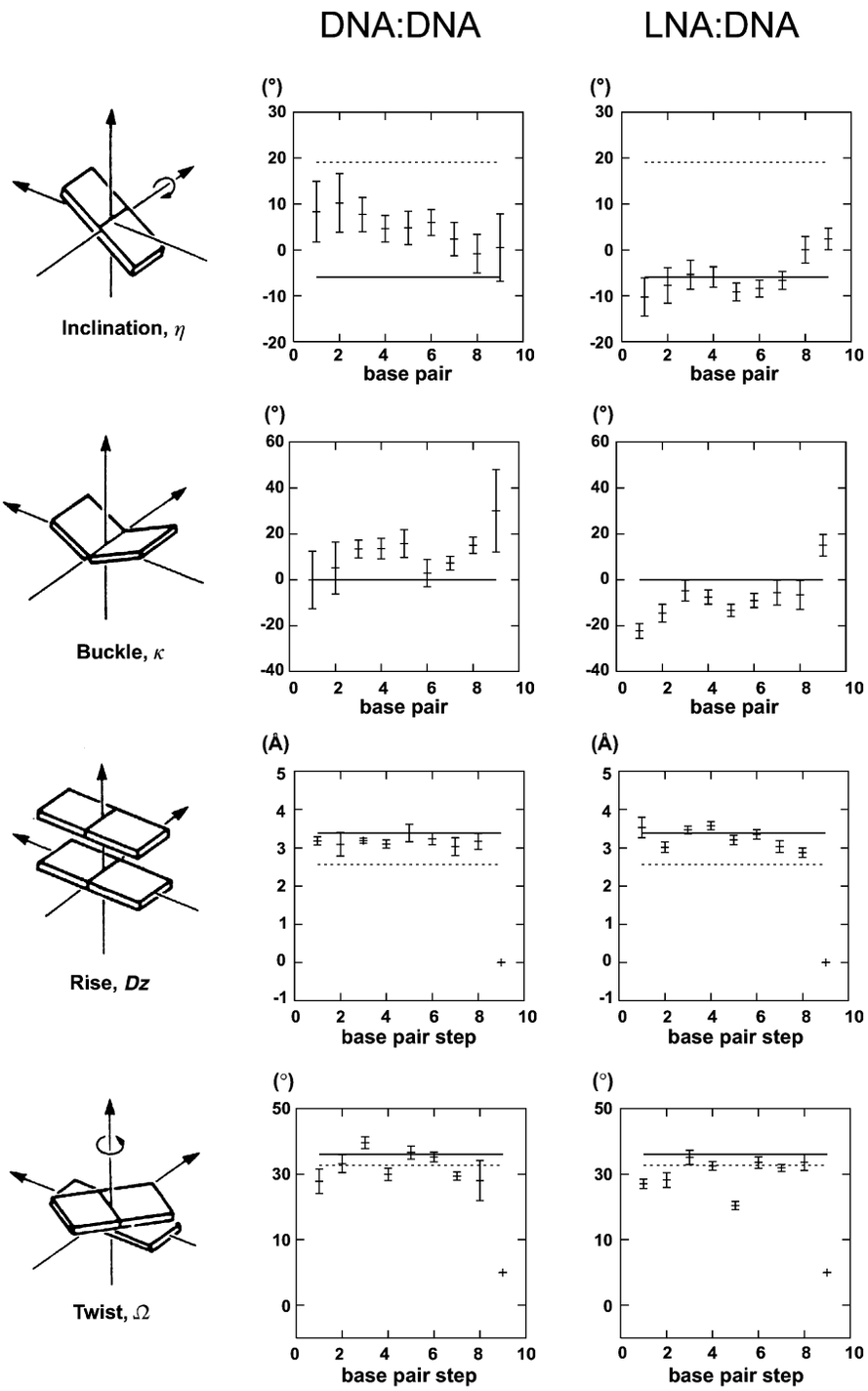


Fig. 4 A few of the helical parameters for the d(CT^{L}GAT^{L}AT^{L}GC):d(GCATATCAG) duplex (*right*) and the unmodified duplex (*left*) calculated using CURVES and compared with canonical A-DNA (--------) and canonical B-DNA (——). The helical parameters not shown do not deviate significantly between the two duplexes.

melting temperature is increased by 5.3 °C per modification compared to the unmodified DNA duplex. This huge increase in melting temperature implies that the LNA:DNA duplex is thermodynamically much more stable than the DNA:DNA duplex. However, this increased stability is not clearly reflected in a large difference between the dynamically averaged structures obtained for the two duplexes.

Earlier, we found that the conformations of the nucleotides in the ssLNA d(CT^{L}GAT^{L}AT^{L}GC) had larger fractions of N-type conformation than in the unmodified ssDNA, and that this ssLNA oligonucleotide was preorganized for the formation of the LNA:DNA duplex.[7] This implies that a smaller change of entropy has to be paid in the duplex formation. However, there are no changes in the structure of the d(CT^{L}GAT^{L}-AT^{L}GC):d(GCATATCAG) duplex that indicate an increased stability due to improved base pair stacking or hydrogen bonding. Consequently, it is tempting to conclude that the stabilisation of the LNA:DNA duplexes is exclusively entropy related.

However, it does not seem obvious that the entropy effect is solely able to account for the extremely large stabilizing effect observed in LNA:DNA duplexes. A combined contribution to the stability of the LNA:DNA duplexes from both preorganization and improved stacking was the most obvious explanation that was deduced from the structural work published earlier.[12] This was supported by the inclusion of this abasic modified nucleotide in an oligonucleotide using melting temperature measurement to conclude that the nucleobases are of pivotal

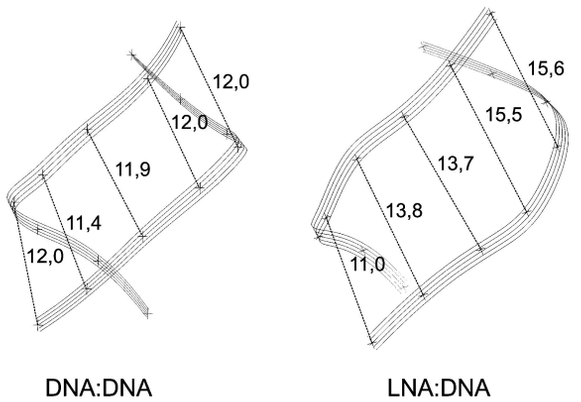


Fig. 5 The P–P distances across the minor groove in the d(C**T**^{L}G**A**^{L}-**A**^{L}**T**^{L}GC):d(GCATATCAG) duplex (*right*) and the unmodified duplex (*left*). The minor groove width is the P–P distance minus the van der Waals radius of the phosphate group (5.8 Å).[6b]

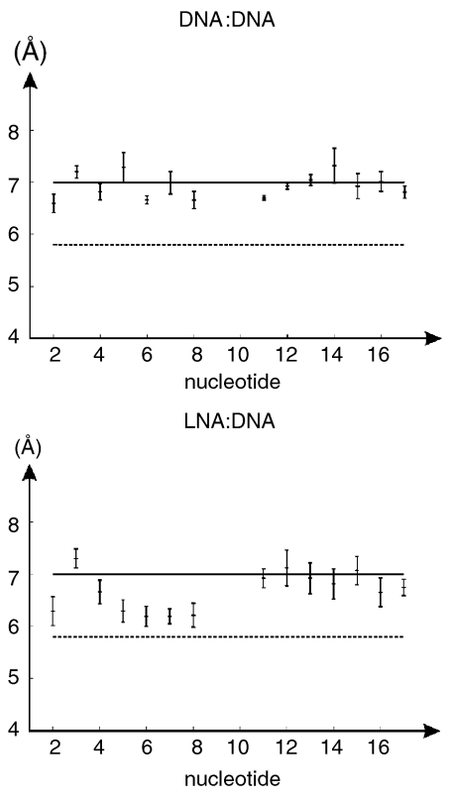


Fig. 6 The intra-strand P–P distances between P5′(*n*) and P3′(*n* + 1) in the d(C**T**^{L}G**A**^{L}**A**^{L}**T**^{L}GC):d(GCATATCAG) duplex (*bottom*) and the unmodified duplex (*top*). The distances are averaged values of the 20 structures calculated for each duplex. The standard deviations are indicated by bars. The corresponding values of the canonical A-DNA (---------) and the canonical B-DNA (——) are indicated.

importance for the stability of the LNA duplexes.[16] However, the structure comparisons between the LNA:DNA and DNA:DNA duplexes in this work do not confirm that the LNA modifications introduce an increased stacking in the duplex. Actually, except for the change in sugar conformation there are only small differences in the structures of the LNA:DNA and the unmodified duplex and these differences are mostly related to an unwinding of the helix and a widening of the minor groove. This finding makes it plausible that it is necessary to include the hybridized water in order to fully understand the stability of the LNA:DNA duplexes on an atomic level. But such a description may also include the dynamics of the LNA:DNA duplexes and is beyond the capability of structural work that is possible at the moment.

Experimental

Sample preparation

The d(C**T**^{L}G**A**^{L}**A**^{L}**T**^{L}GC) oligonucleotide was synthesized as described elsewhere.[4] The unmodified oligonucleotides were purchased from DNA Technology, Århus, Denmark. The oligonucleotides were purified by site-exclusion on a Sephadex G15 column. The samples of both the LNA:DNA and the DNA:DNA duplexes were obtained by dissolving an equimolar amount of the two single strands in 0.5 mL of 10 mM sodium phosphate buffer (pH 7.0), 100 mM NaCl, 0.05 mM NaEDTA, 0.01 mM NaN_3 and 0.1 mM sodium 3-(trimethylsilyl)propane-1-sulfonate (DSS). For experiments carried out in D_2O the duplex solutions were lyophilized three times from D_2O and redissolved in 99.96% D_2O (Cambridge Isotope Laboratories). Mixtures of 90% H_2O and 10% D_2O (0.5 mL) were used for experiments examining exchangeable protons. The mixtures were heated to 80 °C and slowly cooled to achieve hybridization. The final concentrations of the duplexes were 4 mM in each sample.

NMR experiments

NMR experiments were performed on a Varian UNITY 500 spectrometer or a Varian INOVA 800 spectrometer at 25 °C. NOESY spectra with mixing times of 50, 100, 150 and 200 ms were acquired of the LNA:DNA sample in D_2O at 500 MHz using 512 t_1-experiments of 64 scans, 1024 complex points in t_2, a pulse repetition time of 3.4 s and a spectral width of 5000 Hz. NOESY spectra of the DNA:DNA sample in D_2O were acquired with the same mixing times at 800 MHz using 800 t_1-experiments of 32 scans, 2048 complex points in t_2, a pulse repetition time of 4.8 s and a spectral width of 8000 Hz. The States phase cycling scheme was used in both cases and the residual signal from HOD was removed by low-power presaturation. The NOESY spectra in H_2O with 200 ms mixing time were acquired using the WATERGATE NOESY pulse sequence using 2048 complex points with a spectral width of 16000 Hz at 800 MHz and 10000 Hz at 500 MHz.

TOCSY spectra with mixing times of 30, 60 and 90 ms were obtained in the TPPI mode at 500 MHz. Gradient enhanced ^{1}H–^{13}C HSQC spectra were obtained at 500 MHz. Inversion recovery experiments to estimate the T_1 relaxation rates were also obtained for both samples at the relevant resonance frequencies. DQF-COSY spectra of the DNA:DNA duplex were recorded and analyzed similarly to the spectra of the LNA:DNA duplex as described earlier.[7] *J*-scaled ^{1}H–^{31}P HMBC spectra[17] were obtained with the scaling parameter $\kappa = 5$ in 56 t_1 experiments using 2048 complex points in t_2. A gradient selection of both N-type and P-type spectra were used with a spectral width of 600 Hz in the ^{31}P domain.

The acquired data were processed at optimal conditions using FELIX (version 97.2, MSI, San Diego, CA).

Distance restraints

The NOESY build-up experimental data set for each sample was further processed in the following way: the upper and lower diagonal part of each of the four NOESY build-up spectra were integrated separately with FELIX, yielding a total of eight peak intensity sets that were corrected for minor saturation effects. The RANDMARDI procedure[8a] of the complete relaxation matrix analysis method MARDIGRAS[8b] was used to calculate interproton distance bounds from the eight intensity sets. In the calculations, an absolute noise level of the same order of magnitude as the smallest integrated cross peak was used. In the RANDMARDI procedure, 30 different intensity sets were generated from each experimental data set based on the given noise levels, and MARDIGRAS calculations were performed on all of them. Resulting distances from all eight

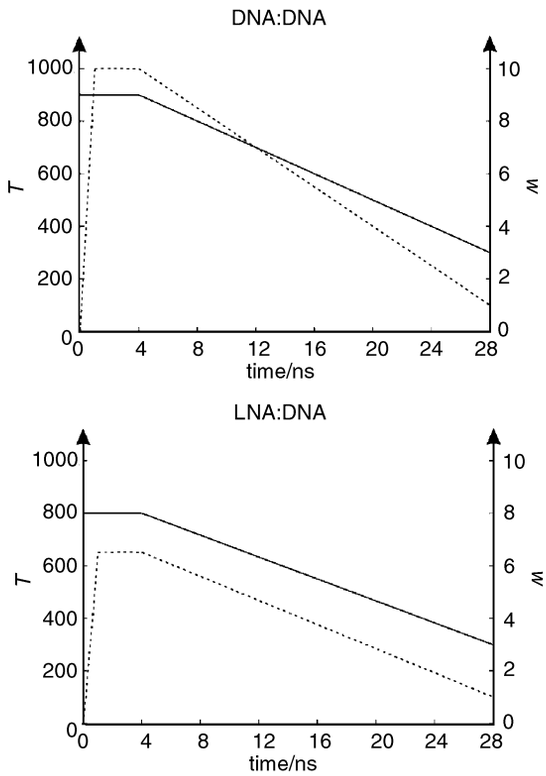


Fig. 7 The simulated annealing protocols used in the structure determination of the d(C**T**^{L}G**A**^{L}**T**^{L}**A**^{L}G**C**):d(GCATATCAG) duplex (*bottom*) and the unmodified duplex (*top*). The solid lines are the temperatures as a function of time. The dashed lines are the weights of the force constants as a function of time.

times 30 intensity sets were combined into a single bounds file from which the rMD restraint file was generated. Upper and lower bounds in the bounds file were average interproton distances ± one standard deviation calculated from all of the MARDIGRAS runs. An additional 0.3 Å was added to the upper bounds. The structure used in the RANDMARDI procedure was as close to the final structure as possible. This was obtained by repeating the RANDMARDI procedure followed by rMD calculations until convergences of both procedures were obtained simultaneously.

Restrained molecular dynamics

The distance restraints obtained in the RANDMARDI calculations were incorporated into the rMD procedure using AMBER5.[18] The two starting models for structure refinement of the DNA:DNA duplex were either A-form or B-form DNA built in AMBER5 by the *nucgen* and *nukit* modules. The similar starting models for the LNA:DNA duplex were modified to include the LNA modifications using the *leap* module of AMBER5. The charges of the atoms in the modified nucleotide were obtained by RESP calculations as described by Bayly *et al.*[19]

All the distance restraints were incorporated into an rMD procedure for structure refinement. An initial energy minimization of the starting models was followed by 28 ns of restrained molecular dynamics. The simulating annealing protocols used in the rMD calculations by the *sander* module in the AMBER5 program are illustrated in Fig. 7.

The standard pseudo-energy terms were used to enforce the distances and the experimental value of an H3′(*i*)–P5′(*i* + 1) coupling constant.[20] A Karplus equation was used to translate the coupling constants to the dihedral angle.[21] No lower bounds for the coupling constants were introduced.

Helix parameters were calculated with the program CURVES 5.1.[22]

Acknowledgements

This work was supported by grants from The Danish Natural Science Research Council and The Danish Technical Research Council. We thank The Instrument Center for NMR-Spectroscopy of Biological Macromolecules at The Carlsberg Laboratory, Copenhagen for providing spectrometer time on the 800 MHz spectrometer. Ms Britta M. Dahl is thanked for synthesis of the LNA oligonucleotides.

References

1 E. Uhlmann and A. Peyman, *Chem. Rev.*, 1990, **90**, 543; N. T. Thuong and C. Hélène, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 666; P. E. Nielsen, *Bioconjugate Chem.*, 1991, **2**, 1; C. Hélène and J.-J. Toulmé, *Biochem. Biophys. Acta*, 1990, **1049**, 99; C. Hélène, *Anti-Cancer Drug Design*, 1991, **6**, 569; W. F. Flanagan, *Cancer Metastasis Rev.*, 1998, **17**, 169; A. De Mesmaeker, R. Häner, P. Martin and H. E. Moser, *Acc. Chem. Res.*, 1995, **28**, 366.
2 O. Y. Fedoroff, M. Salazar and B. R. Reid, *J. Mol. Biol.*, 1993, **233**, 509; J.-J. Toulmé and D. Tidd, *Ribonucleases H*, eds. R. J. Crouch and J.-J. Toulmé, INSERM, Paris, 1999, p. 225.
3 C. Wahlestedt, P. Salmi, L. Good, J. Kela, T. Johnsen, T. Hökfelt, C. Broberger, F. Porreca, A. Koshkin, M. H. Jacobsen and J. Wengel, *Proc. Natl. Acad. Sci., USA*, 2000, **97**, 3344.
4 S. K. Sing, P. Nielsen, A. A. Koshkin and J. Wengel, *Chem. Commun.*, 1998, 455; S. K. Singh and J. Wengel, *Chem. Commun.*, 1998, 1247; A. A. Koshkin, S. K. Singh, P. Nielsen, V. K. Rajwanshi, R. Kumar, M. Meldgaard, C. E. Olsen and J. Wengel, *Tetrahedron*, 1998, **54**, 3607; A. A. Koshkin, V. K. Rajwanshi and J. Wengel, *Tetrahedron Lett.*, 1998, **39**, 4381; A. A. Koshkin, P. Nielsen, M. Meldgaard, V. K. Rajwanshi, S. K. Singh and J. Wengel, *J. Am. Chem. Soc.*, 1998, **120**, 13252.
5 S. Obika, D. Nanbu, Y. Hari, K. Morio, Y. In, T. Ishida and T. Imanishi, *Tetrahedron Lett.*, 1997, **38**, 8735.
6 (*a*) R. E. Dickerson, M. Bansal, C. R. Calladine, S. Diekmann, W. N. Hunter, O. Kennard, E. von Kitzing, H. C. M. Nelson, R. Lavery, W. K. Olson, W. Saenger, Z. Shakked, D. M. Soumpasis, C.-S. Tung, H. Sklenar, A. H. J. Wang and V. B. Zhurkin, *EMBO J.*, 1989, **8**, 1; (*b*) W. Saenger *Principles of Nucleic Acid Structure*, Springer Verlag, New York, 1984.
7 M. Petersen, C. B. Nielsen, K. E. Nielsen, G. A. Jensen, K. Bondensgaard, S. K. Singh, V. K. Rajwanshi, A. A. Koshkin, B. M. Dahl, J. Wengel and J. P. Jacobsen, *J. Mol. Recognit.*, 2000, **13**, 44.
8 (*a*) H. Liu, H. P. Spielmann, N. A. Ulyanov, D. E. Wemmer and T. L. James, *J. Biol. NMR*, 1995, **6**, 390; (*b*) B. A. Borgias and T. L. James, *J. Magn. Reson.*, 1990, **87**, 475; B. A. Borgias, M. Gochin, D. Kerwood and T. L. James, *Progr. Nucl. Magn. Reson. Spectrosc.*, 1990, **22**, 83.
9 K. Wüthrich, *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York, 1986; D. R. Hare, D. E. Wemmer, S.-H. Chou, G. Drobny and B. R. Reid, *J. Mol. Biol.*, 1983, **171**, 319; J. Feigon, W. Leupin, W. A. Denny and D. R. Kearns, *Biochemistry*, 1983, **22**, 5943.
10 R. Boelens, R. M. Scheek, K. Dijstra and R. Kaptein, *J. Magn. Reson.*, 1985, **62**, 378.
11 H. P. Spielmann, D. E. Wemmer and J. P. Jacobsen, *Biochemistry*, 1995, **34**, 8542; C. H. Gotfredsen, H. P. Spielmann, J. Wengel and J. P. Jacobsen, *Bioconjugate Chem.*, 1996, **7**, 680; M. Petersen and J. P. Jacobsen, *Bioconjugate Chem.*, 1998, **9**, 331; F. Johansen and J. P. Jacobsen, *J. Biomol. Struct. Dyn.*, 1998, **16**, 205; C. H. Gotfredsen, G. Wiswanadham, P. N. Jørgensen, J. Wengel and J. P. Jacobsen, *Asian Chem. Lett.*, 2000, **4**, 171; L. B. Jørgensen, P. Nielsen, J. Wengel and J. P. Jacobsen, *J. Biomol. Struct. Dyn.*, 2000, **18**, 45; K. Bondensgaard and J. P. Jacobsen, *Bioconjugate Chem.*, 1999, **10**, 735.
12 (*a*) K. E. Nielsen, S. K. Singh, J. Wengel and J. P. Jacobsen, *Bioconjugate Chem.*, 2000, **11**, 228; (*b*) C. B. Nielsen, S. K. Singh, J. Wengel and J. P. Jacobsen, *J. Biomol. Struct. Dyn.*, 1999, **17**, 175.
13 A. Mujeeb, S. M. Kerwin, G. L. Kenyon and T. L. James, *Biochemistry*, 1993, **32**, 13419.
14 J. M. Aramini, A. Mujeeb and M. W. Germann, *Nucleic Acids Res.*, 1998, **26**, 5644.
15 R. Lavery and H. Sklenar, *J. Biomol. Struct. Dyn.*, 1989, **7**, 655; R. Lavery and H. Sklenar, *J. Biomol. Struct. Dyn.*, 1988, **6**, 63.
16 L. Kvaernø and J. Wengel, *Chem. Commun.*, 1999, 657.
17 C. H. Gotfredsen, A. Meissner, J. Ø. Duus and O. W. Sørensen, *Magn. Reson. Chem.*, 2000, **38**, 692.
18 D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham III,

W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, D. M. Ferguson, R. J. Radmer, G. L. Seibel, U. C. Singh, P. K. Weiner and P. A. Kollman, *AMBER 5*, University of California, San Francisco, 1997.
19 C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269.
20 A. Majumdar and R. V. Hosur, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1992, **24**, 159.
21 S. S. Wijmenga and B. N. M. van Buuren, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1998, **32**, 287; R. Nibedita, R. A. Kumar, A. Majumdar, R. V. Hosur and G. Govil, *Biochemistry*, 1993, **32**, 9053.
22 R. Lavery and H. Sklenar, *J. Biomol. Struct. Dyn.*, 1989, **7**, 655; R. Lavery and H. Sklenar, *J. Biomol. Struct. Dyn.*, 1988, **6**, 63.