Solution Structure of a Steroid−DNA Complex with Cholic Acid Residues Sealing the Termini of a Watson−Crick Duplex†

Jennifer Tuma and Clemens Richert*

Department of Chemistry, Universität Konstanz, D-78457 Konstanz, and Institut für Organische Chemie, Universität Karlsruhe (TH), D-76131 Karlsruhe, Germany

Received March 5, 2003

ABSTRACT: The three-dimensional structure of a covalent hybrid between cholic acid and the self-complementary DNA hexamer 5′-TGCGCA-3′ was solved via two-dimensional NMR and restrained torsion angle molecular dynamics. In the complex, refined to a pairwise rmsd of 0.64 Å, the steroid binds to the terminal T:A base pairs via extensive van der Waals contacts but without any hydrogen bonds or detectable dipole-dipole interactions. The contacts involve the methyl groups as well as one edge of the streoid's sterane skeleton and both nucleobases and the deoxyriboses of the terminal base pair of the DNA. The surprising shape complementarity between steroid and the undisturbed DNA termini explains the increase in fidelity and affinity observed for hybridization probes bearing bile acid residues. Since the hydroxyl groups of the steroid do not contribute to the binding of the DNA, they may be derivatized, potentially giving access to a new class of specific binders for blunt ends of Watson−Crick duplexes.

Steroids, most notably steroid hormones, are known to induce or suppress gene expression (*1*−*3*) but are not generally thought of as DNA-binding molecules. Instead, they are known to induce their effect on genomic DNA through nuclear receptors (*4*−*6*). Some reports on direct molecular interactions between steroids and DNA exist, but most of these are on steroidal amines (*7*−*12*), not hormones (*13*), sterols, or bile acids without cationic groups. When steroid derivatives were appended to oligonucleotides to be employed for biotechnological or biomedical applications (*14*−*20*), it was steroid−steroid interactions (*21*, *22*), or interactions with receptors (*23*, *24*) or membranes (*25*−*27*), that were envisioned, and not steroid−DNA interactions. A phosphoramidite building block of a cholesterol derivative suitable for covalently modifying oligonucleotides is now commercially available (*28*), but the supplier cites enhanced penetration of 2′-*O*-methyl oligoribonucleotides and thus improved transport into cells as key feature of oligonucleotides prepared with this reagent (*29*). However, the NMR structure of an estrone-tethered tandem DNA duplex where two steroid moieties interact not just with each other but also with DNA has been published (*30*). Further, stabilizing effects of covalent steroid appendages on DNA duplexes and triplexes devoid of other steroids have been observed (*31*, *32*).

Among the steroids, the bile acids, including their best known representative, cholic acid, were perhaps the least obvious ligands for nucleic acids, since they exert their known biological role in emulsifying dietary fats in the small intestine, i.e., not inside cells, where genomic DNA is found. It was therefore interesting to see that DNA aptamers can bind cholic acid tightly (*33*). When bile acids covalently appended to the 5′-terminus of oligonucleotides were found to stabilize DNA duplexes and to enhance DNA/RNA selectivity and mismatch-discrimination (*34*), these findings came as a surprise. They were not the result of a specific design, based on early reports on interactions between steroids with nucleobases (*35*), but were obtained in the context of a combinatorial study (*36*). The melting point increase observed for short duplexes bearing the cholic acid cap is 11 °C per modification and thus more pronounced than for other steroidal appendages (*30*, *31*). Further work showed that covalently appending cholic acid to template strands also accelerates nonenzymatic single nucleotide extension reactions (*37*), an effect conceptually similar to that of polymerases catalyzing the extension of primers in duplexes with templates.

Much like polymerases, cholic acid residues appended to oligonucleotides with a 5′-terminal 5′-amino-2′,5′-dideoxynucleotide are also found to exhibit their duplex stabilizing and mismatch-suppressing effect on more than one terminal Watson−Crick base pair (*37*, *38*). Namely, T:A- and C.G.-terminated duplexes have elevated UV melting points when stabilized by cholic acid residues (*34*, *38*), and T:A, C:G, and A:T base pairs all have been shown to be formed faster in nonenzymatic single-nucleotide extension reactions (*37*). This makes the steroid residues attractive as "caps" for high-fidelity hybridization probes to be employed for biotechnological applications, such as DNA chips. A recent study confirmed the duplex-stabilizing effect of cholic acid caps on hybridization probes of oligonucleotide arrays (*39*). This prompted our reinvestigation of the structural information gleaned from earlier work on oligonucleotides with 5′-

† Supported by Deutsche Forschungsgemeinschaft (Grants No. RI 1063/1-2 and FOR 434) and Fonds der Chemischen Industrie (Account No. 164431). J.T. is a recipient of a predoctoral fellowship from the State of Baden-Württemberg.

* Corresponding author. Address: Institut für Organische Chemie, Universität Karlsruhe (TH), Fritz-Haber-Weg 6, D-76131 Karlsruhe, Germany. Tel. ++49 (0) 721 608 2091. Fax: ++49 (0) 721 608 4825. E-mail: cr@rrg.uka.de.

10.1021/bi030056g CCC: $25.00 © 2003 American Chemical Society
Published on Web 07/09/2003

appended cholic acid residues (*34*) and the elucidation of the three-dimensional structure of a representative cholic acid-capped complex via two-dimensional NMR and restrained molecular dynamics. Here we report the high-resolution structure of the bile acid−DNA complex (Chl-TGCGCA)$_2$ (Figure 1),[1] where the steroid is in close contact with the terminal base pair of the duplex.

MATERIALS AND METHODS

Synthesis and Sample Preparation. A sample of Chl-TGCGCA (**1**) was prepared on 10 μmol scale using a methodology described previously (*34*), on an ABI 380 DNA synthesizer according to the manufacturer's recommendations, except that the capping steps were omitted and the 5′-MMT-protected 3′-phosphoramidite of 5′-amino-5′-deoxythymidine (*40*) was employed in the last phosphoramidite coupling step. The cholic acid residue was coupled manually as described previously (*34*). The crude steroid−DNA hybrid was HPLC purified on a Macherey−Nagel Nucleosil C4 column, using a gradient of CH_3CN in 0.1 M triethylammonium acetate, pH 7, with detection at 260 nm. The purified product was lyophilized twice from H_2O, four times from 10% aqueous NH_3 to displace residual triethylammonium counterions, and twice from D_2O. The residue was dissolved in D_2O containing 150 mM NaCl and 10 mM phosphate buffer pH 7, uncorrected for deuterium effect, to a final volume of 200 μL and a strand concentration of 6 mM. For the acquisition of spectra detecting exchangeable protons, the sample was lyophilized to dryness and immediately taken up in H_2O/D_2O (9:1). The sample was stored at 4 °C between acquisitions without preservatives and showed no signs of decomposition after 14 months.

NMR Spectroscopy. NMR spectra were recorded in NMR microtubes susceptibility-matched to D_2O (Shigemi Co., Tokyo, Japan) on Bruker DRX 600 and DRX 500 spectrometers at 286 K. Two-dimensional spectra were acquired with 2048 data points in f2 and 512 increments in f1. For samples in D_2O, suppression of the excess solvent peak was achieved via presaturation during the recycle delay. For samples in H_2O/D_2O (9:1), the WATERGATE pulse sequence was employed (*41*). NOESY spectra (*42*) were acquired at mixing times of 100 and 250 ms, of which the latter was used for generating distance constraints. DQF-COSY (*43*), TOCSY (spin lock time 70 ms) (*44*), HMQC (either ^{1}H coupled or decoupled to determine C−H coupling constants), and HMBC experiments (*45*, *46*) were recorded for resonance assignment of the steroid moiety. The average C−H coupling constant measured in the nondecoupled HMQC experiment was 140 Hz, leading to a delay (D2) of 3.57 ms for the decoupled experiment used to identify directly linked coupling partners. For the detection of long-range couplings, a delay of 60 ms, corresponding to an average ^{3}J coupling constant of 10 Hz, was used for the HMBC experiment. Spectra were processed using XWINNMR (Bruker Biospin).

Peak Assignments. A brief account of the assignment of selected resonances of (**1**)$_2$ has been published previously (*34*). In short, the starting point for the assignment of the DNA portion of (**1**)$_2$ was the resonance of the methyl group of thymidine at 1.74 ppm. The NOESY cross-peak of this singlet with its integration for three protons to the resonance of H6 of the same residue at 7.30 ppm provided an entry into sequential assignment of all H1′ and nonexchangeable nucleobase resonances of the nucleotides (*47*−*49*). Starting from the H1′ frequencies thus determined, the remaining protons of the deoxyribose spin systems were assigned based on cross-peaks in DQF-COSY and TOCSY spectra. Peak overlap prevented the unambiguous assignment of all H5′/5″ resonances and the diasterotopic position of some geminal coupling partners was not resolved. The H5 resonances of the deoxycytidine resonances were identified on the basis of DQF-COSY cross-peaks to their H6 neighbors. Exchangeable protons at N3 of T2 and N4 of C4 and C6 were identified on the basis of NOESY cross-peaks to H2 of A7 and the H5 resonances of the cytosines, respectively. The frequencies of the NH1 signals for G3 and G5 were assigned based on NOESY cross-peaks to the NH41 resonances of the Watson−Crick base pairing partners C4 and C6. The assignment procedure for the cholic acid residue is described in the text (vide infra).

Generation of Restraints. Cross-peaks in a NOESY spectrum of (**1**)$_2$, acquired at 600 MHz in buffered D_2O at a mixing time of 250 ms, were integrated. The intensities of cross-peaks between protons with known distances, such as H5 and H6 of the cytosine residues, were used to sort the cross-peaks into classes of very strong (interproton distance of 2.5 ± 1.0 Å), strong (3.0 ± 1.0 Å), medium (3.5 ± 1.0 Å), weak (4.0 ± 1.0 Å), very weak (4.5 ± 1.0 Å), and very very weak (5.0 ± 1.0 Å) intensity. For cross-peaks that were difficult to integrate, such as those that are overlapping or of substantially different intensity on the two sides of the diagonal, the boundaries of the distance constraints were set to ±2.0 Å. For very well defined, intense cross-peaks from protons of interior nucleotides, boundaries of ± 0.5 Å were used during refinement. Diastereotopic protons that could not be stereoselectively assigned were treated with the geometric center pseudoatom approach (*50*). Trivial distance constraints, resulting inherently from the covalent structure of **1**, were not entered to avoid less successful calculations.

The NH1 and NH3 protons of the deoxyguanosines and thymidine in (**1**)$_2$ are observed in the chemical shift range typical for hydrogen bonding in Watson−Crick base pairs (*34*). Further, the NOEs between protons of neighboring nucleotides are those expected for an undisrupted B-form duplex (*51*). Together, this information was used as the basis for introducing base pairing restraints (hydrogen bonding and base-pair planarity), using default values from CNS and constraint boundaries of ±0.1 Å for the hydrogen bonds. All phosphorus resonances of the ^{31}P NMR spectrum of the solution of (**1**)$_2$ appear within 1.7 ppm units. Since all H1′ resonances also appear as triplets with ^3J coupling constants of 7.0 ± 1.0 Hz, as expected for the South or 2′-endo conformation of the deoxyribose rings (*51*), dihedral angle constraints for B-form DNA from the literature (*52*) were entered for the backbone with constraint boundaries of ±20°. To aid the refinement stage of this work, restraint violations were detected using either a visualization program written in-house (*53*) or VMD-XPLOR (*54*).

Structure Generation and Restrained Molecular Dynamics. Topology and parameter files for molecular dynamics calculations involving the cholic acid residues were generated

[1] Abbreviations: dsDNA, double-stranded DNA; Chl, cholic acid residue.

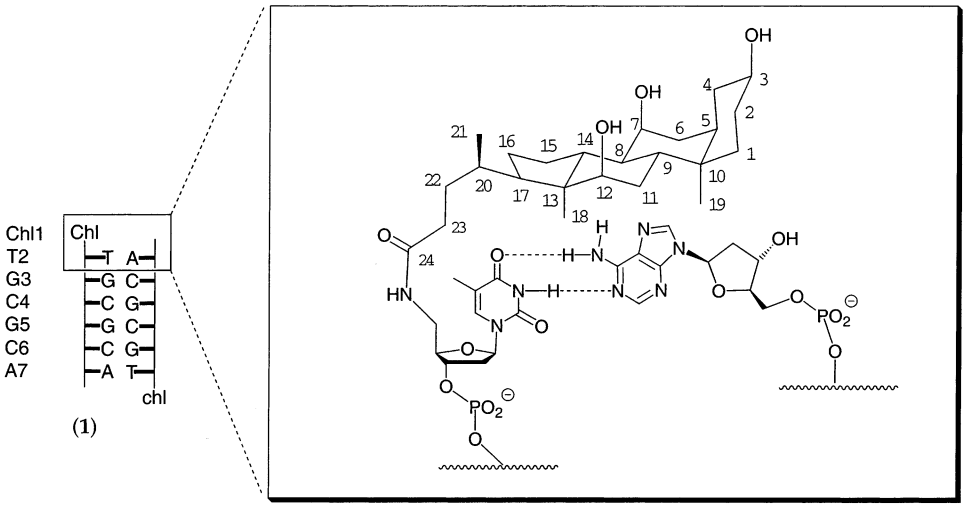

FIGURE 1: Schematic drawing of the duplex (**1**)$_2$ studied in this work. The highlighting box shows the terminus with the numbering scheme used.

with the help of XPLO2D (version 3.0.1) (*55*), using structural data from an X-ray crystal structure of cholic acid (*56*). The structural data used include the bond lengths, bond angles, and the connectivity. The output format of XPLO2D generating files in CNS syntax was used. The output was manually corrected after visualizing the structure in VMD. In particular, the file was edited to reduce C−H bond lengths over 1.1 to 1.08 Å and to include missing dihedral angles. The parameter and topology files thus obtained were employed for structure generation in X-PLOR (version 3.851), using a patch from the cholic acid residue linked to the DNA strand written as recommended in the X-PLOR manual (*50*). The DNA portion of the strands was generated in X-PLOR using default parameters. Restrained molecular dynamics calculations with parameters thus obtained were performed with CNS (version 1.0) (*57*) on LINUX or IRIX platforms, using the torsion angle molecular dynamics protocol (*58*). The number and order of computational steps during the different phases of the structure generation was identical to those reported in our earlier work (*59*) and is detailed in Table S1 (Supporting Information). The coordinates of a representative refined structure of (**1**)$_2$ has been deposited in the PDB database. The PDB ID code is 1ON5.

RESULTS

Resonance Assignment. The assignment of the DNA portion followed established protocols and confirmed the chemical shifts reported earlier (*34*). Assignment of the resonances of the steroid moiety of (**1**)$_2$ proved more challenging. Peak overlap in the region from −0.2 to 2.5 ppm of the ^{1}H-spectrum was compounded by magnetic anisotropies caused by the nucleobases of the terminal base pair, inducing chemical shift changes that made assignment strategies based on analogy to cholic acid derivatives previously assigned (*60*) unreliable. Six resonances provided potential entries into the spin systems of the steroid rings. Three of these were the resonance of CH protons with geminal hydroxyl groups at positions 3, 7, and 12 (see Figures 1 and S1 for the numbering scheme used). Their chemical shift at lower field put them outside the most crowded region of the ^{1}H NMR spectrum. Of these, only one gave two strong DQF-COSY cross-peaks to vicinal

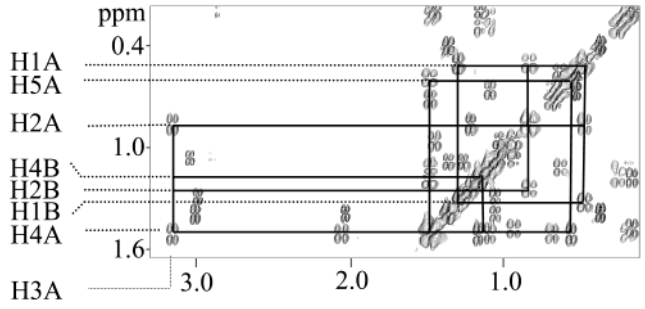

FIGURE 2: Expansion of the high field region of the DQF-COSY spectrum of (**1**)$_2$, acquired at 600 MHz in H_2O/D_2O (9:1), showing the assignment of proton resonances of ring A of the cholic acid residue. See Figure S1 for the naming of atoms.

neighbors (Figure 2), identifying it as the one at position 3. The geminal coupling partners of the strongly coupling, antiperiplanar vicinal protons (H2A and H4A) could then be identified in an HMQC spectrum (H2B and H4B, Figure S2, Supporting Information), since they are linked to the same carbon atom. Confirming evidence came from the weak cross-peaks of H2B and H4B to H3 in the COSY spectrum (Figure 2). While this only established the relative position to H3, the identities of each methylene group (position 2 or 4) were established on the basis of the number of additional vicinal coupling partners. As expected, H4A gives a cross-peak to H5A, for which no other strong scalar coupling partners to vicinal neighbors is found, but H2A couples strongly to H1A, which in turn gives an additional COSY cross-peak to H1B. These assignments were confirmed via the TOCSY spectrum (Figures S3, Supporting Information).

Of the possible entries from the resonances of ring A to those of ring B, namely, coupling between H5A and H6A/B and long-range heterocorrelation between C1 and H19A−C, the latter proved easier to realize unambiguously. After the resonance of C1 had been identified in the HMQC and the delays in the HMBC experiment (*45*) had been optimized for three-bond correlations, the expected three cross-peaks originating from H19 were observed (Figure 3). Two of these were known (C1 and C5), leading to C9, whose directly linked proton (H9) was identified via the HMQC. Starting from H9, the remaining resonances in ring B could be assigned via DQF-COSY and TOCSY.

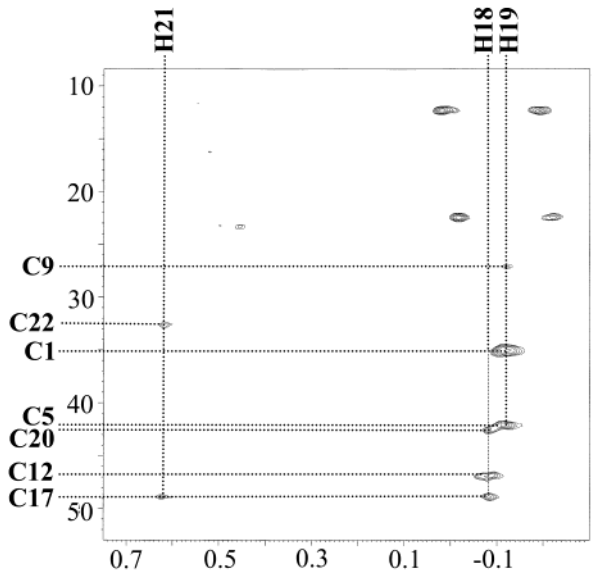

FIGURE 3: Expansion of a portion of an HMBC spectrum of $(\mathbf{1})_2$, acquired at 600 MHz in H_2O/D_2O (9:1), showing cross-peaks due to $^3J_{C-H}$ coupling between protons of methyl groups of the cholic acid residue and quaternary carbons of the same residue.

The other methyl group-based entry points to assignment then became obvious, since CH_3-21 has a vicinal coupling partner and appears as a doublet, identifying CH_3-18 by deletion analysis. With this and the only remaining H−C−OH starting point established at position 8, rings C and D were assigned using the same strategy as detailed above. NOESY cross-peaks from the methyl group at position 7 of T2 helped to confirm the assignments. All but four protons of the steroid moiety were thus assigned, though the diastereotopic identity was not established for all protons of methylene groups due to peak overlaps. In these cases, the distance constraints generated in subsequent steps were introduced via the pseudoatom approach (*50*). A list of the resonances is given in Table S3 (Supporting Information).

Restrained Molecular Dynamics. With the assignment of the proton resonances in hand, distance constraints were generated from NOESY cross-peaks, as detailed in Materials and Methods. A qualitative discussion of the NOEs can be found in our earlier work (*34*). Restrained molecular dynamics with the torsion angle protocol (*58*) then led to a first set of structures displaying the features of a DNA duplex, which was refined by introducing base pairing and backbone constraints. The former were based on the observation (*34*) of NH3 and NH1 resonances of the thymidine and deoxyguanosine residues, respectively, in the ppm range typical for Watson−Crick base pairing (*51*). The latter were introduced since coupling constants for the deoxyribose resonances fell within the range expected for a B-form duplex (*51*). The refinement of the structure was largely based on the comparison between experimental and back-calculated NOESY spectra. The overlay of these for a representative

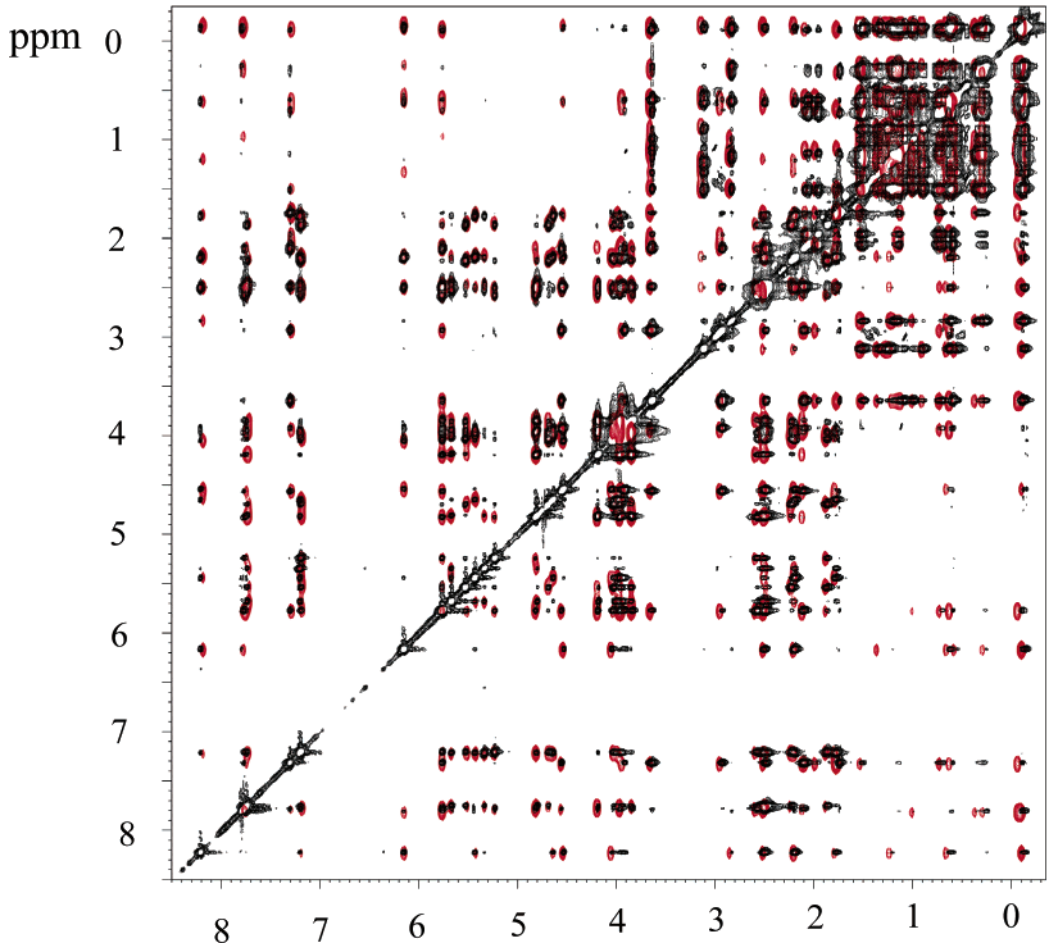

FIGURE 4: Overlay of the experimental NOESY spectrum of $(\mathbf{1})_2$ acquired at 600 MHz and 13 °C, with a mixing time of 250 ms (black) and the NOESY spectrum back-calculated from the lowest energy structure of $(\mathbf{1})_2$ obtained via molecular dynamics (red). The back-calculated spectrum was generated in GIFA (*75*) using the two-spin approximation and cross-peaks intensities calculated in X-PLOR (*50*).

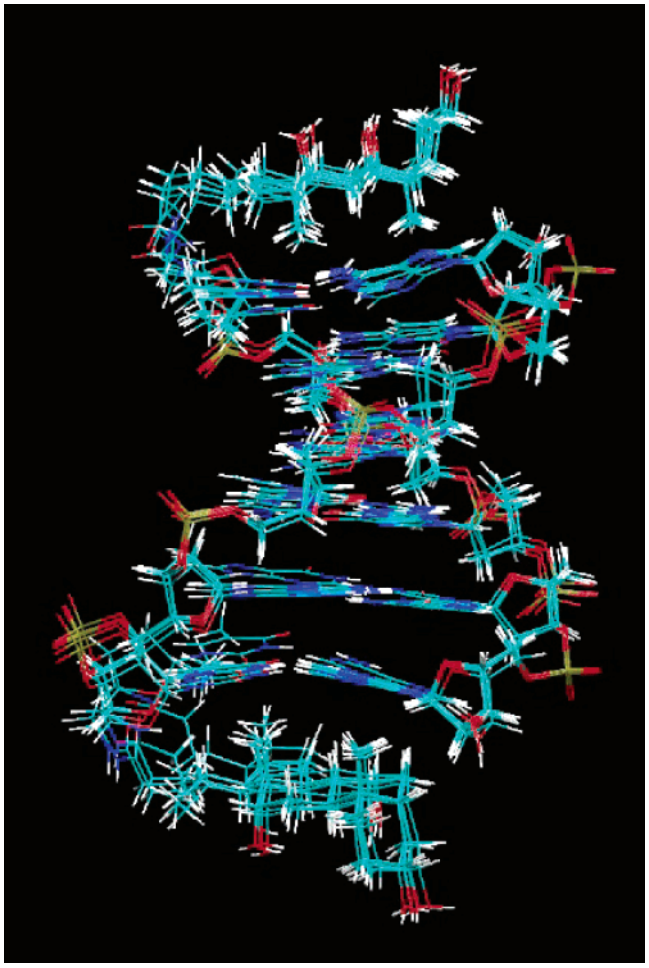

FIGURE 5: Overlay of the 10 violation-free structures of (**1**)$_2$ of lowest energy, obtained via restrained torsion angle molecular dynamics calculation in CNS. This and the subsequent graphic were generated using the program VMD (*76*). Coloring scheme: hydrogen, white; carbon, light blue; nitrogen, blue; oxygen, red; phosphate, green.

refined structure is shown in Figure 4. To compensate in part for the two-spin approximation used in the backcalculation of the NOESY spectrum, the experimental NOESY spectrum was symmetrized for this overlay. Figure 5 shows the 10 lowest energy structures of (**1**)$_2$, as obtained from torsion angle molecular dynamics. All low-energy structures are of the same fold, where a largely undisturbed B-form DNA duplex is capped by the steroid residues. Table 1 provides key data on constraints used and the quality of the structures obtained. The number of the former is smaller than in studies where NOESY cross-peaks are converted to constraints automatically, since trivial cross-peaks, resulting necessarily from the covalent structure, were not entered in our case to avoid biasing the calculations toward "overdefined" regions of the structure.

A detailed view of the terminal region from the backbone of the 3′-terminus (Figure 6a) shows the shape complementarity between the steroid and the 3′-terminal nucleotide. The van der Waals representation of the terminal region (Figure 6b) confirms that the acyclic linker portion of the steroid also packs well against the methyl group of T2, explaining why introducing an additional amino acid residue between cholic acid residue and T2 leads to decreased duplex stability (*34*). Given the puckered shape of the cholic acid residue, it

Table 1: Statistical Data on the Structure of (**1**)$_2$ as Determined by Restrained Molecular Dynamics

Constraints	
NOE-based total	130
interresidue	40
intraresidue	90
dihedral angle constraints	52
hydrogen bonding constraints	16
base pair planarity constraints	6
Ten Lowest-Energy Structures	
NOE constraint violations (>0.5 Å)	
dihedral angle constraint violations (>30°)	
bond lengths (>0.05 Å)	0
bonding angles (>5°)	0
rmsd from average (all coordinates)	0.39 Å
pairwise rmsd (all coordinates)	0.64 Å
rmsd from average (residue 1)	0.51 Å
rmsd from average (residue 2)	0.41 Å
rmsd from average (residue 3)	0.26 Å
rmsd from average (residue 4)	0.22 Å
rmsd from average (residue 5)	0.13 Å
rmsd from average (residue 6)	0.13 Å
energy	−346 ± 2 kcal/mol

is surprising how snugly it fits onto the terminal T:A base pair, generating a folded structure with neither any exposed side chains nor any holes for solvent molecules. A compact cap is formed by the steroid that shields the termini, without either disrupting hydrogen bonds in the DNA portion or forming any new ones with the donor functionalities of the steroid. Since the hydroxyl groups are on the α-face of the steroid ring, they protrude out into the solvent and cannot engage in hydrogen bonding with the DNA, in agreement with earlier findings that these hydroxyl groups are not critical for duplex stability (*34*). The DNA helix itself is slightly bent, but retains the feature typical of a B-form duplex (Figure 7; Table S4, Supporting Information).

DISCUSSION

The high-resolution structure of (Chl-TGCGCA)$_2$ helps to explain the duplex stabilizing effect of cholic acid residues attached to the termini of oligonucleotides (*34*, *38*). The structure also helps to explain why cholic acid residues, when appended to the template strand, increase the rate and fidelity of nonenzymatic single nucleotide extension reactions, an effect not observed for other acyl groups (*37*). The steroid packs with all four rings and its alkyl side chain against the T2:A7 base pair, thus allowing for extensive van der Waals contacts with both terminal nucleotides. This bridging of the terminus buries a large lipophilic surface area and thus also contributes to duplex stability through a hydrophobic effect. Finally, the shielding of the terminal base pair from solvent can be expected to lead to stronger hydrogen bonds between T2 and A7 when compared to the unmodified duplex, where their base pair is exposed to water.

Specific interactions are also noteworthy. For example, the A-ring of the steroid packs against the 2′-methylene group of A7, and thus senses differences between 2′-endo and 3′-endo conformations of the deoxyribose of this residue. The former conformation is to be expected for DNA:DNA duplexes (B-form), whereas DNA:RNA duplexes favor the latter conformation (A-form). Cholic acid-capped hybridization probes accordingly show enhanced DNA/RNA selectivity (*34*). Further, the tight fit between steroid and the terminal

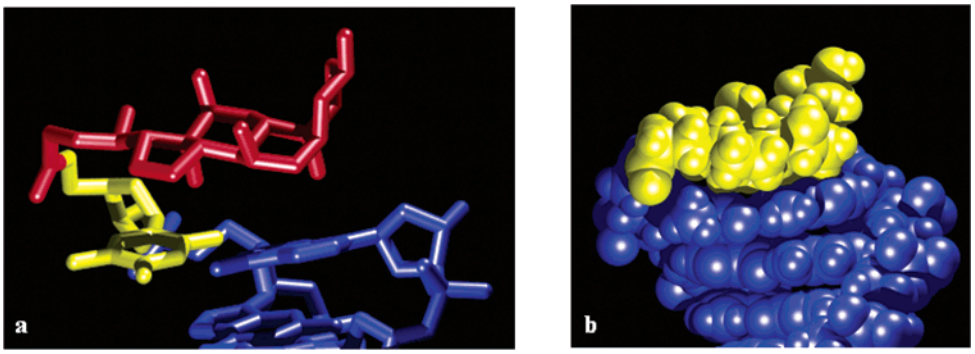

FIGURE 6: Details of the three-dimensional structure of $(\mathbf{1})_2$. (a) View of the terminus with the strand displaying its 3′-terminus in the foreground with hydrogens omitted. The cholic acid residue is shown in red, residue T2 is in yellow, and the remainder of the DNA is in blue. (b) View of the terminus from the major groove (van der Waals representation); the cholic acid residue is in gold and the DNA in blue.

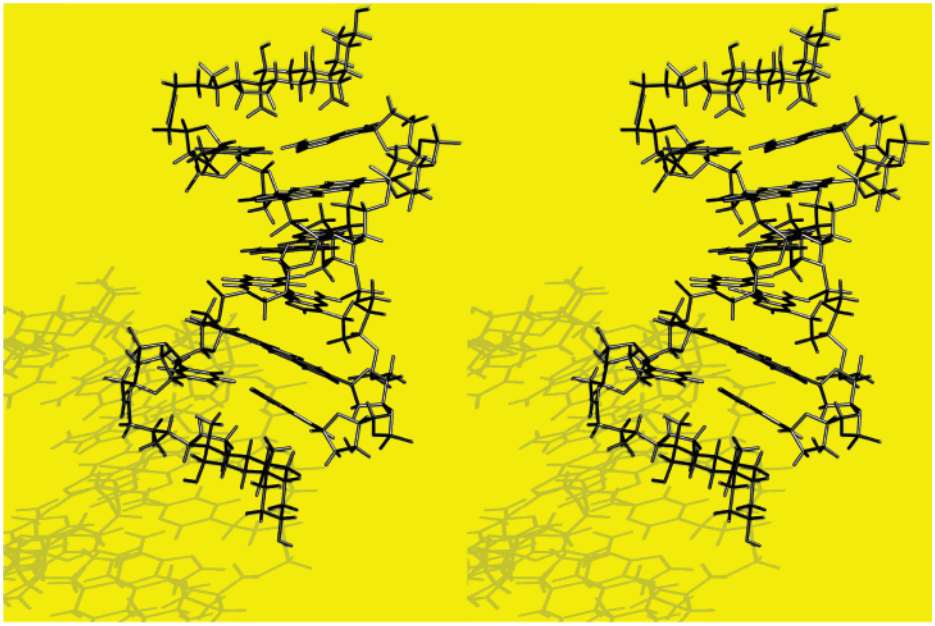

FIGURE 7: Stereo image of the lowest energy structure of $(\mathbf{1})_2$, as obtained by restrained torsion angle molecular dynamics. The shadows were introduced to aid visualization of the structure.

nucleotides of the duplex helps to rationalize the enhanced mismatch discrimination (*34*, *38*). Mismatched base pairs, including pyrimidine:pyrimidine, purine:purine base pairs, and wobble base pairs (T:G), will not fit the convex side of the steroid and thus fail to realize the duplex-stabilization that its binding provides. Since the complex of the L-shaped steroid with the terminal nucleotides is limited to the inner "ridges" set up by the deoxyriboses and the "valley" defined by the nucleobases but does not include outside "walls" for the base pair, hypotheses on the role of shape complementarity (*61*) in base pairing and single nucleotide extension reactions (*37*) may benefit from detailed structural information on this simple, well-defined system.

Taking a more general view of the molecular interactions, one notes how strongly duplex-stabilizing the interactions between the aliphatic sterane scaffold and the nucleobases is. Aromatic DNA binders, whose potential to interact with nucleobases has long been known, do not provide substantially greater duplex stabilization when employed as "molecular caps" (*36*, *38*, *59*), except when C-nucleotides, benefiting from the added interactions (and preorganizing effect) of the deoxyribose unit, are employed (*62*). While aromatic molecules tend to mimic the central structural motif of nucleic acid structures, namely, stacking flat lipophilic molecules regardless of lipophilicity displayed at the terminus, the cholic acid−DNA complex buries as much hydrophobic surface as possible and presents polar groups to the solvent. In this sense, it is reminiscent of protein structures or protein−DNA complexes, particularly those of structural proteins and genomic DNA (*63*).

In the field of bioorganic chemistry, bile acids have become versatile building blocks for generating artificial binding motives (*64*−*66*). Further, steroidal diamines, compounds that are structurally similar to hormonal steroids, are known DNA-binders (*67*). This raises the somewhat academic question is "why does nature not employ direct steroid−DNA interactions", even though feedback regulation of bile acid synthesis links these two classes of biomolecules (*68*) and requires specific molecular interactions? The structure presented here provides one answer to this question: such direct interactions would require exposed terminal base pairs, and these are found only at the termini of chromosomes in genomic DNA. The mode of interaction between steroid and DNA observed here will not be possible in the interior of Watson−Crick duplexes, where the L-shaped cholic acid residue would have to intercalate. The

thickness of the flat part of the "L" in itself (approximately 6 Å) presents an obstacle for intercalation. But the concave side of the steroid ring system would also offer too little packing surface for a second base pair.

The difference between termini and the interior of long double-stranded DNA must be borne in mind when comparing the current structure with the earlier structural proposals for steroidal amines binding to DNA (*67*, *69*). These proposals include structures with direct stacking of sterane-bound methyl groups on T:A base pairs (*70*). Further, a number of steroidal diamines, such as dipyrandium (*71*), have a trans-connected A-ring rather than the cis-connected A-ring of the bile acids and are therefore more prone to interact with kinked dsDNA via a process resembling intercalation than cholic acid. Steroidal diamines can induce hyperchromicity and duplex destabilization, effects that are the opposite of what is observed for the covalently appended bile acids, as expected for a binding mode separating base pairs and kinking DNA and not the capping of undisturbed DNA observed here.

Though bile acid-terminal dsDNA interactions may just be an accident in molecular recognition between fundamental classes of biomolecules, the strength and specificity of these interactions may allow us to put them to good use. Bile acid−nucleic acid complexes are being used to target therapeutic oligonucleotides, and duplex stabilization is a welcome side effect of the modification introduced (*72*). Attempts to design improved caps, with even stronger duplex-stabilizing effects, may benefit from the results of the current study. For example, the contact between CH_3-21 and O4′ of A7 may be used to introduce a hydrogen bond. In addition, the hydroxyl group at position 7 (ring B) is positioned well to deliver a ligand to the major groove of the DNA duplex. Finally, the detailed structural information available through the current study may help to refine structural proposals for aptamer-cholic acid complexes (*73*), which may be employed as sensors for biomedical applications (*74*). Other applications may emerge for ligands that bind preferentially to blunt ends of DNA, a binding site fundamentally different from the groves, the backbone, and intercalation sites.

ACKNOWLEDGMENT

The authors wish to thank A. Friemel, A. Geyer, A. Rapp, and P. Lang for help with the acquisition of NMR spectra, K. Siegmund and J. Bühler for help with computer issues, and S. Narayanan, N. Griesang, and G. Beck for comments on the manuscript.

SUPPORTING INFORMATION AVAILABLE

Figures of structures and spectra and tables of peak assignments and resonances. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Aranda, A., and Pascual, A. (2001) Nuclear hormone receptors and gene expression, *Physiol. Rev. 81*, 1269−1304.
2. Klinge, C. M. (2001) Estrogen receptor interaction with estrogen response elements, *Nucleic Acids Res. 29*, 2905−2919.
3. Dickson, R. B., and Stancel, G. M. (2000) Estrogen receptor-mediated processes in normal and cancer cells, *J. Natl. Cancer Inst. Monogr. 27*, 135−145.
4. Whitfield, G. K., Jurutka, P. W., Haussler, C. A., and Haussler, M. R. (1999) Steroid hormone receptors: evolution, ligands, and molecular basis of biological function, *J. Cell. Biochem. S32−33*, 110−122.
5. Kumar, R., and Thompson, E. B. (1999) The structure of the nuclear hormone receptors, *Steroids 64*, 310−319.
6. Makishima, M., Lu, T. T., Xie, W., Whitfield, G. K., Domoto, H., Evans, R. M., Haussler, M. R., and Mangelsdorf, D. J. (2002) Vitamin D receptor as an intestinal bile acid sensor, *Science 296*, 1313−1316.
7. Mahler, H. R., Goutarel, R., Khuong-Huu, Q., and Ho, M. T. (1966) Nucleic acid interactions. VI. Effects of steroidal diamines, *Biochemistry 5*, 2177−2191.
8. Mahler, H. R., Green, G., Goutarel, R., and Khuong-Huu, Q. (1968) Nucleic acid-small molecule interactions. VII. Further characterization of deoxyribonucleic acid-diamino steroid complexes, *Biochemistry 7*, 1568−1582.
9. Gabbay, E. J., and Glaser, R. (1971) Topography of nucleic acid helices in solutions. Proton magnetic resonance studies of the interaction specificities of steroidal amines with nucleic acid systems, *Biochemistry 9*, 1665−1674.
10. Hsieh, H.-P., Muller, J. G., and Burrows, C. J. (1994) Structural effects in novel steroidal polyamine-DNA binding, *J. Am. Chem. Soc. 116*, 12077−12078.
11. Hsieh, H.-P., Muller, J. G., and Burrows, C. J. (1995) Synthesis and DNA binding properties of C3-, C12-, ad C24- substituted amino-steroids derived from bile acids, *Bioorg. Med. Chem. 6*, 823−838.
12. Muller, J. G., Ng, M. M. P., and Burrows, C. J. (1996) Hydrophobic vs Coulombic interactions in the binding of steroidal polymanines to DNA, *J. Mol. Recogn. 9*, 143−148.
13. Cohen, P., Chin, R.-C., and Kidson, C. (1969) Interactions of hormonal steroids with nucleic acids. II. Structural and thermodynamic aspects of binding, *Biochemistry 8*, 3603−3609.
14. De Smidt, P. C., Doan, T. L., de Falco, S., and van Berkel, T. J. C. (1991) Association of antisense oligonucleotides with lipoproteins prolongs the plasma half-life and modifies the tissue distribution, *Nucleic Acids Res. 19*, 4695−4700.
15. Stein, C. A., Pal, R., Devico, A. L., Hoke, G., Mumbauer, S., Kinstler, O., Sarngadharan, M. G., and Letsinger, R. L. (1991) Mode of action of 5′-linked cholesteryl phosphorothioate oligonucleotides in inhibiting syntia formation and infection by HIV-1 and HIV-2 in vitro, *Biochemistry 30*, 2439−2444.
16. Zhou, J. H., Pai, B. S., Reed, M. W., Gamper, H. B., Lukhtanov, E., Podyminogin, M., Meyer, R. B., and Cheng, Y. C. (1994) Discovery of short, 3′-cholesterol-modified DNA duplexes with unique antitumor activity, *Cancer Res. 54*, 5783−5787.
17. Marasco, C. J., Angelino, N. J., Paul, B., and Dolnick, B. J. (1994) A simplified synthesis of acridine and/or lipid containing oligonucleotides, *Tetrahedron Lett. 35*, 3029−3032.
18. Vu, H., Schmaltz Hill. T., and Jayaraman, K. (1994) Synthesis and properties of cholesteryl-modified triple-helix forming oligonucleotides containing a triglycyl linker, *Bioconjugate Chem. 5*, 666−668.
19. Manoharan, M., Johnson, L. K., Bennett, C. F., Vickers, T. A., Ecker, D. J., Cowsert, L. M., Freier, S. M., and Cook, P. D. (1994) Cholic-acid oligonucleotide conjugates for antisense applications, *Bioorg. Med. Chem. Lett. 4*, 1053−1060.
20. Manoharan, M., Tivel, K. L., Inamati, G., Monia, B. P., Dean, N., and Cook, P. D. (1997) Cholesterol conjugated uniform and gapmer phosphorothioate oligonucleotides targeted against PKC-alpha and C-raf gene expression, *Nucleosides Nucleotides 16*, 1139−1140.
21. Gryaznov, S. M., and Lloyd, D. H. (1993) Modulation of oligonucleotide duplex and stability via hydrophobic interactions, *Nucleic Acids Res. 21*, 5909−5915.
22. Letsinger, R. L., Chatuvedi, S. K., Farooqui, F., and Salunkhe, M. (1993) Use of hydrophobic substituents in controlling self-assembly of oligonucleotides, *J. Am. Chem. Soc. 115*, 7535−7536.
23. Rebuffat, A. G., Nawrocki, A. R.; Nielsen, P. E., Bernasconi, A. G., Bernal-Mendez, E., Frey, B. M., and Frey, F. J. (2002) Gene delivery by a steroid-peptide nucleic acid conjugate, *FASEB J. 16*, 1426−1428.
24. Bijsterbosch, M. K., Manoharan, M., Tivel, K. L., Rump, E. T., Biessen, E. A. L., DeVrueh, R. L. A., Cook, P. D., and van Berkel, T. J. C. (1997) Modification of a phosphorothioate oligonucleotide with cholesterol induces association of the oligonucleotide to serum lipoproteins and affects its biological fate, *Nucleosides Nucleotides 16*, 1165−1168.

25. Letsinger, R. L., Zhang, G., Sun, D. K., Ikeuchi, T., and Sarin, P. S. (1989) Cholesteryl-conjugated oligonucleotides: synthesis, properties, and activity as inhibitors of replication of human immunodeficiency virus in cell culture, *Proc. Natl. Acad. Sci. U.S.A. 86*, 6553−6556.
26. MacKellar, C., Graham, D., Will, D. W., Burgess, S., and Brown, T. (1992) Synthesis and physical properties of anti-HIV antisense oligonucleotides bearing terminal lipophilic groups, *Nucleic Acids Res. 20*, 3411−3417.
27. Zhang, G., Farooqui, F., and Letsinger, R. F. (1996) Informational liposomes: complexes derived from cholesteryl-conjugated oligonucleotides and liposomes, *Tetrahedron Lett. 37*, 6243−6246.
28. From Glen Research, Sterling, VA 20164. See: http://www.glenres.com.
29. Oberhauser, B., and Wagner, E. (1992) Effective incorporation of 2′−O-methyl-oligoribonucleotides into liposomes and enhanced cell association through modification with thiocholesterol, *Nucleic Acids Res. 20*, 533−538.
30. Denisov, A. Y., Sandström, A., Maltseva, T. V., Pyshnyi, D. V., Ivanova, E. M., Zarytova, V. F., and Chattopadhaya, J. (1997) The NMR structure of estrone (Es)-ethered tandem DNA duplex [d(5′pCAGCp3′)-Es] + [Es-d(5′pTCCA3′)]: d(5′pTGGAGCTG3′), *J. Biomol. Struct. Dyn. 15*, 499−516.
31. Gryaznov, S. M., and Lloyd, D. H. (1993) Modulation of oligonucleotide duplex and triplex stability via hydrophobic interactions, *Nucleic Acids Res. 21*, 5909−5915.
32. Bhatia, D., Yue-Ming, L., and Ganesh, K. N. (1999) Steroid-DNA conjugates: improved triplex formation with 5-amido-(7-deoxycholic acid)-dU incorporated oligonucleotides, *Bioorg. Med. Chem. Lett. 9*, 1789−1794.
33. Kato, T., Yano, K., Ikebukuro, K., and Karube, I. (2000) Interaction of three-way DNA junctions with steroids, *Nucleic Acids Res. 28*, 1963−1968.
34. Bleczinski, C. F., and Richert, C. (1999) Steroid-DNA interactions increasing stability, sequence-selectivity, DNA/RNA discrimination, and hypochromicity of oligonucleotide duplexes, *J. Am. Chem. Soc. 121*, 10889−10894.
35. Molinari, G., and Lata, G. F. (1962) Interaction of steroids with some pyrimidine and purine derivatives, *Arch. Biochem. Biophys. 96*, 486-490.
36. Altman, R. K., Schwope, I., Sarracino, D. A., Tetzlaff, C. N., Bleczinski, C. F., and Richert, C. (1999) Selection of modified oligonucleotides with increased target affinity via MALDI-monitored nuclease survival assays, *J. Combin. Chem. 1*, 493−508.
37. Rojas Stütz, J. A., and Richert, C. (2001) A steroid cap induces enhanced selectivity and rates in nonenzymatic single nucleotide extensions of an oligonucleotide, *J. Am. Chem. Soc. 123*, 12718−12719.
38. Mokhir, A. A., Tetzlaff, C. N., Herzberger, S., Mosbacher, A., and Richert, C. (2001) Monitored selection of DNA-hybrids forming duplexes with capped terminal C:G base pairs, *J. Combin. Chem. 3*, 374−386.
39. Dombi, K. L., Griesang, N., and Richert, C. (2002) Oligonucleotide arrays on aldehyde-bearing glass with coated background, *Synthesis* 816−824.
40. Tetzlaff, C. N., Schwope, I., Bleczinski, C. F., Steinberg, J. A., and Richert, C. (1998) A convenient synthesis of 5′-amino-5′-deoxythymidine and preparation of peptide-DNA hybrids, *Tetrahedron Lett. 39*, 4215−4218.
41. Sklenar, V., Piotto, M., Leppik, R., Saudek, V. (1993) Gradient Tailored water Suppression for 1H-15N HSQC Experiments Optimozed to Retain Full Sensitivity, *J. Magn. Res., Ser. A 102*, 241−245.
42. Kumar, A., Ernst, R. R., Wüthrich, K. (1980) A Two-dimensional Nuclear Overhauser Enhancement (2D NOE) Experiment for the Elucidation of complete Proton−Proton Cross-Relaxation Networks in Biological Macromolecules, *Biochem. Biophys. Res. Comm. 95*, 1−6.
43. Piantini, U., Sørensen, O. W., and Ernst, R. R. (1982) Multiple Quantum Filters for Elucidating NMR Coupling Networks, *J. Am. Chem. Soc. 104*, 6800−6801.
44. Bax, A., and Davis, D. (1985) MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy, *J. Magn. Reson. 65*, 355−360.
45. Bax, A., and Summers, M. (1986) 1H and 13C Assignments from Sensitivity-Enhanced Detection of Heteronuclear Multiple-Bond Connectivity by 2D Multiple Quantum NMR, *J. Am. Chem. Soc. 108*, 2093−2094.
46. Willker, W., Leibfritz, D., Kerssebaum, R., and Bermel, W. (1993) Gradient Selection in Inverse Heteronuclear Correlation Spectroscop, *Magn. Res. Chem. 31*, 287−292.
47. Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., and Kaptein, R. (1984) Sequential resonance assignments in proton NMR spectra of oligonucleotides by two-dimensional NMR spectroscopy, *Biochemistry 23*, 1371−1376.
48. Feigon, J., Leupin, W., Denny, W. A., and Kearns, D. R. (1983) Two-dimensional proton nuclear magnetic resonance investigation of the synthetic deoxyribonucleic acid decamer d(ATATC-GATAT), *Biochemistry 22*, 5943−5951.
49. Patel, D. J., Shapiro, L., and Hare, D. (1986) Sequence-dependent conformation of DNA duplexes. The AATT segment of the d(G-G-A-A-T-T-C-C) duplex in aqueous solution, *J. Biol. Chem. 261*, 1223−1229.
50. Brünger, A. T. (1992) *X-PLOR Manual*, Yale University Press, New Haven, CT.
51. Wijmenga, S. S., Mooren, M. M. W., and Hilbers, C. W. (1993) in *NMR of Macromolecules, A Practical Approach* (Roberts, G. C. K., Ed.) pp 217−288, Oxford University Press, Oxford, U.K.
52. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer, New York.
53. Siegmund, K. (2002) Diploma Thesis, pp 38−43 and 93−97, University of Konstanz, Konstanz, Germany.
54. Schwieters, C. D., and Clore, G. M. (2001) The VMD-XPLOR visualization package for NMR structure refinement, *J. Magn. Res. 149*, 239−244.
55. Kleywegt, G. J., and Jones T. A. (1997) Model-building and refinement practice, *Methods Enzymol. 277*, 208−230.
56. Miki, K., Kasai, N., Shibakami, S., Chirachanchai, S., Takemoto, K., and Miyata, M. (1990) Crystal structure of cholic acid with no guest molecules, *Acta Crystallogr. C46*, 2442−2445.
57. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system (CNS): a new software system for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
58. Stein, E. G., Rice, L. M., and Brünger, A. T. (1997) Torsion angle molecular dynamics: a new, efficient tool for NMR structure elucidation, *J. Magn. Res. Ser. B 124*, 154−164.
59. (a) Ho, W. C., Steinbeck, C., Richert, C. (1999) Solution structure of the aminoacyl-capped oligodeoxyribonucleotide duplex (W-TGCGCAC)2, *Biochemistry 38*, 12597−12606, (b) Tuma, J., Connors, W. H., Stitelman, D. H., Richert, C. (2002) On the effect of covalently appended quinolones on termini of DNA-duplexes, *J. Am. Chem. Soc. 124*, 4236−4246.
60. Wollborn, U., and Leibfritz, D. (1991) Untersuchung des stereochemischen Einflusses von Substituenten in 3-Stellung an Cholsäurederivaten mit Hilfe 1D- und 2D-NMR-Methoden, *Liebigs Ann.*, 1237−1239.
61. Morales, J. C., and Kool, E. T. (1998) Efficient replication between non-hydrogen-bonded nucleoside shape analogues, *Nat. Struct. Biol. 5*, 950−954.
62. Guckian, K. M., Schweitzer, B. A., Ren, R. X. F., Sheils, C. J., Paris, P. L., Tahmassebi, D. C., and Kool, E. T. (1996) Experimental measurement of aromatic stacking affinities in the context of duplex DNA, *J. Am. Chem. Soc., 118*, 8182−8183.
63. Werner, M. H., Gronenborn, A. M., and Clore, G. M. (1996) Intercalation, DNA kinking, and the control of transciption, *Science 271*, 778−784.
64. Barry, J. F., Davis, A. P., Perez-Payan, M. N., Elsegood, M. R. J.; Jackson, R. F. W., Gennari, C., Piarulli, U., and Gude, M. (1999) A trifunctional steroid-based scaffold for combinatorial chemistry, *Tetrahedron Lett. 40*, 2849−2852.
65. Yuan, C., Toshiro, S., and Clark, S. W. (1996) Sequence-selective peptide binding with a peptide-A,B-trans-steroidal receptor selected from an encoded combinatorial library, *J. Am. Chem. Soc. 118*, 1813−1814.
66. Bowe, C. L., Mokhtarzadeh, L., Venkatesan, P., Babu, S., Axelrod, H. R., Sofia, M. J.; Kakarla, R., Chan, T. Y., Kim, J. S.; Lee, H. J., Amidon, G. L., Choe, S. Y., Walker, S., and Kahne, D. (1997) Design of compounds that increase the absorption of polar molecules, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12218−12223.
67. Waring, M. J., and Henley, S. M. (1975) Stereochemical aspects of the interactions between steroidal diamines and DNA, *Nucleic Acids Res. 2*, 567−586.

68. Kerr, T. A., Saeki, S., Schneider, M., Schaefer, K., Berdy, S., Redder, T., Shan, B., Russell, D. W., and Schwarz, M. (2002) Loss of nuclear receptor SHP impairs but does not eliminate negative feedback regulation of bile acid synthesis, *Dev. Cell 2*, 713−720.
69. Sobell, H. M., Tsai, C.-C., Jain, S. C., and Gilbert, S. G. (1977) Visualization of drug-nucleic acid interactions at atomic resolution, *J. Mol. Biol. 114*, 333−365.
70. Similarly proposed for steroidal diamines in: Gourevitch, M., Puidomenech, P., Cave, A., Etienne, G., Mery, J., and Parello, J. (1974) Model studies in relation to molecular structure if chromatin, *Biochimie 56*, 967−985.
71. Hui, X., Gresh, N., and Pullman, B. (1989) Modelling basic features of specificity in the binding of a dicationic steroid diamine to double-stranded oligonucleotides, *Nucleic Acids Res. 17*, 4177−4187.
72. Lehmann, T. J., and Engels, J. W. (2001) Synthesis and properties of bile acid phosphoramidites 5′-tethered to antisense oligodeoxynucleotides against HCV, *Bioorg. Med. Chem. 9*, 1827−1835.
73. Kato, T., Takemura, T., Yano, K., Ikebukuro, K., and Karube, I. (2000) In vitro selection of DNA aptamers which bind to cholic acid, *Biochim. Biophys. Acta 2000, 1493*, 12−18.
74. Kato, T., Yano, K., Ikebukuro, K., and Karube, I. (2000) Bioassay of bile acids using an enzyme-linked DNA aptamer, *Analyst 125*, 1371−1373.
75. Pons, J. L., Malliavin, T. E., and Delsuc, M. A. (1996) Gifa V4: a complete package for NMR data-set processing, *J. Biomol. NMR 8*, 445−452.
76. Humphrey, W., Dalke, A., Schulten, K. (1996) VMD − Visual Molecular Dynamics, *J. Mol. Graphics* **1996**, *14*, 33−38.

BI030056G