# Effects of nucleoside analog incorporation on DNA binding to the DNA binding domain of the GATA-1 erythroid transcription factor

Matthew Foti[a,c], James G. Omichinski[d], Stephen Stahl[d], Donna Maloney[a], John West[a], Barry I. Schweitzer[a,b,*]

[a] *Walt Disney Memorial Cancer Institute at Florida Hospital, 12722 Research Parkway, Orlando, FL 32826, USA*
[b] *Department of Chemistry, University of Central Florida, Orlando, FL 32816, USA*
[c] *Department of Chemistry, University of South Florida, Tampa, FL 33620-8800, USA*
[d] *Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, MD 20892-0520, USA*

Received 16 November 1998

**Abstract We investigate here the effects of the incorporation of the nucleoside analogs araC (1-β-D-arabinofuranosylcytosine) and ganciclovir (9-[(1,3-dihydroxy-2-propoxy)methyl] guanine) into the DNA binding recognition sequence for the GATA-1 erythroid transcription factor. A 10-fold decrease in binding affinity was observed for the ganciclovir-substituted DNA complex in comparison to an unmodified DNA of the same sequence composition. AraC substitution did not result in any changes in binding affinity. $^{1}H$-$^{15}N$ HSQC and NOESY NMR experiments revealed a number of chemical shift changes in both DNA and protein in the ganciclovir-modified DNA-protein complex when compared to the unmodified DNA-protein complex. These changes in chemical shift and binding affinity suggest a change in the binding mode of the complex when ganciclovir is incorporated into the GATA DNA binding site.**


*Key words:* Ganciclovir; Cytosine arabinoside; GATA-1; Binding; Nuclear magnetic resonance

## 1. Introduction

Nucleoside analogs have found widespread use in modern medicine but relatively little information is available concerning the molecular mechanisms that account for their potency. Many of these analogs are incorporated into DNA but do not act as direct chain terminators, suggesting a structural origin for their inhibitory effects. Despite the lack of information about these structural effects, very few studies have been aimed at determining the effects of these analogs on protein-DNA interactions. In the present study, we look at two clinically relevant nucleoside analogs and their effects on a DNA binding protein when incorporated into DNA oligonucleotides.

The nucleoside analog 9-[(1,3-dihydroxy-2-propoxy)-methyl]guanine (ganciclovir, DHPG) is an effective antiviral agent used in the treatment of a variety of herpes simplex virus (HSV) and cytomegalovirus (CMV) infections. Ganciclovir has been used most frequently to treat patients with immune system deficiencies such as AIDS patients and those undergoing chemotherapy and bone marrow transplants [1–6]. Recently, ganciclovir has found additional use in the form of a gene therapy agent in the treatment of malignant tumors [7–15].

Although ganciclovir has found many clinical uses, a complete understanding of its mechanism of action is not available. Ganciclovir is a simple analog of deoxyguanosine that lacks the 2′ carbon atom of normal deoxyguanosine but still retains a 3′ and 5′ hydroxyl group, allowing it to be incorporated into internal positions in DNA. Ganciclovir is converted to the chiral (*S*)-triphosphate form in virally infected cells where it may competitively inhibit DNA polymerases. The analog can also become incorporated into newly synthesized DNA where it interferes with further DNA replication. The level of ganciclovir incorporation into DNA has been correlated with its toxic effects on treated cells [16,17]. DNA polymerases incorporate ganciclovir into primer strands followed by an additional nucleotide before chain extension is strongly inhibited [18,19].

We have recently determined the solution structure of a DNA duplex containing ganciclovir at an internal position of the sequence [20]. Significant structural distortions localized around the ganciclovir site were observed, including a kink in the sugar-phosphate backbone at the site of ganciclovir incorporation. Other structural abnormalities included changes in the conformation of the residue immediately 3′ to the ganciclovir residue. This deoxyribose ring exhibited a C2′ endo conformation in an unmodified control DNA of the same sequence but assumed an A-type C3′ endo conformation in the modified DNA. In addition to the unusual sugar conformation, an A-type glycosidic torsion angle was also observed at this residue.

The nucleoside analog arabinosylcytosine (1-β-D-arabinofuranosylcytosine, araC) is structurally equivalent to deoxycytosine but possesses a hydroxyl group at the 2′ position of the deoxyribose ring in an arabinose orientation. It is used clinically to treat various forms of leukemia [21,22]. AraC is converted to araCTP and may competitively inhibit DNA polymerases. This analog is also incorporated into DNA primer strands and inhibits chain extension [23–28]. The solution structure of a DNA duplex containing araC has recently been described [29]. Although no major distortions were observed in the structure when compared to a control sequence, $^{31}P$ relaxation studies indicated a decrease in conformational mobility around the araC site [30].

While several laboratories, including ours, have examined the structural effects of these and other nucleoside analogs in free duplex DNA, it cannot be assumed that the structural changes observed in the free DNA predominate in a DNA-protein complex. Although numerous studies have been carried out on the biochemical effects of araC and ganciclovir on DNA polymerase function, structural studies on polymerase-

*Corresponding author. Molecular Staging, Inc., 4 Pearl Court, Allendale, NJ 07401, USA. Fax: (1) (201) 818-4561.
E-mail: barrys@molecularstaging.com


PII: S0014-5793(99)00026-5

DNA complexes have proven to be quite challenging. In addition, it is unlikely that the consequences of incorporation of these analogs into DNA are limited to DNA polymerases. For example, it has recently been shown that incorporation of the nucleoside analog 1(2-deoxy-2-fluoro-β-D-arabinofuranosyl)-5-iodouracil (FIAU) into DNA had significant effects on the binding of the transcription factors AP-1 and TFIID to their target DNA sequences [31]. Thus, the biological activity and/or cytotoxicity of nucleoside analogs may be due, in part, to the ability of these compounds to alter DNA-protein contacts and thereby modulate gene expression.

We have chosen to examine the effects of nucleoside analogs in the context of a model DNA binding protein – the DNA binding domain of the chicken erythroid transcription factor GATA-1 (cGATA-1). The GATA-1 transcription factor is required for transcription in erythroid cell lines and has been shown to be required for normal hematopoietic development [32–34]. Interestingly, ganciclovir has been shown to inhibit normal hematopoiesis [35,36]. GATA-1 functions by binding a (A/T)GATA(A/G) motif within regulatory regions of promoters and enhancers to activate transcription. The GATA-1 protein contains two zinc finger domains; the carboxyl terminal domain (designated herein the cGATA-1 DNA binding domain) has been shown to be sufficient for DNA binding [37]. The solution structure of the complex between the cGATA-1 DNA binding domain and a 16 bp duplex DNA (Fig. 1A) has recently been determined at high resolution [38].

In the present study, we have incorporated ganciclovir at the G position of the GATA recognition site of cGATA-1 and have compared the binding affinities of the complex of the cGATA-1 DNA binding domain with the modified DNA and the complex with the unmodified (control) DNA. In addition, we have qualitatively probed the structural effects of ganciclovir incorporation on the DNA-protein complex with heteronuclear NMR chemical shift analysis of protein backbone proton/nitrogen and homonuclear chemical shift analysis of DNA exchangeable proton resonances. We have also examined the effects of including the antineoplastic drug araC in the binding region of the DNA duplex.

## 2. Materials and methods

### 2.1. Preparation of the cGATA-1 DNA binding domain

The $^{15}N$ labeled DNA binding domain of the chicken GATA-1 protein was expressed in *Escherichia coli* in minimal medium containing [$^{15}N$]$NH_4Cl$ as the sole nitrogen source. The crude protein was purified on DEAE-Sepharose Fast Flow (Pharmacia) followed by S-Sepharose Fast Flow (Pharmacia). The protein was then purified by reverse phase HPLC in TFA/acetonitrile. HPLC fractions were pooled and lyophilized to dryness. The lyophilized protein was then suspended as 0.15 mM solution with 1.1 eq. of zinc in 0.05% TFA. The pH of the sample was slowly raised to 6.5 with dilute NaOH. A complete description of the expression and purification of the cGATA-1 DNA binding domain can be obtained from Omichinski et al. [37,38].

### 2.2. Oligonucleotide synthesis and DNA duplex preparation

The araC phosphoramidite was supplied by Glen Research (Sterling, VA). The enantiomerically pure ganciclovir phosphoramidite was synthesized as described in Marshalko et al. [39]. All oligonucleotides were synthesized by Oligos etc. (Wilsonville, OR). The DNA control duplex strands 5′-GTTGCAGATAAACATT-3′ and 5′-AATGTTTATCTGCAAC-3′ were synthesized on a 10 μmol scale and purified by reverse phase HPLC before and after detritylation. The modified araC strand 5′- AATGTTTATXTGCAAC-3′ (X = araC) was synthesized on a 0.2 μmol scale, detritylated, and gel purified. The ganciclovir-containing oligonucleotide strand 5′-GTTGCAZATAAACATT-3′ (Z = ganciclovir) was synthesized in two 1.0 μmol scale syntheses, and purified by reverse phase HPLC before and after detritylation. The annealing of both the control and ganciclovir duplexes were monitored by $^1H$ NMR. Single strands were titrated with their complementary strand until NMR signals from each strand became equally intense at 65°C, then cooled slowly to room temperature. The araC duplex was annealed by combining equal quantities of each strand in $H_2O$. The solution was then heated to 90°C and cooled slowly to room temp. The annealed product was checked by gel electrophoresis.

### 2.3. Preparation of radiolabeled probes

Double stranded oligonucleotides were end-labeled with $^{32}P$ in a reaction mixture containing 100 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 5 mM DTT, 10 units of T4 polynucleotide kinase (Stratagene), 10 μCi [γ-$^{32}P$]ATP (Amersham), and 10 pmol of oligonucleotide in a total volume of 20 μl. Unincorporated [γ-$^{32}P$]ATP was removed from the reaction with Sephadex G-25 spin columns (Boehringer Mannheim).

### 2.4. Gel mobility shift assays

Unless otherwise specified, gel mobility shift assays were performed with 6 pmol of peptide in a total volume of 20 μl containing 25 mM HEPES pH 7.9, 50 mM KCl, 6.25 mM $MgCl_2$, 0.5 mM DTT, 0.1 mM EDTA, 0.1% NP40, and 5% glycerol. DNA was added to the reaction mixture prior to a 15 min incubation on ice. After incubation samples were loaded onto a 5% non-denaturing polyacrylamide gel (29:1 acrylamide:bis-acrylamide, running buffer: 0.5×TBE, pH 8.3). To ensure the binding reactions entered the gel as quickly as possible, samples were loaded onto a running gel. Electrophoresis was carried out 4°C for 45 min at 15 V/cm. Gels were dried under vacuum at 80°C for 1 h, and radioactivity in DNA bands was quantified with a Betascope (Betagen). All binding experiments were carried out in triplicate.

### 2.5. NMR spectroscopy

The ganciclovir-modified DNA-protein complex was prepared by titrating a DNA solution (1.3 ml, ∼200 μM DNA, 10 mM NaCl) with the cGATA-1 DNA binding domain (∼0.14 mM, 0.05% TFA, pH 6.5) until a 1:1 ratio was obtained (a precipitate formed if the peptide was added in excess). The solution was then concentrated with a Centricon-3 ultrafiltration device (Amicon). The final solution was adjusted to a volume of 250 μl, and was placed in a Shigemi 5 mm microtube. The final sample contained 1.0 mM complex and 10 mM NaCl in 10% $D_2O$/90% $H_2O$. The unmodified DNA-protein complex was prepared in the same manner.

All NMR experiments were performed on a Varian 600 MHz Unity Plus spectrometer at a regulated temperature of 20°C. Pulsed field gradient sensitivity-enhanced $^1H$-$^{15}N$ HSQC [40] spectra were collected for both the control and ganciclovir-containing DNA-protein complexes. The water resonance was suppressed with a water flipback selective pulse. The data were acquired in the phase-sensitive mode of States-TPPI with 2048 complex points in $t_2$ ($^1H$) and 256 complex points in $t_1$ ($^{15}N$). The $^1H$ and $^{15}N$ spectral widths were 7500 Hz and 2000 Hz, respectively. GARP $^{15}N$ decoupling was utilized during the acquisition period. 200 ms NOESY spectra for the control and ganciclovir-DNA complex were acquired with an excitation sculpting pulse for solvent suppression [41]. Selective 180° pulses were applied

A Control Duplex:

5′-$G_1$ $T_2$ $T_3$ $G_4$ $C_5$ $A_6$ $G_7$ $A_8$ $T_9$ $A_{10}A_{11}A_{12}C_{13}A_{14}T_{15}T_{16}$-3′
3′-$C_{32}A_{31}A_{30}C_{29}G_{28}T_{27}C_{26}T_{25}A_{24}T_{23}T_{22}T_{21}G_{20}T_{19}A_{18}A_{17}$-5′

B Ganciclovir Duplex:

5′-$G_1$ $T_2$ $T_3$ $G_4$ $C_5$ $A_6$ **$Z_7$** $A_8$ $T_9$ $A_{10}A_{11}A_{12}C_{13}A_{14}T_{15}T_{16}$-3′ Z= ganciclovir
3′-$C_{32}A_{31}A_{30}C_{29}G_{28}T_{27}C_{26}T_{25}A_{24}T_{23}T_{22}T_{21}G_{20}T_{19}A_{18}A_{17}$-5′

C AraC Duplex:

5′-$G_1$ $T_2$ $T_3$ $G_4$ $C_5$ $A_6$ $G_7$ $A_8$ $T_9$ $A_{10}A_{11}A_{12}C_{13}A_{14}T_{15}T_{16}$-3′ X = araC
3′-$C_{32}A_{31}A_{30}C_{29}G_{28}T_{27}$**$X_{26}$**$T_{25}A_{24}T_{23}T_{22}T_{21}G_{20}T_{19}A_{18}A_{17}$-5′

Fig. 1. Sequence composition of DNA duplexes containing the GATA binding site. A: Unmodified duplex. B: Ganciclovir-substituted duplex. C: AraC-substituted duplex.

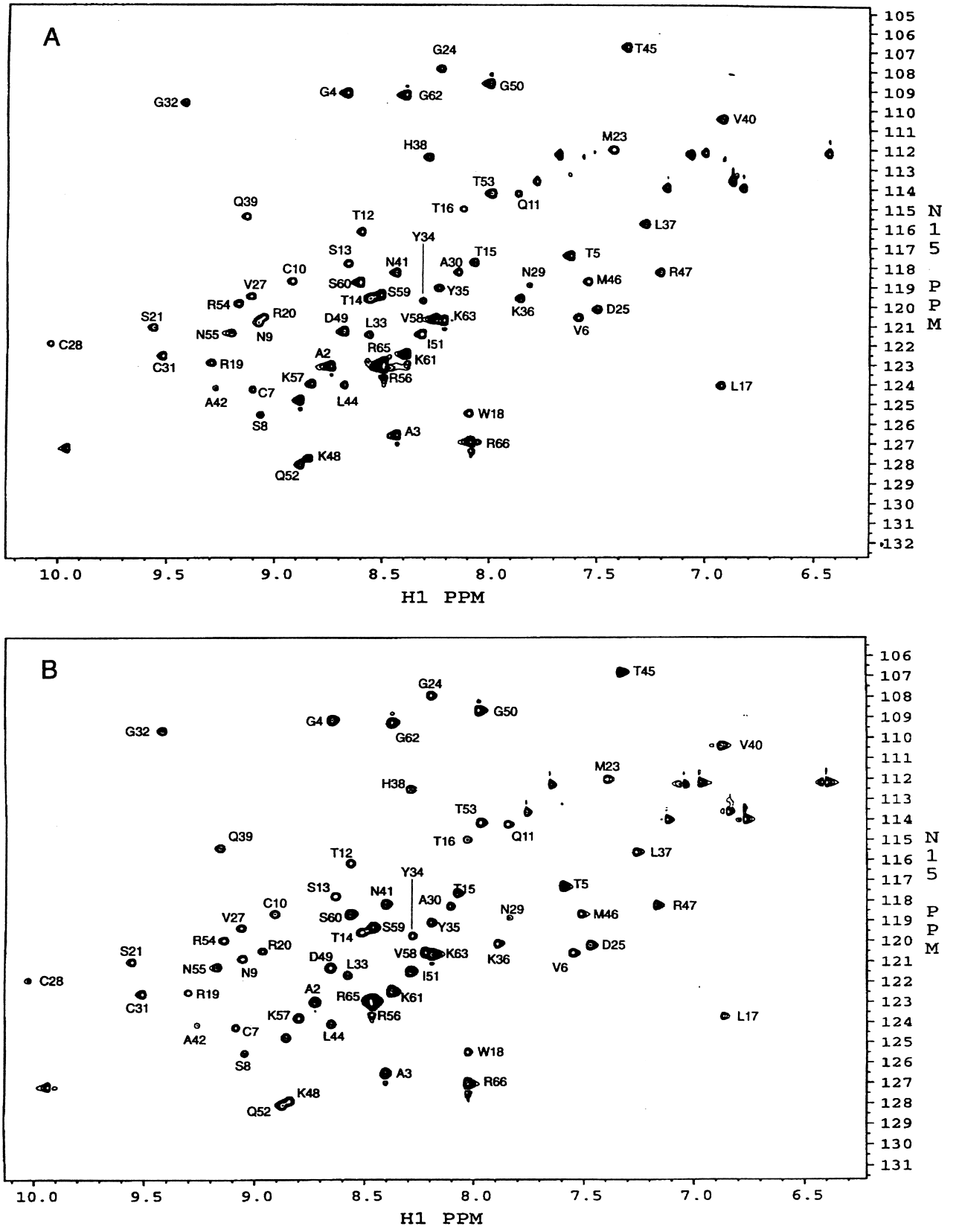


Fig. 2. $^{1}H$-$^{15}N$ HSQC 2D contour plots of the control unmodified DNA-protein complex (A) and the ganciclovir DNA-protein complex (B). Assignments were made as described in Section 3.

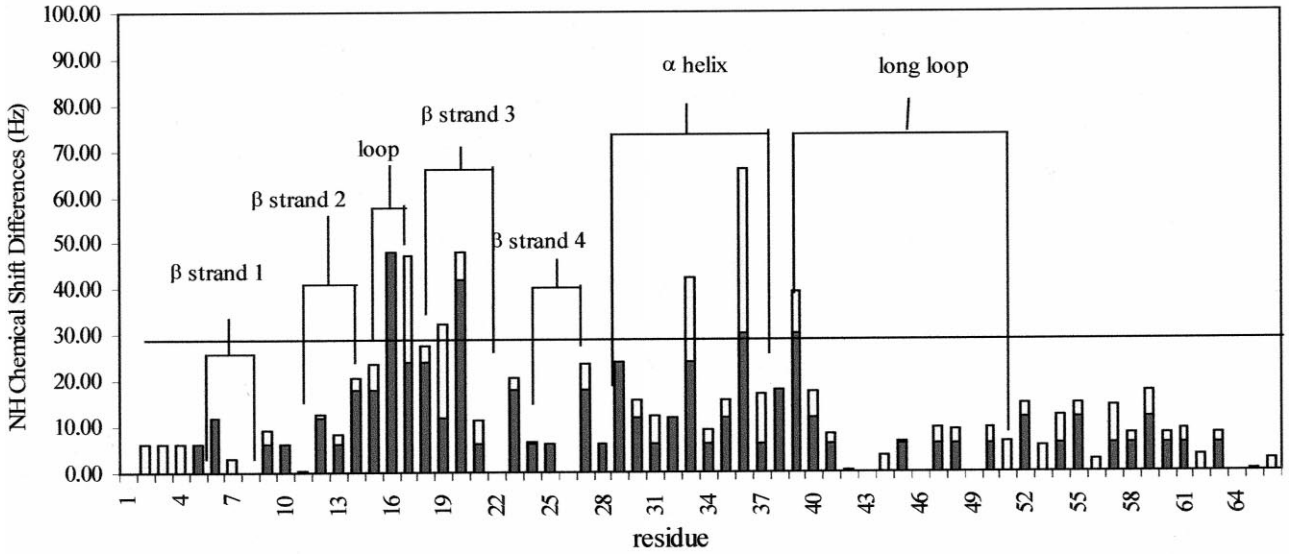


Fig. 3. Sum of the absolute value of $\Delta\delta^1$H and $\Delta\delta^{15}$N chemical shift differences for backone resonances of the ganciclovir DNA-protein complex and control unmodified DNA-protein complex. $^1$H chemical shift differences are in black and $^{15}$N chemical shift differences are in white. Protein secondary structure is indicated for selected regions. The horizontal segmented line represents twice the standard deviation of the mean chemical shift difference.

with a SEDUCE profile to suppress the water resonance. Gradient pulses were applied along the $z$ axis for 1 ms at 15 and 3 G/cm with 50 μs delays before and after the gradient pulses. These experiments were collected with States-TPPI phase cycling with 2048 complex points in $t_2$ and 1024 complex points in $t_1$, a relaxation delay of 2.5 s, and a spectral width of 12 500 Hz.

NMR data were processed with NMRpipe software [42]. The HSQC and NOESY experiments were processed with a cosine bell squared window function in the direct dimension and a squared-sine bell shifted 60° in the indirect dimension. Data in the indirect dimension for both sets of experiments was zero-filled to 2048 points. After zero filling, the NOESY spectra had a digital resolution of 6 Hz/pt and the HSQC 3.4 Hz/pt in proton and 1 Hz/pt in nitrogen.

## 3. Results

### 3.1. DNA binding

The consequences on DNA binding affinity of nucleoside analog incorporation into the DNA binding sequence of GATA-1 were assessed with a gel mobility shift assay. Three shift assays were performed in this study: (1) a control unmodified DNA-cGATA-1 protein assay, (2) a ganciclovir-modified DNA-cGATA-1 protein assay, and (3) an araC-modified DNA-cGATA-1 protein assay. Radioactivity in bands representing bound and free DNA were quantified and analyzed by Scatchard analysis to provide dissociation constants ($K_d$) for the DNA-protein interactions. The dissociation constant for the unmodified control DNA-cGATA-1 protein complex was determined to be $10 \pm 5$ nM, which is equivalent to the value that was previously observed for this sequence [37]. Scatchard analysis for the ganciclovir-modified DNA-cGATA-1 protein complex gave a dissociation constant of $100 \pm 20$ nM, 10-fold larger than the control. In contrast to the effect on DNA binding produced by the ganciclovir modification, the introduction of araC into the DNA duplex did not significantly affect the dissociation constant ($K_d = 10 \pm 1$ nM).

### 3.2. Chemical shift analysis

The reduction in binding affinity produced by the ganciclovir modification prompted us to examine the protein for structural changes using the $^1$H-$^{15}$N HSQC experiment. The results of $^1$H-$^{15}$N HSQC experiments on both the control and ganciclovir-modified DNA-cGATA-1 protein complex are shown in Fig. 2. The chemical shifts of the backbone $^1$H-$^{15}$N resonances in the control DNA-cGATA-1 protein complex experiment were essentially identical to those that were previously reported [38]; consequently, assignments of the peaks were made from this earlier work. Assignments of the peaks in the HSQC spectrum of the ganciclovir-modified DNA-cGATA-1 protein complex were made by comparison with the control complex spectrum since only relatively small changes were observed (proton $< 60$ Hz, nitrogen $< 40$ Hz). The control DNA-cGATA-1 protein complex and the ganciclovir-modified DNA-cGATA-1 protein complex resonances were nearly superimposable except for several changes that were observed at key protein-DNA contact points. The absolute value of the backbone $^1$H-$^{15}$N chemical shift difference between the control and ganciclovir-modified DNA-cGATA-1 protein complexes are plotted in Fig. 3. Although the chemical shift changes are dispersed throughout the protein sequence, the significant changes ($> 2 \times$S.D.) are concentrated in two distinct regions of the protein. The first of these two regions includes residues 15–19, which make up β-strand 3 and the loop connecting it to β-strand 2. The second region includes residues 28–38, which make up the α-helix that lies in the major groove of the DNA. Both of these regions are involved in base contacts in the major groove of the DNA.

Seven residues in the ganciclovir-modified DNA-cGATA-1 protein complex exhibit combined $^1$H-$^{15}$N chemical shift changes greater than twice the standard deviation of the mean difference. These residues include threonine-16, leucine-17, arginine-19, arginine-20, leucine-33, lysine-36, and glutamine-39. The combined backbone $^1$H-$^{15}$N chemical shift differences for these residues are 48, 47, 32, 48, 42, 66, and 39 Hz, respectively. With the exception of glutamine-39 all the residues either interact directly with the DNA or are part of

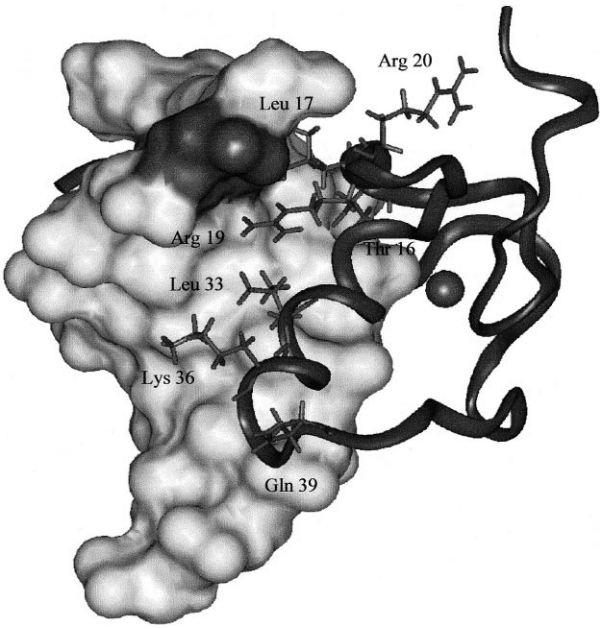


Fig. 4. Model illustrating the interaction of the DNA binding domain of cGATA-1 with unmodified DNA containing the GATA binding site [37]. The DNA is depicted as a molecular surface with site of ganciclovir incorporation in the ganciclovir DNA-protein complex (residue G7) highlighted. The protein is represented as a backbone ribbon encapsulating the zinc atom. The side chains of protein residues exhibiting backbone $^{1}H$-$^{15}N$ chemical shift changes in the ganciclovir DNA-protein complex are shown. The model was generated with Insight95.0 (Molecular Simulations, Inc.).

secondary structure that interacts with the DNA in the control unmodified DNA-cGATA-1 protein complex [38].

In order to provide information on the structural integrity of the DNA, a chemical shift comparison was performed for the exchangeable and H2 DNA protons in the control and ganciclovir-modified DNA-protein complex. Assignment of the guanosine H1, cytosine NH′/H″, thymine H3, and adenine H2 protons in both complexes were made from 2-D NOESY spectra utilizing an excitation sculpting technique to suppress water [41]. Chemical shift changes in the imino and amino protons of the DNA in both complexes are consistent with one another with the exception of three residues, including the ganciclovir (G7) site. In the ganciclovir-modified DNA-protein complex, upfield shifts of 150 Hz and 120 Hz are observed for the imino protons of T25 and T27, respectively. The imino proton of ganciclovir (G7 residue site) exhibits a downfield shift of 60 Hz.

## 4. Discussion

Both araC and ganciclovir have been studied extensively from the standpoint of the biochemical consequences of these analogs on DNA polymerase function. Much is also known about the structural effects of these analogs when incorporated into DNA duplexes. While it is important to fully understand the effects of the analogs in free duplex DNA, it cannot be assumed that these same structural consequences predominate in a DNA-protein complex. The biophysical experiments presented herein represent the first results obtained with these analogs and a protein that is not a DNA polymerase.

When araC, an analog of cytosine, is incorporated into DNA, few significant structural distortions are observed in a free DNA duplex in solution [29]. In the solution structure of the control DNA-cGATA-1 protein complex, the side chains of arginine-56 and serine-59 have hydrophobic interactions with the furanose ring of C26 in the minor groove of the DNA [38]. The 2′-hydroxyl group of araC, however, protrudes into the major groove of the DNA [29] in a location where no interactions occur with the protein. Given the above information, the lack of a significant change in binding affinity in the cGATA-1 DNA binding domain complex when araC is incorporated at position 26 of the DNA binding sequence was not unexpected. Although the araC modified DNA-cGATA-1 protein complex was not included in the present NMR studies, it may be worthwhile in future studies to examine whether the hydroxyl group on araC participates in hydrogen bonding or alters the dynamics in the complex. The 2′ hydroxyl group of araC appears to be capable of forming intramolecular hydrogen bonds with backbone phosphate groups in self-complementary duplexes [29] and in primer-template duplexes (unpublished results).

Ganciclovir, an analog of deoxyguanosine, lacks a 2′ carbon atom resulting in an acyclic sugar. Using high resolution NMR, we have demonstrated that when incorporated into a DNA decamer duplex, ganciclovir significantly perturbs the local structure of the double helix [20]. More specifically, ganciclovir induces a conformational change in the residue 3′ to the incorporation site and creates a kink in the backbone. The base of ganciclovir exhibits some additional stacking interactions in the helix but generally is not perturbed [20]. Despite these significant changes in the structure of free DNA, it was still not clear to us how ganciclovir incorporation would affect the structure of a DNA-protein complex.

When bound to DNA, the cGATA-1 DNA binding domain protein folds into a structure containing four β-strands (residues 5–27), an α-helix (residues 28–38), and a long loop (residues 39–51). The α-helix and the loop connecting strands β2 and β3 of the protein orient in the major groove of the DNA while the carboxyl-terminal tail wraps around the DNA and lies in the minor groove [38]. Many of these interactions involve hydrophobic contacts with the furanose rings and bases of various residues within the DNA duplex, including residue G7, the site of ganciclovir incorporation. Consequently, it was not very surprising to find that ganciclovir incorporation had a negative impact on the DNA binding affinity to this zinc finger protein. Due to the greater degrees of freedom allowed in DNA containing ganciclovir, the decrease in binding affinity could have been entirely due to entropic effects. As described in more detail below, the NMR data that we have obtained implies that this change in binding affinity has a structural basis.

Seven residues in the ganciclovir-modified DNA-cGATA-1 protein complex exhibit backbone $^{1}H$-$^{15}N$ chemical shift changes $>29$ Hz (twice the standard deviation of the mean difference) compared to the unmodified DNA-cGATA-1 protein complex (Fig. 3). These residues include threonine-16, leucine-17, arginine-19, arginine-20, leucine-33, lysine-36, and glutamine-39. Threonine-16, leucine-17, and arginine-19 are part of β-strand 3 and the loop connecting it to β-strand 2 (Fig. 4) [38]. Threonine-16 makes hydrophobic contacts with DNA residues A24 and T25 in the control DNA-protein complex [38]. A combined $^{1}H$-$^{15}N$ shift of 48 Hz is observed for

this amino acid in the ganciclovir-modified DNA-cGATA-1 protein complex. The side chain of leucine-17, which exhibits a combined amide $^{1}H$-$^{15}N$ shift of 47 Hz, forms hydrophobic interactions with the bases of A6, G7 (ganciclovir site), and T25 (complementary to A8) in the major groove of the DNA. The side chain of arginine-19 forms salt bridges and/or hydrogen bonds with the 5′ phosphate of the G7 residue in the control DNA-cGATA-1 protein complex. The 32 Hz upfield change in chemical shift of the amide nitrogen in arginine-19 in the ganciclovir-modified DNA-cGATA-1 protein complex could be due to a conformational change in the residue in response to an altered backbone shape in DNA containing ganciclovir.

The combined amide $^{1}H$-$^{15}N$ shifts of leucine-33 and lysine-36 were 42 Hz and 66 Hz, respectively. Both residues are part of the α-helix that lies in the major groove of the DNA and interacts primarily through hydrophobic contacts with the bases and sugars in the GATA recognition sequence (G7 to T9; T22 to T26). In the case of leucine-33, the side chain makes contacts with the sugars of T23 and A24 in the strand opposite the site of ganciclovir incorporation. The amide $^{1}H$-$^{15}N$ of lysine-36 exhibits the largest change in chemical shift but was not shown to contact the DNA in the solution structure of the control DNA-cGATA-1 protein complex [38], although it lies very close (ca. $<7$ Å) to the base of DNA residue T9.

Arginine-20 and glutamine-39 exhibit absolute amide $^{1}H$-$^{15}N$ shifts of 48 Hz and 39 Hz, respectively. Neither residue was previously reported to contact the DNA in the unmodified DNA-cGATA-1 protein complex. However, arginine-20 is part of β-strand 3 and the loop connecting it to β-strand 2, and glutamine-39 is the N-terminal residue connecting to the α-helix which probes the DNA major groove (Fig. 4) [38].

Chemical shift changes of the exchangeable and H2 DNA protons are limited to the ganciclovir residue and the A-T base pairs adjacent to the ganciclovir residue (Fig. 1B). The 60 Hz downfield shift of the ganciclovir imino proton was not previously observed in free duplex DNA [20]. Whether the changes in chemical shift observed in the DNA exchangeable protons around the ganciclovir residue are the result of the ganciclovir substitution on the structure of the bound DNA or on interactions with the protein is therefore uncertain. In a free decamer DNA duplex containing ganciclovir, an upfield shift of 84 Hz is observed for the imino proton of the thymine in the same relative location as T25 is in the 16-mer duplex in this study (i.e. one residue 3′ to ganciclovir in the complementary strand). Analysis of the solution structure of the decamer suggested that this chemical shift was probably due to the markedly increased opening and slide observed at this base pair [20]. Although the chemical shift in the free duplex is not as large as the 150 Hz shift seen for the T25 imino proton in the protein-bound DNA, the changes may be of the same origin. The side chain of leucine-17 forms hydrophobic interactions with the base of T25 as previously discussed. The chemical shift in the amide group of leucine-17 may also be related to the increased shift observed at T25. The imino proton of T27 in the ganciclovir-DNA complex shows an upfield shift of 120 Hz. An A-T base pair was not located 5′ to ganciclovir residue in the free decamer duplex previously studied [20], so no direct comparison between that duplex and residue T27 in the DNA-protein complex can be made.

The results that we have presented suggest a change in the binding interaction between the cGATA-1 DNA binding domain and DNA containing ganciclovir. Amide proton and nitrogen chemical shift analysis suggests that the changes in the interaction are localized to the α-helix, β-strand 3 and the loop connecting it to β-strand 2 of the protein. The amino acid residues in these regions probe the major groove in the control DNA-protein complex at the site of ganciclovir incorporation (G7) and 3′ to it [38]. $^{1}H$-$^{15}N$ chemical shift changes observed in residues 17 and 19 suggest the presence of a local distortion in protein structure in response to the ganciclovir substitution in the GATA site. These residues interact with the ganciclovir residue directly and it would not be unreasonable to postulate some conformational changes in these protein residues. However, other chemical shift data suggests a larger or more global effect such as a change in the binding mode of the complex when ganciclovir is incorporated into the GATA site. The decrease in binding affinity observed for the ganciclovir DNA-protein complex may be the result of an altered or secondary mode of binding whereby the normal secondary structure of the protein is altered to accommodate a distorted DNA structure.

*Acknowledgements:* We would like to thank to Dr. Kevin Gardner (University of Toronto) for his critical review of the manuscript, and Dr. Lewis Kay (University of Toronto) for providing us with the pulsed field gradient sensitivity-enhanced $^{1}H$-$^{15}N$ HSQC pulse sequence.

## References

[1] Erice, A., Jordan, M.C., Chace, B.A., Fletcher, C., Chinnock, B.J. and Balfour, H.H. (1987) J. Am. Med. Assoc. 257, 3082–3087.
[2] Zaia, J.S. (1993) Clin. Infect. Dis. 17, S392–399.
[3] Buckner, F.S. and Pomeroy, C. (1993) Clin. Infect. Dis. 17, 644–656.
[4] Nevins, T.E. and Dunn, D.L. (1992) J. Am. Soc. Nephrol. 2, S270–S273.
[5] Faulds, D. and Heel, R.C. (1990) Drugs 39, 597–638.
[6] Emmanuel, D. (1990) Semin. Hematol. 27, 22–27.
[7] Culver, K.W., Ram, Z., Walbridge, S., Ishii, H., Oldfield, E.H. and Blaise, R.M. (1992) Science 256, 1550–1552.
[8] Culver, K.W., Van Gilder, J., Link, C.J., Carlstrom, T., Buroker, T., Yuh, W., Koch, K., Schabold, K., Doornbas, S., Wetjen, B. and Blaise, R.M. (1994) Hum. Gene Ther. 5, 343–379.
[9] Smythe, W.R., Hwang, H.C., Amin, K.M., Eck, S.L., Davidson, B.L., Wilson, J.M., Kaiser, L.R. and Albelda, S.M. (1994) Cancer Res. 54, 2055–2059.
[10] Chen, S.H., Shine, H.D., Goodman, J.C., Grossman, R.G. and Woo, S.L. (1994) Proc. Natl. Acad. Sci. USA 91, 3054–3057.
[11] Hasegawa, Y., Emi, N. and Shimokata, K. (1995) J. Mol. Med. 73, 107–112.
[12] Ido, A., Nakata, K., Katau, Y., Nakao, K., Murata, K., Fujita, M., Ishii, N., Tamaoki, T., Shiku, H. and Nagataki, S. (1995) Cancer Res. 35, 3100–3109.
[13] Yoshida, K., Kawami, H., Yamaguchi, Y., Kuniyasu, H., Nishiyama, M., Hirai, T., Yanagihara, K., Tahara, E. and Toge, T. (1995) Cancer 75, 1467–1471.
[14] Manome, Y., Abe, M., Hagen, M.F., Fine, H.A. and Kufe, D.W. (1994) Cancer Res. 54, 5408–5413.
[15] DiMaio, J.M., Clary, B.M., Bia, D.F., Coveney, E., Pappas, T.N. and Lyerly, H.K. (1994) Surgery 116, 205–213.
[16] Cheng, Y.C., Grill, S., Dutschman, G.E., Nakayama, K. and Bastow, K.F. (1983) J. Biol. Chem. 258, 1260–1264.
[17] Frank, K.B., Chiou, J.F. and Cheng, Y.C. (1984) J. Biol. Chem. 259, 1566–1569.
[18] Reardon, J.E. (1989) J. Biol. Chem. 264, 19039–19044.
[19] Reid, R., Mar, E.-C., Huang, E.-S. and Topal, M.D. (1988) J. Biol. Chem. 263, 3898–3904.

[20] Foti, M., Marshalko, S., Schurter, E., Kumar, S., Beardsley, G.P. and Schweitzer, B.I. (1997) Biochemistry 36, 5336–5345.
[21] Frei, E., Bickers, J., Lane, M., Leary, W. and Talley, R. (1969) Cancer Res. 29, 1325–1352.
[22] Bodey, G., Freireich, E., Monto, R. and Hewlett, J. (1969) Cancer Chemother. Rep. 53, 59–66.
[23] Furth, J.J. and Cohen, S.S. (1968) Cancer Res. 28, 2061–2067.
[24] Momparler, R. (1972) Mol. Pharmacol. 8, 362–370.
[25] Yoshida, S., Yamada, M. and Masaki, S. (1977) Biochim. Biophys. Acta 477, 144–150.
[26] Dicioccio, R.A. and Srivastava, B.I.S. (1977) Eur. J. Biochem. 79, 411–418.
[27] Fridland, A. (1977) Biochemistry 16, 5308–5312.
[28] Cozzarelli, N.R. (1977) Annu. Rev. Biochem. 46, 641–684.
[29] Schweitzer, B.I., Mikita, T., Kellogg, G.W., Gardner, K.H. and Beardsley, G.P. (1994) Biochemistry 33, 11460–11475.
[30] Schweitzer, B.I., Gardner, K.H. and Tucker-Kellogg, G. (1995) J. Biol. NMR 6, 180–188.
[31] Staschke, K.A., Richardson, K.K., Mabry, T.E., Baxter, A.J., Scheuring, J.C., Huffman, D.M., Smith, W.C., Richardson, F.C. and Colacino, J.M. (1996) Nucleic Acids Res. 24, 4111–4116.
[32] Pevny, L., Simon, M.C., Robertson, E., Klein, W.H., Tsai, S.F., D'Agati, V., Orkin, S.H. and Costantini, F. (1991) Nature 349, 257–260.
[33] Tsai, S.F., Martin, D.I., Zon, L.I., D'Andrea, A.D., Wong, G.G. and Orkin, S.H. (1989) Nature 339, 446–451.
[34] Pandolfi, P.P., Roth, M.E., Karis, A., Leonard, M.W., Dzierzak, E., Grosveld, F.G., Engel, J.D. and Lindenbaum, M.H. (1995) Nature Genet. 11, 40–44.
[35] Nassiri, M.R., Emerson, S.G., Devivar, R.V., Townsend, L.B., Drach, J.C. and Taichman, R.S. (1996) Br. J. Haematol. 93, 273–279.
[36] Dornsife, R.E. and Averett, D.R. (1996) Antimicrob. Agents Chemother. 40, 514–519.
[37] Omichinski, J.G., Trainor, C., Evans, T., Gronenborn, A.M., Clore, G.M. and Felsenfeld, G. (1993) Proc. Natl. Acad. Sci. USA 90, 1676–1680.
[38] Omichinski, J.G., Clore, G.M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S.J. and Gronenborn, A.M. (1993) Science 261, 438–446.
[39] Marshalko, S.J., Schweitzer, B.I. and Beardsley, G.P. (1995) Biochemistry 34, 9235–9248.
[40] Kay, L.E., Keifer, P. and Saarinen, R. (1992) J. Am. Chem. Soc. 114, 10665–10669.
[41] Callihan, D., West, J., Schweitzer, B.I. and Logan, T.M. (1996) J. Magn. Reson. Ser. B 112, 82–86.
[42] Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) J. Biol. NMR 6, 277–293.