AN ABSTRACT OF THE DISSERTATION OF

Monika Ivancic for the degree of Doctor of Philosophy in Biochemistry and Biophysics presented on June 15, 2001. Title: The Asymmetric Binding of Actinomycin D to DNA Hexamers

Redacted for Privacy

Abstract approved: _____

Victor L. Hsu

The solution structure determination of DNA molecules has been an important part of structural biology. NMR solution structures are a complement to structures solved via X-ray crystallography; the two methods are the only ways of obtaining three dimensional coordinates of macromolecules. Because of the nature of the molecule, the solution structure determination of DNA has been a challenging task. Assignments are the first and most important part of NMR structures, and can be simplified for DNA with the use of the rotating frame Overhauser spectroscopy (ROESY) experiment. The ROESY technique can be used for unambiguous assignments of H2′ and H2″ protons and for distinguishing the three main forms of DNA duplexes: A-form, B-form and Z-form.

Many types of DNA have been examined using NMR spectroscopy, including drug-bound DNA complexes. Most previous studies of complexes of the anti-cancer drug Actinomycin D (ActD) and DNA used self-

complementary sequences to identify stabilizing features. The studies presented in this thesis use non-self-complementary DNA hexamers to identify the two orientations in the binding of the asymmetric ActD drug. The largest preference of asymmetric binding was found for the d(CCGCCG)•d(CGGCGG) sequence; however, NMR spectral complications prevented the structure elucidation of this complex. Instead the solution structure was determined for the complex with the next largest orientational preference, ActD:d(CTGCGG)•d(CCGCAG), which has 67% of ActD molecules intercalated with the benzenoid side of ActD in the first strand. The solved structure identifies unusual DNA features, which could be due to the bound drug inducing structural changes to the B-DNA duplex or the presence of conformational motion.

For seven of the eight sequences, the orientation of ActD intercalation within the DNA duplex was identified. The largest preference occurs when the benzenoid intercalation site is followed by a guanine. When this guanine is replaced by an inosine, a reduction in the asymmetric binding of ActD is observed, indicating that the guanine NH_2 group plays a role in the intermolecular contacts. Thus, the two orientations of ActD binding are not present in equal concentrations although their structures are similar, and the preference of orientation is influenced by the asymmetric DNA sequence flanking the intercalation site.

The Asymmetric Binding of Actinomycin D to DNA Hexamers

by

Monika Ivancic

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Doctor of Philosophy

Presented June 15th, 2001
Commencement June 2002

Doctor of Philosophy dissertation of Monika Ivancic presented on June 15, 2001

APPROVED:

Redacted for Privacy

Major Professor, Representing Biochemistry and Biophysics

Redacted for Privacy

Chair of the Department of Biochemistry and Biophysics

Redacted for Privacy

Dean of Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Redacted for Privacy

Monika Ivancic, Author

ACKNOWLEDGEMENTS

The work presented in this thesis would not have been possible without the help, guidance and support of many people. I would like to begin by thanking my mentor, Dr. Victor Hsu, who made this research possible with the acquisition of the Bruker DRX600 instrument and the continuous funding for my work. I would like to thank past and present members of Dr. Hsu's lab, including Dr. Sik Lok Lam and Joshua Hicks for their helpful discussions. A special thank you to Dr. Michael Kennedy and Dr. Kate McAteer at the Environmental Molecular Sciences Laboratories at PNNL in Richland, Washington for their help and guidance. Their love and dedication to science inspired my continued interest in the field.

I would like to extend my gratitude to Dr. P. Shing Ho for his helpful suggestions and insights. Many thanks to Dr. P. Andy Karplus for his curious and realistic view of science, as well as his camaraderie. Thank you to my other committee members, Dr. Michael Schimerlik and Dr. Bruce Coblentz, for their time and guidance in helping me attain my goals. My gratitude goes to all my fellow graduate students who provided support on an emotional, as well as scientific level. In addition, I would like to thank Dr. Lynn Cominsky in the Physics Department at Sonoma State University for being an excellent role model in striving for higher achievements.

This work would not have been possible without the love, friendship and enormous support of my mother, Susan, and my sister, Kati Ivancic. My warmest thanks also to my father, Marko Ivancic, for his love, encouragement and a listening ear. Thank you to my late grandmother, Maria Ivancic, whom by striving for justice and equality in humankind proved to be a great symbol of strength.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

LIST OF APPENDIX TABLES

DEDICATION

I dedicate my thesis to my best friend and husband, George Christian Sliter. I thank him for his enormous patience, helpful suggestions and heartfelt inspiration.

The Asymmetric Binding of Actinomycin D to DNA Hexamers

Chapter I

Introduction

Scientists have been intrigued by the purpose and function of the DNA molecule since before the discovery of its structure by James Watson and Francis Crick in 1953 (Watson & Crick, 1953). It's been known that nucleic acid molecules play a central role in the biological processes of a cell. DNA is the carrier of genetic information and is transcribed into RNA. RNA then communicates the genetic information to ribosomes for protein synthesis. DNA also plays an essential role in cell division, during which it becomes duplicated. RNA and DNA's basic structural features are well established (Arnott *et al.*, 1974, 1980), since it is known that naturally occurring RNA may either be single-stranded or form an A-type helix (Voet & Voet, 1990), while DNA primarily forms a B-type helix (Voet & Voet, 1990), with a few examples of single-stranded, triplex and quadruplex structures. The number of nucleic acid structures solved via X-ray crystallography and NMR spectroscopy is small compared with the number of solved protein structures (Wijmenga & van Buuren, 1998). In NMR spectroscopy, this is due to the problem of extensive resonance overlap in spectra of these compounds (Wijmenga & van Buuren, 1998), the limitations on molecular size of the macromolecule under investigation (Wüthrich, 1986), the inherent nature and geometry of nucleic acids as well as computational problems (Wijmenga *et al.*, 1993). Solution structures are important in DNA and RNA structural biology as a complement

to crystallography, especially in cases where crystal packing forces affect DNA and RNA structures (Wijmenga & van Buuren, 1998). Most solution structures of DNA duplexes are in the B-form conformation (Wüthrich, 1986), while many crystal structures of DNA duplexes are in the A-form (Fuller *et al.*, 1965). Because of limitations in DNA structures derived via X-ray crystallography, the solution structure elucidation of these molecules is of utmost importance. Determining the structure and conformation of DNA molecules in general, and studying the dynamics associated with DNA, provides useful and necessary information for the complete understanding of DNA's function within the cell. The relationship between the structure of a biomolecule and its function is well established (Luisi, 1995; Grosschedl, 1995); knowing the exact three dimensional shape helps identify the active sites, the possible binding sites and places where the molecule interacts with other molecules. In this thesis two aspects of DNA solution structure are covered. The first is the use of the ROESY (rotating frame Overhauser spectroscopy) experiment (Bothner-By *et. al*, 1984) in aiding proton assignments of DNA molecules (Ivancic & Hsu, 2000), assignments being the first essential step in solution structure determination. The second aspect of this thesis deals with two solution structures of a DNA hexamer complexed with Actinomycin D (ActD), the drug adopting two orientations of binding. ActD binds to double stranded DNA to inhibit transcription and has been used in the past as an anti-cancer agent. These studies involve the elucidation of DNA structures in solution, many forms of which have been solved using this technique.

NMR is particularly suited for identification of nucleic acid regions containing higher conformational flexibility (Wijmenga & van Buuren, 1998). Many different types of DNA molecules with unusual architecture have been studied via NMR spectroscopy, including hairpins (Hare & Reid, 1986; Ikuta *et*

al., 1986), triplexes (Sun *et al.*, 1996; Wang & Feigon, 1999) and quadruplexes (Feigon, 1993; Feigon *et al.*, 1995; Kettani *et al.*, 1995; Gilbert & Feigon, 1999) in addition to regular duplexes. The introduction of synthetic methods for preparing well defined DNA sequences enabled the determination of several solution DNA hairpin structures, in addition to other unusual structures (Hare & Reid, 1986; Ikuta *et al.*, 1986). The main folding principles of hairpin loops could thus be determined (Hilbers *et al.*, 1994; Van Dongen *et al.*, 1996). Within the last decade, many structural studies have been undertaken of DNA and RNA helices containing more than two strands. The potential use of nucleic acids as therapeutics (Gee & Miller, 1992) and evidence that alternative structures may have specific functional roles *in vivo* (Frank-Kamenetskii & Mirkin, 1995) have kindled interests in such structures. Both the parallel motif and the antiparallel motif of DNA triplexes have been solved by NMR spectroscopy (Sun *et al.*, 1996; Wang & Feigon, 1999). Most of the detailed structural information available on triplexes comes from NMR studies of intramolecular triplexes (Gilbert & Feigon, 1999). Guanine quadruplex structures have been the subject of great interest during the past several years as well. Interests in such structures have evolved since the discovery that telomeres at the ends of linear chromosomes are comprised of repeats of guanine-rich sequences and such sequences form G-quartets *in vitro* (Williamson *et al.*, 1989). NMR solution structures show that the G-quartets form with different topologies and strand orientations, depending on the sequence and number of strands (Feigon *et al.*, 1995; Kettani *et al.*, 1995). Cytosine-rich sequences were also found to form a tetrameric four-stranded structure, and for the d(TCCCCC) sequence one parallel $C^+ \bullet C$ duplex with hemiprotonated $C^+ \bullet C$ pairs intercalates into a second duplex, which is oriented antiparallel to the first duplex and is called the I-motif (Feigon, 1993). These cystosine-rich sequences found in centromeres are interesting because

they are the complements of the guanine-rich telomere sequences (Gilbert & Feigon, 1999). Numerous NMR-solved structures of triplexes and quadruplexes have emerged in the literature, yet the number of solution structures of regular duplexes far surpasses them. Rather than discuss the numerous NMR studies and examples of specific duplex DNA molecules, a general approach in solution structure determination is described.

The solution structure determination of DNA molecules has been challenging, because of the limited spatial information associated with these molecules. It is well known that the NOESY (nuclear Overhauser effect spectroscopy) experiment is the most important NMR experiment in three dimensional structure determination, correlating protons that are spatially proximate and providing interproton distance constraints. In the NOESY experiment short distances are observed (under 5 Å), thus only local intranucleotide interactions and internucleotide interactions are detected. Since DNAs are effectively linear, only short range sequential constraints can be obtained and the distance ranges need to be determined more accurately than for folded proteins. In structural studies of proteins, it is sufficient to interpret the NOE strengths as weak, medium and strong and classify them into appropriate distance "bins", such as ranges of 1.8 Å to 2.5 Å, 1.8 Å to 3.5 Å and 1.8 Å to 5.0 Å. On the other hand, for nucleic acids it is necessary to derive more accurate interproton distances from the volumes of the NOE crosspeaks. The total relaxation matrix approach to structure elucidation takes into account spin diffusion effects and several of these programs have been developed (Borgias & James, 1990; Zhu & Reid, 1995) which determine reliable distance information from NOE spectra. Another reason for the necessity for accurate distance information is that DNAs are proton poor and the locations of the protons within the molecule are not ideal. An abundance of non-exchangeable protons

on the sugar rings of the backbone exist, however the bases only possess one or two non-exchangeable protons each. Besides the H6/8 base proton, adenines have an additional H2 proton, cytosines have an H5 proton and thymines have the additional methyl protons. Additional exchangeable exocyclic amine protons exist on cytosines, guanines and adenines. Global features such as bending of the helix are difficult to determine directly via NMR spectroscopy, since DNA does not have regions that fold back on themselves, as do RNA and proteins. In DNA structure elucidation, only local features are directly determined, and via the local features a global structure is evaluated. Most NMR structural studies have focused on local DNA features, such as the bases and a number of the base pair structural elements, since they can be determined very accurately.

Another limitation in solution structure determination is that NOE intensities may be affected by conformational fluctuations occurring on the time scale of the NMR experiment (Tonelli & James, 1998). Depending on the type of motion, the NOE intensity must be treated accordingly, since the intensity of such NOE peaks is a weighted average of the intensities given by the atoms in each of the conformers (Tonelli & James, 1998). This internal motion would complicate three dimensional structure determination, since a single distance corresponding to each NOE intensity is derived under the assumption of a single 'rigid' conformation. Many limitations exist in the solution structure determination of DNA molecules, and some are difficult to overcome. However, NMR spectroscopists are constantly developing new strategies and approaches to address these limitations.

The advantage of ^{1}H NMR spectroscopy of DNA molecules is that the protons on the molecule resonate at characteristic chemical shifts, as shown in Figure I.1. When 2D ^{1}H NMR is employed, the different regions are separated

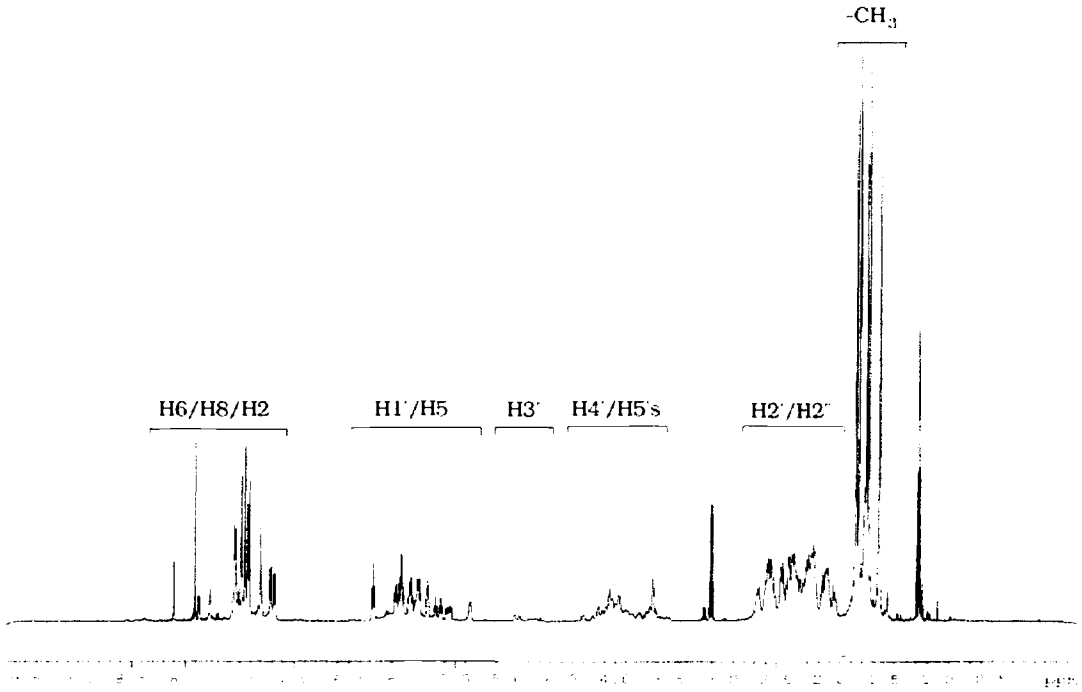

Figure I.1. The ^{1}H 1D NMR spectrum of a DNA hairpin molecule, d(TCGCGTTTTCGCGA), in $^{2}H_2O$. Protons within the DNA molecule resonate at particular regions, an advantage in nucleic acid NMR spectroscopy. The specifics of each region are described in the text.

by a second dimension, with the most structurally informative regions being the DNA "walking regions", including the aromatic base proton to H2', H2" proton region and the aromatic base proton to H1' proton region.

Through-bond experiments, such as the COSY (correlated spectroscopy) and TOCSY (total correlation spectroscopy) experiments, are good starting points for the proton assignments of DNA molecules. These experiments enable the identification of the thymine methyl proton resonance which shows a through-bond interaction with the thymine base (H6) proton resonance, as well as the correlation of the cytosine H5 and H6 base proton resonances. These experiments are also useful in identification of the geminal H2', H2" proton pair resonances, since these two protons interact with only one H1' proton within the same nucleotide. The information gained in the through-bond experiments is helpful in the identification of the thymine and cytosine H6 resonances in the through-space NOESY spectra. Knowing which resonances belong to cytosines and thymines and knowing the DNA sequence under study, sequential assignments are made plausible. When a DNA sequence does not contain many cytosines or thymines, sequential assignments are often more difficult. Typically the adenine base protons resonate the farthest downfield, while the guanine base protons resonate between the adenine base protons and the cytosine and thymine base proton resonances, the latter most often located farthest upfield in that region (Scheek *et al.*, 1984).

The above mentioned protons are found in ^2H_2O solutions, while the exchangeable protons become detectable upon dissolving the sample into H_2O and widening the spectral width of the selected NMR experiment. The guanine and thymine imino protons resonate between 11 ppm and 13 ppm, with the guanine iminos typically downfield from the thymine iminos (Reid, 1986). The cytosine amino protons resonate between 6 ppm and 8.5 ppm and NOEs are

often detected between the guanine imino and cytosine amino protons of a base pair. The guanine amino protons resonate between 6 ppm and 8.5 ppm as well, however only in few examples have NOEs been detected between the guanine imino proton and amino protons (Brown *et al.*, 1994). Imino to imino NOEs are only observed between guanine iminos when the opposite strand has a guanine located in the 3' direction.

As previously mentioned most solution structures of DNA duplexes are in the B-form conformation. The A-form helix is a dehydrated form of DNA and to date there has only been one report of this conformation in solution (Sarma *et al.*, 1986). A distinction between A-form and B-form DNA is the type of sugar pucker associated with each, 3'-*endo* (N-type) in A-form and 2'-*endo* (S-type) in B-form. The sugar pucker has traditionally been determined from $^3J_{HH}$ coupling constants in the sugar ring, relating the experimentally measured coupling constant with the torsion angle via the Karplus equation, $^3J_{HH} = 10.2 \cos^2\theta - 0.80 \cos\theta + 0.0$ (Hosur *et al.*, 1988). For the S-type sugar pucker, the H1' to H2' $^3J_{HH}$ value is about 9 Hz, while for the N-type sugar pucker this value is near 6 Hz (Hosur *et al.*, 1988). However, ribose sugar rings are not rigid but readily interconvert between N- and S-type conformations. The sugar puckering states are described by their pseudorotation angles ($-10° < P_N < 20°$ and $120° < P_S < 180°$) and amplitudes ($\phi_m{}^N$ and $\phi_m{}^S$, between 32 and 40 for both) (Wijmenga & van Buuren, 1998). The relative population can be found via the fraction S conformer, pS. By plotting $^3J_{3'4'}$ against $^3J_{1'2'}$, a straightforward check of an equilibrium between N-type and S-type conformers can be applied (Van de Ven & Hilbers, 1988; Wijmenga *et al.*, 1993).

The type of sugar pucker is not always correlated with the same specific conformation of DNA. Recent studies of $d(C_nG_n)$ DNA fragments show that the base pairs can stack in an A-like fashion while the puckering of the deoxyribose

ring is B-like (Trantirek *et al.*, 2000). On the other hand, the transposed sequence $d(G_nC_n)$ displays A-like guanine stacking and the other half of the helix displays B-like stacking (Stefl *et al.*, 2001). These studies show "intermediate" forms of DNA in the solution state and indicate that the type of stacking associated for each form of DNA does not have to absolutely correlate with the same form of DNA's sugar pucker. As presented in the second chapter of this thesis, the ROESY experiment may be used as a direct probe to distinguish between the three main forms of DNA helices, A-form, B-form and Z-form (Ivancic & Hsu, 2000). This experiment is straightforward in that only a distinction between positive and negative crosspeaks and the assignment of the H2' and H2" protons needs to be made to differentiate among the three conformations. An unusual feature of Z-form DNA is that the H5' proton resonance of cytosine residues shifts dramatically upfield into the H2', H2" proton region, which is another characteristic that has been used to identify that form of DNA (Orbons & Altone, 1986; Orbons *et al.*, 1986).

In the above paragraphs, homonuclear experiments were discussed in the acquisition of structural NMR constraints. Heteronuclear experiments are advantageous as well, particularly when it comes to resonance assignments. An important heteronuclear experiment in DNA solution structure determination is the through-bond proton-phosphorous experiment. This experiment yields correlations due to scalar couplings occurring between a backbone phosphorous nucleus and the H3' proton on the 3' side of the phosphodiester bond and the H4' proton on the 5' side of the phosphodiester bond. Thus, the ^{1}H-^{31}P HSQC (heteronuclear single quantum correlation) experiments are valuable for confirmation of complete sequential connectivities (Sklenar *et al.*, 1986). The combination of ^{31}P chemical shifts and ^{1}H-^{31}P NMR coupling constants, obtained via the ^{1}H-detected 2D J-spectrum (Nikonowicz & Gorenstein, 1990; Sklenar &

Bax, 1987), are necessary for relating these values to the backbone dihedral angles using the Karplus equation. The J_{HP} couplings give information on the ε dihedral angle (C4'-C3'-O3'-P i+1) and the ζ dihedral angle (C3'-O3'- P i+1 - O5' i+1) (Nikonowicz & Gorenstein, 1990). Experimentally, the ^{31}P shift exhibits a linear correlation with changes in ε, and quantum mechanical chemical shift calculations were performed to demonstrate that the change in ^{31}P shift is linearly related to changes in the torsion angle ζ (Wijmenga & van Buuren, 1998). Thus, it was inferred that a change in ε affects a corresponding change in ζ. The regular B-DNA helix can exist in a BI conformation, defined by a *gauche* ζ angle and a *gauche* α angle, or in a minor BII conformation with a *trans* ζ angle and a *gauche* α angle. The BII form is only found when unusual features are discovered within the DNA. For example, the backbone of a B-form duplex where two pairs of G:A mismatches paired via a sheared hydrogen bonding scheme, has the BII conformation at that site (Chou, S.-H., *et. al*, 1992). The BII backbone conformation is most often detected with observed downfield shifted ^{31}P chemical shifts, which was also found for a couple of residues in the major complex of the structure presented in the third chapter of this thesis. The advantage of ^{31}P NMR is that the naturally-occurring isotope is NMR active and specific labeling is not required for its detection.

Other techniques are available for obtaining information on the DNA backbone conformation, such as the ^{1}H-^{13}C HSQC experiment collected with a large number of experiments in the second dimension. The H3'-C3' and the H4'-C4' crosspeaks show splittings which allow for the determination of $^{3}J_{HP}$, $^{4}J_{HP}$ as well as $^{2}J_{CP}$ and $^{3}J_{CP}$ coupling constants (Schmieder *et al.*, 1992). The authors used a DNA octamer with ^{13}C in natural abundance, and demonstrated how this experiment can lead to the complete assignment of the carbon resonances of the sugars and bases. Of particular importance is the subsequent

determination of the carbon-phosphorous and proton-phosphorous coupling constant, determined with high accuracy, providing information on the sugar pucker and the backbone torsion angles (Schmieder *et al.*, 1992).

The development of methods for the production of isotopically labeled nucleic acids in quantities suitable for NMR work has made heteronuclear spectroscopy more feasible. One of the most important advantages is the remarkable increase in editing power achieved in three and four-dimensional ^{13}C-edited NOESY experiments (Nikonowicz & Pardi, 1992). The enzymatic synthesis of uniformly $^{13}C^{15}N$-labeled DNA oligonucleotides in milligram quantities was first published in 1995 (Zimmer & Crothers, 1995). A similar enzymatic method was published more recently, using the Taq DNA polymerase and a very efficient protocol that resulted in quantitative polymerization of the template and a higher efficiency of incorporation of the labeled dNTPs (Masse *et al.*, 1998). Chemical synthesis of labeled DNAs using labeled phosphoramidites has been done in the past (Ono *et al.*, 1994; Tate *et al.*, 1994), but this approach is limited by the cost and technical expertise required.

The availability of labeled DNAs make direct detect heteronuclear experiments possible. ^{31}P-^{13}C correlations in the sugar-phosphate backbone ease the assignments of nucleic acid spectra (Aboul-ela & Varani, 1995). Also, triple resonance, (HCP) three dimensional experiments have been introduced to overcome the spectral overlap encountered as nucleic acids of progressively larger molecular weight are investigated (Heus *et al.*, 1994; Marino *et al.*, 1994). The magnetization in these HCP experiments is transferred between neighboring sugar resonances via the intervening phosphorous resonance, enabling the sequential correlation of neighboring nucleotides. Three dimensional ^{13}C-edited COSY and TOCSY experiments provide complementary information to the HCP experiments, although with considerably reduced

sensitivity (Aboul-ela & Varani, 1995). These experiments aid backbone-driven assignments in DNA and RNA molecules.

Triple-resonance experiments have also been used to correlate base and sugar resonances for assignment purposes. The H1'/C1' sugar resonance can be correlated to the N1/N9 base nitrogen, with subsequent transfer of magnetization to the C6/H6 or C8/H8 resonances in pyrimidines and purines, respectively (Farmer *et al.*, 1993; Sklenar *et al.*, 1993; Tate *et al.*, 1994). Correlations involving base-ring nitrogen and carbon resonances have also been accomplished for the assignment of adenosine H2 protons (Legault *et al.*, 1994). An alternative strategy for connecting base and sugar resonances relies on the indirect detection of two- or three-bond correlations using ^{15}N-labeled DNA. Simple ^{1}H-^{15}N correlation experiments can unambiguously distinguish adenosine H2 and H8 resonances and provide correlations between base nitrogens and sugar H1' resonances (Sklenar *et al.*, 1994). Two- or three-dimensional HCCH-COSY or HCCH-TOCSY experiments are most efficient for the identification of all resonances belonging to each individual sugar, since relatively large ^{13}C-^{1}H (150Hz) and ^{13}C-^{13}C (40-50 Hz) scalar couplings exist (Bax *et al.*, 1990). This methodology is extremely powerful, particularly in its three-dimensional extension, although the sensitivity limits its application to molecular weights below 10-15 kD (Aboul-ela & Varani, 1995). The heteronuclear experiments described above provide a complete data set to assign all ^{1}H, ^{13}C, ^{15}N and ^{31}P resonances in a labeled DNA molecule, without the need for NOE-based methodologies.

In recent years new methodologies have been developed in the structure determination of macromolecules in solution. The advent of residual dipolar ^{1}H-^{15}N and ^{1}H-^{13}C couplings (RDCs) in magnetically oriented macromolecules in solution (Tolman *et al.*, 1995), has been shown to enrich protein structure

determination, as well as DNA structure determination in solution (Tjandra *et al.*, 2000; MacDonald *et al.*, 2001). Unfortunately, the latter application is still somewhat impractical because of the limitations in production of selectively labeled samples. RDCs are based on the fact that the anisotropy of the molecular magnetic susceptibility gives rise to a small degree of alignment (Tjandra *et al.*, 1997). Small residual one-bond ^{15}N-^{1}H and ^{13}C-^{1}H dipolar couplings define the orientation of the N-H and C-H bond vectors relative to the molecule's magnetic susceptibility tensor. Thus long-range order information that is not accessible by any of the solution NMR parameters currently used in structure elucidation can be determined directly (Tjandra *et al.*, 1997). The first study utilizing RDCs in DNA molecules employed the Dickerson dodecamer, $d(CGCGAATTCGCG)_2$, in an aqueous liquid crystalline medium containing 5% w/v bicelles (Tjandra *et al.*, 2000). A more interesting study emerged utilizing RDCs in a DNA dodecamer containing an A-tract , with an overall helix axis bend of 19° (MacDonald *et al.*, 2001). In this study, structural features that are different from existing models for A-tract bends were identified (MacDonald *et al.*, 2001). The results from the DNA structures incorporating RDCs are promising, in that the overall helical bend in a DNA duplex can be determined directly via NMR spectroscopy.

New techniques for hydrogen bond detection have also been developed. Residue- and atom- specific ^{15}N labeled DNA oligomers were used to measure ^{15}N-^{15}N scalar coupling constants across base pair hydrogen bonds ($^{2h}J_{NN}$) in the ^{15}N 1D spectrum (Kojima *et al.*, 2000). The temperature and sequence dependence of these coupling constants were examined for the oligomers under study and a decrease of $^{2h}J_{NN}$ values was observed by elevating the temperature, most likely due to the physical decrease in hydrogen bond strength by the increase in the NN distance (Kojima *et al.*, 2000). Additional

experiments using the internucleotide $^{3h}J_{NC}$ coupling in addition to the $^{2h}J_{NN}$ scalar coupling were utilized in G-quartet structure determination. These experiments prove to be invaluable in structure determination of multistranded DNA oligomers and complete characterization of DNA base pairing.

Several methods for obtaining assignments and structure information of DNA molecules have been discussed. The limitations of DNA solution structure determination have also been mentioned, since the nature of the molecule doesn't allow for the accumulation of numerous NMR restraints. Computational problems and limitations also exist in DNA structure determination. Distance geometry algorithms, later renamed to the embed algorithm, have been employed with little success. Due to its sampling properties, this algorithm leads to extended ladder-like DNA structures (Havel, 1990). More favorable sampling properties are exploited in the molecular dynamics (MD) techniques. Restrained molecular dynamics (rMD) has been widely employed in DNA structure determination, with the use of the total relaxation matrix approach in accurate distance determination prominent in the last decade. However, the use of rMD in nucleic acid structure determination has raised the question whether and to what extent the resulting structures are determined by the NMR constraints or by the MD forcefield. It has been shown that the MD forcefield does have an effect on the final DNA structure in unrestrained MD calculations (Cheatham & Kollman, 1996). Still, solution structure determination of DNA is in general more challenging than that of RNA or proteins and a need for the development of better experimental and structure calculation techniques exists. These difficulties exist in studying any form of DNA, including DNA molecules in complex with proteins and ligands. Many solution structures have been solved of DNA in complex with drugs, an advantage being the presence of more NOEs defining the DNA conformation.

Since part of this thesis deals with the binding of an intercalator with peptide residues in the DNA minor groove, Actinomycin D (ActD), it is worthwhile to discuss several other minor groove binders and their effects on DNA structure and function. The most similar drugs would be the bis-intercalators with peptide groups that bind in the minor groove, which include triostin A, echinomycin, TANDEM, and CysMeTANDEM. Triostin A and its analogue echinomycin bind specifically to d(CG) steps (Low *et al.*, 1984; 1986) while CysMeTANDEM and TANDEM recognize d(TA) steps (Addess *et al.*, 1992; 1993) . All of these ligands consist of two quinoxaline chromophores linked by a bicyclic octa-depsipeptide ring and the complexes of the ligands with DNA are very similar in structure (Geierstanger & Wemmer, 1995). Specific hydrogen bonds between the ligands and the DNA bases account for the sequence specificity of each (Geierstanger & Wemmer, 1995).

Other minor groove-binding ligands include chromomycin and mithramycin and, even though both have a chromophore, they are nonintercalating and cause major distortions of the DNA structure, including opening of the minor groove (Gao *et al.*, 1992; Sastry & Patel, 1993). Chromomycin binds to CG-rich sites of at least three base pairs in length, with 5'-GGG-3' and 5'-CGA-3' being the strongest binding sites (Fox & Howarth, 1985; Van Dyke & Dervan, 1983). The two ligands have comparable but not identical sequence specificity and form similar complexes with DNA (Gao *et al.*, 1992; Gao & Patel, 1989; Gao & Patel, 1990; Sastry & Patel, 1993). The complexes retain two-fold symmetry and NOE and coupling data are consistent with a wide and shallow minor groove resulting from complex formation (Sastry & Patel, 1993). Numerous other drugs exist which bind in the minor groove, and the design of sequence-specific minor groove ligands has also been achieved (Dervan, 1986).

In the chapters of this thesis, the solution structure determination of DNA molecules is described and presented in the following manner. In Chapter II, the ROESY experiment is utilized in the assignment strategy of DNA molecules. The ROESY experiment is similar to the NOESY experiment in that it is a through-space detection technique. When applied to DNA molecules, the ROESY spectrum possesses both positive and negative crosspeaks, helpful in identification of particular protons within the DNA oligomer. Specifically, the ROESY experiment enables the unambiguous assignment of H2' and H2" protons, since their correlation with the H1' proton produces crosspeaks of opposite sign. The ROESY experiment may also be used as an editing tool, by making certain crosspeaks disappear from the spectrum at particular mixing times, simplifying spectra of longer oligomers. Another benefit of this experiment is the ability to distinguish different forms of DNA. The specifics of using the ROESY as a direct probe for determination of different forms of DNA is described in Chapter II, in which the technique is demonstrated using a Z-form hexamer.

The third chapter of this thesis deals with two main studies: titrations of ActD with non-self-complementary sequences and the solution structure determination of ActD bound to a hexamer in two orientations. Non-self-complementary hexamers were used in identifying the unequally proportioned binding orientations of Actinomycin D. Two sequences are identified with the largest preference for one orientation above the other, the first being the d(CCGCCG)•d(CGGCGG) sequence, showing a 79% preference for the benzenoid side of the ActD chromophore intercalated in the second strand. Severe linebroadening and spectral overlap prevented the complete characterization and solution structure determination of this complex. The next highest disproportionate binding was found with the d(CTGCGG)•d(CCGCAG)

sequence and a 67% preference for the benzenoid side of the chromophore intercalated in the first strand was identified. This sequence was also used as a template in the incorporation of modified bases, such as a 5-methylated cytosine and an inosine flanking the intercalation site d(GC), to study the effects of simple chemical and structural changes on the orientations of ActD binding. The structure of the ActD:d(CTGCGG)•d(CCGCAG) complex was solved using the total relaxation matrix approach and presented as a family of 10 structures with an rms difference of 0.56 Å for the major orientation of binding and an rms difference of 0.89 Å for the minor orientation of binding. The identification of the orientations of ActD binding for the different hexamer sequences shows the biggest trend for orientational preference is that the benzenoid intercalation site d(GC) is followed by a guanine. Replacement of this guanine by an inosine significantly reduces the preference of the benzenoid side for that strand. Aside from the orientation of the ActD chromophore within the DNA duplex, the two overall complexes exhibited very similar structures.

Chapter II

The ROESY experiment as an aid to determining solution structures of DNA oligomers

Monika Ivancic and Victor L. Hsu

Published in *Biopolymers*
Volume 54, 35-43 (2000)

II.1 Synopsis

Important intrinsic characteristics of the ROESY experiment were found to be advantageous in DNA solution structure determination. In a ROESY experiment, the different mechanisms of relaxation result in different signs of crosspeaks, enabling a clear distinction between H2' resonances and H2" resonances of the DNA sugar backbone. This method is of particular importance in crowded spectra, for purine resonances whose H2', H2" protons typically resonate closely, as well as in conditions where linebroadening makes coupling constants in a COSY experiment impossible to determine. By observing the signs of crosspeaks in the base proton to H2', H2" sugar proton region, the ROESY spectrum can be used to distinguish A-form, B-form and Z-form DNA.

II.2 Introduction

When determining solution structures of biomolecules via NMR spectroscopy, assignments of proton resonances are vital and necessary for structure elucidation. For peptides and proteins, one identifies backbone protons as well as side chain protons. For DNA oligomers, the protons on the sugar moieties are most abundant, while the bases contain very few non-exchangeable protons. Because of such low density of protons on the bases there is a lack of spatial information, making it even more important to assign the proton resonances accurately. The H2' and H2" protons are often difficult to assign as spectral overlap is sometimes observed for these protons, particularly for purine residues. Spectral overlap in all regions also exists in

spectra of longer DNA molecules. We show how the ROESY (rotating frame NOESY) experiment can be used to assign the H2' and H2" protons unambiguously. In addition we also show how the ROESY experiment can be used as an editing tool, by making certain crosspeaks "disappear" from the spectrum.

There are two through-space detection techniques, the NOESY experiment and the ROESY experiment, which experimentally differ by their implementation of the mixing period. In the NOESY experiment, the mixing period begins with a 90° pulse after which the relevant spins evolve and cross-relax, followed by a 90° read pulse. Cross-relaxation occurs parallel to the static magnetic field B_0, and is dependent on longitudinal relaxation. On the other hand, the mixing period in a ROESY experiment involves a strong on-resonance *rf* field (spin-lock) applied orthogonal to B_0. In this case, cross-relaxation takes place perpendicular to B_0 and is dependent on transverse relaxation (Jones, 1966; Solomon, 1955). The most useful advantage of the ROESY experiment is that the cross-relaxation rate in the rotating frame, σ^r, is always positive and monotonically increasing, unlike the laboratory frame (NOESY) cross-relaxation rate, σ^n, which is negative for molecules with $\omega_o\tau_c < 1$, then passes through zero for rigid body isotropic motion when $\omega_o\tau_c = (5)^{1/2}/2$ and becomes positive for molecules with $\omega_o\tau_c > 1$. Thus for molecules with $\omega_o\tau_c \sim (5)^{1/2}/2$ the ROESY experiment is necessary for determining structural conformation (Bothner-By *et al.*, 1984). For slow molecular motion, $\omega_o\tau_c >> 1$, as is the case for biomolecules, laboratory frame cross-relaxation is an energy conserving process which leads to spin diffusion (Macura & Ernst, 1980). Spin diffusion is a phenomenon in which magnetization is relayed between two protons via another proton. This phenomenon is largely undesirable since it makes the NOE between the two

distant protons stronger, making them appear closer than they are in reality. However, in rotating frame cross-relaxation spin diffusion is more strongly attenuated than in laboratory frame cross-relaxation (Farmer II *et al.*, 1987), and thus the ROESY experiment can lead to more accurate measurements of cross-relaxation rates and better structure determination for biomolecules. Even so, NOESY experiments are more commonly used for a number of reasons including the fact that cross-relaxation to a group of chemically equivalent spins is more efficient in a NOESY than a ROESY experiment, yielding higher sensitivity and a higher signal to noise ratio (Farmer II *et al.*, 1988) and because ROESY experiments are often difficult to implement on older spectrometers.

Although for large macromolecules, the ROESY and NOESY spectra are qualitatively similar a prominent difference is that while both types of spectra can contain positive and negative crosspeaks, the sign of the crosspeak in a ROESY spectrum depends on the type of transfer that occurs between two protons. Direct through-space cross-relaxation results in a negative crosspeak relative to the diagonal and spin-diffusion results in a positive crosspeak. The general rule is that the sign of a crosspeak in a ROESY spectrum is $(-1)^m$, where m is the number of transfers in the cross-relaxation pathway (Bax *et al.*, 1986; Farmer II *et al.*, 1987). In addition to cross-relaxation there are other types of transfers that can occur, such as COSY-type and homonuclear Hartmann-Hahn transfer. COSY-type crosspeaks are easily recognized by their antiphase character, while homonuclear Hartmann-Hahn crosspeaks have the same sign as the diagonal. Thus any negative crosspeak in a ROESY spectrum is the result of direct cross-relaxation (three step transfers are rarely observed), while a positive crosspeak can arise from homonuclear Hartmann-Hahn transfer, spin diffusion, or both (Brown & Farmer II, 1989). Also, it

should be noted, that although the intensity of a particular crosspeak may have contributions from more than one transfer mechanism, the contribution from direct Hartmann Hahn effects is very small (Farmer II *et al.*, 1987).

II.3 Materials and Methods

Although it may seem counter-intuitive, for the purpose of unambiguous assignments of the H2' and H2" protons, the ROESY experiment should be optimized to maximize the number of positive crosspeaks for reasons that are described below. For a through-space interaction, the sign of the crosspeak in a ROESY spectrum varies as a function of mixing time and spatial geometry of the participating protons. Initially at low mixing times, a crosspeak shows negative intensity due to direct cross-relaxation between two spatially proximal protons. At longer mixing times, a relay pathway can begin to dominate in which magnetization between two protons gets relayed through a third spatially proximal proton, yielding a positive crosspeak indicative of a two-step transfer (three-spin effect) (Brown & Farmer II, 1989). For our purposes, the ideal mixing time for which the three spin effect is maximal and most DNA crosspeaks appear was found at 350 msec, thus this mixing time was used with various spin-lock power levels (2 kHz - 6 kHz). The lower field strengths are believed to minimize the homonuclear Hartmann-Hahn effect (Bax *et al.*, 1986), however, the intensities of the observed ROEs are also reduced. In our experiments most of the power levels yielded comparable positive and negative crosspeak intensities, though at a power level of 2.6 kHz all

crosspeak intensities are reduced and at 2 kHz many of the positive crosspeaks disappear. With a mixing time of 350 msec, maximum spin diffusion was observed with a spin-lock power level of 4.7 kHz. The effect of varying spin-lock duration is clearly observed in Figure 1, as we go from a short mixing time of 100 msec to a longer mixing time of 400 msec, more positive crosspeaks appear in the spectrum. Our experiments were performed on a 3.3 mM B-form DNA hairpin d(TCGCGTTTTCGCGA), using a DRX600 Bruker NMR spectrometer. ROESY spectra were recorded using various mixing times (50 msec - 500 msec), with a constant rf spin-lock power level of 4.7 kHz.

II.4 Results and Discussion

II.4.1 DNA ROESY spectral characteristics

Two interesting and important regions in a DNA spectrum are the H2', H2" to H1' connectivities, and the H2', H2" to H6/H8 connectivities, which when combined provide information on spatial geometry and helical conformation. The H2' and H2" protons in a DNA sugar have direct, through-bond interactions with the H1' proton. The H2' resonance is usually distinguished from the H2" resonance by the intensity of its crosspeak with the H1' proton in the NOESY spectrum. Since in a DNA molecule, the H2" is located closer to the H1' than the H2' is to the H1', the former crosspeak will have a greater intensity than the latter crosspeak. However, differences in these interproton distances are often small (Figure 2), and sometimes these resonances overlap, especially in purine residues (Scheek, 1984), making it difficult to distinguish the H2' resonance from the H2" resonance. This is especially difficult if the H1' protons of two different residues resonate at

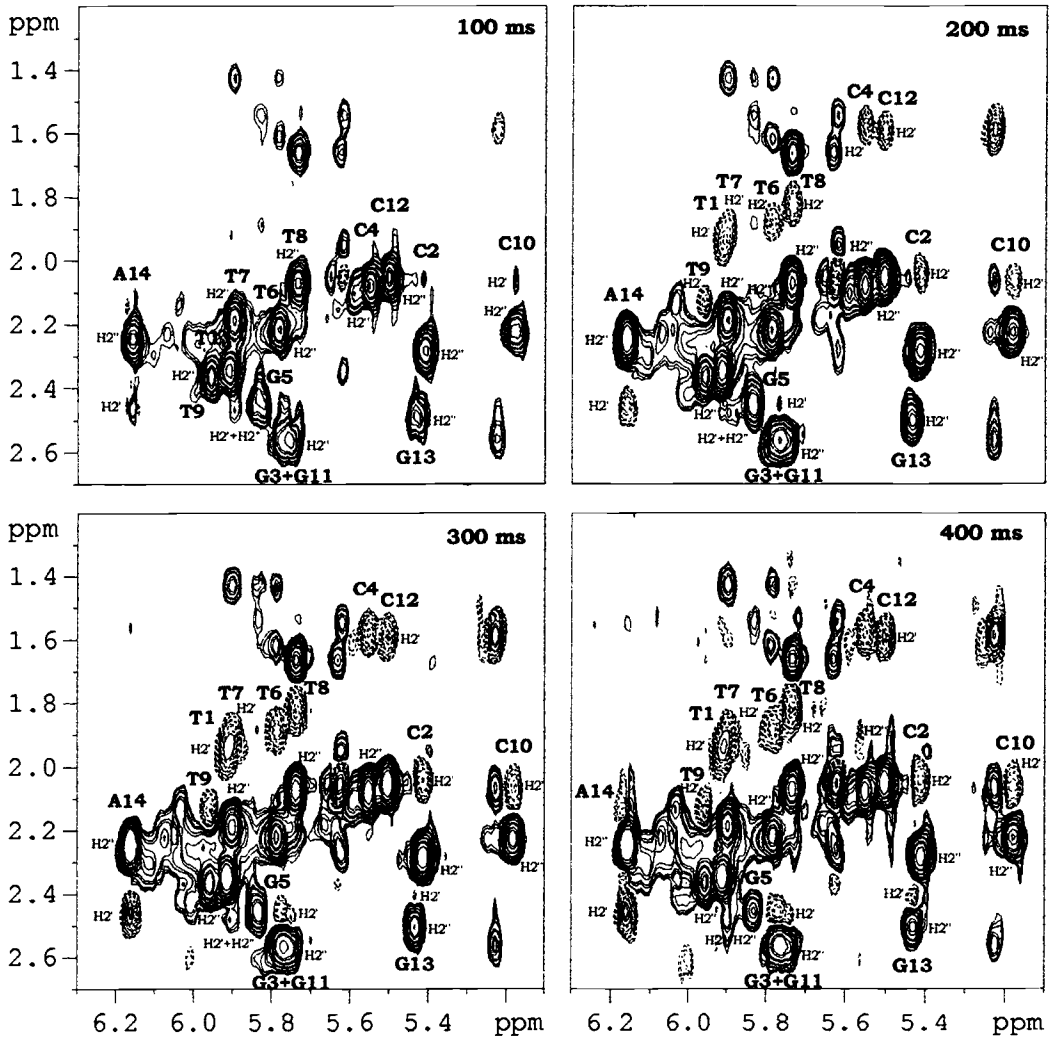

Figure II.1 The H1' to H2", H2' crosspeak region in ROESY spectra of a 3.3 mM hairpin molecule recorded at various mixing times using a spin-lock power level of 4.7 kHz. The contours of the negative crosspeaks are represented as solid lines while the contours of the positive crosspeaks are represented as dashed lines. The sign of the H1' to H2' crosspeak changes depending on the mixing time as described in the text.

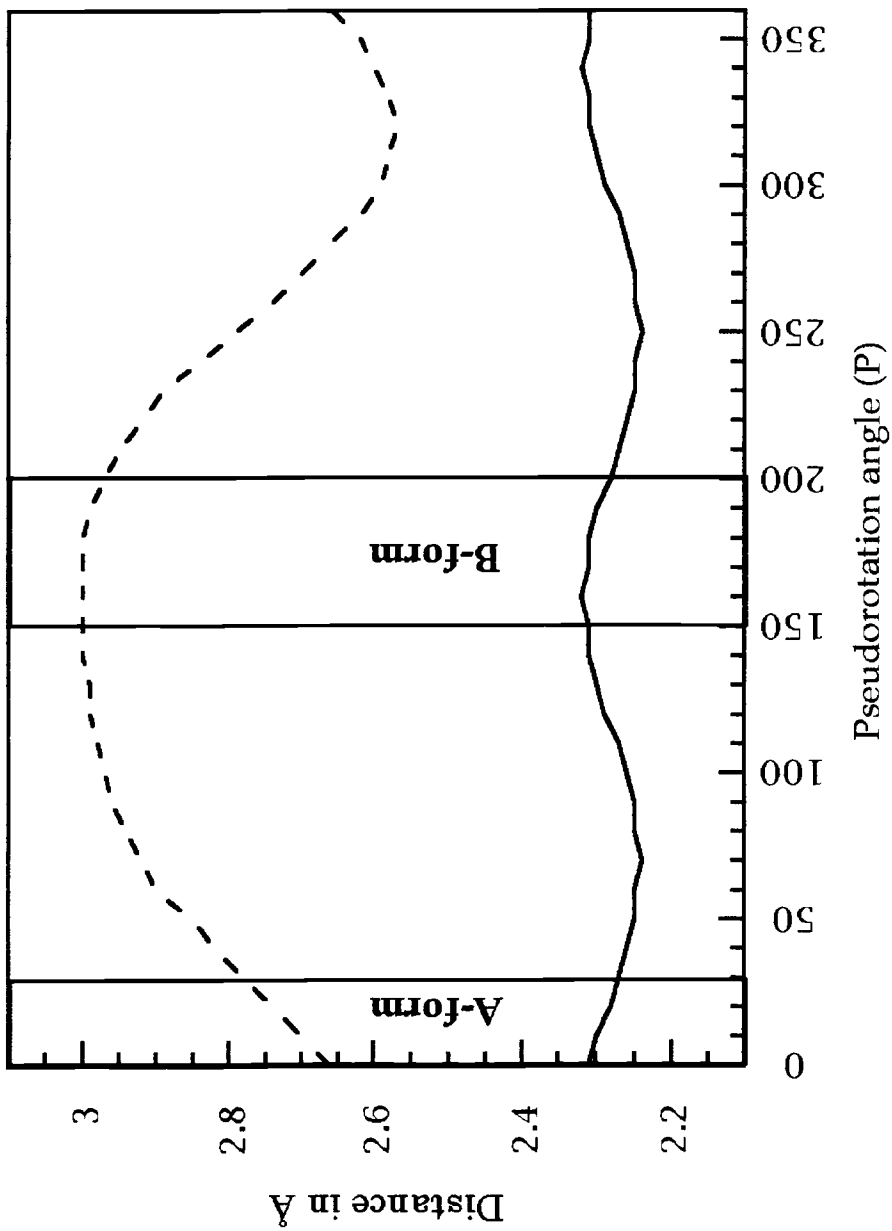

Figure II.2 Interproton distances between H1' and H2', H2" sugar protons as a function of the pseudorotation angle. The H1' to H2" distance (———) is shorter than the H1' to H2' distance (------) at all pseudorotation angle values. For the typical A-DNA sugar pucker (C3'-endo), the H1' to H2" and H1' to H2' distances are more equidistant than for the B-DNA sugar pucker (C2'-endo).

similar frequencies and the H2' proton resonance of one residue is close to the H2" proton resonance of the other residue. This problem can be overcome in the ROESY spectrum. Even though the H1' and H2" protons are scalar-coupled, the intensity of this crosspeak in the ROESY spectrum is negative for all mixing times and spin-lock power levels, (Figure 1), a result of direct cross-relaxation. The same is also observed for the H2' to H2" proton crosspeak, indicating that direct cross-relaxation dominates over scalar-coupling. On the other hand, the sign of the crosspeak between the H1' and H2' protons depends on the chosen mixing time. For a short mixing time of 50 msec this interaction has a negative crosspeak indicative of direct cross-relaxation. At 100 msec, for most residues this crosspeak disappears or its intensity is very low, signifying that both direct cross-relaxation and a two step transfer are occurring, leading to a cancellation of signal intensity. At longer mixing times, t_m > 150 msec, the H1' to H2' crosspeak is positive. At these longer mixing times, this crosspeak has contributions from the three-spin effect with the magnetization being relayed through the H2" proton. For certain residues, this sign alteration occurs at a higher mixing time - 300 msec for residues G3 and G11, and 400 msec for G13. These observations suggest that the sugar puckers of these residues may periodically reside in the C3'-endo conformation for which the H1' to H2' and the H1' to H2" distances are more equivalent than in the C2'-endo conformation (Figure 2). These residues may also exhibit a higher frequency and amplitude of motion and are possibly in rapid exchange between N and S puckers in their sugar rings, necessitating a longer mixing time for the H1' to H2' crosspeak to appear positive in sign. Thus by choosing proper experimental parameters for the ROESY experiment, the sign alteration for crosspeaks in this region makes assignment of the H2' and H2" proton resonances unambiguous. As

mentioned earlier, this is of particular importance for purine residues since their H2' and H2" protons resonate closely and also for longer DNA oligomers exhibiting increased spectral overlap.

The second interesting region in the DNA ROESY spectrum is the H2', H2" to H6/H8 connectivities region, which possesses mostly negative crosspeaks. Each residue exhibits intranucleotide interactions between the H2', H2" sugar protons and its own H6 or H8 base proton, resulting in two crosspeaks for each residue in this region. In a typical NOESY spectrum connectivities from the H2', H2" protons of one residue to the H6/H8 proton of the next residue (on the 3' side) are also observed, resulting in four crosspeaks for each base proton resonance. However, in this region, the effect of spin diffusion on the sign of the crosspeaks in the ROESY spectra can also be advantageous. The intranucleotide crosspeak from the H6/H8 base proton to the H2' proton is negative, a result of direct cross-relaxation during the ROESY spin-lock period. Yet the crosspeak between the H6/H8 proton and the H2" proton appears positive, suggesting the cross-relaxation is relayed through another proton, most likely the H2' proton. For the internucleotide base proton to sugar proton crosspeaks, the opposite trend is observed. The crosspeak between the H2" proton of one residue and the H6/H8 proton of the 3' residue is negative, indicative of direct cross-relaxation, while the crosspeak between H2' proton of one residue and the H6/H8 proton of the 3' residue is positive, a result of relayed cross-relaxation. This is shown in Figure 3 in which the internucleotide crosspeaks between G11 and C12 are indicated with arrows. Since the distance between the H2" of the first residue and the H6/H8 of the following residue is shorter than the distance between the H2' of the first residue and the H6/H8 of the following residue (Table 1), more of the former, negative, crosspeaks appear. By tabulating the distances

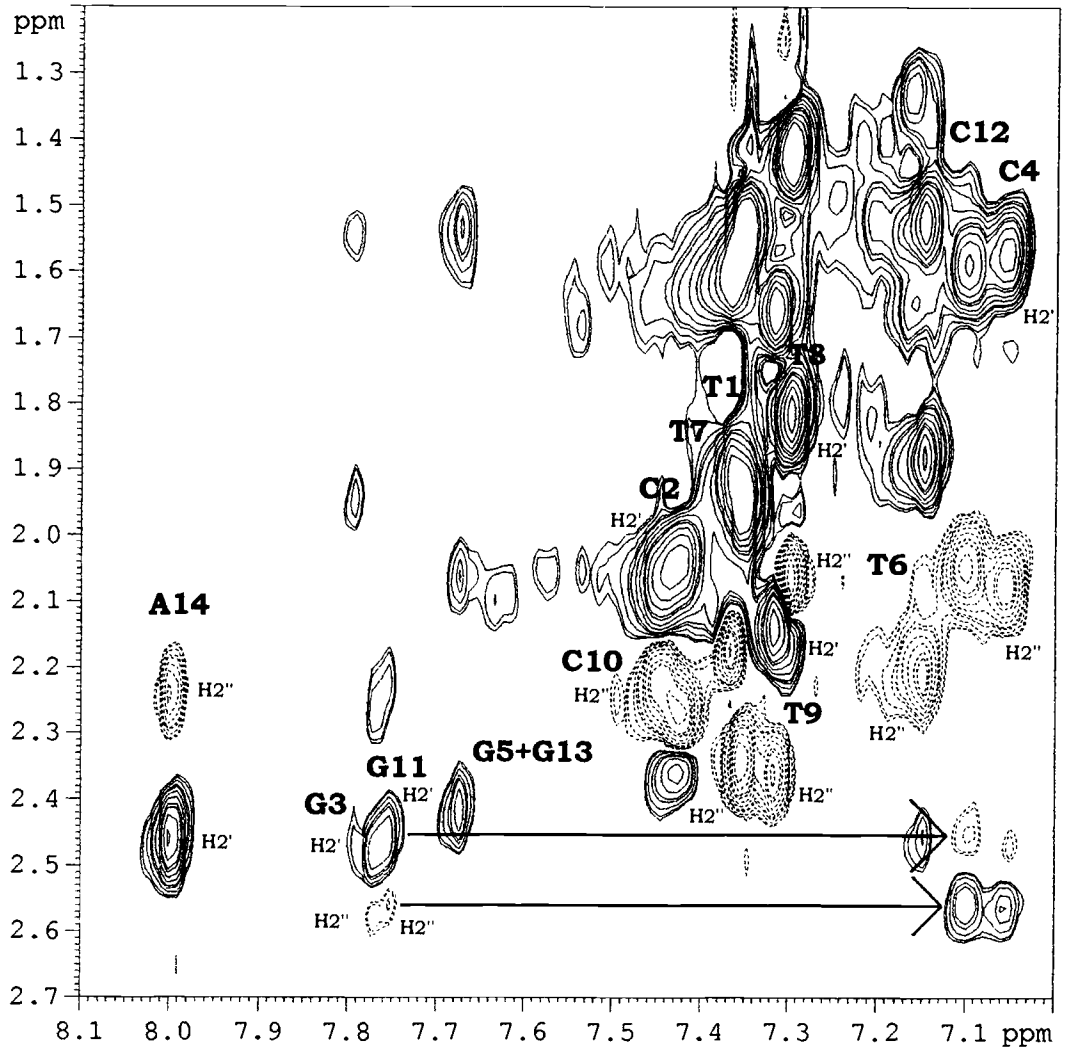

Figure II.3 The H6/H8 to H2', H2" crosspeak region in the ROESY spectrum of a 3.3 mM hairpin molecule recorded with a spin-lock power level of 4.7 kHz and a 350 msec mixing time. The pairs of labeled crosspeaks are the intranucleotide interactions. The contours of the negative crosspeaks are represented as solid lines while the contours of the positive crosspeaks are represented as dashed lines. Very few internucleotide interactions appear; crosspeaks between H2', H2" of G3 and H6 of C4 as well as crosspeaks between H2', H2" of G11 and H6 of C12 are visible with the characteristic crosspeak sign intensities as described in the text.

Table II.1 Interproton distances (Å) for A-, B-, and Z-DNA. The internucleotide distances are between a nucleotide and its 3' neighbor (Westernik *et al.*, 1984).

	form of DNA		
Intranucleotide interactions	A-form	B-form	Z-form
H6/H8 to H1'	3.8	3.8	3.7 (Pyr) 2.6 (Pur)
H6/H8 to H2'	3.8	2.0	3.1 (Pyr) 4.1 (Pur)
H6/H8 to H2"	4.6	3.5	4.3 (Pyr) 4.7 (Pur)
Internucleotide interactions	A-form	B-form	Z-form
H6/H8 to H6/H8	4.7	5.0	6.1 (Pur-Pyr) 5.3 (Pyr-Pur)
H1' to H6/H8	4.0	2.9	3.7 (Pur-Pyr) 6.5 (Pyr-Pur)
H2' to H6/H8	2.0	3.9	3.1 (Pur-Pyr) 7.1 (Pyr-Pur)
H2" to H6/H8	3.0	2.5	4.3 (Pur-Pyr) > 7.5 (Pyr-Pur)

between these protons (Table 1) this crosspeak pattern can easily be determined.

By varying the spin-lock power level and mixing time the ROESY experiment may be used as a tool for editing complex DNA spectra. Since, as described above, some of the crosspeaks between the H1' and the H2' proton can be made to "disappear", by balancing contributions from direct cross-relaxation and spin-diffusion. As shown in Figure 1, at a mixing time of 100 msec, most of the above mentioned crosspeaks are absent from the spectrum, with few weak ones remaining, such as the C10 and A14 H1' to H2' crosspeaks. The number of crosspeaks in this region of the spectrum can be reduced by up to a factor of two, making the spectrum easier to interpret. Likewise, because spin diffusion can be attenuated by the proper choice of spin-lock power levels and mixing times, about
half of the internucleotide crosspeaks in the H2', H2" to H6/H8 region can be rendered "invisible".

II.4.2 Distinguishing Z-form from B-form DNA

The crosspeaks that we have observed in the ROESY spectra of DNA oligomers exhibit a typical absorption pattern, as described above, indicative of B-form conformations. When implemented this way, the ROESY experiment can be used in distinguishing Z-form DNA from B-form DNA since, in Z-form DNA, the internucleotide distances between the H2' proton and the H6/H8 proton of the following residue is shorter than the distance between the H2" proton and the H6/H8 proton of the following residue, the opposite of B-form DNA (Table 1). Because of these geometrical differences in Z-form DNA some unusual spectral characteristics arise. For Z-form DNA, the pyrimidine H4', H5' and H5" resonances display a large upfield shift,

even as much as 1.7 ppm upfield from their normal position in the spectrum of the B-DNA duplex (Orbons, 1986). The upfield shifts agree with the known crystal structure of Z-DNA (Fujii, 1982), since the pyrimidine H4', H5" protons and especially the H5' protons are positioned inside the shielding cone of the 3'-neighboring guanine base (Orbons, 1986). From previous shielding calculations on the Z-forms of d(CG^{m5}CG) and d(CG^{m5}CGCG) (Giessner-Prettre, 1984), it is predicted that the H5' resonance of pyrimidines will display maximum upfield shift. Another typical Z-form DNA characteristic is that the distance from H1' of a pyrimidine to H8 of a purine exceeds 0.6 nm, thus these crosspeaks are absent from the spectrum, as are the crosspeaks between the H2', H2" protons of a pyrimidine and the H8 proton of the following purine. (Table 1)

The ROESY experiment was implemented on a hexamer of the sequence d(^{m5}CG^{m5}CG^{m5}CG) with fully methylated cytosines at the C5 position, previously found to form Z-DNA (Orbons, 1986). ROESY spectra of the 2.9 mM hexamer were collected under conditions of 70% ^2H_2O/30% CH_3OH-d4 and 5mM $MgCl_2$ as previously reported (Orbons, 1986). In these conditions both the B-form and Z-form oligos are present, but the Z-form resonances are easily distinguished from the B-form resonances. The ROESY experiment was recorded using a 350 msec mixing time and a spin-lock power level of 4.7 kHz. In the H1' to H2', H2" region of this Z-DNA hexamer positive and negative crosspeaks appear in the same pattern as for B-DNA, Figure 4. The crosspeak between the H1' proton and the H2" proton exhibits negative intensity, while the crosspeak between the H1' proton and the H2' proton exhibits positive intensity. From this figure it is evident that the G2H2' proton and the ^{m5}C3H2" proton resonate at the same frequency, as do the G4H2' proton and the ^{m5}C5H2" proton. The intranucleotide crosspeaks

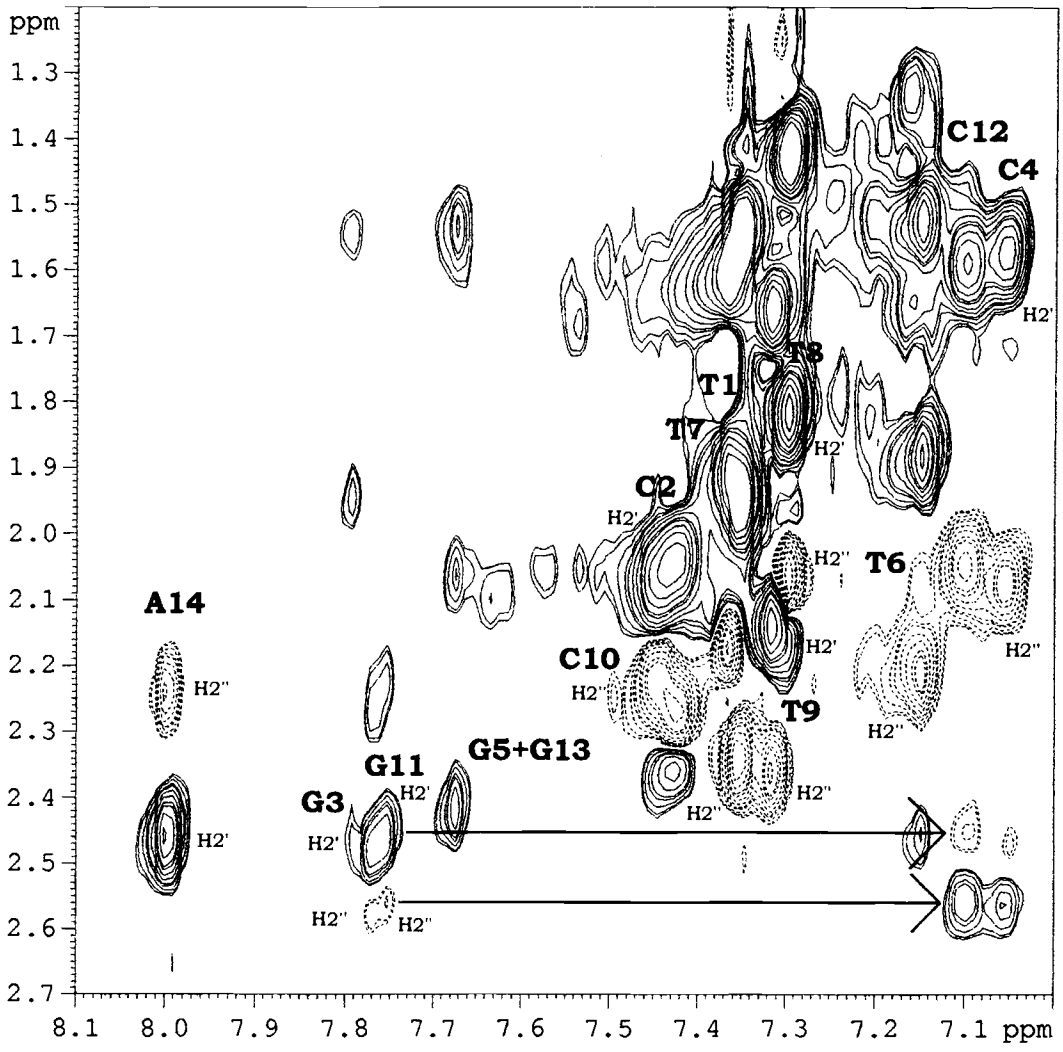

Figure II.4 The H1' to H2', H2" crosspeak region in the ROESY spectrum of a 2.9 mM Z-form DNA hexamer recorded with a spin-lock power level of 4.7 kHz and a 350 msec mixing time. Both B-form and Z-form hexamers are present under the conditions used, and the crosspeaks belonging to B-form are labeled "B". The Z-form crosspeaks are labeled with the nucleotide name and number and the proton interactiong with the H1' proton. The contours of the negative crosspeaks are represented as solid lines while the contours of the positive crosspeaks are represented as dashed lines.

in the H2', H2" to H6/H8 region appear as for B-form DNA as well, with the H6/H8 to H2' crosspeak negative and the H6/H8 to H2" crosspeak positive, while the internucleotide pattern is opposite that observed for B-form DNA. Unfortunately, the internucleotide distances from a pyrimidine sugar to a purine base in Z-form DNA are too large to be observed, as previously discussed (Table 1). However, the internucleotide distance from a purine sugar to a pyrimidine base are observable and the crosspeak between G2H2" and ^{m5}C3H6 exhibits positive intensity, while the crosspeak between G2H2' and ^{m5}C3H6 is negative, as expected (Figure 5). However, the G2H2' proton resonates at the same frequency as the ^{m5}C3H2" proton, thus two interactions produce this negative crosspeak: the interaction between ^{m5}C3H6 and ^{m5}C3H2", expected to result in a positive crosspeak, and the interaction between G2H2' and ^{m5}C3H6. The latter interaction has a shorter distance within the molecule and thus has a stronger dipolar coupling. Since the crosspeak is negative, this interaction dominates over the former interaction, as shown in Figure 5. The crosspeaks between G4 and ^{m5}C5 protons appear in the same way, with spectral overlap of the G4H2' and the ^{m5}C5H2" protons. By observing these and the patterns of the non-overlapping crosspeaks in the H6/H8 to H2', H2" region, we can easily and directly distinguish Z-form from B-form DNA.

II.4.3 Distinguishing A-form from B-form DNA

For A-form DNA it would be expected that the intranucleotide crosspeaks would exhibit the same intensity pattern as B-form DNA, because the H1' to H2" distance is always shorter than the H1' to H2' distance for all values of the pseudorotation angle (Figure 2). However, since A-DNA adopts a C3'-endo sugar pucker, as opposed to the C2'-endo sugar pucker in B-DNA,

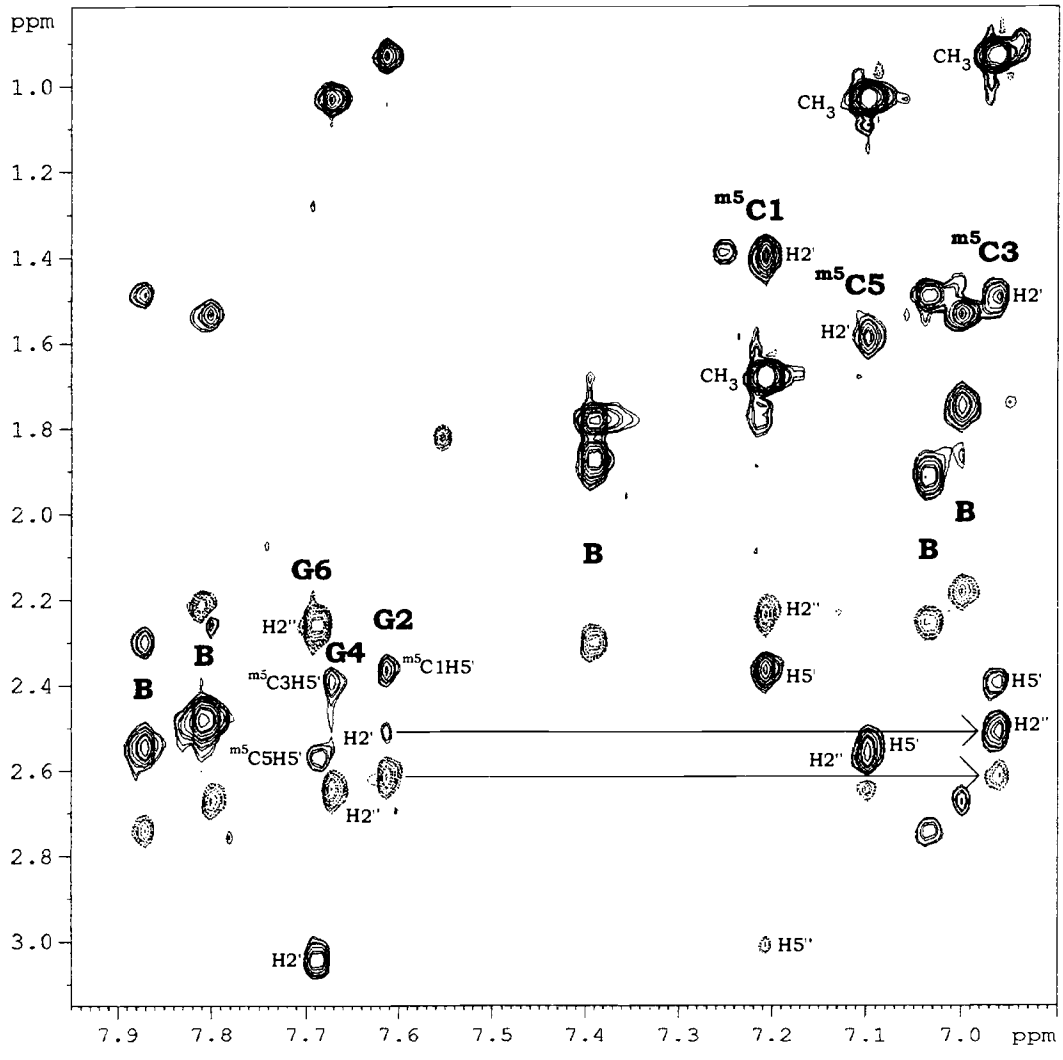

Figure II.5 The H6/H8 to H2', H2" crosspeak region in the ROESY spectrum of a 2.9 mM Z-form DNA hexamer recorded with a spin-lock power level of 4.7 kHz and a 350 msec mixing time. Both B-form and Z-form hexamers are present under the conditions used, and the crosspeaks belonging to B-form are labeled "B". The Z-form crosspeaks are labeled with the nucleotide name and number and the proton interacting with the base proton. The contours of the negative crosspeaks are represented as solid lines while the contours of the positive crosspeaks are represented as dashed lines. The arrows point out the internucleotide interactions in Z-form DNA, see text.

the internucleotide crosspeaks are expected to have the opposite intensity pattern. For A-form DNA, the H2' proton of the 5' residue is closer to the base proton of the following residue than is the H2" proton (Table 1). Thus a negative crosspeak is expected between the H2' proton and the H6/H8 proton of the following residue, a direct cross-relaxation crosspeak, and a positive crosspeak is expected between the H2" proton and the H6/H8 proton of the following residue, a relayed cross-relaxation crosspeak. This variation in the internucleotide distances as a function of sugar pucker is shown in Figure 6. In this manner the ROESY experiment could be used to distinguish between A-form and B-form conformations of DNA. Similarly, A-form DNA can be distinguished from Z-form DNA, even though the crosspeak pattern in the H2', H2" proton to H6/H8 proton region exhibits the same positive and negative intensities. In A-form DNA the internucleotide distances are shorter than the intranucleotide distances, thus the internucleotide ROEs will have a stronger intensity than the intranucleotide ROEs, while for Z-DNA the opposite is true.

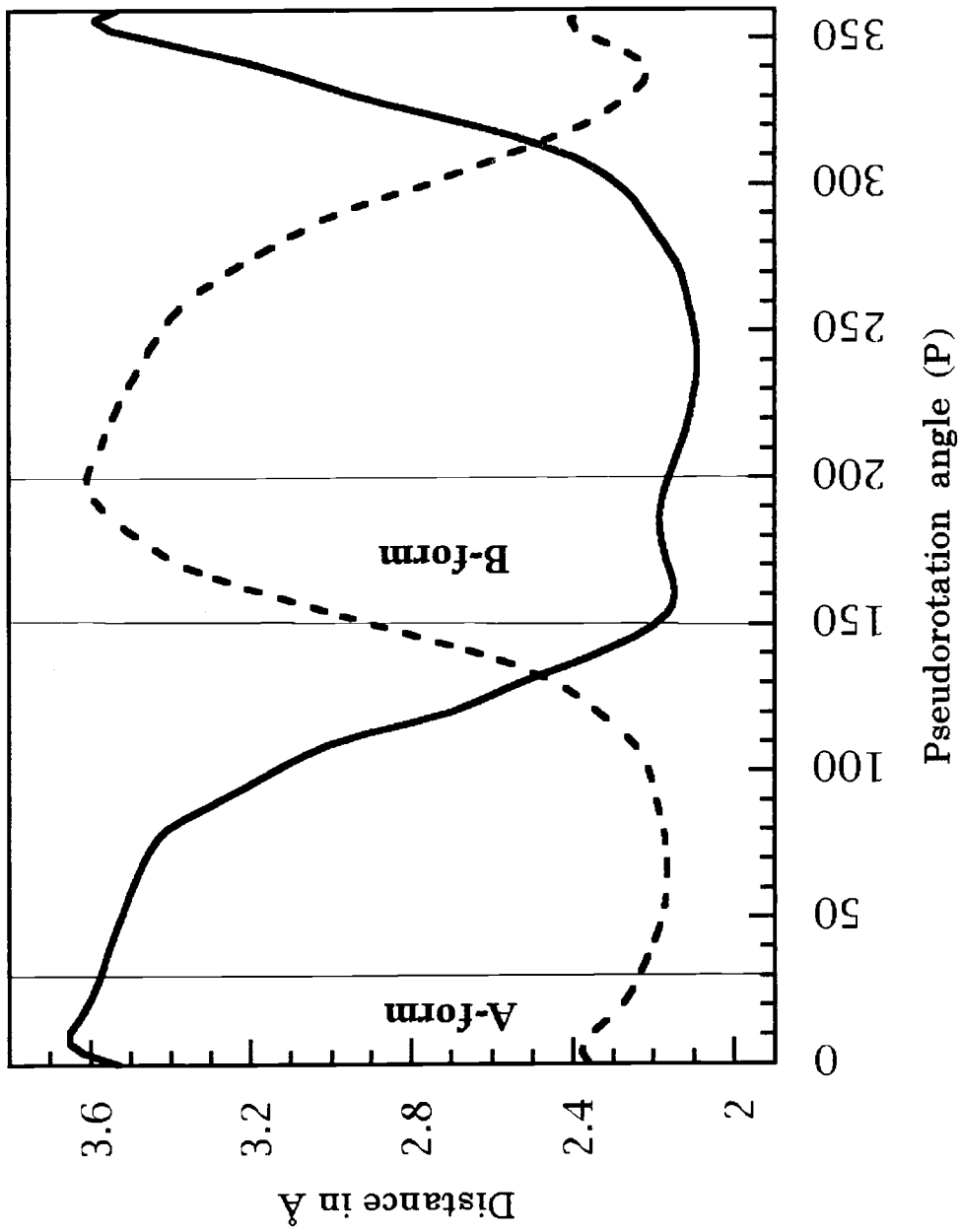

Figure II.6 Internucleotide distances between the H2' of the 5' residue and the base proton (H8) of the 3' residue (------) and between the H2" of the 5' residue and the base proton (H8) of the 3' residue (———). For typical B-form DNA sugar pucker values, the H2" (n) to H8 (n+1) distance is shorter than the H2' (n) to H8 (n+1) distance resulting in a negative ROESY crosspeak. For typical A-form DNA sugar pucker values, the opposite trend should be observed. Values were extracted using a canonical B-DNA dinucleotide created with the 'Biopolymer' module of "InsightII" (Biosym Technologies, Inc.), and overlaying different sugar pucker rings on the first nucleotide, leading to approximate distances.

II.5 Conclusion

The ROESY experiment is an important tool for DNA solution structure determination. The ROESY spectrum and its alternating crosspeak intensities enhances DNA proton resonance assignments, especially when distinguishing the H2' protons from the H2" protons. This method of stereospecifically assigning the H2', H2" resonances is straightforward and unambiguous, particularly in cases where line broadening makes measurements of couplings impossible in a COSY spectrum. Another significant aspect of the ROESY experiment is the ability to make certain crosspeaks "disappear", thus making the experiment a very useful method for editing complex spectra. Lastly, the ROESY technique can also be a direct method to distinguish B-DNA from A-DNA and Z-DNA conformations based on the fact that the internucleotide distances between the H2', H2" and the H6/H8 alternate in each form due to the different sugar puckers, glycosidic torsion angles and helicity handedness.

II.6 Acknowledgements

We would like to acknowledge Jeff Vargason for writing the program to generate sugar pucker coordinates at different values of pseudorotation angle. We would also like to thank the National Institute of Environmental Health Sciences Pilot Project (2PDIES00040-34) for providing funding for this research. Our Bruker DRX 600 MHz NMR spectrometer was purchased with funds provided by the W. M. Keck Foundation, the National Science Foundation (BIR-9413692) and Oregon State University.

Chapter III

The Role of Asymmetric DNA Sequences in the Orientations of Actinomycin D Binding

Monika Ivancic

III.1 Synopsis

The cellular actions of the anti-tumor drug Actinomycin D have been extensively studied in order to fully understand how ActD inhibits DNA transcription. The three dimensional structure of ActD bound to DNA enhances the understanding of ActD's action within the cell. Several high resolution structures of ActD:DNA complexes exist, though only one involves non-self-complementary sequences (Liu, *et.al*, 1991; Kamitori & Takusagawa, 1992, 1994; Brown *et al.*, 1994; Lian *et al.*, 1996). When the complex between ActD and a DNA hairpin molecule was investigated, the solution structure showed that ActD has two orientations of binding and that it prefers one orientation above the other (Brown *et al.*, 1994). In the current study, every possible combination of non-self-complementary base pairs surrounding the d(GC) intercalation site were examined to find the sequences showing maximal disproportionate binding. The d(CCGCCG)•d(CGGCGG) sequence was shown to have the highest preference of orientation, with 79% of ActD molecules binding with the benzenoid side of the chromophore in the d(CGGCGG) strand. NMR spectral complications prevented the structure elucidation of this complex. The d(CTGCGG)•d(CCGCAG) sequence was identified to have 67% of ActD molecules bound with the benzenoid side of the chromophore intercalated in the d(CTGCGG) strand (the major complex), while 33% of the ActD molecules bound with the benzenoid side in the d(CCGCAG) strand (the minor complex). The solution structures of both orientations of binding were solved using the total relaxation matrix approach for structure refinement with a family of 10 structures generated for each complex. The rms difference over all atoms is 0.56

Å for the major complex and 0.89 Å for the minor complex. The structures of the DNA hexamers are in good agreement in both complexes with a pairwise rms difference over all atoms of 0.99 Å. By overlaying the four cyclic pentapeptide rings, the total rms difference over all atoms is determined to be 0.47 Å, and significant structural differences between the major and minor orientations of binding are not detected. The preference of orientation of ActD was identified for seven out of eight sequences, with the highest requirement for asymmetric binding being that the benzenoid intercalation d(GC) site is followed by a guanine.

III.2 Introduction

Actinomycin D (ActD), produced by the bacterium *Streptomyces antibioticus*, is a drug which has been used clinically to treat certain types of cancer (Farber, 1966). ActD binds to DNA and inhibits transcription by hindering the translocation of RNA polymerase (Goldberg *et al.*, 1962; Kirk, 1960; Mauger, 1980; Reich *et al.*, 1961). The drawback in using ActD in clinical treatments is that it binds to DNA sequences of all cell types - it does not differentiate between normal and damaged DNA - thus it affects both healthy and cancerous cells. The chemical composition of ActD is an asymmetric phenoxazone chromophore with a benzenoid and a quinoid side, each connected to identical cyclic pentapeptide lactone rings (Figure III.1). ActD binds to DNA from the minor groove side of the DNA duplex (Figure III.2) (Lian *et al.*, 1996). The chromophore is intercalated between base pairs while the

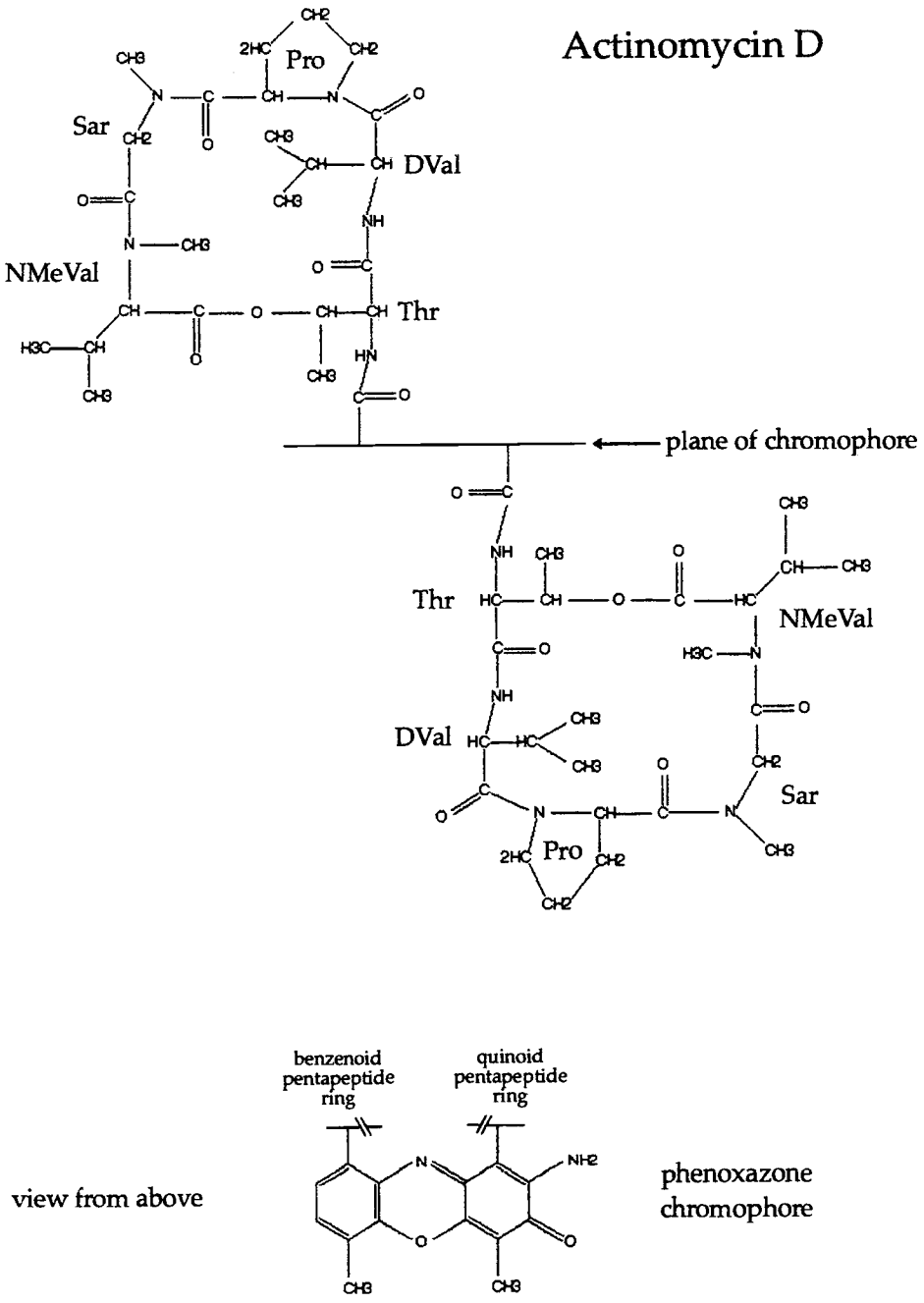

Figure III.1 Chemical composition of ActD, consisting of an asymmetric phenoxazone chromophore with a benzenoid side and a quinoid side. Out of each side protrude identical cyclic pentapeptide lactone moieties. Each cyclic pentapeptide moiety consists of a L-threonine (Thr)-D-valine (DVal)-L-proline (Pro)-sarcosine(Sar)-L-N-methylvaline (NMV) sequence with an ester linkage between γ-O of L-Thr and CO of NMV.

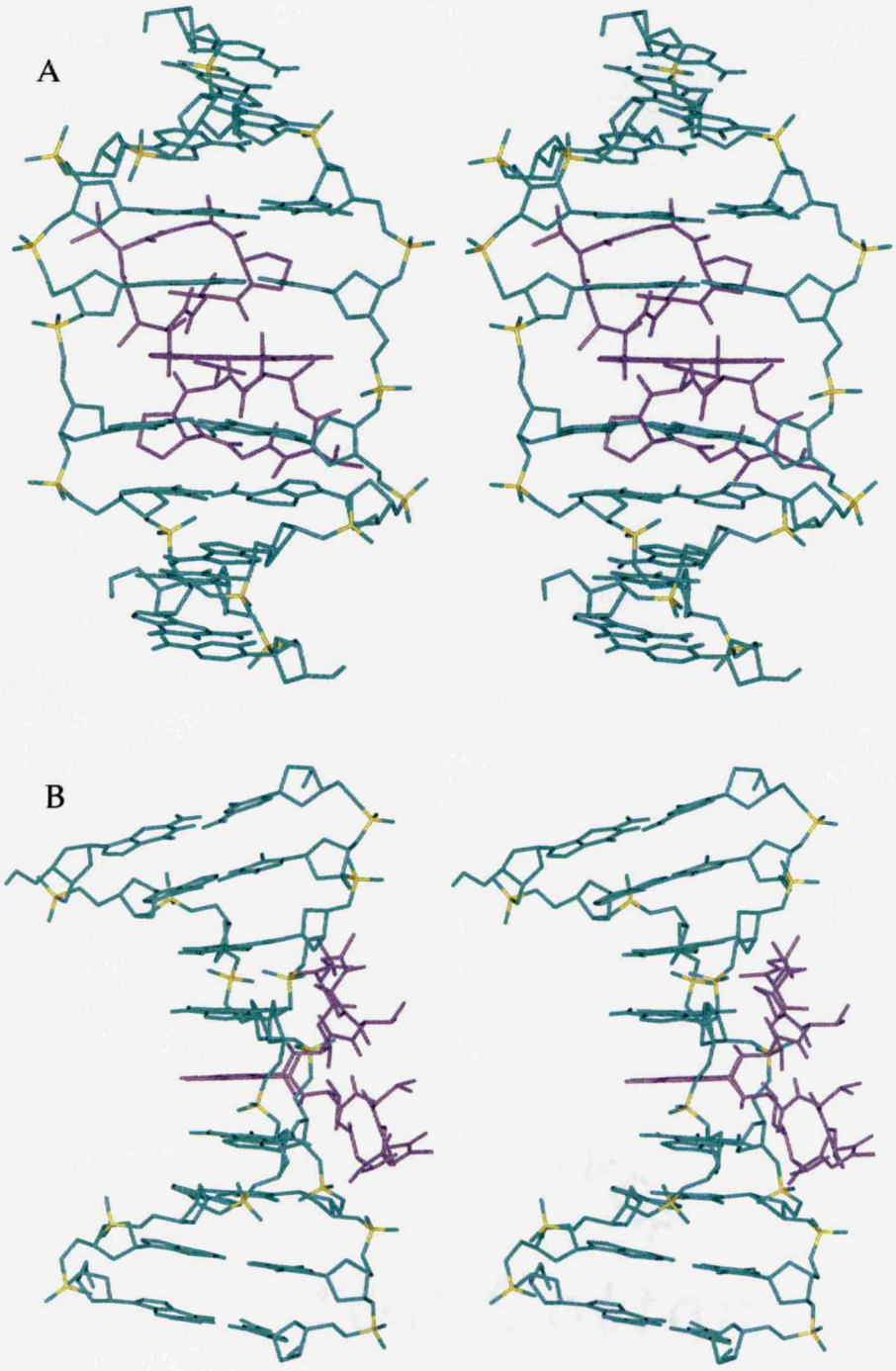

Figure III.2 Stereoviews of the crystal structure of ActD intercalated into the d(GAAGCTTC)$_2$ DNA sequence (Kamitori & Takusagawa, 1992). In a) the view into the major groove is shown, with the ActD drug in purple and the DNA helix in green. b) A side view of the same complex is presented. The chromophore stacks with the DNA base pairs, while the cyclic pentapeptide rings bind in the minor groove.

cyclic pentapeptide lactone rings bind in the minor groove (Brown *et al.*, 1984; Brown *et al.*, 1994) . ActD preferentially intercalates adjacent to guanine residues (Cerami, 1967; Wells & Larson, 1970) , the requirement for an ActD classical mode of binding to double stranded DNA being the exocyclic 2-amino group of guanine and a guanine (3'-5') cytosine step. Its site specificity was further confirmed via DNase I footprinting studies (Fox & Waring, 1984; Goodisman & Dabrowiak, 1992; Goodisman *et al.*, 1992; Waterloh & Fox, 1991).

Researchers are interested in the molecular details of the ActD:DNA interaction to understand the cellular action of ActD and to design a more effective anti-cancer drug (Shinomiya *et al.*, 1995; Takusagawa *et al.* 1996). The first crystal structure of the deoxyguanosine mononucleoside:ActD cocrystal was determined by Jain and Sobell in 1972 (Jain & Sobell, 1972) , and was used to construct a model for how ActD might bind to longer DNA sequences (Sobell & Jain, 1972) . In the proposed 'Sobell and Jain' model the ActD phenoxazone chromophore intercalates between d(GC) steps from the DNA minor groove, and the complex is stabilized by intermolecular stacking between the DNA base pairs and the ActD chromophore, specific hydrogen bonds, and van der Waals interactions between the cyclic pentapeptide rings and the DNA minor groove (Sobell, 1985; Sobell & Jain, 1972) . There is a strong hydrogen bond between the exocyclic guanine amino proton at the intercalation site and an ActD threonine carbonyl oxygen, and a weaker hydrogen bond between the guanine N3 and an ActD threonine amide proton, the quantitation of strength based on the geometry of the hydrogen bond. These hydrogen bonds were hypothesized to be responsible for the guanine specificity and are illustrated in Figure III.3. Single crystal structures were also solved for free ActD (Ginell,

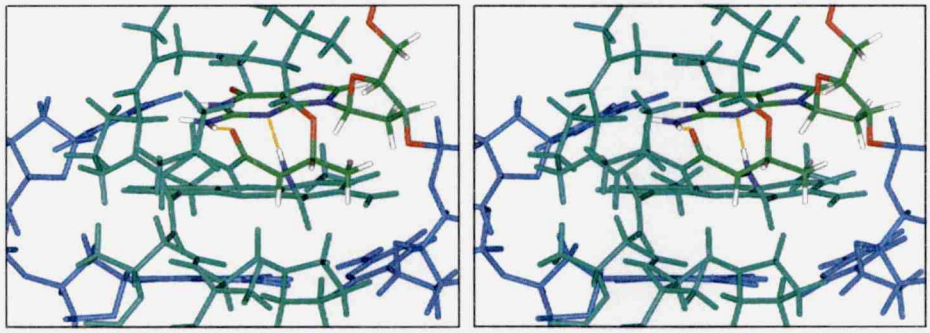

Figure III.3 Hydrogen bonds formed between ActD and the intercalation site guanine base from the NMR structure of ActD:d(AAAGCTTT)$_2$ (Liu, X. *et al.*, 1991). The hydrogen bonds are shown in orange color. A strong hydrogen bond is formed between the Thr carbonyl oxygen and the guanine exocyclic NH_2 group (distance 1.9 Å, angle 170°) and a weaker hydrogen bond is formed between the Thr amide proton and guanine N3 (distance 2.0 Å, angle 163°). Identical hydrogen bonds are formed on the opposite side of the drug.

1988) , which were essential for comparison with the ActD structure in complex with DNA. No significant differences are observed in the free ActD structure and in the structure of ActD bound to the $d(GAAGCTTC)_2$ sequence as solved by NMR (Lian *et al.*, 1996).

Following these studies, ActD was crystallized in complex with $d(GC)_2$ and a pseudo-intercalated complex was identified (Takusagawa *et al.*, 1982) . The only crystal structures of ActD intercalating at the center of a longer DNA helix utilizes the $d(GAAGCTTC)_2$ sequence and identifies three different binding modes for the drug (Kamitori & Takusagawa, 1992; 1994) . Two of the ActD:$d(GAAGCTTC)_2$ complexes found to be symmetric are crystallized in the F222 space group (Kamitori & Takusagawa, 1994) , while the asymmetric complex is crystallized in the C2 space group. In the latter complex, the DNA helix is slightly unwound by rotating one of the base-pairs at the intercalation site, creating a unique asymmetrically wound helix (Kamitori & Takusagawa, 1992). All three structures are similar and portray all aspects of the 'Jain and Sobell' model. The crystallographic structures establish that the cyclic pentapeptide lactone rings of ActD span two base pairs on either side of the intercalation site on the DNA sequence.

The first NMR studies of ActD:DNA complexes showed that the binding curve for poly d(GC) is sigmoidal while the binding curves of all other poly d(NG) dinucleotides is hyperbolic, indicating that ActD binds cooperatively to d(GC) sequences (Krugh, 1972). A similar study using mononucleotides showed that the ActD molecule has two binding domains for guanine nucleotides on either side of the chromophore (Krugh & Neely, 1973). More in depth NMR studies, utilizing intermolecular NOEs, oxygen labeling and involving ActD

complexes with the self-complementary duplex $d(AGCT)_2$ were published (Delepierre, 1989; Gorenstein *et al.*, 1984; Reid *et al.*, 1983) , as were other studies involving complexes with $d(CGCG)_2$ (Delepierre, 1989; Patel, 1976; Petersheim *et al.*, 1984). These studies verified the proposed model for binding -- that the drug chromophore intercalates between the d(GC) base pairs of the double helix while the pentapeptide lactone rings fill the minor groove. Distortions in the helix geometry were also identified, especially when ^{31}P NMR spectroscopy is utilized. It was shown that two ^{31}P resonances shift downfield and that these downfield shifts reflect unwinding of the O-P-O DNA backbone angles (Delepierre, 1989; Gorenstein *et al.*, 1984; Patel, 1976; Petersheim *et al.*, 1984). Complexes with $d(ATGCAT)_2$ were also investigated using both ^{1}H and ^{31}P NMR experiments (Brown *et al.*, 1984; Patel, 1974), in addition to complexes using longer DNA sequences (Brown *et al.*, 1994; Jones, 1988; Lian *et al.*, 1996; Liu *et al.*, 1991; Patel, 1981) . These NMR studies support the 'Sobell and Jain' model (Sobell & Jain, 1972), identify the changes in the DNA helix upon complexation, particularly the unwinding of the helix at the intercalation site, and find no significant conformational change of the pentapeptide lactone rings between bound and free ActD.

All of the above mentioned studies utilize self-complementary DNA sequences and thus identify only one orientation of binding of ActD. The study that led to our investigations involve the binding of ActD to the stem of a DNA hairpin sequence, making this the first non-self-complementary DNA sequence complexed with ActD (Brown *et al.*, 1994) . It was found that ActD bound in two orientations and the drug preferred one orientation of binding over the other.

The major orientation was found to be 59% abundant while the minor orientation was only 41% abundant. This hairpin was designed to have two base pairs between the intercalation site and the tetra-thymine loop (Brown *et al.*, 1994). When one of these base pairs was omitted and the same analysis performed, it was found that the drug preferred the opposite orientation of binding. These results suggest that the ActD binding site may have more sequence specificity than has previously been appreciated. To gain a better understanding of what drives this asymmetric binding and why ActD binds in two different orientations we undertook investigations of ActD complexes with non-self-complementary sequences.

For the current study, DNA hexamers with a central d(GC) intercalation site were designed (Table III.1). All possible combinations of non-self-complementary flanking base pairs were investigated, evaluating the ActD:DNA complexes via 1D and 2D NMR techniques, and resulting in the identification of two sequences that exhibited the largest ratios of preference for one orientation over the other. One of these sequences was d(CTGCGG)•d(CCGCAG) with a preference of 67% for the benzenoid side of the chromophore located between the central d(GC) of the first strand of the DNA helix. The solution structure of this sequence complexed with ActD was solved for both the major and minor orientations of binding. The detailed molecular structure of both complexes identifies whether or not there are differences in how the drug binds in the two orientations. The other strongly asymmetric sequence, d(CCGCCG)•d(CGGCGG), was found to have a preference of 79% for the benzenoid side intercalated between the central d(GC) step on the second strand of the DNA helix. However, resonance overlap and

Table III.1. Sequences with every possible base pair surrounding the central d(GC) intercalation site.

sequences	**A**	**T**	**C**	**G**
CaGCxG	CaGC**A**G	CaGC**T**G	CaGC**C**G	CaGC**G**G
	GTCGTC	GTCGAC	GTCGGC	GTCGCC
CtGCxG	CtGC**A**G	CtGC**T**G	CtGC**C**G	CtGC**G**G
	GACGTC	GACGAC	GACGGC	GACGCC
CcGCxG	CcGC**A**G	CcGC**T**G	CcGC**C**G	CcGC**G**G
	GGCGTC	GGCGAC	GGCGGC	GGCGCC
CgGCxG	CgGC**A**G	CgGC**T**G	CgGC**C**G	CgGC**G**G
	GCCGTC	GCCGAC	GCCGGC	GCCGCC

*Lower case letter – modification of second base
*Bold letter – modification of fifth base

line broadening severely complicated the complete characterization of this complex. The identification of benzenoid preferred strands in all non-self-complementary sequences, allows for the recognition of trends associated with this preferential binding.

Further investigations into the solution structure of the d(CTGCGG)•d(CCGCAG) sequence binding to ActD were performed by modifying the bases flanking the central intercalation site and quantitating any changes in orientation or distribution of ActD binding. A 5-methylcytosine was used and ActD binding to the d(CTGCGG)•d(C^{5me}CGCAG) sequence showed that 66% of ActD preferred having the benzenoid side of the chromophore intercalated in the d(CTGCGG) strand. Some biological significances of CpG methylation are found in its association with transcriptional repression (Razin & Cedar, 1991) and its ability to modulate chromatin structure (Davey *et al.*, 1997; Lewis & Bird, 1991) . Previous NMR investigations show that a methylated cytosine influences the structure of the DNA molecule (Lefebvre *et al.*, 1995; Marcourt *et al.*, 1999) and thus may have an effect on the DNA's drug binding abilities.

Another sequence studied utilized an inosine instead of a guanine as a flanking base. The d(CTGCIG)•d(CCGCAG) sequence lacks a NH_2 group in the minor groove at the flanking site and is found to reduce the preference of ActD orientation. With this sequence, 58% of ActD molecules bound with the benzenoid side of the chromophore in the d(CTGCIG) strand. Researchers are interested in inosine nucleotides, since inosine occurs naturally in the wobble positon of the anticodon of some t-RNA's (Kumar *et al.*, 1992). Poly (rI) and poly (dI) serve as templates for the incorporation of cytosine into products of

DNA and RNA polymerases (Hall *et al.*, 1985). Structural studies have shown that inosine does not have a significant effect on the overall DNA structure (Oda *et al.*, 1991; Uesugi *et al.*, 1987) , however, local distortions exist, which may have an influence on how a drug molecule binds to this locally distorted DNA sequence. The ActD titration studies described in this chapter identify the change in preferred orientation for ActD in the sequence incorporating inosine in replacement of guanine. The simple chemical changes associated with inosine and 5-methylated cytosine incorporation enhances our understanding of what drives the asymmetric binding of ActD.

III.3 Materials and methods

III.3.1 Design of sequences in the ActD binding study

In order to ensure that all possible flanking base pairs surrounding the central d(GC) intercalation site were investigated, a table of possible hexamer sequences was constructed with the following rationale: since the G:C base pair has reduced end fraying in solution (Nonin *et al.*, 1995), each hexamer duplex has terminal G:C base pairs, and each sequence has a central d(GC) intercalation site. As seen in Table III.1, the second and fourth base pair in each sequence is varied. For each sequence, its complement strand is also identified, and the table shows a total of 16 duplex sequences.

In this study non-self-complementary sequences are of interest, thus the four self-complementary sequences in Table III.1 were automatically excluded. For each non-self-complementary sequence there exists an identical sequence in

the complement strand. In the table with twelve non-self-complementary sequences, there are only six unique non-self-complementary sequences. These sequences are d(CAGCAG)•d(CTGCTG), d(CAGCCG)•d(CGGCTG), d(CAGCGG)•d(CCGCTG), d(CTGCCG)•d(CGGCAG), d(CTGCGG)•d(CCGCAG) and d(CCGCCG)•d(CGGCGG). These non-self-complementary sequences encompass all possible combinations of flanking base pairs surrounding the central intercalating d(GC) site and all six sequences were involved in this analysis.

Two additional sequences utilizing modified bases were used in this study as well. The d(CTGCGG)•d(CCGCAG) sequence for which the structure was solved in complex with ActD, was used as the "parent" sequence in incorporating a methylated cytosine base and an inosine base, flanking the central intercalation site. The d(CTGCGG)•d(C^{5me}CGCAG) sequence and the d(CTGCIG)•d(CCGCAG) sequence were used in titration studies to observe whether these simple modifications had an influence on the orientation of ActD binding.

III.3.2 DNA synthesis and purification

DNA sequences were synthesized via the phosphoramidite method at the Center for Gene Research and Biotechnology at Oregon State University. Each DNA sequence was synthesized in 3 μmol quantities and purified using polyacrylamide gel electrophoresis (PAGE). PAGE was followed by electroelution, to extract the DNA sample from the excised gel band. The DNA sample was then further purified using a diethylaminoethyl (DEAE) Sephacel anion exchange column and desalted using a C18 Sep Pak cartridge. The exact

details of this procedure are described in Appendix 1. The purity of the DNA was judged by a ^1H 1D spectrum, and then the oligomer was annealed to its similarly purified complement strand. After this process the sample was dried via an N_2 stream and resuspended in a 10 mM sodium phosphate buffer, pH 7, and 10 mM NaCl.

III.3.3 Complex formation

ActD was purchased from Sigma in 10 mg quantities and dissolved in 2 ml ^2H_2O to give a 3.98 mM ActD stock solution. Purity was checked via a proton 1D experiment and the solution was used without further purification. The DNA duplexes were titrated with ActD, starting with a half-molar ratio of ActD to DNA, and adding ActD with one half the volume of the previous addition, until the ActD:DNA molar ratio reached 1:1. This titration process was monitored using proton NMR spectra until the resonances corresponding to free DNA disappeared (Figure III.4). At the end point of these titrations, the solution of the ActD:DNA complex was dried and redissolved in 0.50 ml of ^2H_2O. The complexes that were formed were between 0.6 mM and 3.7 mM in concentration, all in a 10 mM sodium phosphate buffer, pH 7, and 10 mM NaCl (Table III.2). For NMR experiments in H_2O, the sample was dried under a stream of argon then dissolved in a mixture of 0.45 ml H_2O/0.05 ml ^2H_2O, the ^2H_2O used for the deuterium lock signal.

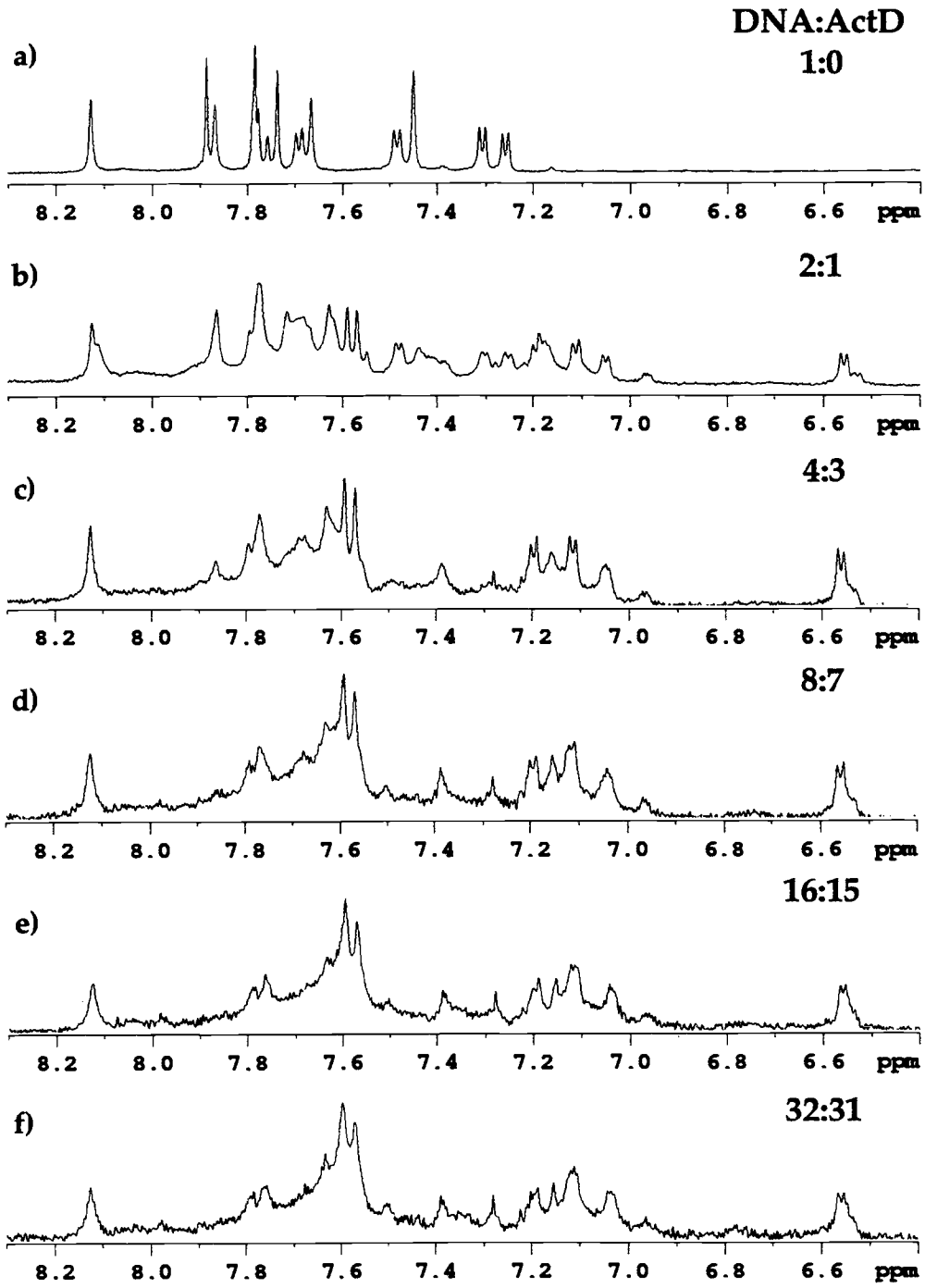

Figure III.4 The 600 MHz ^{1}H 1D aromatic proton region of d(CTGCGG)•d(CCGCAG) upon slow addition of ActinomycinD. Spectrum a) portrays just the DNA aromatic peaks, while spectrum b) has 50% ActD added. In spectrum b) one can still identify peaks due to free DNA, while most peaks shift to their new, complex forming positions. As ActD is added to the sample (spectra c through f) positive complex formation is observed.

Table III.2 Complexes used in our study.

DNA sequence	concentration	T of NMR expts.
$^{5'}$CAGCAG$^{3'}$ $^{3'}$GTCGTC$^{5'}$	0.69 mM	298 K
$^{5'}$CAGCCG$^{3'}$ $^{3'}$GTCGGC$^{5'}$	3.3 mM	298 K
$^{5'}$CAGCGG$^{3'}$ $^{3'}$GTCGCC$^{5'}$	3.7 mM	298 K
$^{5'}$CTGCCG$^{3'}$ $^{3'}$GACGGC$^{5'}$	0.57 mM	298 K
$^{5'}$CTGCGG$^{3'}$ $^{3'}$GACGCC$^{5'}$	0.62 mM	308 K
$^{5'}$CCGCCG$^{3'}$ $^{3'}$GGCGGC$^{5'}$	1.04 mM	293 K
$^{5'}$CTGCGG$^{3'}$ $^{3'}$GACG$\underline{C}$C$^{5'}$	2.50 mM	298 K
$^{5'}$CTGCIG$^{3'}$ $^{3'}$GACGCC$^{5'}$	1.80 mM	298 K

$\underline{C}$ = 5-methyl-cytosine

III.3.4 Homonuclear NMR spectroscopy

NMR spectra were collected on a Bruker DRX600 operating at 14.1 T. Phase-sensitive NOESY, double-quantum filtered COSY (DQF-COSY), and TOCSY spectra were acquired in absorption mode using time-proportional phase incrementation (TPPI) (Marion & Wüthrich, 1983). Phase cycling allowed quadrature detection in t_1 and the spectrometer carrier offset was placed at the water solvent resonance frequency. NOESY spectra of samples in ^2H_2O were recorded with mixing times of 300 ms for the free DNA samples and with mixing times of 50 ms, 100 ms, 150 ms, 300 ms and 500 ms for the complexes. For the NOESY experiments the watergate water suppression scheme (Piotto *et al.*, 1992) was utilized to suppress the residual solvent resonance in the spectra. For the TOCSY experiments continuous radio frequency irradiation was applied during the recycle delay, saturating the residual solvent resonance. For experiments in H_2O, watergate NOESY experiments with mixing times of 50 ms and 300 ms were used for both free and complexed samples. For several of the samples the temperatures of the H_2O experiments were decreased to reduce the rate of imino proton exchange with the solvent. The temperatures which provided optimal dispersion of the proton resonances varied depending on the sample and are listed in Table III.2.

The homonuclear experiments in ^2H_2O were acquired with a spectral width of 10 ppm, while the experiments in H_2O were acquired with a spectral width of 25 ppm. Spectra were acquired with 512 increments in t_1 and 4096 complex points in t_2, for all experiments. For most NOESY spectra, 64 transients were averaged for each t_1 value, while 96 transients were averaged for the DQF-

COSY spectra and 32 transients were averaged for the TOCSY spectra. These values also depended on the sample concentration averaging more transients per t_1 value for the lower concentration samples. For most experiments, a cosine squared window function was used for apodization during processing. Zero-filling of the data produced 2D matrix sizes of 4K by 4K real points. The data was processed using the Bruker XWIN-NMR software package running on Silicon Graphics IRIX workstations.

III.3.5 Heteronuclear NMR spectroscopy

The sensitivity enhanced proton-detected ^{13}C heteronuclear single-quantum coherence experiment (HSQC) of the ActD-DNA complex was obtained using the pulse sequence of Schleucher *et al.* (1994) in the echo-antiecho mode (Kay *et al.*, 1992; Palmer III *et al.*, 1991; Schleucher *et al.*, 1994). A total of 64 transients were averaged for each of the 512 increments, and 2048 complex points in t_2 were recorded. Spectral widths of 10 ppm in the ^1H dimension and 100 ppm in the ^{13}C dimension were employed. The carrier frequency was set to the H_2O resonance frequency in the ^1H dimension and at 50 ppm in the ^{13}C dimension. A proton-carbon coupling constant of 140 Hz was used to optimize coupling evolution during the $1/(2J_{HX})$ delay.

The ^1H-^{31}P scalar correlated 2D spectra were obtained with the same HSQC pulse sequence. A total of 160 transients were averaged for each of the 200 t_1 increments and 2048 complex points in t_2 were recorded for most samples. Spectral widths of 10 ppm in the ^1H dimension and 10 ppm in the ^{31}P dimension were employed. The carrier frequency was set to the H_2O resonance frequency

in the ^{1}H dimension and at -1.0 ppm in the ^{31}P dimension. The $1/(2J_{HX})$ delays were optimized for a proton-phosphorous coupling constant of 21 Hz. The data was processed using linear prediction of 50 points in the t_1 dimension, and a cosine squared window function in both dimensions.

III.3.6 Structure calculations

NOE distance and dihedral angle restrained molecular dynamics and energy minimization calculations were performed on a Silicon Graphics Indigo 2 workstation using the Crystallography and NMR System (CNS) software package (Brünger *et al.*, 1998). The iterative relaxation matrix analysis was employed in the solution structure determination of the major and minor ActD:d(CTGCGG)•d(CCGCAG) complexes. The initial NOE distance constraints were estimated using the 2-spin approximation. These were repeatedly refined via back-calculation of a simulated NOESY spectrum using the BIRDER program (Zhu & Reid, 1995) by minimizing the differences between the NOE build-up curves for the experimental and simulated NMR data. BIRDER calculates cross-relaxation rates using the total relaxation matrix approach, a method that accounts for the effects of spin diffusion. BIRDER uses the ^{1}H assignments and linewidths, the solved solution structure as well as the correlation time of the molecule, the NOE mixing times and the experimental relaxation delay to backcalculate a simulated spectrum. An isotropic correlation time was used for the ActD:hexamer complex estimated at a value of 2.20 ns. The program takes into account incomplete recovery of z-magnetization, necessary when protons with different T1 recovery times are present. The T1 recovery times for the different types of DNA protons were experimentally

determined (Table III.3) using the inversion recovery method and nine partial recovery periods. The T1 inversion recovery experiments for the DNA aromatic protons are shown in Figure III.5. BIRDER also accounts for differential external relaxation, which must be considered in accurate reproduction of experimental NOE buildup curves and has been shown to be helpful in reducing the residual (R) factor. Integration of selected crosspeaks at three mixing times, 100ms, 150ms and 300ms, for both experimental and backcalculated spectra was used for residual factor calculation. The equation $R=(1/N)\,\Sigma_N\,[\,|\,V_{exp}-V_{sim}\,|\,/V_{exp}]$ was used for R factor calculation where N is the number of NOESY crosspeaks selected.

For structure calculations, all non-significantly overlapping NOE restraints were refined using the total relaxation matrix approach as described above and the same crosspeaks were included in the R-factor calculation. A significant matter of complication was the presence of two conformers, and the concomitant overlap of many crosspeaks of the major and minor conformers. Where it was certain that both the major and minor conformer NOEs lie under one peak, the peak was scaled to 0.6 times the volume for the major and 0.4 times the volume for the minor conformer (the ratios determined at the elevated temperature of 308K and using the same crosspeaks as the ratio determination at 298K). Unfortunately, numerous overlapping crosspeaks exist, which have contributions from multiple proton-proton interactions within one conformer. Thus the interactions of these crosspeaks were excluded from the structure calculations as well as the R-factor calculation. To obtain a better idea of which interactions are contributing most to the R-factor, the types of crosspeaks were subdivided into three tables. The first table included only DNA

Table III.3 T1 values determined for DNA protons in the sample.

DNA proton	T1 recovery
H1'	1.18 sec
H2'/H2''	1.22 sec
H3'/H4'	1.25 sec
H8/H8	1.30 sec
CH_3	0.97 sec

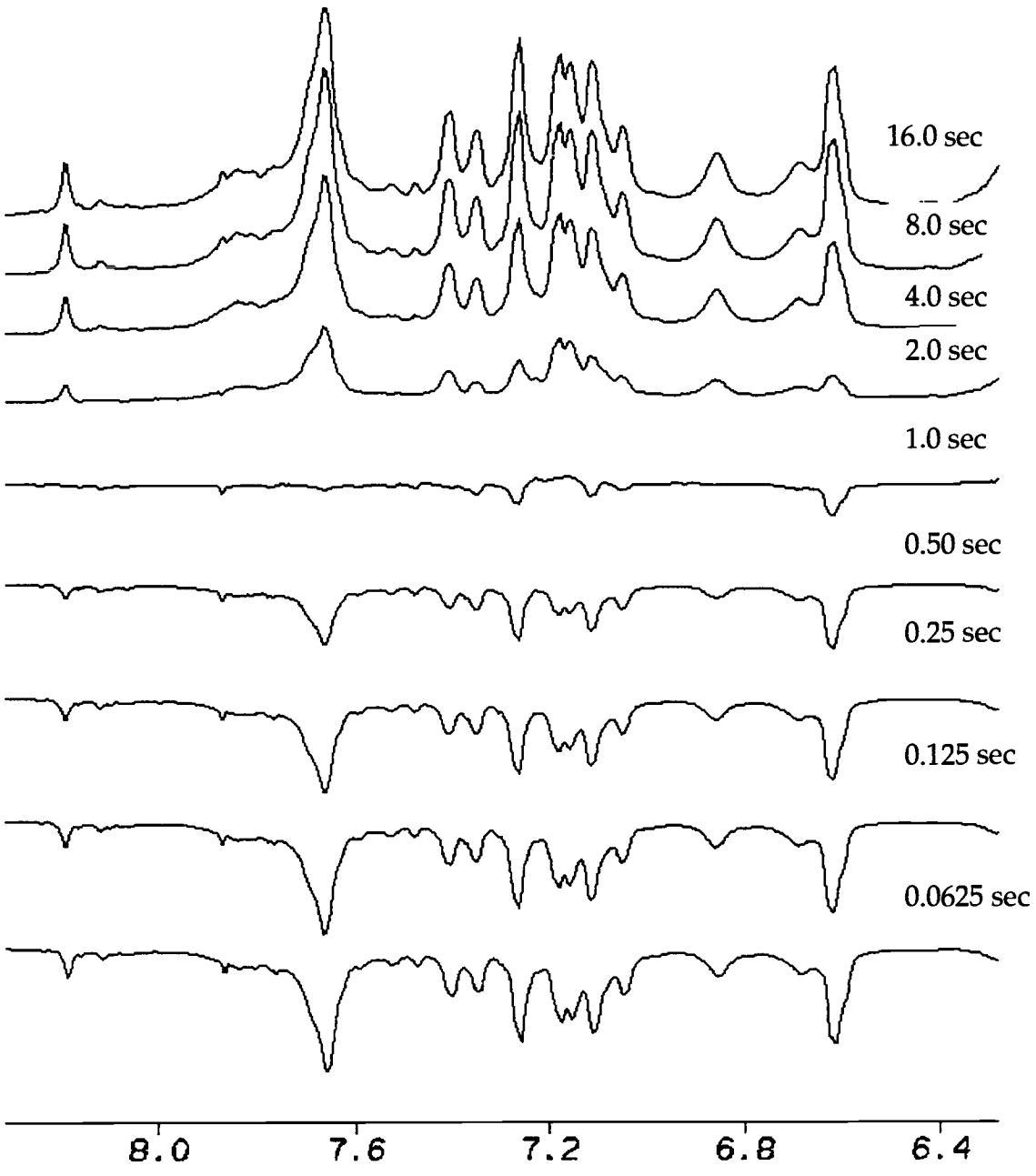

Figure III.5. T1 inversion recovery experiments on the ActD:d(CTGCGG)•d(CCGCAG) sample. The aromatic H6/H8 proton region is shown with the recycle delay times indicated. The average T1 period for protons in this region is determined to be 1.3 sec. The percent recovery is determined with the equation: % recovery = $1 - e^{-t/T1}$, where t is the total relaxation time (RD+AQ) and T1 is the inversion recovery time. The average percent recovery for the H6/8 protons is calculated to be 81%.

crosspeaks, the second table included only intermolecular crosspeaks and the third table included only ActD crosspeaks. In the first table were 59 DNA crosspeaks, of which only a few were scaled to the appropriate volume because of their overlap with the other conformer. The interactions in this table were used as distance restraints in the structure calculation and were used to validate the final DNA structure in the R-factor calculation. The second table includes all intermolecular interactions, though this was subdivided into parts a) and b), with a) listing only non-overlapping crosspeaks, 33 total, that would be used in validating the position of ActD within the DNA hexamer and b) having nearly 60 crosspeaks. The third table involves ActD crosspeaks with a distinction of a) non-overlapping crosspeaks and b) overlapping crosspeaks, the first category being used for structure and R-factor calculations and totaling 52 crosspeaks and the second category totaling about 160 crosspeaks. The distance constraints for each interaction were identified as described and tight bounds of +/- 0.2 Å were placed on each constraint. For the major complex a total of 144 interactions were used in the R-factor and structure calculations. For the minor complex more crosspeaks were used for the structure calculation than for the R-factor calculation, since only 71 peaks were completely resolved and included in the total relaxation matrix refinement. Because of such a low number of resolvable crosspeaks and the need for additional restraints during the structure calculations, the other 73 interactions included in the structure refinement were assumed to have an equal interproton distance as in the major complex. Thus, for these 73 interactions that were observed in the NOESY spectra, the same NOE restraints were used as in the major complex.

Since the data was recorded at 35°C for maximal chemical shift resolution and ease of identification of the major and minor complexes, and since the DNA is of such short length, the presence of local motion within the sample can not be ruled out. At this higher temperature, the molecule tumbles in solution faster, thus sharper resonance signals are observed. A second set of structures was calculated which allow for the possibility of conformational motion. For these structures the same distance constraints were used determined from the NOE volumes as described above, but for the internucleotide distance constraints the upper bounds were opened to a value of 5 Å to allow for exploration of more conformational space. The families of these structures showed more conformational heterogeneity, since loose bounds were employed during the calculations and are discussed in the following section.

Sugar pucker geometries were evaluated using dipolar coupled data. Spectral overlap in the H1′ to H2′, H2″ protons region and linewidths greater than the coupling constants make sugar puckers for individual nucleotides difficult to determine from DQF-COSY data. Sugar pucker dihedral angles were not explicitly restrained in the structure calculation for this reason, however, the H1′ to H4′ NOE was used as a distance constraint and in the R-factor calculation.

Values for the backbone dihedral angles were extracted from the ^{31}P chemical shifts based on the relationship previously described (Nikonowicz & Gorenstein, 1990) . The ε dihedral angle (C4′-C3′-O3′-P_{i+1}) and ζ dihedral angle (C3′-O3′-P_{i+1} – O5′$_{i+1}$) were restrained to –165° +/- 20° and –95° +/- 20°, respectively, at the non-intercalating sites, which are typical B-form DNA values in solution structures. At the phosphate near the benzenoid side of the chromophore the ε dihedral angle was restrained to –55° +/- 40° and the ζ

dihedral angle was restrained to –250° +/- 40°, while on the quinoid side of the chromophore the ε dihedral angle was restrained to –15° +/- 40° and the ζ dihedral angle was restrained to –300° +/- 40°. The values applied for each complex are listed in Table III.4.

The starting structures were generated by using a previously solved crystallographic N8ActD:DNA structure (Shinomiya *et al.*, 1995) and replacing the N8ActD N8 atom with C8 and appropriate bases with bases in d(CTGCGG)•d(CCGCAG) using the InsightII molecular modeling program (MSI/Biosym Technologies, Inc., San Diego, CA). Structures were determined using a simulated annealing protocol implemented in the CNS program (Brünger *et al.*, 1998). The structures were then subjected to molecular refinements in two stages, a high temperature stage and a slow-cooling stage. The high temperature portion of the calculation involved 14 ps of molecular dynamics (7000 steps of 0.002 ps each) to ensure that an adequate amount of conformational space was explored. The slow cooling portion started at a temperature of 1000K, cooling to 300K in 50K steps. The time trajectory of the slow-cooling molecular refinement was 5 ps in length in total (5000 steps of 0.001 ps each). A square-well potential was used for the NOE restraints, and a scaling factor of 150 for the NOE energy term was employed during the molecular dynamics calculation. The method of center averaging was used for all distance restraints with the SUM averaging method developed by Nilges (Nilges, 1993; Nilges, 1995). The final part of the calculation involved 10 cycles of 200 steps each of energy minimization. To maintain DNA base pairing and the double-helical nature of the hexamer, hydrogen bond NOE constraints were

Table III.4 Dihedral angle DNA backbone constraints for both complexes

residue	major complex			minor complex		
	^{31}P shift (ppm)	ε angle (°)	ζ angle (°)	^{31}P shift (ppm)	ε angle (°)	ζ angle (°)
C1	-3.75	-165	-95	-4.08	-165	-95
T2	-4.10	-165	-95	-3.94	-165	-95
G3	-2.30	-55	-250	-1.49	-15	-300
C4	-4.03	-165	-95	-3.91	-165	-95
G5	-3.47	-130	-140	-3.9	-165	-95
G6	n/a	n/a	n/a	n/a	n/a	n/a
C7	-3.75	-165	-95	-3.76	-165	-95
C8	-4.00	-165	-95	-4.04	-165	-95
G9	-1.22	-15	-300	?	-55	-250
C10	-3.22	-120	-170	-3.77	-165	-95
A11	-3.84	-165	-95	-3.88	-165	-95
G12	n/a	n/a	n/a	n/a	n/a	n/a

employed throughout the calculation with the same scaling factor as used in interproton NOE constraints.

Ten calculations were carried out, each with a different seed number, and each generating 10 structures. For each calculation, the best structure of the ten resulting structures was selected based on the lowest energies and a low number of NOE violations. Thus, the best 10 structures were used to represent each complex calculated with 'tight bounds'. The same was done with the structures generated using 'loose bounds'. Structure analysis was performed on these 10 'tight bounds' structures as well as on the best 10 structures calculated using the 'loose bounds' conditions to get a good representation of the conformational heterogeneity occurring within the sample.

III.4 Results

III.4.1 Resonance assignments of free DNAs

The ^{1}H NMR spectra for each of the free DNA sequences were assigned prior to titration with ActD. The ^{1}H and ^{31}P assignment tables are located in Appendix 2, except for the one belonging to the d(CTGCGG)•d(CCGCAG) sequence, which is in Table III.5. Complete sequential connectivities were identified for each strand of the six sequences, and confirmed by use of the ^{1}H-^{31}P HSQC experiments. The latter experiment shows correlations due to scalar couplings occurring between a backbone phosphorous nucleus and the H3' proton on the 3' side of the phosphodiester bond and the H4' proton on the 5'

Table III.5 ^{1}H and ^{31}P assignments of d(CTGCGG)•d(CCGCAG) reported in ppm*.

nucleotide	H8/H6	H5/H2/Me	H1'	H2'/H2"	H3'	H4'	aminos	iminos	^{31}P
C1	7.84	5.95	5.92	2.15/2.57	4.68	4.10	7.03		-3.58
T2	7.52	1.71	5.77	2.22/2.53	4.89	4.18			-3.36
G3	7.95		5.91	2.68/2.73	4.99	4.38		12.89	-3.41
C4	7.32	5.41	5.66	1.84/2.27	4.81	4.15	6.54/8.49		-3.29
G5	7.85		5.63	2.65/2.71	4.95	4.32		13.26	-3.20
G6	7.80		6.15	2.52/2.33	4.63	4.19			
C7	7.75	5.95	5.99	2.08/2.53	4.67	4.11	7.16		-3.37
C8	7.55	5.68	5.58	2.15/2.44	4.86	4.12	7.01/8.68		-3.18
G9	7.93		5.90	2.67/2.72	4.99	4.37		13.05	-3.49
C10	7.38	5.47	5.46	1.94/2.27	4.80	4.12	6.52/8.51		-3.24
A11	8.20	7.81	6.04	2.71/2.88	5.01	4.39			-3.42
G12	7.73		6.02	2.44/2.25	4.62	4.16			

*ppm values are referenced to the solvent resonance signal at the appropriate temperature

side of the phosphodiester bond. This experiment was used to verify that the sequential DNA assignments were made accurately. The 1D spectra of exchangeable and non-exchangeable protons of d(CTGCGG)•d(CCGCAG) are shown in Figures III.6 and III.7, the exchangeable protons giving rise to three signals in the range between 12 ppm and 13 ppm. These were assigned to the guanine imino protons, indicating duplex formation for d(CTGCGG)•d(CCGCAG) recorded at a temperature of 288K. The absence of the terminal base pair iminos and the T:A base pair imino is most likely due to rapid exchange of the imino proton with solvent caused by fraying or "breathing".

III.4.2 Complex formation, stoichiometry and the two orientations of binding

For all six sequences complex formation was monitored using the aromatic proton region of the 1D ^{1}H spectra as described in the Materials and Methods section. Upon gradual addition of the 4 mM solution of ActD, broadening of the 1D NMR peaks is observed as well as a reduction in the signal to noise ratio. This is a result of the decreasing concentration of the sample, since the sample volume is increased with each addition of ActD. Even after drying and rehydrating the sample to the original concentration, the peaks remain broad because of significant overlap between the DNA and ActD protons and because of the formation of two conformers. The presence of two orientations of binding is immediately detected in the two resonance signals of the chromophore's H7 proton at 6.60 ppm (Figure III.6). The two orientations of binding can also be observed in the 1D spectra of the imino proton region (Figure III.7). The two distinct sets of resonance signals represent the major and

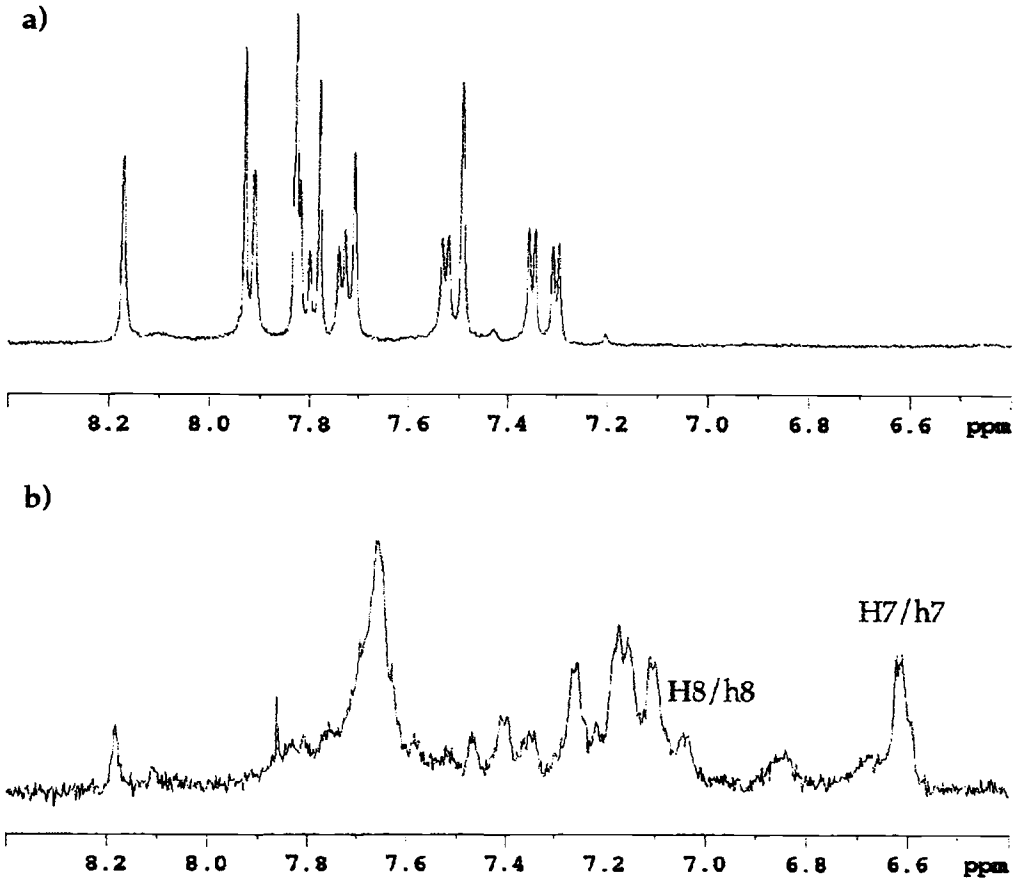

Figure III.6 Aromatic proton region of the 600 MHz 1D ^{1}H spectra of a) d(CTGCGG)•d(CCGCAG) in H_2O, 15°C, pH7, 10mM NaCl, and b) ActD:d(CTGCGG)•d(CCGCAG) (1:1) in $^{2}H_2O$, 15°C, pH7, 10mM NaCl. There is a doubling of the resonances of the complex, compared to free DNA, this is particularly observed for the H8 and H7 protons of the chromophore. This illustrates the formation of two distinct complexes.

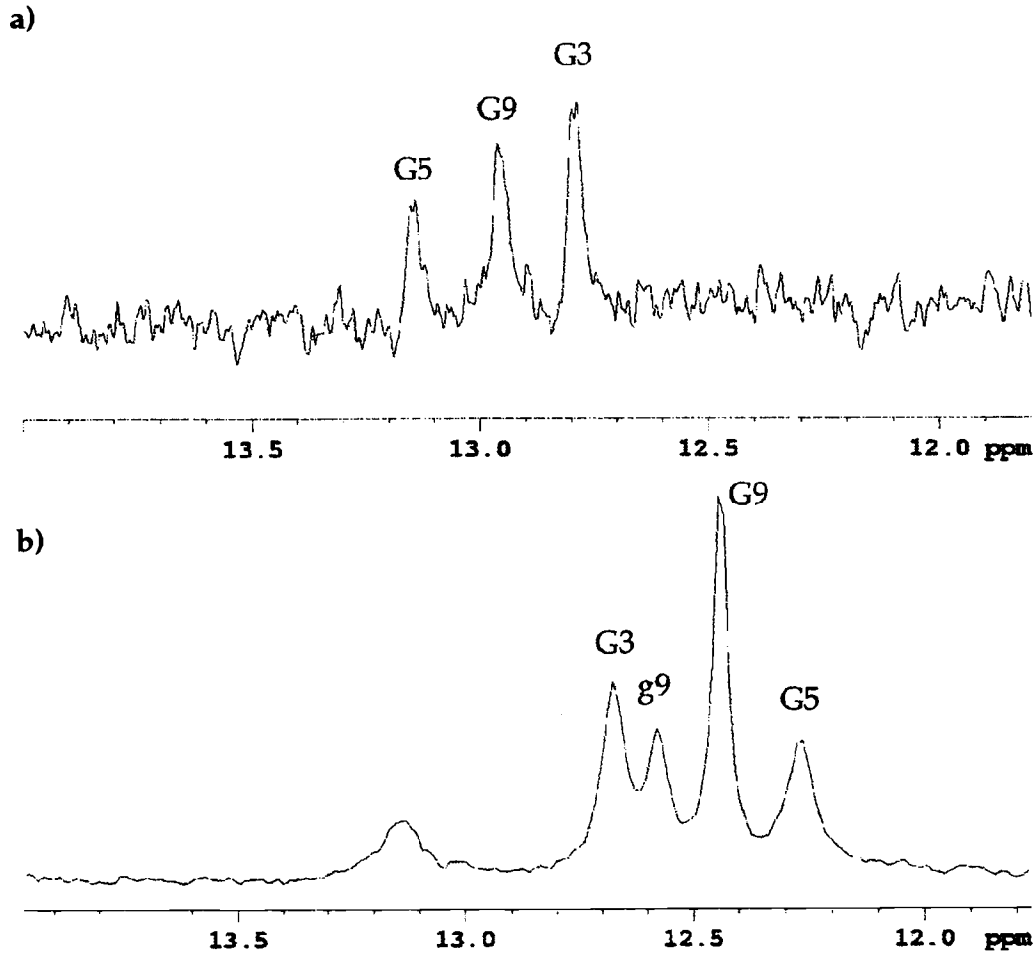

Figure III.7 Imino proton regions of the 600 MHz 1D ^{1}H spectra of a) d(CTGCGG)•d(CCGCAG) in H2O, 15°C, pH7, 10mM NaCl, and b) ActD:d(CTGCGG)•d(CCGCAG) (1:1) in $^{2}H_2O$, 15°C, pH7, 10mM NaCl. Separate resonances are observed for the minor complex, denoted in lowercase letters, and major complex, denoted in uppercase letters.

minor orientation of the drug with respect to the non-self-complementary DNA hexamers. For all the samples, the two orientations of the drug in the complex are in slow exchange, thus both are observable on the NMR timescale. By integration of peak volumes in the NOESY spectrum at 150 msec mixing time and recorded at 298K, the percentages of the two orientations were identified for each sequence (Table III.6). The crosspeaks used for this determination were carefully selected - for most samples this meant choosing peaks that did not overlap and for which a clear distinction between the crosspeaks due to the major complex and the crosspeaks due to the minor complex existed (Table III.7). However, for a few complexes this was not possible, such as for the complex involving inosine. For this complex only, in addition to a few chromophore to DNA crosspeaks, non-overlapping intraDNA crosspeaks from the major and minor conformers were included in the percentage determination. Another criterion for the chosen crosspeaks was based on the strength, as there is more experimental error in determining volumes associated with weak crosspeaks than strong crosspeaks (Liu *et al.*, 1995), thus only strong crosspeaks were chosen for an accurate percentage determination. In most spectra the chosen crosspeaks included ones from the H8 proton of the ActD chromophore to the benzenoid side guanine and cytosine protons as well as the crosspeak from the H8 proton to the H7 proton on the ActD chromophore. The specific crosspeaks used for each complex are listed in Table III.7. By averaging the percent differences of major vs. minor orientation over several crosspeaks for each sample, the percentages for each sample were identified and are listed in Table III.6. The greatest difference in the percentage of major and minor orientation formed was discovered to be 79% vs. 21% for

Table III.6 Preference of orientation of ActD in the DNA sequences investigated.

DNA sequence	no. of peaks integrated	% major	% minor	STD	T for % determination	Benzenoid - major complex
^{5}CAGCAG$^{3'}$ $^{3'}$GTCGTC^{5}	5	57	43	10	298 K	^{5}CTGCTG$^{3'}$
^{5}CAGCCG$^{3'}$ $^{3'}$GTCGGC^{5}	5	55	45	9	298 K	?
^{5}CAGCGG$^{3'}$ $^{3'}$GTCGCC^{5}	10	57	43	9	298 K	^{5}CAGCGG$^{3'}$
^{5}CTGCCG$^{3'}$ $^{3'}$GACGGC^{5}	4	53	47	6	298 K	^{5}CGGCAG$^{3'}$
^{5}CTGCGG$^{3'}$ $^{3'}$GACGCC^{5}	10	67	33	5	298 K	^{5}CTGCGG$^{3'}$
^{5}CCGCCG$^{3'}$ $^{3'}$GGCGGC^{5}	5	79	21	4	298 K	^{5}CGGCGG$^{3'}$
^{5}CTGCGG$^{3'}$ $^{3'}$GACG**C**C^{5}	11	66	34	7	298 K	^{5}CTGCGG$^{3'}$
^{5}CTGCIG$^{3'}$ $^{3'}$GACGCC^{5}	8	58	42	7	298 K	^{5}CTGCIG$^{3'}$

C = 5-methyl-cytosine

Table III.7 Specific crosspeaks used in the ratio determination of each complex

$^{5'}$CAGCAG$^{3'}$ $^{3'}$GTCGTC$^{5'}$	$^{5'}$CAGCCG$^{3'}$ $^{3'}$GTCGGC$^{5'}$	$^{5'}$CAGCGG$^{3'}$ $^{3'}$GTCGCC$^{5'}$	$^{5'}$CTGCCG$^{3'}$ $^{3'}$GACGGC$^{5'}$
PXZ H8 to G_BH2''	PXZ H8 to $THR_BH\gamma$	PXZ H8 to G_BH2'	PXZ H7 to PXZ $^6CH_3$
PXZ H8 to $THR_BH\gamma$	PXZ H8 to C_BH1'	PXZ H8 to G_BH2''	PXZ H8 to C_BH6
PXZ H7 to PXZ $^6CH_3$	PXZ H8 to G_BH1'	PXZ H8 to $THR_BH\gamma$	PXZ H7 to PXZ H8
PXZ H7 to PXZ H8	PXZ H7 to PXZ H8	PXZ H7 to PXZ $^6CH_3$	PXZ H7 to PXZ H8
PXZ H7 to PXZ H8	PXZ H8 to C_BH6	PXZ H8 to G_BH1'	
		PXZ H7 to G_BH1'	
		PXZ H8 to C_BH5	
		PXZ H7 to C_BH5	
		PXZ H7 to PXZ H8	
		PXZ H7 to PXZ H8	

$^{5'}$CTGCGG$^{3'}$ $^{3'}$GACGCC$^{5'}$	$^{5'}$CCGCCG$^{3'}$ $^{3'}$GGCGGC$^{5'}$	$^{5'}$CTGCGG$^{3'}$ $^{3'}$GACGCC$^{5'}$	$^{5'}$CTGCIG$^{3'}$ $^{3'}$GACGCC$^{5'}$
PXZ H8 to G_BH2'	PXZ H8 to $G_BH2'/H2''$	PXZ H8 to G_BH2'	T2 H6 to T2 CH_3
PXZ H8 to G_BH2''	PXZ H8 to $THR_BH\gamma$	PXZ H8 to G_BH2''	T2 H6 to T2 H2'
PXZ H8 to PXZ $^6CH_3$	PXZ H7 to PXZ $^6CH_3$	PXZ H8 to PXZ $^6CH_3$	T2 H6 to T2 H3'
PXZ H8 to $THR_BH\gamma$	PXZ H7 to PXZ H8	PXZ H8 to $THR_BH\gamma$	PXZ H8 to G_BH2''
PXZ H8 to G_BH3'	PXZ H7 to PXZ H8	PXZ H7 to G_BH2''	PXZ H8 to PXZ $^6CH_3$
PXZ H8 to G_BH4'		PXZ H8 to G_BH1'	PXZ H8 to $THR_BH\gamma$
PXZ H8 to C_BH1'		PXZ H8 to C_BH1'	PXZ H8 to G_BH1'
PXZ H8 to G_BH1'		PXZ H8 to G_BH8	PXZ H7 to PXZ H8
PXZ H7 to PXZ H8		PXZ H7 to G_BH8	
PXZ H7 to PXZ H8		PXZ H7 to PXZ H8	
		PXZ H7 to PXZ H8	

* PXZ = phenoxazone chromophore

the ActD:d(CCGCCG)•d(CGGCGG) complex; however, severe line broadening and spectral overlap of several key resonances complicated the complete sequential assignments of the major and minor complex. Thus the complex with the next largest difference in percentage of major to minor orientation, found to be 67% vs. 33% for the ActD:d(CTGCGG)•d(CCGCAG) complex, was used for determining the structures of the two orientations of ActD binding. This sequence was then also used as a template for incorporation of modified bases at the flanking sites.

III.4.3. Results of titration studies using sequences with modified bases

The complex of ActD with d(CTGCGG)•d(C^{5me}CGCAG) had an obvious major and minor orientation of ActD binding identified in the ^{1}H 1D spectra of the aromatic region (Figure III.8). The two orientations were completely identified in the $^{2}H_2O$ NOESY spectrum, particularly in the base proton to H2', H2'' proton region. For the major complex, the benzenoid side of the chromophore is intercalated on the first strand, between G3 and C4, while for the minor complex, the benzenoid side of the chromophore is intercalated on the second strand, between G9 and C10. The crosspeaks due to the ActD chromophore's H7 and H8 protons is of utmost importance in quantitating the orientational preference for binding. By integrating the volumes of these crosspeaks, it was found that 66% of ActD molecules were in the major orientation of binding, while 34% were in the minor orientation.

On the other hand, due to severe line broadening, major and minor orientations were not immediately detected in the complex between ActD and

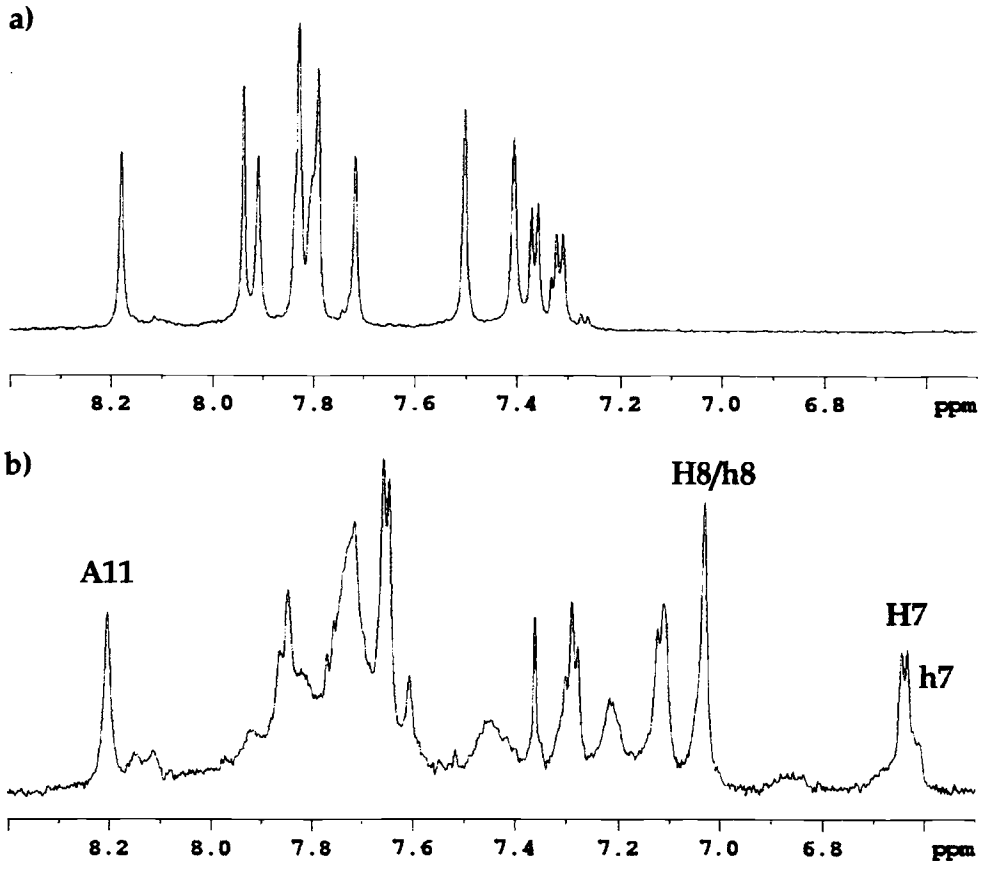

Figure III.8. The 600 MHz ^{1}H 1D NMR spectrum recorded at 298K of the ActD:d(CTGCGG)•d(C^{5me}CGCAG) complex in $^{2}H_2O$. a) The aromatic proton region of free DNA in 10 mM NaPi, pH 7, 10 mM NaCl. b) The aromatic proton region of this DNA complexed with ActD. The presence of a major orientation and minor orientation is detected in the two resonance signals of the chromophore's H7 proton around 6.62 ppm.

d(CTGCIG)•d(CCGCAG). This is observed in the amino proton region in the ^{1}H 1D spectrum of this complex, as seen in Figure III.9. NOESY experiments were recorded at various temperatures ranging between 283K and 318K, without significant improvement. The best spectral dispersion was observed at 308K, and by processing only the "walking region" of the NOESY spectrum to a final matrix size of 16k by 16k data points, it was possible to detect major and minor orientations. Yet for the chromophore's H8 and H7 protons, the major resonances still overlap the minor resonances. Because of severe overlap and the inability of obtaining accurate volume integration results, careful selection of specific crosspeaks to be used in quantitative determinations had to be made as described in the previous section. The major complex with the benzenoid side of the chromophore between G3 and C4 was determined to be 58% of the total, while the minor is determined to be 42% of the total. The major and minor complex of the sequence containing inosine were found closer to being equal than the original sequence containing guanine, which was 67% in the major complex and 33% in the minor complex. The characterization of the complexes utilizing modified bases is summarized in the last two rows of Table III.6.

The DNA strands for both the major and minor orientation complex were almost completely assigned, excluding several of the minor resonances that were difficult to detect. The DNA protons in the ActD:d(CTGCGG)•d(C^{5me}CGCAG) complex was assigned in a standard manner and are listed in Table A3.1 of Appendix 3. The resonances corresponding to the major and minor orientations were more difficult to distinguish in the ActD:d(CTGCIG)•d(CCGCAG) complex. Reprocessing selected regions at higher digital resolution proved to be of utmost importance in assigning the

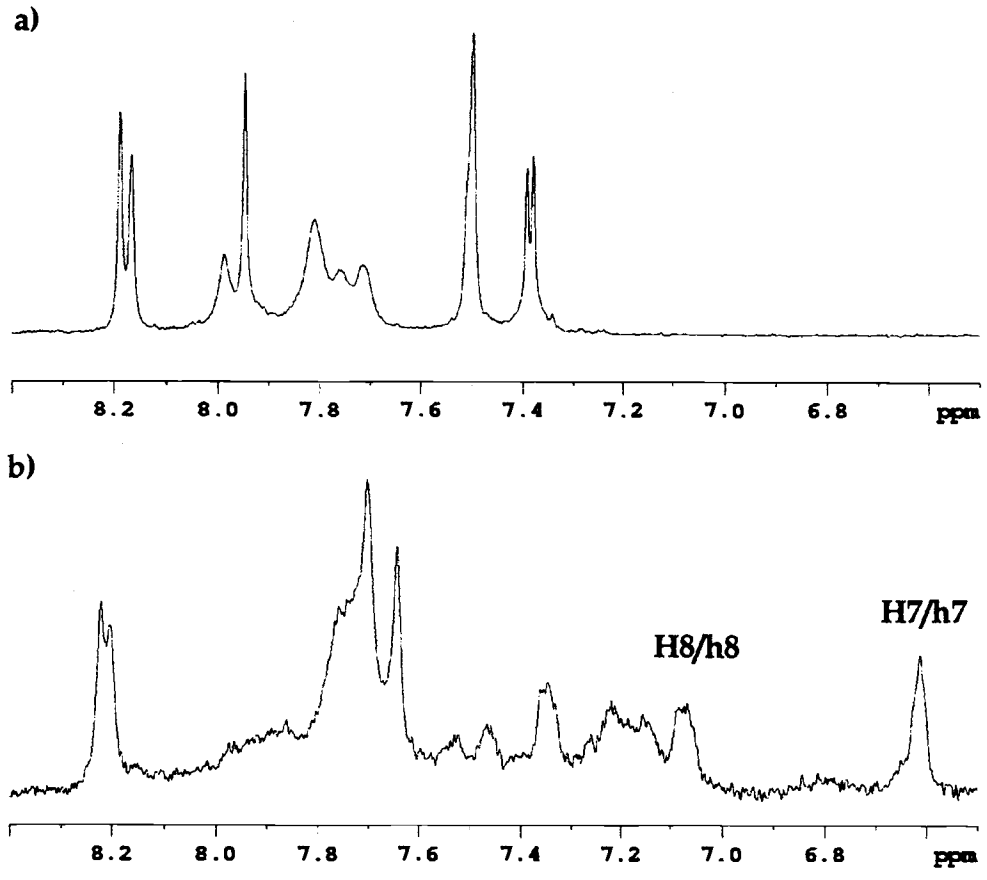

Figure III.9. The 600 MHz ^{1}H 1D NMR spectrum recorded at 298K of the ActD:d(CTGCIG)•d(CCGCAG) complex in $^{2}H_2O$. a) The aromatic proton region of free DNA in 10 mM NaPi, pH 7, 10 mM NaCl. b) The aromatic proton region of the DNA complexed with ActD. Due to linebroadening, the major and minor orientation can not be distinguished from each other in this spectrum.

resonances of this sample and the DNA proton resonances of the DNA are listed in Table A3.2 of Appendix 3.

The ActD drug in both complexes was only partly assigned for identification of orientation within the duplexes. In both complexes, the phenoxazone chromophore protons are identified in a straightforward manner. The 6-CH_3 resonances are identified through their interaction with the H7 and H8 protons, while the 4-CH_3 resonances are found via their interaction with the 6-CH_3 protons. The Thr γ-methyl protons of the benzenoid cyclic pentapeptide ring are identified via their interaction with the chromophore's H8 proton. This proton resonance can then be used to identify all the other proton resonances of the Thr spin system, though only the above mentioned ActD proton resonances were identified in the modified base complexes and are listed in Tables A3.3 and A3.4 of Appendix 3.

III.4.4 Resonance assignments of the DNA in the ActD:d(CTGCGG)•d(CCGCAG) complex

For ActD:d(CTGCGG)•d(CCGCAG), the presence of an A:T base pair facilitated resonance assignments, thus the major and minor orientations of binding were readily identified. The major complex has the benzenoid side of the chromophore located at the GC step between G3 and C4, while the minor complex has the benzenoid side of the chromophore located between G9 and C10 on the complementary strand. The chemical shifts of the corresponding protons in each complex are very similar due to the nearly symmetric orientation between the major and the minor complex (Table III.8). The chemical shift values reported in this table are referenced to the solvent

Table III.8 ^{1}H and ^{31}P assignments of d(CTGCGG)•d(CCGCAG) in complex with ActD reported in ppm.

nucleotide	complex	H8/H6	H5/H2/Me	H1'	H2'/H2''	H3'	H4'	aminos	iminos	^{31}P
C1	MAJOR	7.65	5.86	6.01	1.88/2.36	4.62	4.07			-3.75
	minor	7.68	5.85	6	1.74/2.27	4.61	3.95			-4.08
T2	MAJOR	7.15	1.75	5.64	1.52/1.76	4.72	3.96			-4.10
	minor	7.08	1.81	5.65	1.31/1.61	4.69	3.79			-3.94
G3	MAJOR	7.63		5.64	2.61/2.74	4.94	3.85	7.21/7.69	12.62	-2.30
	minor	7.69		5.64		4.71	3.96			-1.49
C4	MAJOR	7.25	5.88	6.02	1.27/1.75	4.4	4.16	6.80/7.90		-4.03
	minor	7.24	5.74	5.9	1.74/2.07	4.51	3.73	6.78/7.99		-3.91
G5	MAJOR	7.80		5.27	2.59/2.46	4.53	4.23	7.22/7.54	12.22	-3.47
	minor	7.83		5.37	2.46/2.54	4.87	3.88			-3.90
G6	MAJOR	7.84		5.37	2.54/2.44	4.88	3.97			
	minor	7.75		6.05	2.57/2.33	4.62	3.97			
C7	MAJOR	7.65	5.87	6.01	1.88/2.38	4.61	4.07			-3.75
	minor	7.64	5.87	6	1.87/2.27	4.6	4.07			-3.76
C8	MAJOR	7.18	5.75	5.65	1.33/1.71	4.68	3.94	6.81/7.85		-4.00
	minor	7.14	5.75	5.63	1.28/1.75	4.71	3.95			-4.04
G9	MAJOR	7.66		5.63	2.54/2.66	4.9	3.89	7.32/8.02	12.39	-1.22
	minor	7.62		5.62	2.59/2.71	4.92	3.85	7.34/7.71	12.53	?
C10	MAJOR	7.40	5.8	5.46	1.93/2.14	4.53	4.02	6.64/8.02		-3.22
	minor	7.34	5.96	5.86	1.36/1.72	4.41	3.72			-3.77
A11	MAJOR	8.18		5.63	2.66	4.95	4.19			-3.84
	minor	8.18		5.78	2.65	4.91	4.18			-3.88
G12	MAJOR	7.69		5.65	2.68/2.58	4.9	3.97			
	minor	7.93			(2.63/2.55)	4.63	4.06			

resonance signal at 308 K. Because the DNA unwinds at the intercalation step, it is not possible to sequentially assign the entire DNA sequence using NOEs, since the distance between G3 and C4 as well as G9 and C10 becomes greater than 6Å. However, the observed NOEs between the DNA nucleotides and the phenoxazone chromophore can be used to confirm the DNA assignments at the intercalation site. Figure III.10 shows the H2′, H2″ to base proton connectivities of ActD:d(CTGCGG)•d(CCGCAG), indicating how the strand which has the intercalated benzenoid side of the phenoxazone chromophore is identified. A standard sequential assignment is possible from nucleotides C1 through G3, from C4 through G6, as well as from C7 through G9 and C10 through G12.

All the nonexchangeable protons of the DNA in the major complex were easily identified using this strategy and are tabulated in Table III.5. Identifying the resonances of the minor complex proved to be more challenging because of the weaker crosspeak intensities, which result from the low 33% abundance. The resonances of the base proton of the terminal residues in the minor complex were not observed and are likely to be hidden under stronger peaks. This assumption is valid because all the other crosspeaks in this region have been assigned and all other minor complex resonances were identified. For some residues (A11) the base proton of the major and minor complex overlap, thus it is impossible to distinguish the H2′, H2″ resonances, yet connectivites to the H2′, H2″ protons of the previous residue (C10) are easily identified. The H1′ protons of the minor complex were difficult to identify, because of severe overlap around 5.63 ppm.

The exchangeable DNA protons were found upon exchanging the sample into H_2O and widening the spectral width to 25 ppm. The guanine imino

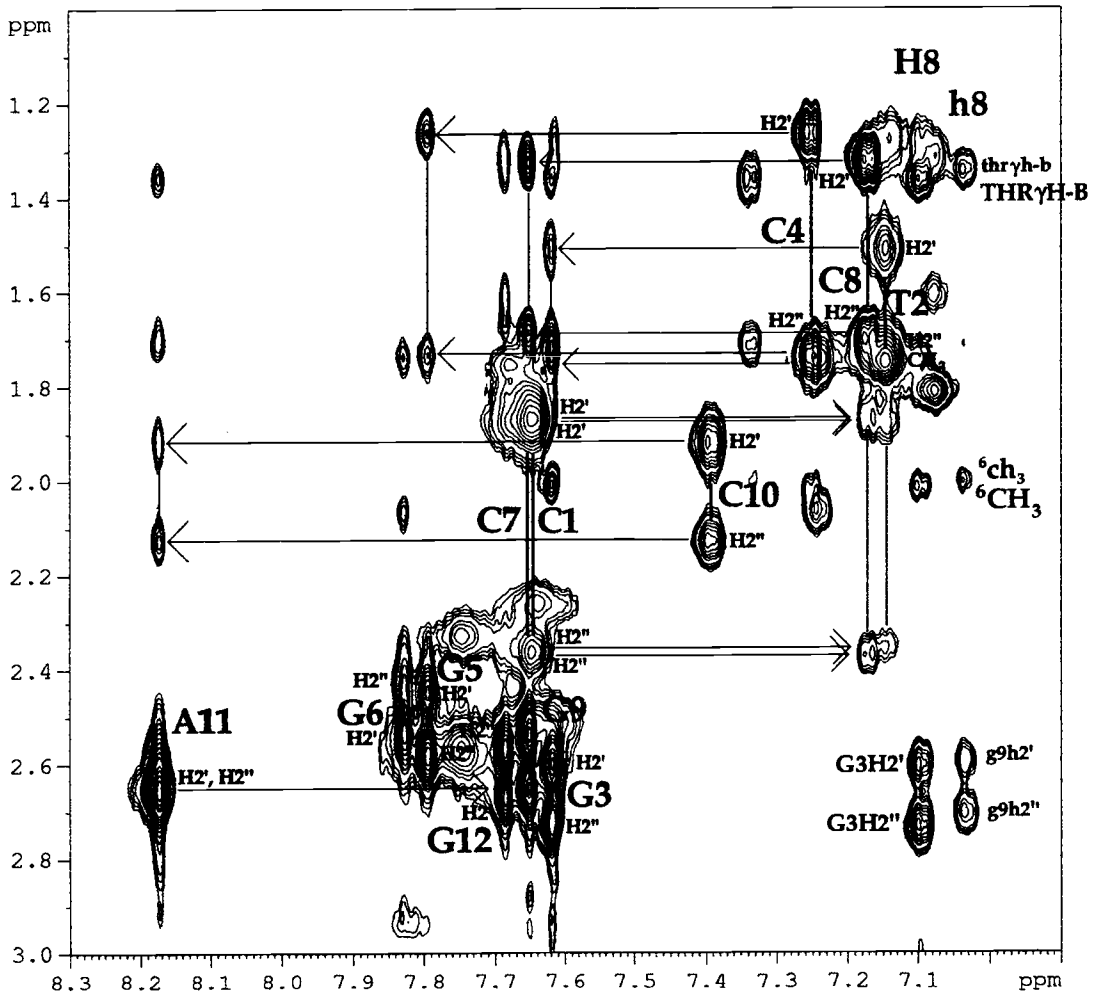

Figure III.10 The 600 MHz ^{1}H NOESY spectrum of ActD:d(CTGCGG)•d(CCGCAG) (1:1) in $^{2}H_2O$, 35°C, pH7, 10mM NaCl displaying the base proton to H2′, H2″ proton region. From this region, one can determine which strand the benzenoid side of the chromophore intercalates, because of the H8 to benzenoid Guanine H2′, H2″ NOEs. The major complex has its B-side between G3 and C4, while the minor complex has its b-side between g9 and c10.

protons resonate between 12 ppm and 13 ppm, while the guanine and cytosine amino protons resonate between 6 ppm and 8 ppm. The cytosine amino protons are identified via a crosspeak between CH5 and the non-hydrogen bonded amino proton. This latter amino proton has a strong crosspeak with the hydrogen-bonded amino proton, which in turn has a crosspeak across the base pair to the guanine imino proton. In fact, both the cytosine amino protons exhibit crosspeaks to the guanine imino proton, and by connecting these back to the appropriate cytosine H5 proton, one can assign the imino resonances to a particular guanine. For the complex, the guanine amino protons are identified as well, since both show an NOE crosspeak to the imino proton, unlike for the free DNA hexamers. Usually the exocyclic guanine amino protons are exchange broadened by rotation about the C-N bond, and these crosspeaks are not observed (Rajagopal *et al.*, 1988). However, since in the ActD-DNA complex the rotation of these amino protons located in the minor groove seems to be hindered by hydrogen bonding with ActD (Brown *et al.*, 1994), they appear in the spectrum and can be seen in Figure III.11.

III.4.5 Resonance assignments of ActD in the complex

The ActD molecule (Figure III.1) consists of two cyclic pentapeptide lactone rings which are covalently connected to the benzenoid and quinoid side of the phenoxazone chromophore. Each pentapeptide moiety consists of a L-threonine (Thr)-D-valine (DVal)-L-proline (Pro)-sarcosine (Sar)-L-N-methylvaline (NMV) sequence with an ester linkage between γ-O of L-Thr and CO of NMV, and each pentapeptide ring is designated either benzenoid and quinoid depending on which side of the chromophore it is linked to. Because of

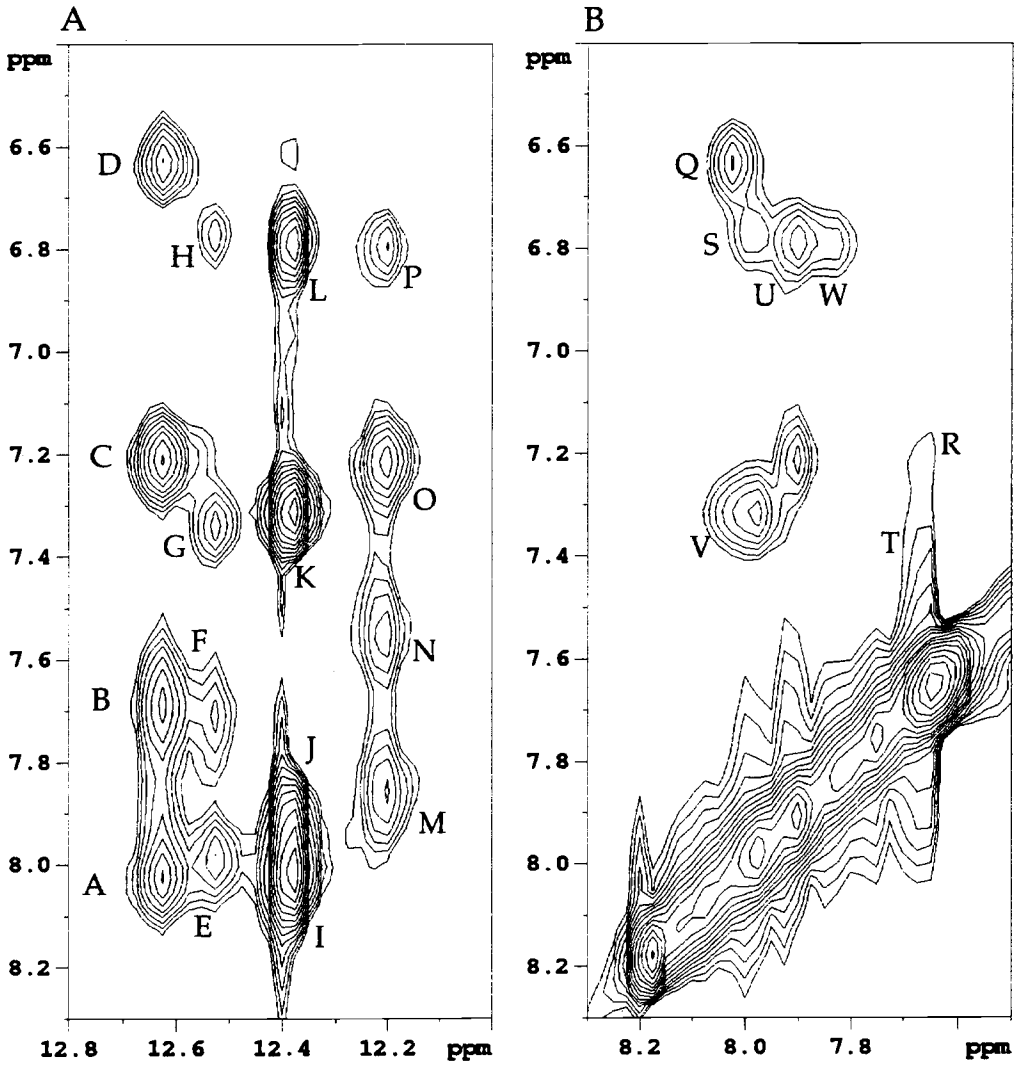

Figure III.11 The imino to amino region of the 600 MHz ^{1}H NOESY spectrum (mixing time 300 ms) of the ActD:d(CTGCGG)•d(CCGCAG) (1:1) complex in H_2O, 15°C, pH7, 10mM NaCl. correlating a) distance connectivities between guanosine imino protons, 12ppm to 12.8ppm, and amino protons, 6.4ppm to 8.1ppm, and b) distance connectivities between 7.5ppm to 8.3ppm aminos and 6.4ppm to 8.1ppm aminos. Crosspeaks **A** through **P** are designated as follows: **A**. G3H1 to C10H42, **B**. G3H1 to G3H22, **C**. G3H1 to G3H21, **D**. G3H1 to C10H41, **E**. g9h1 to c4h42, **F**. g9h1 to g9h22, **G**. g9h1 to g9h21, **H**. g9h1 to c4h41, **I**. G9H1 to G9H22, **J**. G9H1 to C4H42, **K**. G9H1 to G9H21, **L**. G9H1 to C4H41, **M**. G5H1 to C8H42, **N**. G5H1 to G5H22, **O**. G5H1 to G5H21, **P**. G5H1 to C8H41. The crosspeaks **Q** through **Y** are designated as follows: **Q**. C10H42 to C10H41, **R**. G3H22 to G3H21, **S**. c4h42 to c4h41, **T**. g9h22 to g9h21, **U**. C4H42 to C4H41, **V**. G9H22 to G9H21, **W**. C8H42 to C8H41.

the presence of a major and minor orientation of binding, two sets of proton resonances are observed for ActD in the complex, one set exhibiting a stronger intensity than the other. The intensity ratios of the major and minor complex correspond to the intensity ratios of the major and minor complex observed in the DNA crosspeaks. Resonance assignments are complicated by the fact that the cyclic pentapeptide lactone rings connected to the benzenoid and quinoid sides of the phenoxazone chromophore each exhibit nearly identical chemical shifts (see Table III.9). The chemical shift values reported in this table are referenced to the solvent resonance signal at 308 K.

The aromatic 6-CH_3, H7 and H8 protons of the chromophore are easily assigned based on the presence of crosspeaks between the 6-CH_3 and H7 protons, and between the H7 and H8 protons in the TOCSY and NOESY spectra. The 4-CH_3 group was assigned by its NOE crosspeak to the 6-CH_3 group in the complex and each amino acid spin system was identified by analysis of the TOCSY spectra of the complex. Sequential connectivities within the individual cyclic pentapeptide lactone rings were identified from the NOE crosspeak patterns in the NOESY spectrum of the complex.

Differences between the chemical shifts of the chromophore H8/H7 protons in the two complexes (Figure III.12) make it possible to distinguish the two complexes formed with d(CTGCGG)•d(CCGCAG). The major complex exhibits several correlations between the chromophore H8 proton (7.10 ppm) and the G3/C4 residues of the DNA, while the minor complex exhibits connectivities between the chromophore H8 proton (7.03 ppm) and the G9/C10 residues of the DNA. The benzenoid and quinoid pentapeptide lactone rings can be assigned via the Thr methyl groups on the benzenoid and quinoid side,

Table III.9 Actinomycin D ^{1}H assignments for the free drug and the major and minor complexes.

residue	complex	Hα	Hβ	Hγ	Hδ	N-CH_3/NH
Threonine	free (B)	4.81	5.31	1.34		
	free (Q)	4.72	5.28	1.35		
	MAJOR (B)	4.87	5.17	1.36		7.91
	MAJOR (Q)	4.60	5.10	1.33		7.99
	minor (B)	4.71	5.13	1.34		7.84
	minor (Q)	4.70	5.10	1.31		7.93
D-Valine	free (B)	3.60	2.13	1.06/0.83		8.12
	free (Q)	3.62	2.14	1.06/0.83		8.07
	MAJOR (B)	3.52	2.10	1.04/0.81		8.07
	MAJOR (Q)	3.55	2.11	1.03/0.80		8.13
	minor (B)	3.50	2.09	1.01/0.78		8.14
	minor (Q)	3.54	2.10	1.02/0.78		8.06
Proline	free (B)	6.15	2.57/1.99	2.06	3.92/3.70	
	free (Q)	6.18	2.53/2.00	2.06	3.83/3.72	
	MAJOR (B)	6.17	2.88/1.70	2.06	3.95/3.89	
	MAJOR (Q)	6.23	3.02/1.80	2.03	3.97/3.87	
	minor (B)	6.14	2.89/1.69	2.07		
	minor (Q)	6.20	2.93/1.81	1.99		
Sarcosine	free (B)	4.74/4.21				2.86
	free (Q)	4.73/4.19				2.85
	MAJOR (B)	4.57/4.17				2.92
	MAJOR (Q)	4.58/4.13				2.85
	minor (B)	4.48/3.87				2.75
	minor (Q)	4.55/3.92				2.78
Me-Valine	free (B)	3.30	2.47			3.02
	free (Q)	3.28	2.47			2.99
	MAJOR (B)	2.99	2.44	0.91/0.80		2.88
	MAJOR (Q)	2.96	2.43	0.90/0.79		2.88
	minor (B)	3.05	2.39	0.88/0.76		2.97
	minor (Q)	3.05	2.41	0.86/0.73		2.97
		H7	H8	$4CH_3$	$6CH_3$	
Chromo-phore	free	7.41	7.44	1.64	2.45	
	MAJOR	6.61	7.10	1.68	2.02	
	minor	6.59	7.03	1.69	2.01	

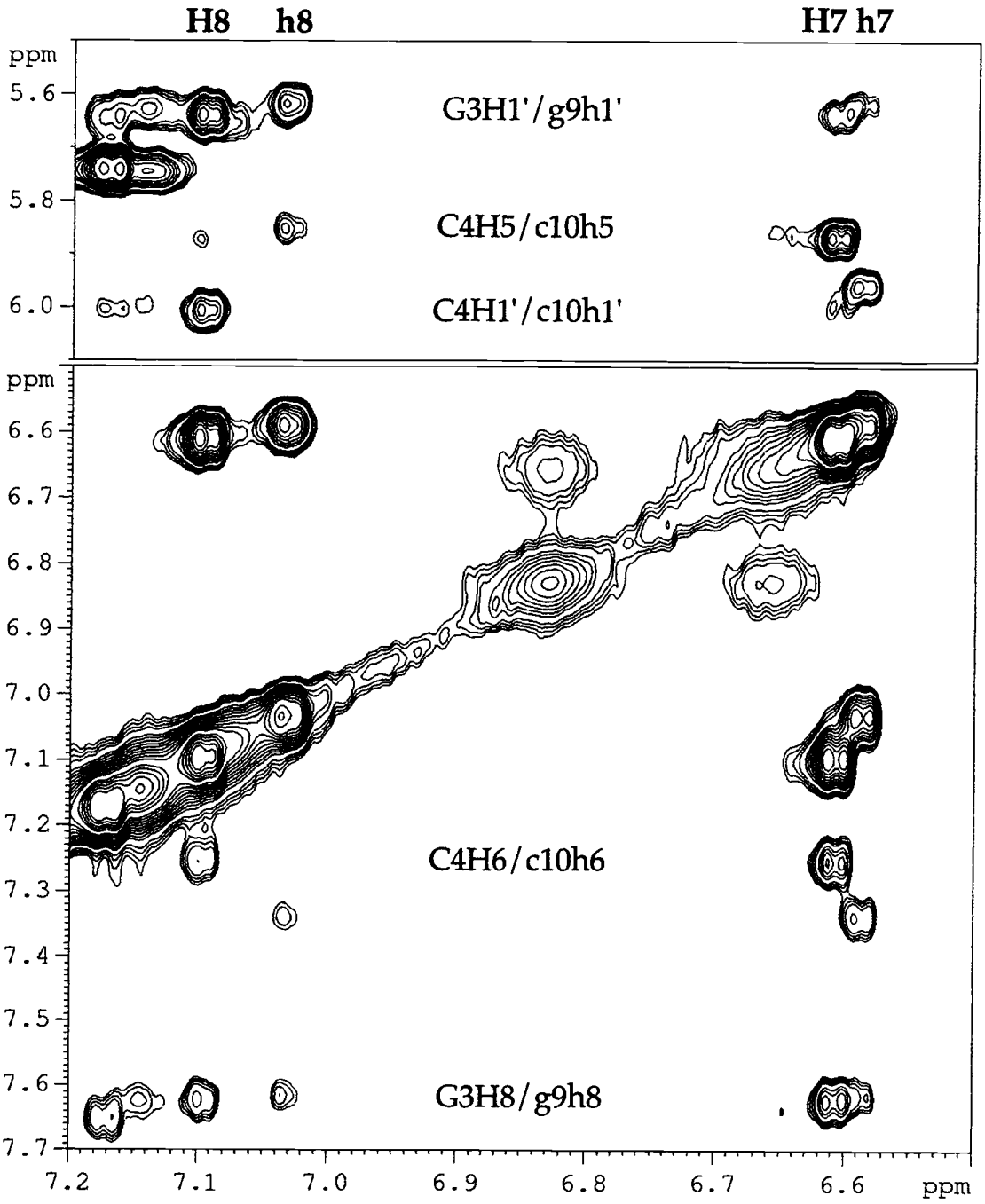

Figure III.12a (legend on next page)

b)

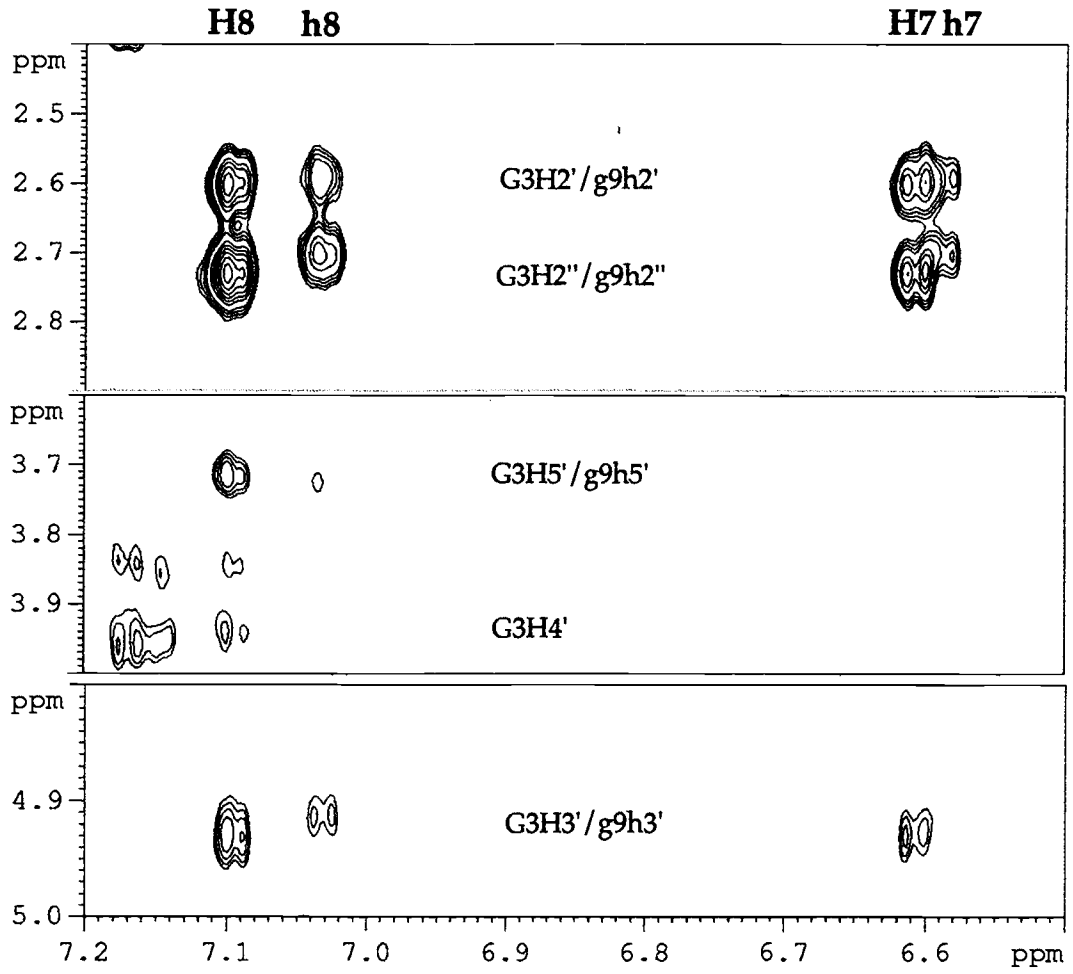

Figure III.12 Regions of the 600 MHz ^{1}H NOESY spectrum of ActD:d(CTGCGG)•d(CCGCAG) (1:1) in ^2H_2O, 15°C, pH7, 10mM NaCl, displaying ActD to DNA contacts between the phenoxazone chromophore protons (H8/h8 and H7/h7) and protons on the nucleotides at the intercalation site. a) NOEs to the G3, C4 base protons and H1′ protons of the major complex and to the g9 and c10 base protons and h1′ protons of the minor complex are shown. b) For the major complex NOEs to the G3 H2′, H2″, H3′, H4′ and H5′ protons are observed while for the minor complex NOEs to the g9 h2′, h2″, h3′, h4′ and h5′ are shown.

which are distinguished by observing the NOE between the chromophore's H8 and H7 protons and the methyl group of Thr which occurs on the benzenoid ring only. Thus the two Thr γ-methyl resonances are distinguished from each other and are also used to identify the Thr Hα chemical shifts and the Thr Hβ chemical shifts of each pentapeptide ring. Other NOEs exist between the Thr γ-methyl groups and other protons of the peptide chain. Thus the DVal γ-methyl groups and the NMV γ-methyl groups as well as the NMV Hα and Hβ protons are identified. Additional NOEs are observed across the benzenoid and quinoid pentapeptide lactone rings, from the benzenoid Thr γ-methyl to the quinoid Pro Hδ, Hγ and Hα, and vice versa. In this manner all the amino acid spin systems of both the benzenoid and quinoid pentapeptide rings are identified. This procedure is described in detail in Appendix 4 as well as in the hairpin structure publication (Brown *et al.*, 1994).

III.4.6 Intermolecular interactions

All of the 89 intermolecular interactions for the major complex and 90 intermolecular interactions for the minor complex between ActD and d(CTGCGG)•d(CCGCAG) listed in Tables III.10 and III.11 have been identified in the NOESY spectrum in ^2H_2O. The most prominent and easily distinguished crosspeaks between ActD and the DNA helix are the crosspeaks between the H8 and H7 protons on the phenoxazone chromophore and the DNA nucleotide protons at which the benzenoid side of the chromophore is intercalated. These NOEs are shown in Figure III.12a, where crosspeaks between H8/H7 and G3H8, C4H6 and C4H5 are observed. Crosspeaks from the chromophore H8/H7 protons to G3H1', C4H1' (Figure III.12a), G3H3', G3H4' and G3H2'/H2''

Table III.10 Intermolecular contacts in the major complex of ActD:d(CTGCGG)•d(CCGCAG).

phenoxazone protons	DNA protons
4-CH_3	G9H8, H2', H1, C10H6, H5, G3H1
6-CH_3	G3H8, H2', H2", H1, C4H6, H5, H1', G9H1
H8	G3H8, H2', H2", H1', H3', C4H6, H5, H1', H3'
H7	G3H8, H2', H2", H1', H3', C4H6, H5, H1', H3'
benzenoid pp protons	DNA protons
Thr ($CH_3\gamma$)	G3H8, H1', H2', C4H1', H4'
Thr (NH)	G3H1', H21
Pro (Hα)	G3H1
Pro (HβA)	C10H1', A11H1', H4'
Pro (HβB)	A11H1', H4'
Pro (Hγ)	A11H4'
Sar (HαA)	A11H4'
Sar (HαB)	A11H1'
Sar (NCH_3)	A11H1', H4'
NMV (Hα)	G3H1', H3', H4'
NMV ($CH_3\gamma$A)	T2H1', H4', G3H1', H4'
NMV ($CH_3\gamma$B)	T2H1', H4', G3H1', H4'
NMV (NCH_3)	T2H1', G3H1', H4'
quinoid pp protons	DNA protons
Thr (Hα)	G9H8, H1', H2", H3', H4'
Thr (Hβ)	G9H4'
Thr ($CH_3\gamma$)	G9H2"
Thr (NH)	G9H1', H21
Pro (Hα)	G9H1
Pro (HβA)	C4H1', G5H1', H4'
Pro (HβB)	G5H1', H4'
Pro (Hγ)	C4H2', G5H4'
Sar (HαA)	G5H4'
Sar (NCH_3)	G5H1', H4'
NMV (Hα)	G9H1', H3', H4'
NMV ($CH_3\gamma$A)	G9H1', H4', C10H4'
NMV ($CH_3\gamma$B)	C8H1', G9H1', H4', C10H4'
NMV (NCH_3)	C8H1', G9H1', H3', H4'

Table III.11. Intermolecular contacts in the minor complex of ActD:d(CTGCGG)•d(CCGCAG).

phenoxazone protons	DNA protons
4-CH_3	G3H8, H2', H2", C10H6, H5, G9H1
6-CH_3	G9H8, H2', H2", H1, C10H6, H5, H1'
H8	G9H8, H2', H2", H1', H3', C10H6, H5, H1', H3'
H7	G9H8, H2', H2", H1', H3', C10H6, H5, H1', H3'
benzenoid pp protons	DNA protons
Thr (Hα)	G9H1'
Thr (CH_3γ)	G9H8, H2', H2", H1', C10H1', H4'
Thr (NH)	G3H1'
Pro (Hα)	G5H4'
Pro (HβA)	G5H4'
Pro (HβB)	G5H4'
Pro (Hγ)	G5H4'
Sar (HαA)	G5H4'
Sar (HαB)	G5H1'
Sar (NCH_3)	G5H2', H2", H4'
NMV (Hα)	G9H1', H3', H4'
NMV (CH_3γA)	C8H1', H4', G9H1', H4',
NMV (CH_3γB)	C8H1', H4', G9H1', H4',
NMV (NCH_3)	C8H1', G9H1', H3', H4'
quinoid pp protons	DNA protons
Thr (Hα)	G3H1'
Thr (Hβ)	G3H4'
Thr (CH_3γ)	G3H8, H1', H3', H4'
Pro (Hα)	C10H1', A11H4'
Pro (HβA)	C10H1', A11H1', H4'
Pro (HβB)	A11H1', H4'
Pro (Hγ)	C10H2', A11H4'
Sar (Hα)	A11H4'
Sar (NCH_3)	A11H1', H4'
NMV (Hα)	G3H1', H4'
NMV (CH_3γA)	G3H1', H4', C4H4'
NMV (CH_3γB)	T2H1', G3H1', H4', C4H4'
NMV (NCH_3)	T2H1', G3H1', H3', H4'

protons are also observed (Figure III.12b). The same crosspeaks are observed for the minor complex (protons designated with lower case letters), however, in this complex the corresponding interactions are between h8/h7 and g9h8, c10h6, and c10h5. The crosspeaks between h8/h7 and the sugar protons of these nucleotides are observed as well. Crosspeaks between the chromophore's 6-CH_3 protons and G3H8 as well as C4H6 and C4H5 encompass some of the intermolecular NOEs. The related NOEs are observed for the minor complex. Because similar intermolecular NOE patterns are observed between the drug chromophore's protons and the DNA protons for both major and minor complexes, it can be concluded that the conformation at the binding site of the minor complex is very similar to that of the major complex. Most contacts between the cyclic pentapeptide lactone chains and the DNA are concentrated in the G3/C4 and G9/C10 nucleotides on the DNA helix. While T2/G3/C4 of the DNA molecule exhibit most of their intermolecular crosspeaks to the Thr and NMV on the benzenoid pentapeptide chain, C8, G9 and C10 exhibit most of their intermolecular crosspeaks to the Thr and NMV on the quinoid pentapeptide chain. For the Hβ protons of Pro the situation is reversed. The quinoid Pro Hβ is close to the C4H1′ proton while the benzenoid Pro Hβ proton is close to C10H1′. Other intermolecular contacts include NOEs from the benzenoid Thr γ-methyl protons to the G3H8 proton, the G3H1′ proton, the G3H2′, H2″ protons and the C4H1′ and H4′ protons. Numerous NOEs exist between the NMV protons in both cyclic pentapeptide lactone rings and DNA protons (Table III.10). For the minor complex, the pattern of NOEs implies the opposite orientation of binding in that all of the contacts observed in the major complex between the drug and G3/C4 are now observed between the drug and

g9/c10 (Table III.11). The same is true for the intermolecular contacts to the opposite strand.

The 10 exchangeable proton NOEs of the major complex of ActD:d(CTGCGG)•d(CCGCAG) are detected in the H_2O NOESY spectrum and include the imino and amino protons of the nucleotides as well as the peptide amide protons of Thr and DVal. The exchangeable protons of ActD exhibit several NOEs to the DNA helix. The benzenoid Thr NH proton shows a connectivity to G3H1′ as well as to the G3 exocyclic amino proton. Meanwhile the quinoid Thr NH proton shows a connectivity to G9H1′ and to the G9 exocyclic amino proton. These NOEs are particularly strong since the Thr NH is involved in a hydrogen bond to the guanine N3 groups at the intercalation site. Both of the methyl groups on the phenoxazone chromophore exhibit NOE crosspeaks to the imino protons of G3 and G9, which confirms the formation of two complexes differing by the orientation of the drug with respect to the DNA. All the intermolecular NOEs for the major and minor complex are listed in Tables III.10 and III.11, respectively.

III.4.7 ^{31}P NMR

Numerous previous investigations have shown that the chemical shift of the phosphorous nuclei in the DNA backbone linking the nucleotides forming the d(GC) step shift downfield upon chromophore intercalation (Brown *et al.*, 1994; Gorenstein *et al.*, 1984; Patel, 1976; Petersheim *et al.*, 1984; Wilson *et al.*, 1986a). This shift can be observed in Figure III.13, the ^{31}P 1D spectrum of free DNA and DNA in complex with ActD. In the latter spectrum the shifting of at least three phosporous signals to a lower field is clearly observed. A ^{1}H-^{31}P

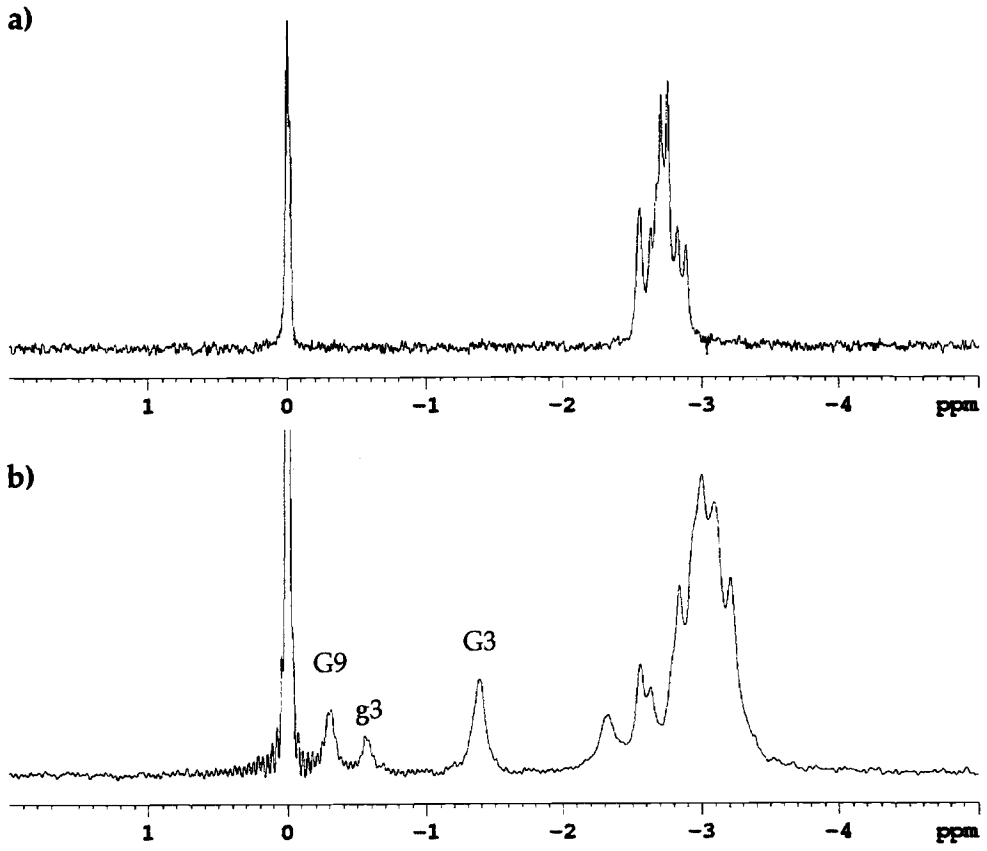

Figure III.13 The 242 MHz 1D ^{31}P spectra of a) d(CTGCGG)•d(CCGCAG) in H_2O, 25°C, pH 7, 10 mM NaCl, and b) ActD:d(CTGCGG)•d(CCGCAG) (1:1) in ^2H_2O, 25°C, pH7, 10mM NaCl. The downfield shifts of the phosphate groups at which intercalation occurs is illustrated. It is also apparent from these resonances that the two complexes formed are present in unequal concentrations.

HSQC spectrum was recorded (Figure III.14) and was used to assign these shifted peaks, which were found to belong to the G3 phosphodiester group and the G9 phosphodiester group. In Table III.8 it can be seen that the phosphorous nuclei on the quinoid intercalation site appear at lower field than do those on the benzenoid intercalation site. This method of confirmation of proton resonance assignments in the DNA molecule is valuable, making the ^{1}H-^{31}P HSQC experiment a powerful technique in DNA solution structure determination.

III.4.8. Structure calculation analysis of ActD:d(CTGCGG)•d(CCGCAG)

For each complex, a total of 100 structures were calculated, out of which the 10 best structures were selected as described in the Materials & Methods section. The final R-factors for each complex are presented in Tables III.12 and III.13 and are separated into different categories, ones involving DNA interactions, ones involving DNA and intermolecular interactions, and ones involving DNA, intermolecular and selected intradrug interactions. The different R-factor categories give an indication of which interactions within the structure contribute most to a large R-factor value. The 10 best calculated structures using tight constraints for the major complex are presented in Figure III.15 and for the minor complex in Figure III.16. The tight constraints are distances defined by the relaxation matrix with a variation of +/- 0.2 Å. Out of a total of 100 structures, the 10 best structures converge to an rms difference of 0.56 Å +/- 0.08 Å for the major complex and 0.89 +/- 0.11 Å for the minor complex.

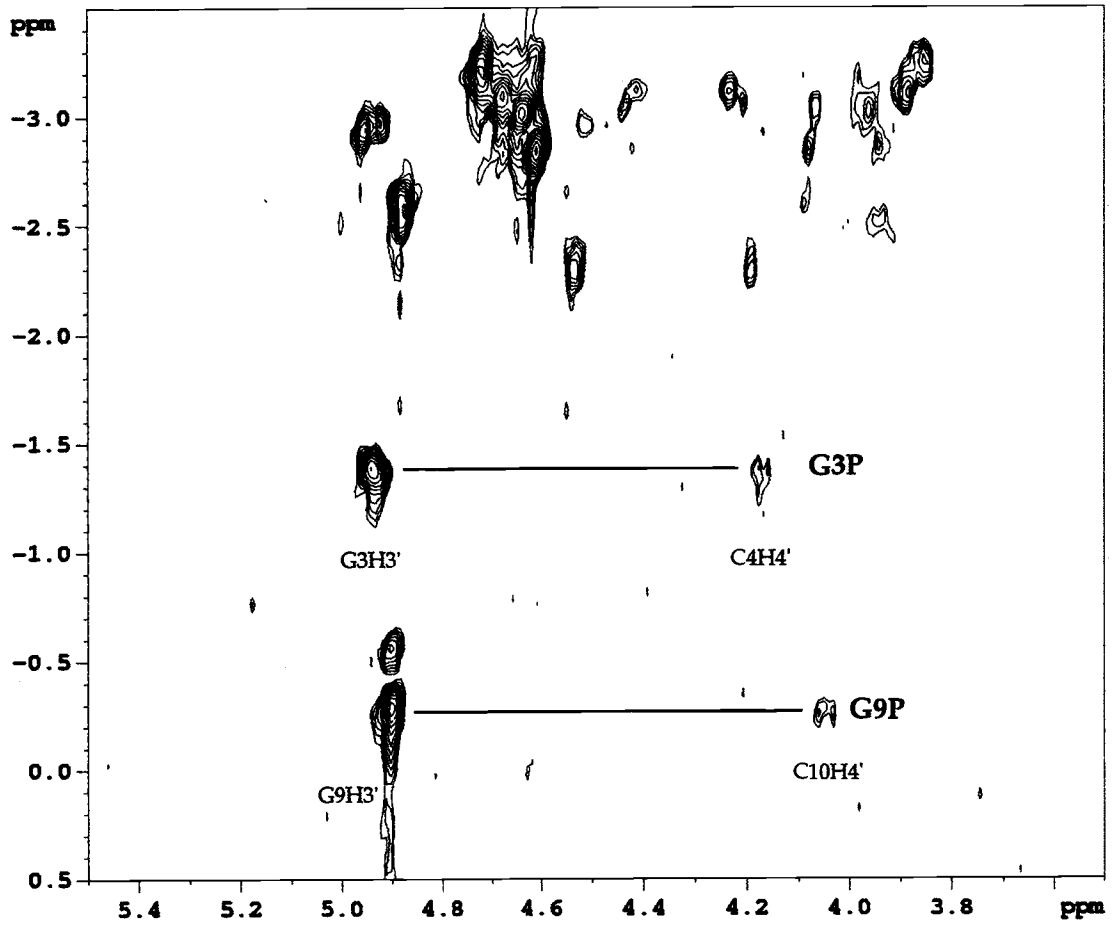

Figure III.14 The ^{1}H-^{31}P HSQC spectrum of the ActD:d(CTGCGG)•d(CCGCAG) (1:1) complex in $^{2}H_2O$, 25°C, pH 7, 10 mM NaCl. This spectrum is used to confirm the ^{1}H resonance assignments of the DNA molecule. This technique is of utmost importance in DNA assignments and structure determination. The downfield shifted resonances at the intercalation site are assigned using this spectrum. Each phosphorous resonance is correlated via scalar coupling to an H3′ proton in the 3′ direction and to the H4′ proton in the 5′ direction.

Table III.12 Residual factors determined for the major complex

R-factors	Major starting	Major - tight bounds	Major-loose bounds
most DNA peaks (59 peaks)	70%	23%	28%
most DNA and DNA to drug peaks (92 peaks)	74%	28%	30%
most DNA, DNA to drug, drug to drug (144 peaks)	71%	37%	39%

Table III.13 Residual factors determined for the minor complex

R-factors	Minor starting	Minor - tight bounds	Minor - loose bounds
most DNA peaks (29 peaks)	104%	44%	41%
most DNA and DNA to drug peaks (53 peaks)	88%	83%	81%
most DNA, DNA to drug, drug to drug (71 peaks)	149%	140%	140%

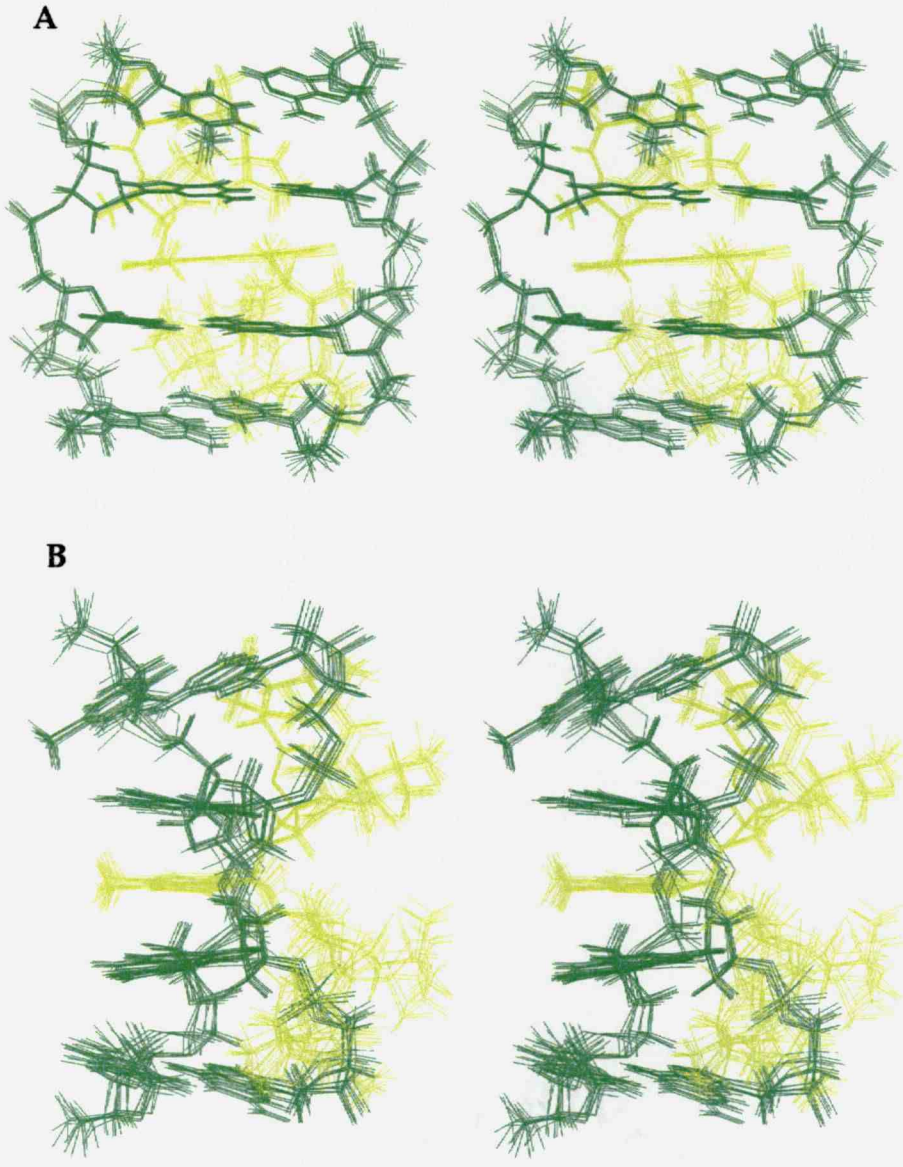

Figure III.15 Stereoview of the 10 superpositioned structures of the central four base pair region of the major complex of ActD:d(CTGCGG)•d(CCGCAG) calculated with tight constraints for the internucleotide interactions. The rms difference for the family of 10 structures is 0.56 Å. a) view into the major groove; b) side view with the minor groove on the right and the major groove on the left. For this structure all the DNA NOESY crosspeaks were treated equally and the distance constraints were refined using the total relaxation matrix approach described in the text.

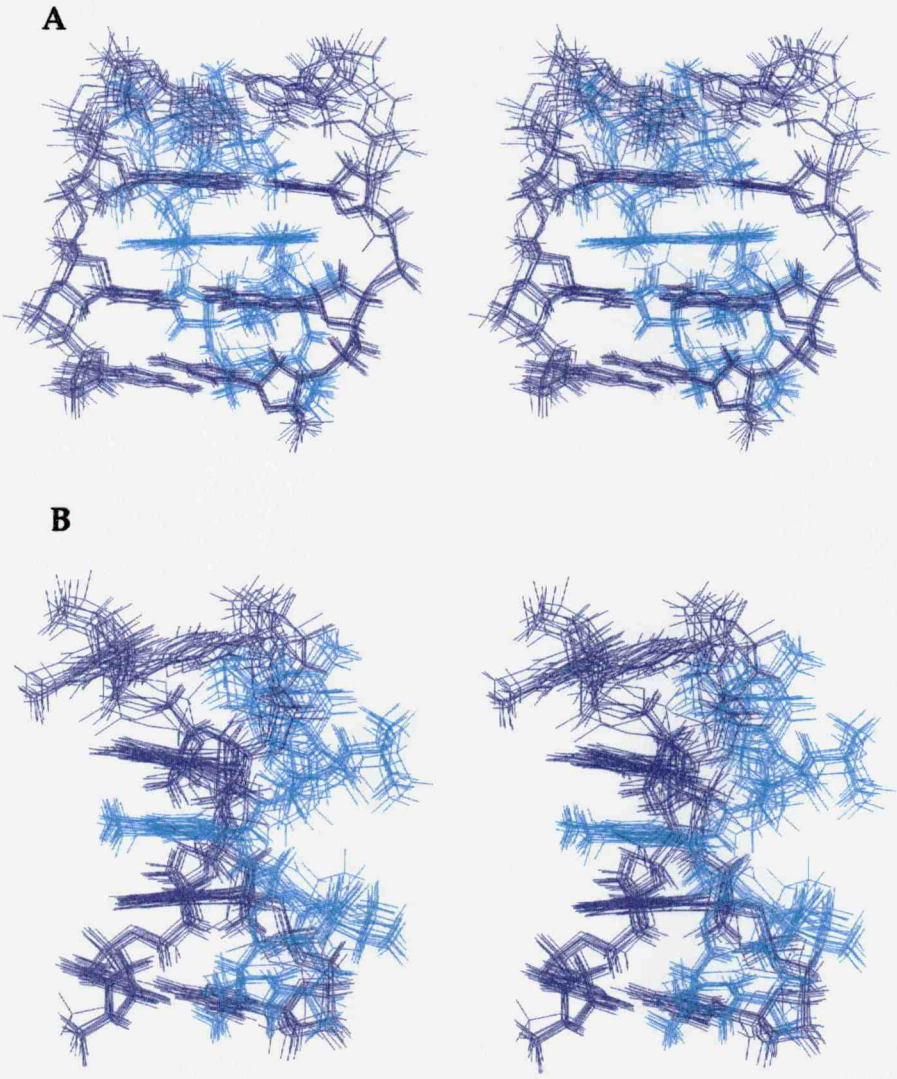

Figure III.16 Stereoview of the 10 superpositioned structures of the central four base pair region of the minor complex of ActD:d(CTGCGG)•d(CCGCAG) calculated using tight constraints for the internucleotide interactions. The rms difference for the family of 10 structures is 0.89 Å. a) view into the major groove; b) side view with the minor groove on the right and the major groove on the left. For this structure all the DNA NOESY crosspeaks were treated equally and the distance constraints were refined using the total relaxation matrix approach described in the text.

A view of the 10 superpositioned structures using loose internucleotide constraints for the major complex is presented in Figure III.17. The 10 superpositioned structures using loose internucleotide constraints for the minor complex are presented in Figure III.18. The ActD chromophore is intercalated into the central $d(GC)_2$ step of the complex. The solution structure is well defined for the central segment of the complex around the intercalation site with a greater spread among the refined structures observed at the ends of the DNA hexamers, reflecting the looser bounds applied for the internucleotide interactions and thus greater conformational heterogeneity. Even though loose bounds were used for all internucleotide interactions, the DNA structures are still restricted to a tight conformational space thus the presence of excessive motion is not evident within these families of structures. With such loose bounds for the internucleotide NOEs, it is somewhat surprising that a broader range of structures do not result from the simulated annealing calculations. One possible explanation is that the DNA to drug NOEs are restricting the DNA molecule from exploring all possibilities of conformational sampling. In order to test this, calculations were repeated in the absence of these intermolecular NOE restraints resulting in a family of 5 DNA structures with a rms difference of 2.4 Å. This result confirms that intermolecular NOEs restrict conformational sampling of the DNA. The distribution of pairwise rms difference values over all atoms for the 10 refined structures including all NOEs is 1.02 Å +/- 0.07 Å for the major complex and 1.30 Å +/- 0.12 Å for the minor complex. In the major complex structures an average of 20 NOEs were violated out of a total of 179 NOEs, while an average of 18 NOEs out of a total of 179 NOEs were violated for the minor structure. An NOE is not considered violated unless the interproton

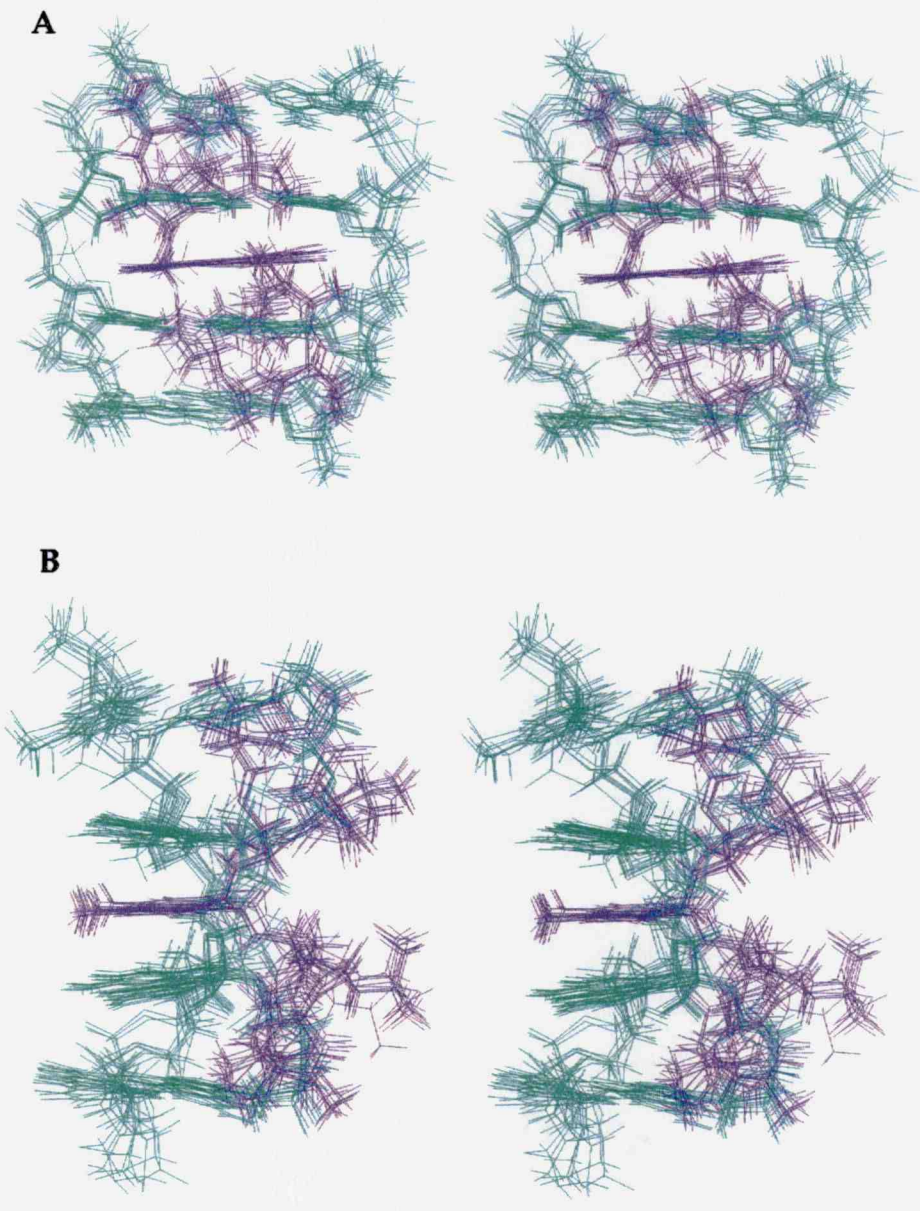

Figure III.17 Stereoview of the 10 superpositioned structures of the central four base pair region of the major complex calculated using loose restraints are displayed with an rms difference over all atoms of 1.02 Å. The DNA is in green color while the ActD drug is in violet. a) A view into the major groove is shown and in b) the side view is presented with the cyclic pentapeptide rings in the minor groove and the chromophore stacking with the base pairs.

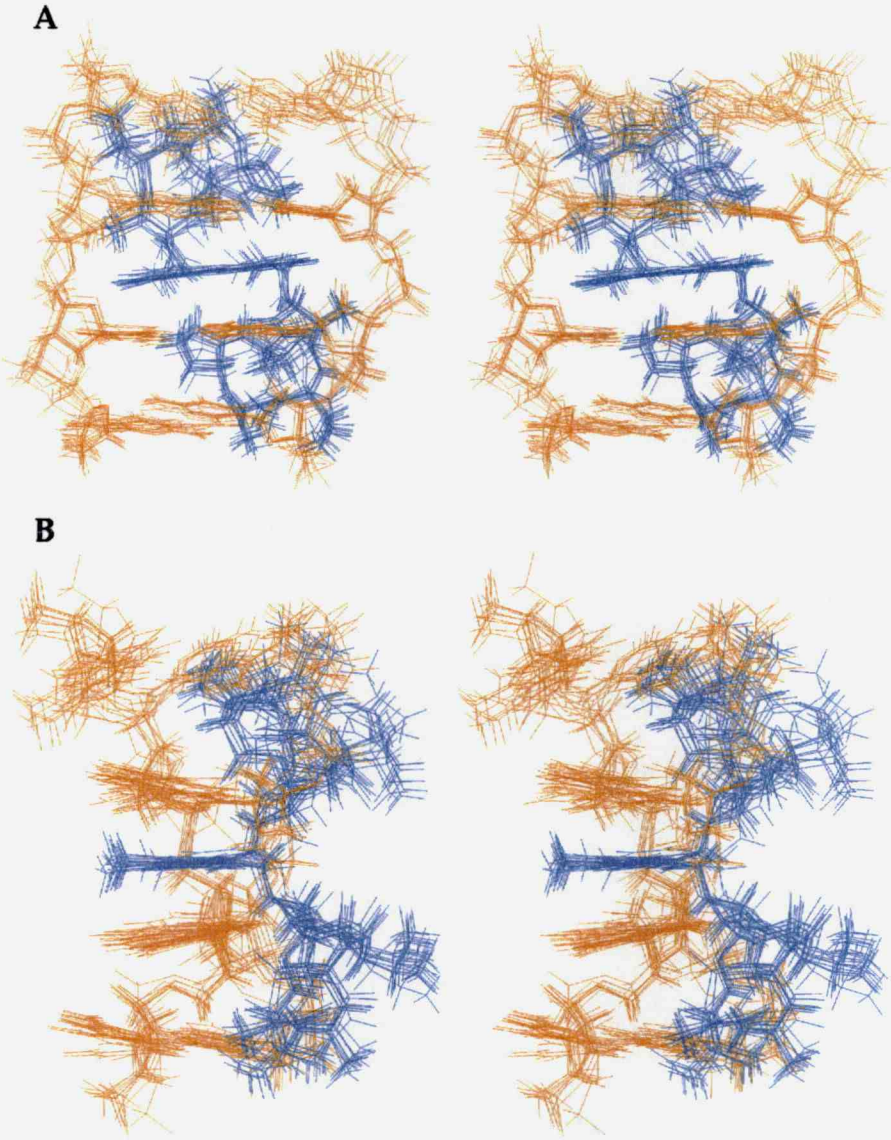

Figure III.18 Stereoview of the 10 superpositioned structures of the central four base pair region for the minor complex calculated using loose internucleotide restraints are displayed with an rms difference over all atoms of 1.30 Å. The DNA is in orange color while the ActD drug is in blue. a) A view into the major groove is shown and in b) the side view is presented with the cyclic pentapeptide rings in the minor groove and the chromophore stacking with the base pairs.

distance in the calculated structure is more than 0.2 Å out of the distance constraint. The views shown in Figures III.15-18 of the complexes highlight the intermolecular contacts between the DNA molecule and the ActD drug.

III.5 Discussion

III.5.1 The influence of non-self-complementary flanking base pairs on the orientation of Actinomycin D binding

The results from the titration studies of ActD binding to non-self-complementary sequences show that ActD binds in two distinct orientations and that it prefers one orientation above the other. The specific orientation of the ActD drug was identified for seven out of eight sequences (Table III.6) and some trends for preferential binding were identified. Unfortunately, for the second sequence in the table the preferred strand was not identified because of unresolved crosspeaks and significant overlap, making the sequential NOEs of the DNA strand impossible to resolve. Of the remaining seven sequences, two are variations of an existing sequence, thus this discussion will revolve around the orientation determined for the five original duplexes, listed in the first six rows of Table III.6. The most obvious observation is that the benzenoid side of the chromophore is found to prefer strands in which a guanine is adjacent to the intercalation site d(GC), and the quinoid side prefers strands incorporating cytosines. In the three sequences with the highest preference, $^{5'}$CGGCGG$^{3'}$, $^{5'}$CTGCGG$^{3'}$ and $^{5'}$CAGCGG$^{3'}$, the benzenoid intercalation site is followed by a guanine, making this the greatest determinant for preference of orientation.

Thymine and adenine bases are found on both sides of the intercalation site for both the benzenoid and quinoid side of the phenoxazone chromophore. A straightforward analysis of the preference for the benzenoid side of the chromophore for a particular sequence is to use the $^{5'}$XGCY$^{3'}$ notation. For a strong orientational preference, it was found that Y = G and that X = G > T > A. As previously mentioned, cytosines are not found on either side of the benzenoid side intercalation site.

The titration data shows that the two distinct orientations of ActD binding exist and that ActD prefers one orientation above the other, which is governed by the nature of the sequence flanking the intercalation site. A particular correlation between sequence and preference of orientation is identified in the six sequences studied, however the specifics of what drives the preferential binding could not be elucidated in detail.

III.5.2 The strongly asymmetric complex, ActD:d(CCGCCG)•d(CGGCGG)

ActD:d(CCGCCG)•d(CGGCGG) is the drug:DNA complex in which ActD shows the strongest preference for one orientation over the other. When the NOESY crosspeaks of this complex were analyzed, it was found that 79% of the ActD molecules had the benzenoid side of the phenoxazone chromophore in the d(CGGCGG) strand. However, because of linebroadening and spectral overlap of key residues, it was impossible to assign the resonances for the entire complex. Only a few resonances were readily identifiable, including the imino protons of G8, G3 and G6.

Complications in studying this complex may be due to a bulky guanine NH_2 group located in the minor groove that interferes with ActD binding, since

it was shown that the binding affinity for $^{5'}$GGCC$^{3'}$ sequences was significantly reduced (Chen, 1988). When sequence-specific inhibition of RNA elongation by ActD was investigated to determine apparent binding sites for ActD, it was shown that ActD preferentially interrupts elongation in the vicinity of the tetranucleotide $^{5'}$XGCY$^{3'}$, where X≠G and Y≠C (Aivasashvilli & Beabealashvilli, 1983). In our study, the central d(GC) binding site is followed by a cytosine in the first strand and preceded with a guanine in the second strand. Possibly the binding of ActD to the d(CCGCCG)•d(CGGCGG) sequence is weak and its orientations of binding thus can not be characterized. Gel mobility shift assays were used to verify whether ActD is indeed binding to the d(CCGCCG)•d(CGGCGG) sequence, and, it could be concluded that the ActD:d(CCGCCG)•d(CGGCGG) complex does have a slightly different gel mobility than d(CCGCCG)•d(CGGCGG) alone and a complex is definitely forming. Binding is also observed in the NOESY spectra, since we observe the H7/H8 protons on the chromophore interacting with the benzenoid guanine H2', H2'' protons, as well as H1' and the benzenoid cytosine H1' and H5 protons. Most of the resonances for this complex were assigned to protons in the molecule, however, some important resonances of intercalation site base protons overlapped and a complete assignment was not possible. Since accurate assignments are the first and one of the most important steps in solution structure determination, the solution structure for this complex was unfortunately impossible to elucidate.

III.5.3 Use of modified bases in ActD binding

The methyl group of the 5-methylcytosine is located in the major groove and is on the opposite side of the base from the Watson-Crick (W-C) base pairing edge (Figure III.19a). Thus this methyl group does not interfere with the standard W-C hydrogen bonding of a cytosine:guanine base pair, although the addition of this electron donating group slightly increases the association constant for base pair formation (Kyogoku *et al.*, 1967) . A bulky methyl group added to the C5 position in cytosine in CpG steps was found to affect DNA structure by promoting helical transitions from B to Z (Behe & Felsenfeld, 1981; Fujii *et al.*, 1982) or from B to A (Frederick *et al.*, 1987; Mooers *et al.*, 1995; Tippin *et al.*, 1997) in crystal structures. Few NMR studies on cytosine methylation have been published, although a joint ^{1}H and ^{31}P NMR, quantitative NOE-derived distance, sugar pucker and ε-torsion angle study on DNA octamers with a central d(CG) step showed that the effect of methylation depends on the initial conformation of the CpG site, governed by the nature of the flanking dinucleotides (Lefebvre *et al.*, 1995) . A more recent study involving dodecamers with a central d(CG) step show that a local variation of the structural parameters is observed and that the grooves of the helix are reorganized, with a severe pinching of the minor groove observed at the central CG subunit (Marcourt *et al.*, 1999) . These studies show that the addition of a methyl group does affect the structure of the DNA both in the crystal as well as in solution.

This methyl group does not have a significant effect on the orientational binding of ActD. Upon integration of several crosspeaks due to the major and minor complexes (Table III.7), we find that 66% of ActD is in one orientation

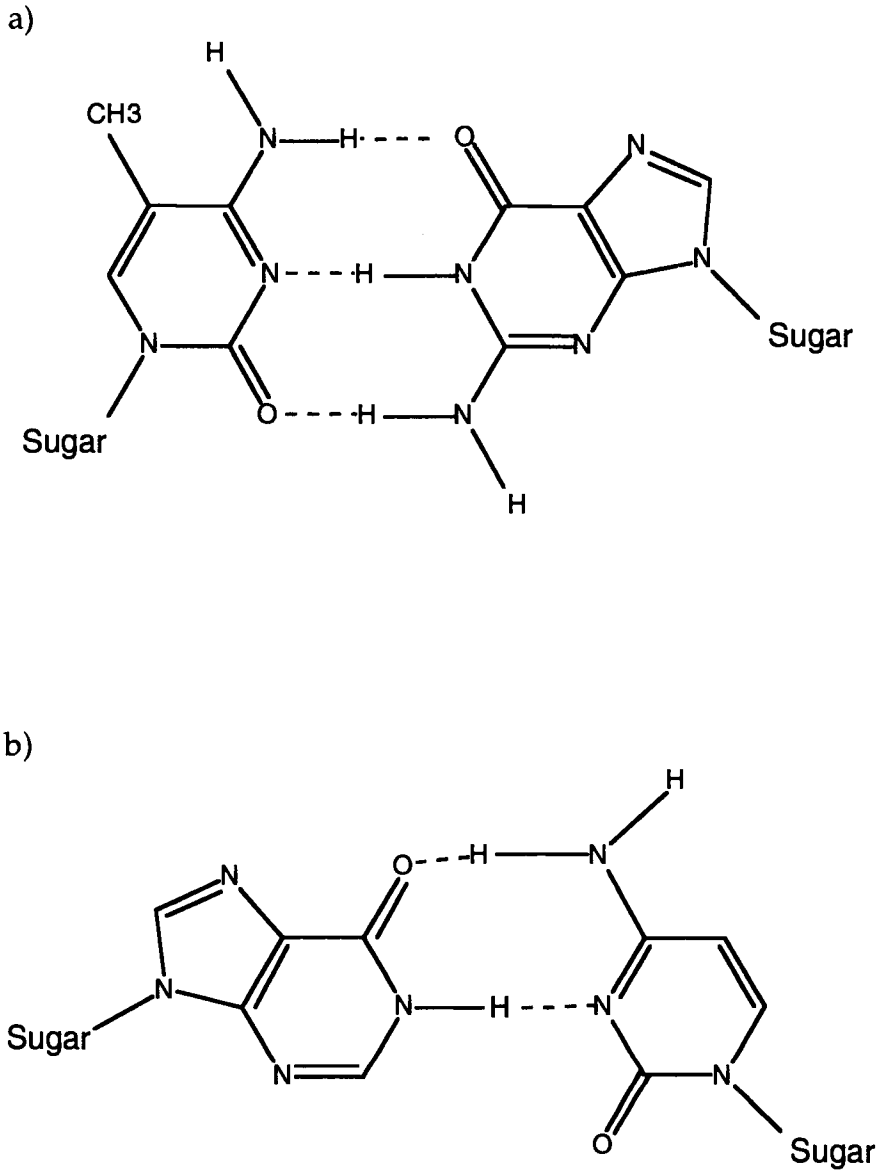

Figure III.19. a) The chemical composition of a 5-methylcytosine:guanine base pair. Three hydrogen bonds stabilize this base pair, with the additional methyl group in the major groove. b) The chemical composition of an inosine:cytosine base pair. In the unmodified G:C base pair, three hydrogen bonds are formed, thus the missing NH_2 group on the minor groove edge results in only two hydrogen bonds in an I:C base pair.

while 34% is in the opposite orientation. These ratios are similar to the ones found for the original d(CTGCGG)•d(CCGCAG) sequence, which had 67% of ActD in one orientation and 33% in the opposite orientation. We may conclude that the DNA in the ActD: d(CTGCGG)•d(C^{5me}CGCAG) complex has very similar interactions with the ActD as in the original complex, and that the addition of a bulky CH_3 group in the major groove does not interfere with ActD binding from the minor groove.

The other modified base sequence involves an inosine replacing a guanine. The inosine base is similar to the guanine base, except that the NH_2 group located in the minor groove is missing. This NH_2 group is involved in the W-C bonding of a guanine:cytosine pair; thus, elimination of this NH_2 group does preclude the formation of three hydrogen bonds in an I:C base pair (Figure III.19b). Crystallographic studies show that incorporation of an inosine base in replacement of guanine in the Dickerson dodecamer produces an isomorphous structure with that of the parent compound, with the largest difference being a high propeller twist for the I:C base pairs (Xuan & Weber, 1992) . Another crystallographic study finds the $d(CGCICG)_2$ structure in Z-form, like the parent compound utilizing guanine, yet differs in the relative orientation, position and crystal packing interactions, and identifies a narrower minor groove and compression along the helical axis (Kumar & Weber, 1993) . Solution studies incorporating inosine in a DNA duplex involve the dI:dG mismatch (Oda *et al.*, 1991) and the dI:dA mismatch (Uesugi *et al.*, 1987) . The first study shows that the dI(syn):dG(anti) base pair in a DNA dodecamer is accommodated in the B-DNA duplex with only a subtle distortion of the local conformation (Oda *et al.*, 1991) . The dI:dA mismatch study shows both bases in

the *anti* glycosidic conformation within a DNA hexamer. A comparison of ^{31}P spectra with a DNA hexamer involving the dI:dC base pair shows the backbone structure of the sequence involving dI:dA is disturbed by the presence of purine:purine mismatches (Uesugi *et al.*, 1987) . These structural studies show that inosine does not have a significant effect on the overall DNA structure, however local distortions exist which may have an influence on how a drug molecule binds to this locally distorted DNA sequence.

The absence of this NH_2 group, located in the minor groove, influences the orientations of binding of ActD. This is shown in the significantly reduced orientational preference of ActD binding to this sequence, since 58% of the complex has ActD in the major orientation and 42% in the minor orientation: elimination of the NH_2 group in the minor groove equilibrates the orientational preference upon ActD complex formation. As seen with the original six sequences, the benzenoid side intercalation site must be followed by a guanine for a strong preference for this strand. Perhaps guanine's exocyclic amino group is important in the formation of H_2O-mediated hydrogen bonds between ActD and the DNA. The shortest distance between the G5 exocyclic amine and an ActD group is found to the quinoid Sar N-methyl group and in fact, a weak NOE is observed in the ActD:d(CTGCGG)•d(CCGCAG) complex. The intermolecular contacts between the G5 exocyclic amine and the Sar N-methyl group could be responsible for the preference of one orientation above the other.

III.5.4 DNA Structure Analysis in the solved ActD:d(CTGCGG)•d(CCGCAG) complex

The DNA structure in the major and minor complex of the solved solution structure of ActD:d(CTGCGG)•d(CCGCAG) for which tight constraints were implemented was analyzed. For the most part, the DNA in both complexes resembles a right-handed helix with features comparable to those of B-form DNA. NOE-derived distances from the H1′ to H4′ proton and from the H2″ to the H4′ proton on the individual sugar rings are listed in Table III.14, and all agree with a predominantly C2′-*endo* sugar pucker. In the major complex there is a large roll present between the T2:A11 and G3:C10 base pairs, with T2:A11 tilting strongly towards the major groove. This roll is less extreme on the opposite side of the chromophore as well as in the minor complex DNA structure. The DNA parameters for both complexes are presented in Figure III.20 and were determined using the program CURVES, by R. Lavery and H. Sklenar (Lavery & Sklenar, 1988). It should be kept in mind that this is a short piece of DNA and the terminal base pairs were not constrained because of the effects of fraying. In addition, a drug intercalator is positioned between the central d(GC) base pairs. This leaves only a dinucleotide on either side of the intercalator for comparison with typical B-form DNA parameters. In Figure III.20a, the base pair rise has a large value above 4 Å on both sides of the intercalation site. The typically large internucleotide constraint distances obtained from the NOE volumes lead to the large base pair rise. The twist (Figure III.20b) indicates the degree of unwinding associated with base pair steps. Major unwinding is observed at the intercalation site, which is typical for the accommodation of the ActD drug. An unwinding of approximately the

Table III.14 NOE distances used for implementing standard B-DNA ν_0 to ν_4 dihedral angle values.

residue	complex	H1'/H4' NOE	H2''/H4' NOE
C1	MAJOR	3.7-3.3 Å	3.7-3.3 Å
	minor	3.7-3.3 Å	4.1-3.7 Å
T2	MAJOR	3.3-2.9 Å	3.7-3.3 Å
	minor	3.7-3.3 Å	6.0-4.1 Å
G3	MAJOR	3.3-2.9 Å	3.7-3.3 Å
	minor	3.3-2.9 Å	n/a
C4	MAJOR	6.0-4.1 Å	4.1-3.7 Å
	minor	4.1-3.7 Å	3.7-3.3 Å
G5	MAJOR	4.1-3.7 Å	4.1-3.7 Å
	minor	6.0-4.1 Å	4.1-3.7 Å
G6	MAJOR	6.0-4.1 Å	4.1-3.7 Å
	minor	4.1-3.7 Å	6.0-4.1 Å
C7	MAJOR	3.7-3.3 Å	3.7-3.3 Å
	minor	2.9-2.5 Å	6.0-4.1 Å
C8	MAJOR	3.3-2.9 Å	3.7-3.3 Å
	minor	3.7-3.3 Å	4.1-3.7 Å
G9	MAJOR	3.3-2.9 Å	3.7-3.3 Å
	minor	2.9-2.5 Å	4.1-3.7 Å
C10	MAJOR	3.7-3.3 Å	3.7-3.3 Å
	minor	4.1-3.7 Å	4.1-3.7 Å
A11	MAJOR	3.7-3.3 Å	4.1-3.7 Å
	minor	4.1-3.7 Å	4.1-3.7 Å
G12	MAJOR	6.0-4.1 Å	3.7-3.3 Å
	minor	6.0-4.1 Å	6.0-4.1 Å

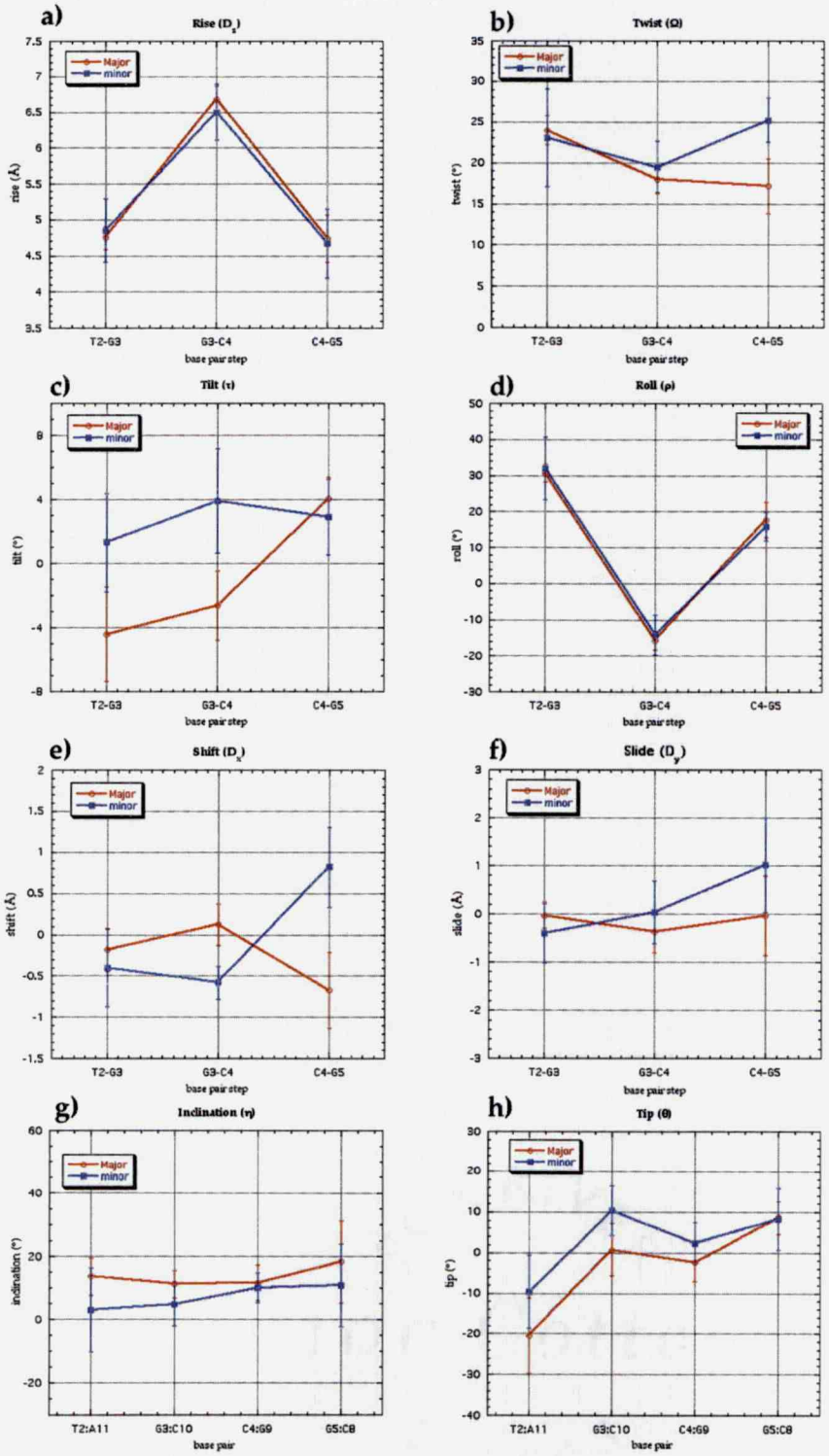

Figure III.20 DNA parameters determined for the 'tight bounds' structure of the major ActD:d(CTGCGG)•d(CCGCAG) complex compared with those of the minor complex. The following parameters are presented: a) base pair rise (D_z), b) base pair twist (Ω), c) base pair tilt (τ), d) base pair roll (ρ), e) shift (D_x), f) slide (D_y), g) base pair inclination (η), and h) base pair tip (θ).

same amplitude is observed at the C4-G5 step in the major complex, which is about 7° greater than on the opposite side of the chromophore. Interestingly, the minor complex DNA has a greater unwinding on the opposite side of the chromophore. The tilt and roll values, shown in Figure III.20c and d, are extreme at all base pair steps in both complexes in comparison with typical B-form values. The shift displays extreme values at all sites for the minor complex, Figure III.20e, while the values for slide are less extreme in the major complex than in the minor complex, Figure III.20f.

By comparing the parameter plots in Figure III.20 it is evident that the DNA structure in both complexes has some unusual characteristics. These unusual features do not resemble those of canonical B-form DNA or of previous solution structures of DNA oligomers. Considering the major complex DNA, the base pair rise at the non-intercalating steps is about 4.7 Å. The shift between base pairs, which is the displacement between the long axes of two base pairs, shows exceptional values, especially between the C4-G5 step. On the other hand, the slide between the base pairs, which is displacement between the short axes of adjacent base pairs, display less extreme values. These three parameters, rise, slide and shift, contribute most to the observation of weak internucleotide NOEs. Two explanations for these unusual features are possible. Either the ActD drug is inducing the DNA to adopt such an unusual conformation, or because of the short length of DNA, local motion exists between the base pairs. The simulated NOESY spectrum produced from the central six nucleotides of the crystal structure of ActD:d(GAAGCTTC)$_2$, displays similarly weak NOEs for the internucleotide interactions. By analyzing the DNA parameters of this structure in the same manner as the DNA structure of the major and minor complexes,

using the program CURVES, some unusual features are identified. The base pair rise on one side of the chromophore is 4.3 Å, while on the opposite side it is 2.7 Å. The side with the large rise has a shift of –0.6 Å and a slide of –0.8 Å, while the side with the short rise has a 0.9 Å shift and a 1.4 Å slide. These values for the major and minor complex are shown in Figure III.20. All of these features contribute to the weak simulated internucleotide NOEs for the crystal structure. Also some unusual sugar puckers exist in the crystal structure, such as a C2′-*exo* and C3′-*exo* at the intercalation site of one strand and C3′-*endo* and C1′-*exo* at the intercalation site of the opposite strand, with C1′-*exo* sugar puckers at flanking nucleotides. The crystal structure was solved at a resolution of 3 Å, thus the sugar pucker values may not be accurately determined. The values for shift and slide are more extreme in the crystal structure than in the solution structure of the major complex. On the other hand, the large rise values are more extreme in the solution structure. Similarities and differences exist between the two structures, and perhaps the consequence of the drug binding to the hexamer sequence results in these unusual DNA features.

Unusual characteristics were identified in the NOESY spectrum of the ActD:d(CTGCGG)•d(CCGCAG) DNA hexamer. These include unusually weak internucleotide NOEs throughout the DNA sequence, as well as reversal of the non-terminal G5 residue's H2′ and H2″ chemical shifts. In contrast, for typical B-DNA the sequential internucleotide NOE from H8/6 (i) to H2" (i-1) should be the second strongest NOE in the set of four interacting with H8/6. The corresponding distances for canonical B-DNA are: H6/8 to H2′ is 2.0 Å, H6/8 to H2″ is 3.2 Å, H6/8 to (i-1) H2′ is 3.9 Å and H6/8 to (i-1) H2″ is 2.2 Å. With free DNA in identical conditions as the complex and at this temperature, sequential

NOEs are not observed (Figure III.21). As the temperature is lowered for free DNA, sequential NOEs begin to be observed at 15°C, though even at 5°C the sequential NOEs do not resemble those of typical B-form DNA (Figure III.21d). The consistently weak internucleotide NOEs found for the ActD:d(CTGCGG)•d(CCGCAG) complex correlate with the uniformly large rise in the solved structure. The canonical B-form DNA rise is 3.4 Å, with the largest rise of 3.7 Å found in Z-form DNA. One possible reason for weak internucleotide NOEs throughout the sequence is that the DNA is poorly structured, since it is a short piece of DNA and fraying occurs near the ends. The DNA has increased motion between the base pairs in the presence of fraying and it is poorly stacked and underwound. In this case, the shortest distance that two protons get to each other is limited by what is observed in the NOE, which is consistent with the $\langle r^{-6} \rangle^{-1/6}$ averaging invoked for this type of motion (Neuhaus & Williamson, 1989). This means that if two protons spend 10% of the time at a separation of 3.0 Å and 90% of the time at a separation of 5.0 Å, then 90% of the NOE intensity comes from the 3.0 Å separation. In this case we open up the outer bounds for the internucleotide restraints to allow for the additional motion. The consequence is an increased rms difference in the families of the structures, which shows a better representation of conformational sampling. Previous solution studies have established that terminal base pairs are fraying (Nonin *et al.*, 1995; Ulrich *et al.*, 1983). In most cases, spectra of the imino protons are used to detect this. The imino proton spectrum of the ActD:d(CTGCGG)•d(CCGCAG) complex (Figure III.7) shows that the G5 imino is broader than the G9 imino and the G3 imino, while the T2 imino does not appear in the spectrum. This fraying motion is often propagated to the

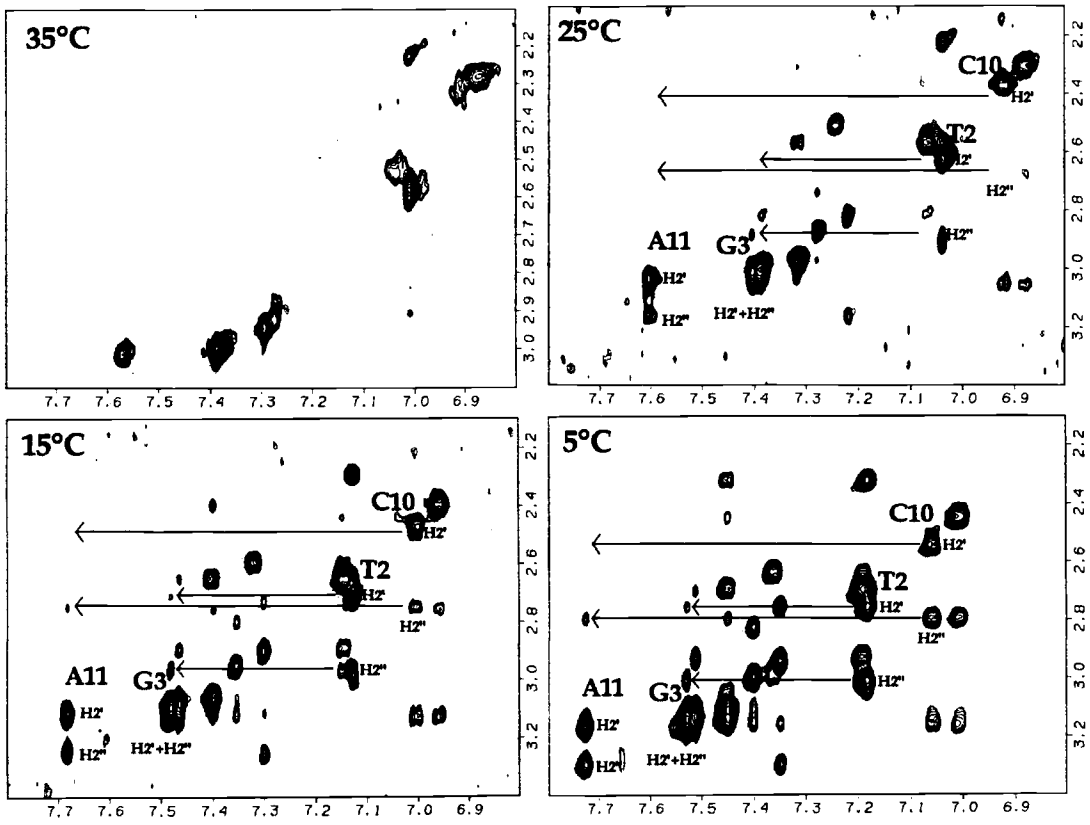

Figure III.21 Aromatic to H2′, H2″ proton regions of NOESY spectra recorded at different temperatures of free d(CTGCGG)•d(CCGCAG) in ^2H_2O, pH7, 10mM NaCl, a) T = 35°C, no sequential NOEs are observed; b) T = 25 °C, few sequential NOEs begind to appear in the spectrum; c) T = 15 °C some sequential NOEs are observed and d) T = 5 °C and the sequential NOEs still do not resemble those of typical B-form DNA.

penultimate base pairs, and often in published solution structures of DNA molecules, the terminal bases are excluded from the final structure (Liu *et al.*, 1991). The terminal bases are omitted in the structures presented here, yet increased fraying is a possible explanation for the penultimate bases as well. Solution studies of DNA hexamers are rare and are usually reported at a low temperature (Lam & Au-Yeung, 1997). The data for this study was recorded at a temperature of 35°C because the best spectral resolution attained at this temperature was necessary for the assignment of both the major and minor complexes. The elevated temperature is expected to increase the conformational dynamics associated with the DNA hexamer, though the presence of ActD stabilizes the duplex, enabling the observation of sequential NOEs. A previous NMR study of ActD binding to $d(ATGCAT)_2$ found that base proton to 5'-neighboring nucleotide deoxyribose 1', 2', 2" contacts are diminished, which was interpreted as a shift in the stacking of the bases (Brown *et al*, 1984). Unfortunately three dimensional atomic coordinates were not available for this complex and a direct comparison with the structure presented in this chapter was not possible.

The DNA parameters for the 10 solved structures using loose constraints for each complex are presented in Figure III.22 for the major complex and Figure III.23 for the minor complex. The base pair rise is centered between 6 Å and 7 Å at the intercalation site, an expected value for accommodation of the ActD chromophore (Figure III.22a). Flanking the intercalation site, the base pair rise varies, with the smallest rise being 3.6 Å and the largest being 5.2 Å, a much larger value than that found in typical DNA. The twist values correspond to those found previously in solved structures of ActD:DNA complexes, with the

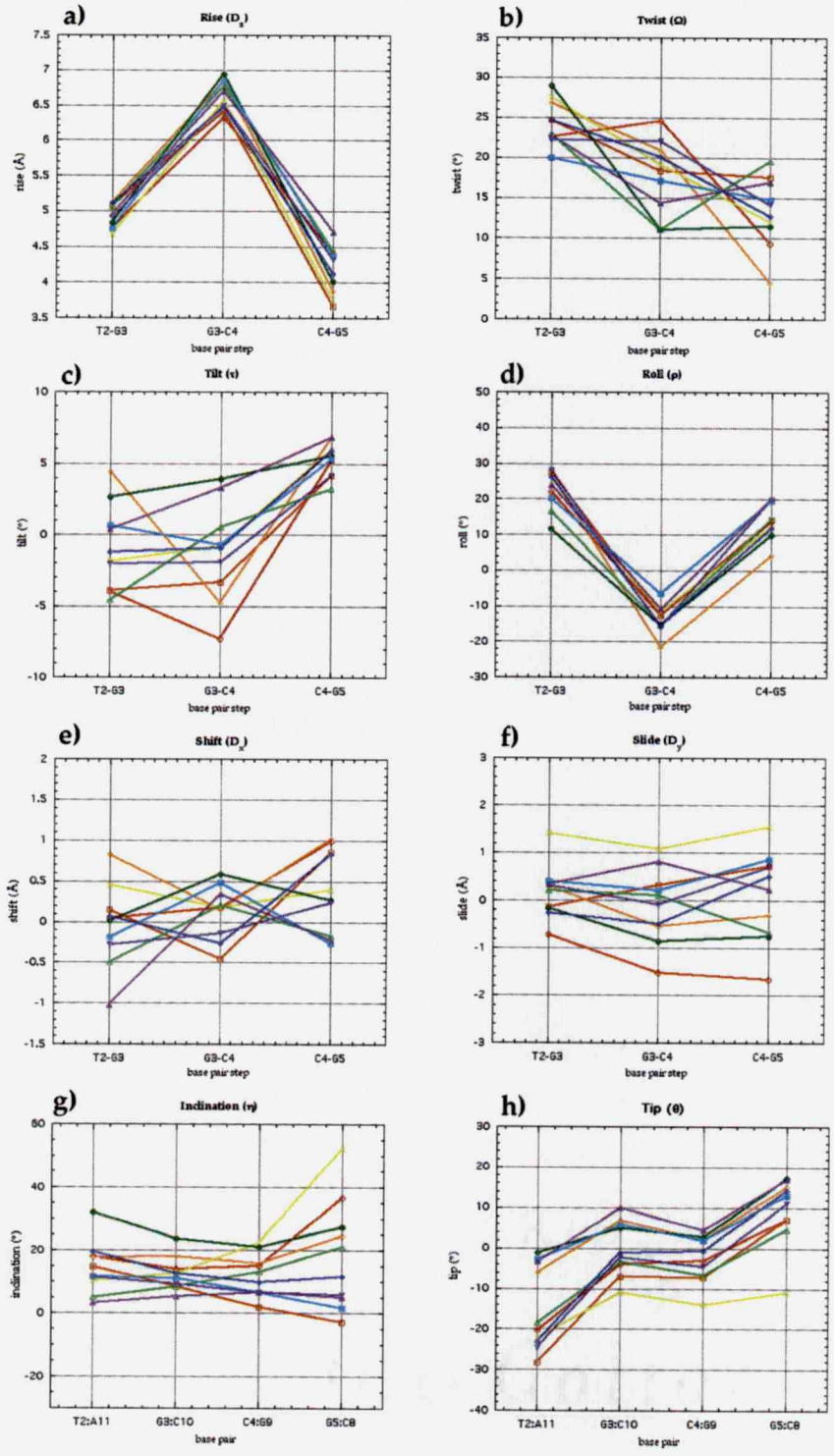

Figure III.22 DNA parameters determined for the 10 final structures of the major ActD:d(CTGCGG)"d_CCGCAG) complex with an rms difference of 1.02 +/- 0.07 Å. The following parameters are presented: a) base pair rise (Dz_y, b) base pair twist (W)_ c) b ase pair tilt (ȶ), d) base pair roll (ρ), e) shift (D_x), f) slide (D_y), g) base pair inclination (η), and h) base pair tip (θ).

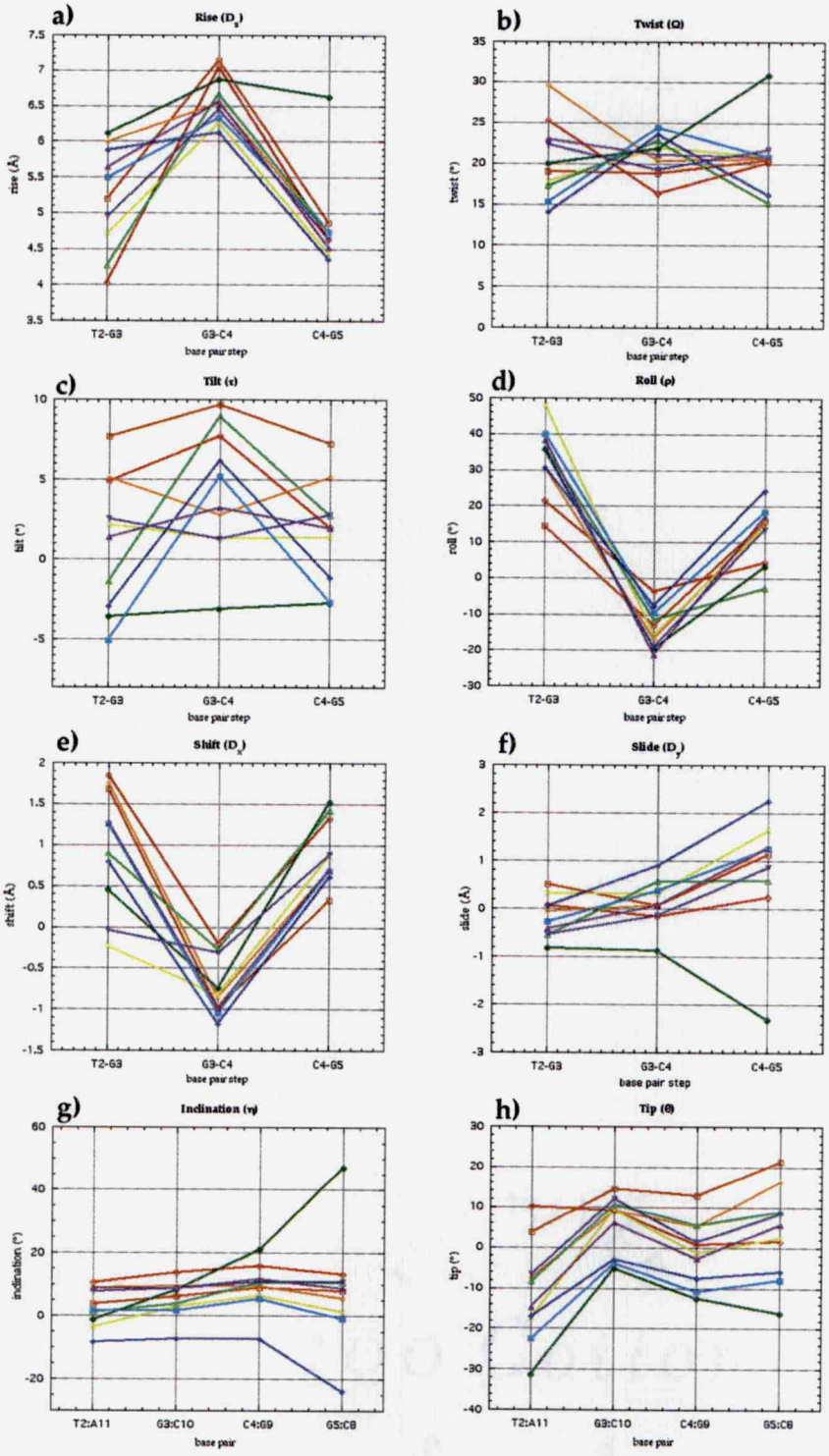

Figure III.23 DNA parameters determined for the 10 final structures of the minor ActD:d(CTGCGG)•d(CCGCAG) complex with an rms difference of 1.30 +/- 0.12 Å. The following parameters are presented: a) base pair rise (D_z), b) base pair twist (Ω), c) base pair tilt (τ), d) base pair roll (ρ), e) shift (D_x), f) slide (D_y), g) base pair inclination (η), and h) base pair tip (θ).

twist at the intercalation site just above 10°, necessary to accommodate the chromophore with the protruding pentapeptide rings between the base pairs (Figure III.22b). The DNA unwinding varies significantly on either side of the chromophore for these 10 structures, indicative of conformational heterogeneity. The roll values resemble those of B-DNA most at the intercalation site, where the chromophore stacks with the bases, and the G:C base pairs are parallel with the chromophore (Figure III.22d). This is the most well-defined region of the structure based on the rms difference calculations. The roll between base pairs increases significantly on either side of the chromophore. The values found for base pair tilt (τ) are spread for all the base pairs in the sequence (Figure III.22 c), as are the values for shift and slide, and the values for the inclination and the tip angles (Figure III.22e through h). Similar observations are found in the minor complex as well, with the DNA parameters portrayed in Figure III.23. However, a higher final R-factor was found for this complex and thus the final DNA parameters may not be very accurate.

Comparing the best structure found for the major and minor complexes, the DNA hexamers are in good agreement with each other, with a pairwise rms difference of 0.99 Å over all atoms (Figure III.24). The major differences are found in the intercalation site ε and ζ backbone angles and are shown in Figure III.25, since the constraints for those angles differed based on the ^{31}P chemical shifts at that site. The position of the T2:A11 base pair and the G5:C8 base pair differ in both the complexes. For both the major and minor complexes the cyclic pentapeptide rings bind more snugly in the minor groove of the G5:C8

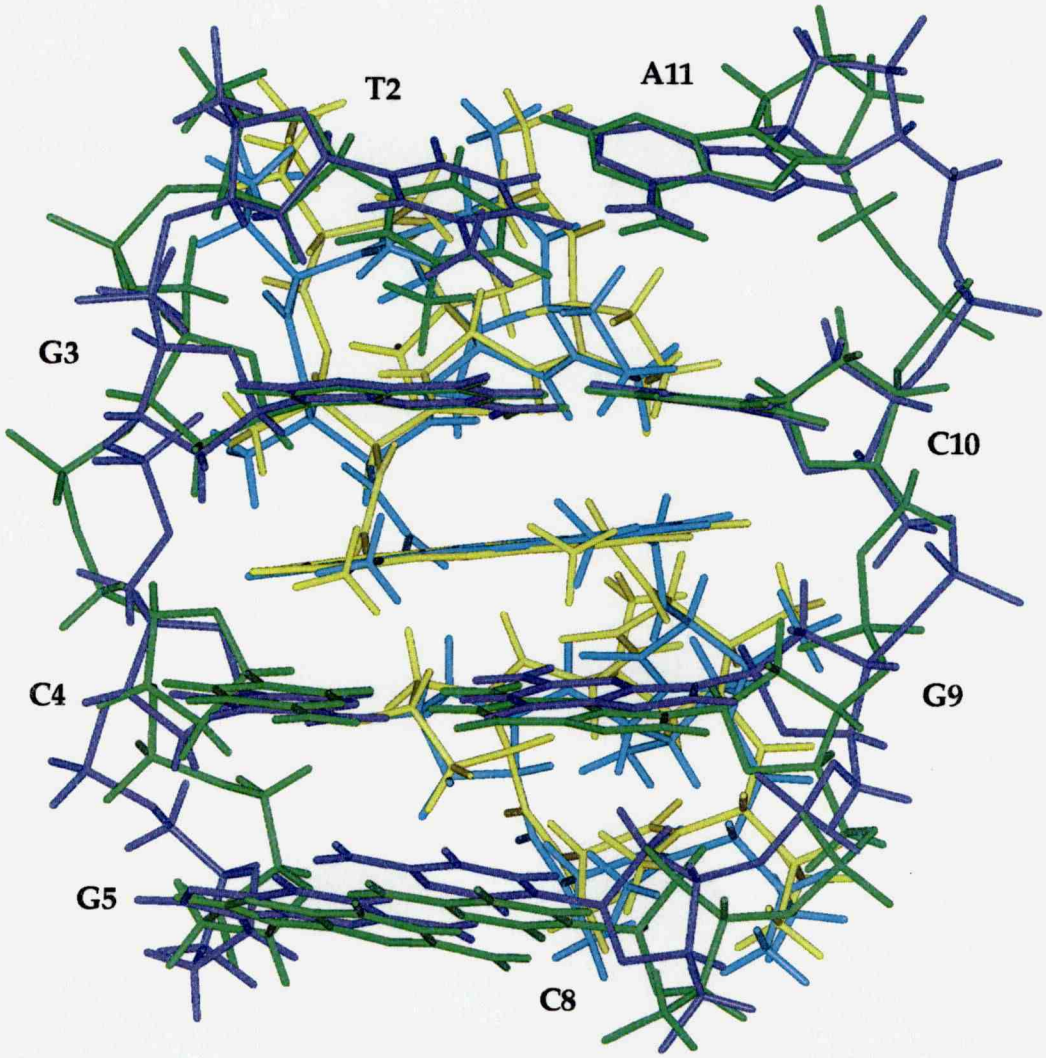

Figure III.24 A comparison of the 'tight bounds' structure of the major (DNA in green, ActD in yellow) and minor (DNA in dark blue, ActD in light blue) complexes. The pairwise rms differences between the DNA molecules in the two structures is 0.99 Å. By superimposing the four cyclic pentapeptides, the total rms difference among the four is 0.47 Å.

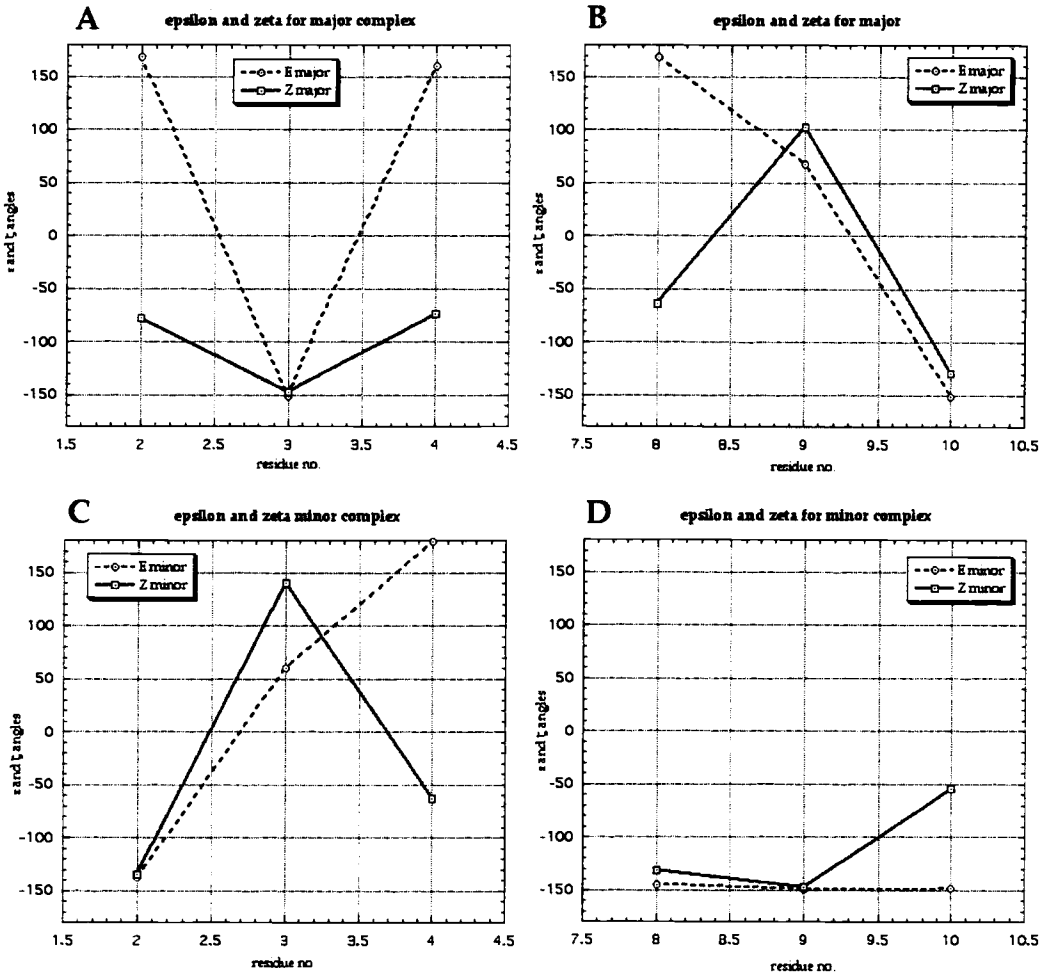

Figure III.25 ε and ζ angles for the central four base pair region of the DNA in the major and minor complexes of ActD:d(CTGCGG)•d(CCGCAG). a) major complex residues 2, 3 and 4; b) major complex residues 8, 9 and 10; c) minor complex residues 2,3 and 4; d) major complex residues 8, 9, 10. The ε angles are shown with dashed lines and the ζ angles are shown in solid lines.

base pair, than in the minor groove of the T2:A11 base pair. Also more helix unwinding is observed at the C4-G5 step in both complexes.

III.5.5 ActD structure analysis in the solved ActD:d(CTGCGG)•d(CCGCAG) complex

Comparing the proton chemical shift values for the cyclic pentapeptide lactone moieties in the free and bound forms of ActD shows that no significant changes occur upon binding (Table III.9). Free ActD in solution shows both the DVal-Pro peptide bonds and the Pro-Sar peptide bonds in the *cis*-configurations (Angerman *et al.*, 1972). The same is found for ActD in both the major and the minor complexes, which is verified by the strong NOEs between the Pro α and DVal α protons and the NOEs between the Pro α and Sar α protons. These *cis*-peptide linkages were previously established in solution (Brown *et al.*, 1994; Liu *et al.*, 1991) and crystal structures (Kamitori & Takusagawa, 1992, 1994) of ActD bound to isolated G-C sites on DNA oligomer duplexes. The same *cis*-peptide linkages were observed in the crystal structures of free ActD (Ginell, 1988), the ActD:dG complex (Sobell *et al.*, 1971) and the ActD:d(GC) complex (Takusagawa *et al.*, 1982). The observed weak NOEs between the DVal α proton and β proton identify the *trans* -conformation for the side chain of this peptide residue, the same being true for the NMV residue. The stronger NOEs between the Thr α and β protons indicate that Thr displays a preference for the *gauche* -orientation. The same was shown for previous ActD:DNA complexes, such as the ActD:hairpin complexes preceding this work as well as the ActD:d$(CGCG)_2$ complexes which used results from $^3J(H\alpha/H\beta)$ coupling constants and NOE measurements (Delepierre, 1989).

The cyclic pentapeptide rings of the ActD molecule in the complex project in opposite directions from the intercalated phenoxazone chromophore. The planes described by the cyclic pentapeptide lactone rings are perpendicular to the planes of the phenoxazone chromophore, with the benzenoid and quinoid rings completely occupying the widened DNA minor groove of both complexes. The two pentapeptide rings are related by a pseudo-2 fold axis of symmetry and overlaying the benzenoid and quinoid rings in the minimized average structure gives an rms difference over all atoms of 0.68 Å for the major complex and 0.62 Å for the minor complex. Overlaying the four pentapeptide rings in the major and minor complexes gives an rms difference of 0.47Å, indicating that the pentapeptide rings are nearly identical in structure. When analyzing the ActD conformation in its complex with DNA, a distinct orientation of polar and non-polar functional groups with respect to the DNA is apparent. The carbonyl groups of Pro, Sar and NMV are oriented perpendicular to the helix axis and are directed towards the solvent, while the carbonyl group of DVal is parallel to the helix axis and directed towards the other cyclic pentapeptide ring. The carbonyl group of Thr is pointed away from the solvent and towards the DNA and acts as a hydrogen bond acceptor, while the Thr amide also points towards the DNA and acts as a hydrogen bond donor. The hydrophobic methyl groups of DVal and NMV and the N-methyl group of Sar are pointed towards the exterior of the complex in both solution structures. The N-methyl group of NMV and the Cα protons of Sar are directed towards the floor of the minor groove and are protected from solvent. Specifically the Cα protons of Sar are pointed towards the A11 H2 proton on the benzenoid side of the major complex and toward the G5 exocyclic amines on the quinoid side. The

N-methyl group of NMV points towards the T2 O2 group on the benzenoid side of the major complex and towards the C8 O2 group on the quinoid side. The Pro ring protons and the Thr methyl groups are directed towards the sugar phosphate backbone.

III.5.6 Intercalation site geometry and intermolecular contacts

The intercalation site NOEs reflect the alignment of the phenoxazone chromophore with respect to the binding site G:C base pairs in the refined structures (Figure III.26). The long axis of the phenoxazone chromophore is aligned parallel to the long axis of the flanking G:C base pairs. The phenoxazone chromophore sits in the center of the intercalation cavity and stacks with the guanosine but not the cytosine bases of flanking G:C base pairs, as shown in Figure III.26b, and in agreement with previously solved structures. The positioning of the chromophore is symmetric with respect to the surrounding bases and does not show an indication for preferential binding. The increased separation at the intercalation site is noted by the interproton separation between the G3 sugar protons and the C4 base proton, as well as between the G9 sugar protons and the C10 base protons. In addition, the G3:C10 base pair at the intercalation site is parallel to the G9:C4 base pair on the other side of the ActD chromophore. Unwinding of the double helix by 11° is observed at the intercalation site. Helix unwinding is observed in all ActD:DNA complexes and is essential for accommodating the drug chromophore.

A complementary fit is observed between the cyclic pentapeptide lactone rings and the minor groove on both sides of the intercalation site of the ActD:d(CTGCGG)•d(CCGCAG) complex. A strong intermolecular hydrogen

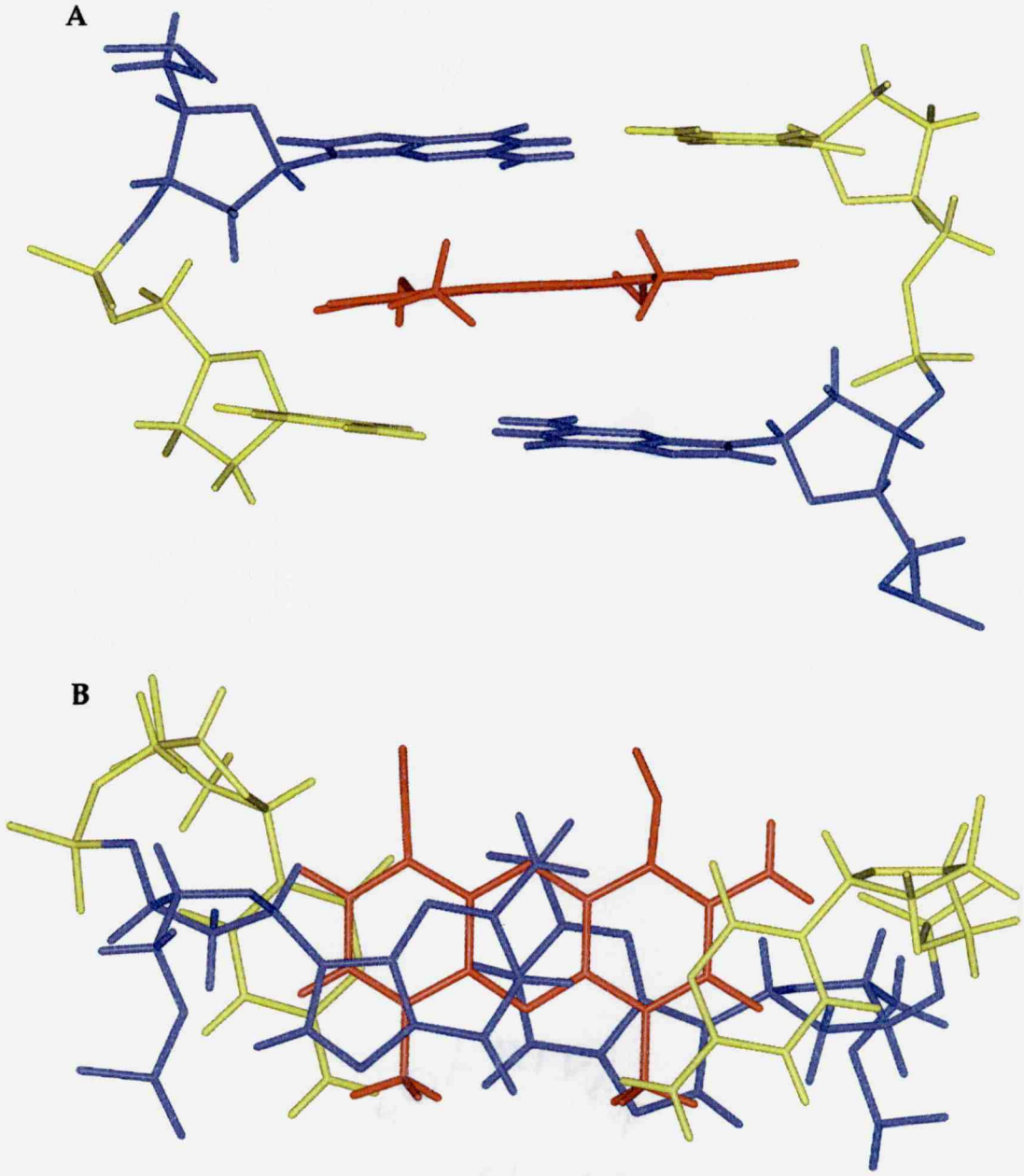

Figure III.26 A view of the intercalation site geometry in the major complex. The guanine residues are in blue, the cytosine residues are in yellow and the ActD chromophore is displayed in red. In a) we observe that the ActD chromophore sits centrally in the intercalation cavity and there is a parallel alignment of the base pairs flanking the intercalation site. b) This view emphasizes the overlap geometry between the phenoxazone chromophore and flanking base pairs.

bond between the NH_2 proton of the intercalation site guanine and the carbonyl oxygen of the benzenoid Thr (distance 1.9 Å, angle 169° for the major complex; distance 1.8 Å, angle 162° for the minor complex), and a weaker hydrogen bond between the N3 of guanine and the amide proton of benzenoid side Thr (distance 1.7 Å, angle 165° for the major complex, distance 1.6 Å, angle 158° for the minor complex) account for the sequence specificity of ActD for $d(GC)_2$ sites on duplex DNA. The above mentioned hydrogen bonds are observed on the opposite or quinoid side and are shown in Figure III.27. The NH_2 proton of guanine to the quinoid carbonyl oxygen distance is 2.0 Å in the major complex with an angle of 146° and 2.4 Å in the minor complex with an angle of 156°. The distance between N3 of guanine and the quinoid Thr amide proton is 2.4 Å in the major complex with an angle of 137° and 2.7 Å in the minor complex with an angle of 124°. The hydrogen bonds were measured from the hydrogen donor atom to the acceptor atom. These intermolecular hydrogen bonds are an important part of the 'Sobell and Jain' model (Sobell & Jain, 1972) and our results support this proposed model. However, these hydrogen bonds were not included explicitly as 'hydrogen bond' constraints in the molecular dynamics calculations of the structure.

III.5.7 Comparisons with previously solved structures

The previous solution structures of DNA:ActD complexes involve self-complementary sequences and as a result the two orientations of binding can not be identified using such sequences. In the solution structure of ActD-$d(A_3GCT_3)_2$, symmetric unwinding of the DNA on either side of the intercalation site is identified and the chromophore is centrally located in the

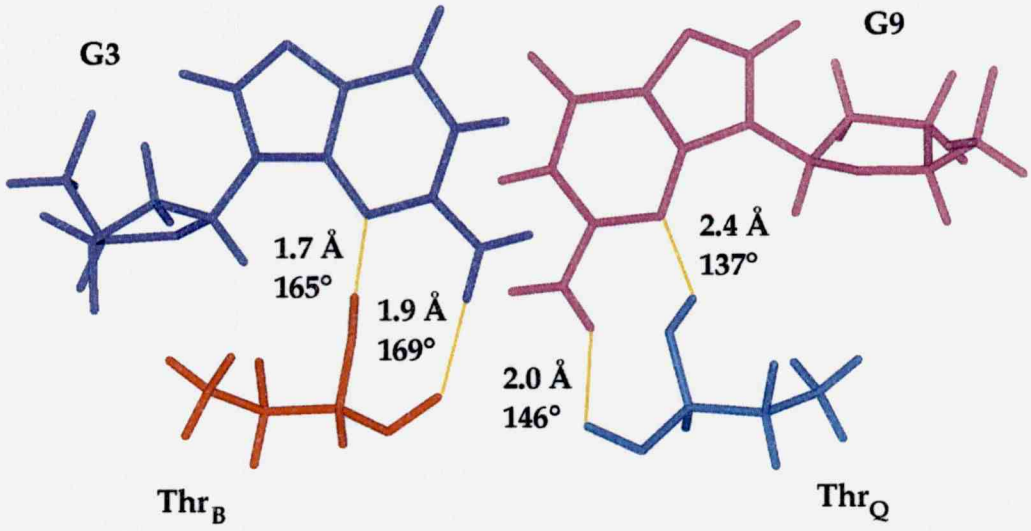

Figure III.27 Hydrogen bonds formed between ActD and d(CTGCGG)•d(CCGCAG) in the major conformer. Two pairs of intermolecular hydrogen bonds are detected in the minor groove of the DNA. One pair is between the benzenoid side Thr in red and G3 of the DNA in blue, the other pair is between the quinoid side Thr in light blue and G9 of the DNA in purple. The hydrogen bonds are from the exposed NH_2 of guanine to the Thr CO, and from the guanine N3 to the Thr NH.

intercalation site cavity and stacks with the guanine bases but not the cytosine bases (Liu *et al.*, 1991) . The solution structure of ActD:d(GAAGCTTC)$_2$ shows that ActD binds in and widens the minor groove and maintains its pseudo-2-fold symmetry (Lian *et al.*, 1996) . In this structure the amino group on the phenoxazone chromophore causes the backbone between the G and C nucleotides where it is intercalated to move by about 2 Å and is in a good position to form a hydrogen bond to the O4' of cytosine. This offset of ActD in the intercalation cavity destroys the 2-fold symmetry of the complex and by closer inspection of the two strands, the dihedral backbone angles differ starting with the A3 ζ angle and continuing through to the T6 ε angle (Lian *et al.*, 1996). Some of the first 2D NMR studies show that the DNA protons located close to the benzenoid side of the chromophore have a different chemical shift than the same protons located close to the quinoid side of the chromophore (Brown *et al.*, 1984; Liu *et al.*, 1991) , an indication that the two-fold symmetry of self-complementary DNA is broken upon ActD complex formation. However, if 50% of ActD molecules bind in one orientation and 50% bind in the opposite orientation, since the DNA strands are identical, there is no distinction between these two orientations. Non-self-complementary DNA sequences are essential in identifying the two orientations of binding.

The intercalation site geometry for the current structure shows that the ActD chromophore is evenly located with respect to the intercalation site guanines in both the major and minor complexes (Figure III.26). As in previously solved structures, the chromophore stacks with the guanine but not the cytosine bases (Liu *et al.*, 1991). The structures presented in this research do not show a significant difference between the intercalation site of the major and

minor complexes. The intercalation site is similar to ones previously identified in ActD complexes with self-complementary sequences (Liu *et al.*, 1991; Lian *et al.*, 1996). The intercalation site geometry does not suggest why the drug prefers one orientation above the other.

III.5.8 Flanking base pairs affect the orientation of ActD binding

In the preceding study, the asymmetry of the DNA sequence with the presence of a hairpin loop showed that ActD has two orientations of binding and that it prefers one orientation above the other. The conclusions state that either a sequence dependence for base pairs neighboring the d(GC) intercalation site exists or the proximity of the hairpin loop may induce steric constraints which exaggerate the electrostatic effects of the intercalation site (Brown *et al.*, 1994). The study presented in this chapter shows that there is a sequence dependence for base pairs flanking the d(GC) site which drives the asymmetric binding of ActD. A hairpin loop is not necessary to obtain the two orientations of ActD binding, though it is an element of asymmetry that led to the observation of the two orientations of binding. By studying every possible combination of non-self-complementary nucleotides surrounding the d(GC) intercalation site, I have shown that certain sequences lead to a stronger preference of orientation than others. The particular causes of asymmetric binding and why ActD prefers one orientation above the other, were, however, not revealed by these studies.

Hypotheses on how the intercalation process between ActD and DNA occurs have been proposed in the past (Hamilton *et al.*, 1963; Kamitori & Takusagawa, 1994). Some plausible interactions between ActD and the minor

groove of DNA are necessary and ActD most likely approaches DNA with the negatively charged quinoid (O3) and the positively charged amino group (N_2) of the chromophore pointed toward the minor groove of DNA (Hamilton *et al.*, 1963). These groups can interact with the positively charged amino group of guanine (N2) and negatively charged oxygen atoms (O4′, O5′) in the minor groove (Kamitori & Takusagawa, 1994). After formation of these initial hydrogen bonds, ActD swings around the O4′ and O5′ atoms of the intercalation site cytosine to insert the chromophore between the base pairs (Kamitori & Takusagawa, 1994). The ActD N_2 to cytosine O4′ and O5′ hydrogen bonds are maintained, while the ActD O_3 to guanine N_2 hydrogen bond is broken. The *trans-trans-trans* conformation of the backbone between the intercalation site guanine and cytosine, typically found in A-from DNA, is important in this recognition process since the O5′ atom is brought to the surface of the minor groove (Kamitori & Takusagawa, 1994). This then becomes the hydrogen bond acceptor for the amino group of the chromophore of ActD. Further experimentation is necessary to test this hypothesis, though support of it comes from the fact that the ActD analogue, 2-chloro-ActD, which has a chloro group replacing the amino group at the 2-position on the chromophore, does not bind to DNA and is inactive (Meienhofer & Atherton, 1977). Incorporating this information with the analysis of the results from the experiments carried out in this chapter may lead to a better understanding of the asymmetric orientational binding. It was shown that the quinoid side of the chromophore does not 'like' the intercalation site d(GC) to be followed by a guanine. Perhaps the 5′-GCG-3′ sequence is less likely to adopt the *trans-trans-trans* backbone conformation between the guanine and cytosine, making the backbone O5′ of cytosine less

available for hydrogen bond formation with the N2 group of ActD. Thus the quinoid side of the chromophore is less likely to approach and be able to bind to this sequence. When G5 is replaced by an inosine, we observe a significant reduction in the benzenoid preference for this strand. Hydrogen bonds between the ActD N2 group and the O4′ and O5′ atoms of the intercalation site cytosine are claimed to be observed in existing structures (Liu *et al.*, 1991; Kamitori & Takusagawa, 1994), though in the present structure these hydrogen bonds are not detected. There is a possibility of a weak hydrogen bond between the ActD N2 group and the backbone O5′ of the intercalation site cytosine (2.4 Å from the proton to the oxygen, angle of 138°). Unfortunately, this region of the complex is proton poor and specific distances between the quinoid side of the chromophore and the DNA backbone are difficult to identify. In addition, the ActD NH_2 protons are not detected in the H_2O NOESY spectra. These protons are also not assigned in previous NMR publications.

Significant structural differences between the major and minor orientation were not identified, though the titration studies tell us what types of sequences the benzenoid side of the chromophore prefers over the quinoid side of the chromophore. Mainly, the benzenoid side prefers guanines, while the quinoid side prefers cytosines adjacent to the intercalation site. After guanines, the benzenoid side prefers thymines on the 5′ side of the intercalation site, followed by adenines. The cyclic pentapeptide rings fit a little less tightly on the benzenoid side than on the quinoid side, and the proline is located slightly closer to the cytosine O2 and guanine exocyclic amino protons in the quinoid ring than in the benzenoid ring in both complexes. The van der Waal's contacts between the minor groove of the T2:A11 base pair and the pentapeptides are not as good

as on the opposite side of the chromophore. For the major complex, a gap exists between the benzenoid Cα protons of Sar and A11 H2, while the same protons on the quinoid ring come much closer to the G5 NH_2 group. Similarly, the quinoid N-methyl group of NMV comes much closer to the C5 O2 atom, than does the benzenoid N-methyl to the T2 O5 atom. A similar gap is found between the minor groove of the T2:A11 base pair and the quinoid pentpapetides in the minor complex, though in this complex it is not as pronounced. The ActD drug is a mostly symmetric molecule, having a benzenoid and a quinoid side of the chromophore. The crystal structure of ActD solved at 0.94 Å resolution (PDB entry 1A7Y) shows that the angles at which the cyclic pentapeptide rings protrude from the chromophore are not identical for the two sides (Table III.15). Perhaps these angles and the ability of particular sequences to unwind more than others drives the orientational preference of ActD. The free energy difference between the major and minor complex is calculated to be 419 cal/mol. Considering such a small value for the difference in free energy, it is not surprising that significant structural differences between the two complexes are not observed.

The solved solution structure of this complex shows that the DNA molecule adopts an unusual conformation, in that the base pair rise on either side of the chromophore is larger than 4 Å. It is difficult to compare the current structure with previously solved ones, since previous structures utilize longer DNA sequences, such as octamers (Liu *et al.*, 1991; Lian *et al.*, 1996) and hairpin molecules which self associate (Brown *et al.*, 1994). In the current study the terminal base pairs were not restrained, leaving dinucleotide steps on either side of the chromophore for structural analysis. Perhaps the presence of the drug

Table III.15 ActD dihedral angles for the cyclic pentapeptides protruding from the chromophore

dihedral angle	benzenoid side	quinoid side
C8-C9-C18-O18	-44.7	-31.3
C8-C9-C18-N_{Thr}	131.8	141.7
C14-C9-C18-O18	129.1	147.4
C14-C9-C18-N_{Thr}	-54.5	-39.6
C9-C18-N_{Thr}-$C\alpha_{Thr}$	-177.7	179.3
O18-C18-N_{Thr}-$C\alpha_{Thr}$	-1.4	-7.6

forces the DNA to adopt this unusual conformation with a higher base pair rise, though the same is not found in existing NMR structures of this complex (Lian, *et al.*, 1996). Even with the 10 conformationally heterogeneous structures, the rms difference is low, approximately 1 Å, most likely because the intermolecular restraints are limiting the conformational sampling of the DNA molecule during molecular dynamics. Because the solution structure of the free DNA was not solved, a direct comparison with the ActD bound DNA structure can not be made. As discussed, the data collected on the free DNA at different temperatures does not indicate canonical B-form for these experimental conditions.

Because of unusual features within the spectra and because of the short length of the DNA, I hypothesize that the structure presented here is dynamic and the base pairs adjacent to the intercalation site $d(GC)_2$ are thought to be undergoing motion. Indeed, overall a higher unwinding is found in the C4 to G5 step, as opposed to the T2 to G3 step in the major complex, though the same is found for the minor complex. If this were to be due to the different angles at which the quinoid rings protrude from the chromophore, then the greater unwinding would be found on the opposite side in the minor complex. The greater unwinding found in the C4 to G5 step could be due to the ability of these steps to unwind more than the T2 to G3 step. However, this goes against previous knowledge of DNA flexibility, since T:A base pairs are known to be more flexible and able to open the minor groove more than G:C base pairs. By careful analysis of the two structures, it is evident that the cyclic pentapeptide rings fit more tightly in the G5:C8 minor groove than in the T2:A11 minor groove of both the major and minor complexes. Thus it is not surprising that a

greater unwinding is observed on this side of the chromophore. Aside from the intercalation site, the structure presented with this research demonstrates the difficulties in DNA structure determination and validation.

The work presented in this chapter is the continuation of the research begun with the ActD:hairpin studies and is critical in investigating why the two nearly identical complexes were not formed in equal concentrations. With the previous publication and ActD binding to non-self-complementary sequences, we have shown that non-self-complementary sequences are of utmost importance in identifying the two orientations of ActD binding via NMR spectroscopy. The greatest preference for one orientation above the other was identified for the d(CCGCCG)•d(CGGCGG) sequence, however complications in the NMR spectra prevented the complete characterization and structure calculation of this complex. The NMR solution structure is presented for the ActD:d(CTGCGG)•d(CCGCAG) complex as families of 10 structures for both the major and minor conformers. Intermolecular NOEs are observed between ActD and the four base pairs surrounding the intercalation site, with similar NOE intensities on both sides of the intercalation site. The binding of ActD affects the local environment of the terminal base pairs as well, evidenced by the different resonance values for these nucleotides in each complex. The results of this research are an essential step in determining the forces involved behind the two orientations of ActD binding and in identifying what causes the preference of one orientation above another.

Chapter IV.

Summary

In the first chapter of this thesis various aspects of DNA solution structure determination are examined. For short oligomers, proton resonance assignments are fairly straightforward and viable. A direct verification of proton assignments is possible through ^{1}H-^{31}P correlation experiments. The challenges of DNA structure determination via NMR spectroscopy were discussed, as well as the need for accurate interproton distance estimation. There also exists a need for better structure calculation algorithms, since both distance geometry and restrained molecular dynamics calculations have their downfalls. If site-specific labeling of DNA molecules was more available, the use of residual dipolar couplings in magnetically aligned molecules would be a favorable method for DNA solution structure determination, since the global bend can be directly determined.

The use of the ROESY experiment with DNA molecules is explained in the second chapter of this thesis. We have shown that this experiment can aid in the stereospecific assignments of the H2' and H2" protons (Ivancic & Hsu, 2000). This is especially useful with purine residues and longer DNA molecules for which the H2', H2" spectral regions become very crowded. By varying the spin-lock power level and mixing time, the ROESY experiment may be used as a tool for editing complex DNA spectra. The crosspeaks that we have observed in the ROESY spectra of our DNA oligomers exhibit a typical absorption pattern, particularly in the H6/8 to H2', H2" proton region, indicative of B-form conformations. When implemented this way, the ROESY experiment can be

used in distinguishing Z-form DNA from B-form DNA since the internucleotide distances between the H2' proton and the H6/H8 proton of the following residue is shorter than the distance between the H2" proton and the H6/H8 proton of the following residue, the opposite of B-form DNA. This was demonstrated on the ROESY spectrum of a DNA hexamer with both the B-form and Z-form oligos present. This method could also be used in the direct identification of A-form DNA, since A-DNA adopts a primarily C3'-*endo* sugar pucker, as opposed to the predominant C2'-*endo* sugar pucker in B-DNA, thus the internucleotide crosspeaks are expected to exhibit the opposite pattern of ROE intensities. In this way, the ROESY experiment can be used as a tool in identification of the different forms of DNA.

In the third chapter of this thesis, the solution structure of a DNA hexamer with a bound intercalator, ActD, was presented. ActD binds to non-self-complementary sequences in unequal proportions. Difficulties with the solution structure determination of the two conformers were encountered, the first being the significant spectral overlap of the proton resonances. Distinguishing the DNA proton resonances of the major and minor conformations was somewhat less problematic than distinguishing the ActD pentapeptide proton resonances. Since ActD has two identical cyclic pentapeptide lactone moieties, the benzenoid ring and the quinoid ring pentapeptides, difficulty exists in distinguishing the protons on each one, especially since several of them completely overlap. With the two sets of pentapeptide protons due to the major conformation, there exist two more sets of pentapeptide protons due to the minor conformation. Many of the pentapeptide protons of the minor conformation have slightly different chemical shifts than those of the major conformer, but several are degenerate with those of the major conformer. Yet another matter of complication was that the DNA crosspeaks exhibit intensities not typical of canonical B-form DNA. In

the NOESY spectra of the complex, the intranucleotide NOEs were strong, while the internucleotide NOEs were significantly weaker than expected for properly stacked DNA. This is especially evident in the base proton to H2', H2" proton region, where both the H6/8 proton to H2' and H2" intranucleotide proton interactions are strong, while the H6/8 proton to the previous nucleotide's H2', H2" protons are weak. This NOE intensity pattern is present throughout the DNA sequence. Two explanations for these unusual intensities were discussed; either the bound drug is inducing the unusual DNA conformation with a significantly large rise between base pairs, or conformational motion between the base pairs due to the short sequence and high experimental temperature is present, or a combination of both.

The first structures were calculated using the isolated spin pair approach (ISPA) (Gronenborn & Clore, 1990), in which the NOE intensities are judged as if the interacting spins exist in isolation and neighboring protons do not contribute to the crosspeak intensity. At short mixing times (t_m) the dependence of crosspeak intensities is close to linear and the NOE intensities are less subject to the effects of spin diffusion. Also with increasing correlation time, the region of linear dependence becomes smaller, and the maximum t_m value for which the ISPA method works decreases (Barsukov & Lian, 1993). Thus the smaller the mixing time the better ISPA will work, however the poor signal-to-noise at short t_m makes intensity measurements difficult. However, when considering a particular experimental situation, it is difficult to decide whether the limit of short mixing is applicable.

The complete relaxation matrix approach is a more accurate method of interproton distance determination, since the method simultaneously accounts for multiple polarization transfer pathways. Interproton distances are derived from the corresponding off-diagonal elements of the relaxation matrix, however,

the measurement of all the intensities including diagonal peaks is an unreachable goal in real cases (Barsukov & Lian, 1993). As a result, several iterative programs have been developed to deal with incomplete intensity information. Two general approaches have been followed; one that utilizes the self-consistency of the relaxation matrix, modifying the matrix until there is agreement between calculated and measured intensities (Borgias & James, 1990), and the other iteratively modifies the structure (Boelens, *et.al*, 1988). In the structures presented in the third chapter, the second iterative approach is employed with the use of the BIRDER program (Zhu & Reid, 1995). This method uses a comparison between the calculated and experimental intensities to define how close the structure is to the experimental data, while refining the structure using the molecular dynamics algorithm (CNS). The actual difference between theoretical (V_{sim}) and experimental (V_{exp}) volumes can be determined via the R-factor calculation, $R=(1/N)\ \Sigma_N\ [\ |\ V_{exp} - V_{sim}\ |\ /\ V_{exp}]$, where N is the number of NOESY crosspeaks considered.

By comparing the initially calculated structures with the structures presented in the third chapter using the total relaxation matrix, several differences were detected. The pairwise rms difference over all atoms between the major complex structure calculated using ISPA and the one using BIRDER is 1.49 Å and the two structures are superimposed in Figure IV.1. For the BIRDER structure of the major complex, the roll between the T2-G3 base pair step is even greater than for the old structure and the intercalation cytosines are placed slightly closer to the chromophore, with a smaller intercalation cavity. The pairwise rms difference over all atoms between the minor complex structures calculated via the two methods is 1.71 Å. The size of the intercalation cavity is the same, while the rise between T2 and G3 is smaller in the ISPA structure and the position of the G5:C8 base pair differs in both structures. The R-factors were

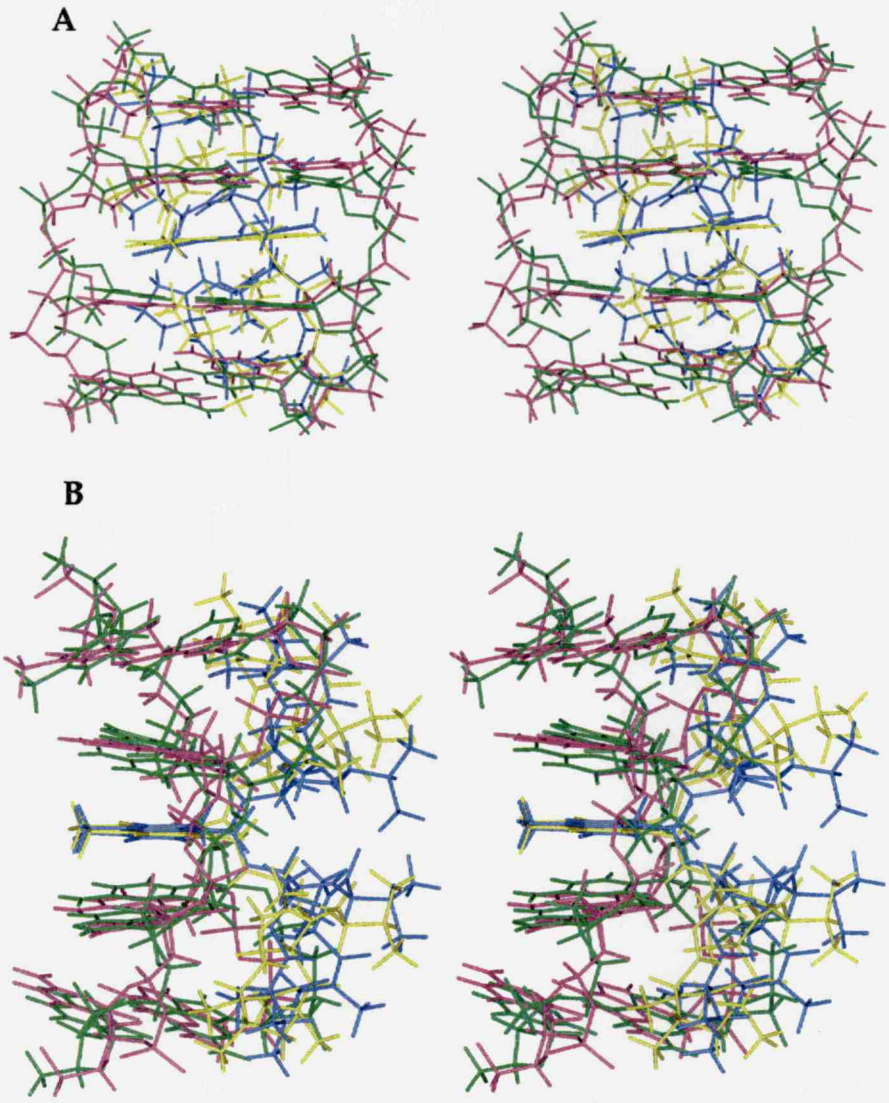

Figure IV.1 A comparison of the ISPA and the BIRDER calculated structures for the major complex of ActD:d(CTGCGG)•d(CCGCAG). The ISPA structure is in violet and blue, while the BIRDER structure is in green and yellow. A stereoview is presented. a) A view into the minor groove. b) A side view of the complex with the minor groove on the right and the major groove on the left. The differences between the two are discussed in the text.

calculated in the same manner for both sets of structures and are listed in Table IV.1. For the ISPA structure of the major complex, the R-factor for the DNA peaks is nearly three times that of the newly presented structure, while the overall R-factor for chosen peaks is almost twice as large. While for the minor complex a similar increase in R-factors is found with the ISPA calculated structures. The R-factor provides a means of quantitating the level of agreement between generated structures and experimental NOE data. Performing a backcalculation considering multiple relaxation pathways provides a quantitation of the magnitude of disagreement with the ISPA determined restraints. The complete relaxation matrix structures presented in Chapter III are in better agreement with the NOE experimental crosspeaks, and this is a good method of validation of the final calculated structures for both complexes. The structures are qualitatively similar, but quantitatively the differences as judged by R-factors and rms differences are significant. The DNA molecule in both sets of structures has some unusual features, including an unusually large rise between the base pairs. This large rise is a result of the weak internucleotide crosspeaks observed in the NOESY spectra.

One way of dealing with the ISPA approach is to subdivide the intensities into general groups, for example strong, medium and weak all with a lower bound equal to the van der Waals distance between two atoms (Barsukov & Lian, 1993). This approach was utilized in the original structure determination of the two complexes, and the crosspeaks were subdivided into five groups of distance restraints. A larger number of interactions was utilized in the ISPA calculated structures, because crosspeaks with significant overlap were included. For the overlapped peaks an open upper bound of 5 Å was used, since it is difficult to know what volume contribution results from one interaction under an overlapped crosspeak. When inspecting crosspeak intensities, an attempt was

Table IV.1 R-factors calculated for the ISPA structures and for the complete relaxation matrix structures

R-factors	ISPA structures	relaxation matrix structures
MAJOR- most DNA peaks (59 peaks)	67%	23%
MAJOR- most DNA and DNA to drug peaks (92 peaks)	66%	28%
MAJOR- most DNA, DNA to drug, drug to drug (144 peaks)	64%	37%
MINOR- most DNA peaks (29 peaks)	133%	44%
MINOR- most DNA and DNA to drug peaks (53 peaks)	156%	83%
MINOR- most DNA, DNA to drug, drug to drug (71 peaks)	274%	140%

made to assess contributions from spin-diffusion by inspecting crosspeak volumes at different mixing times. If the NOE volume built up too quickly between two mixing times, an indication of crosspeak domination by spin-diffusion, an appropriate distance restraint was assigned to the interacting protons. However, it is virtually impossible to account for all possible polarization transfer pathways using the ISPA method. This leads to significant errors in setting the restraint values. For the final ISPA structures, approximately 20 NOEs were violated out of a set of over 400 NOEs, yet if errors exist within the restraint values themselves, the NOE violations become meaningless. For example, a structure can be calculated which satisfies all the restraints, however if the restraints themselves have errors, then the structure does not agree with the experimental data. With the relaxation matrix approach, restraints are established with the benefit of accounting for multiple polarization pathways and indirect NOE contributions. In the relaxation matrix approach fewer interactions were utilized, because of the problem with significant overlap discussed previously. The non-overlapping crosspeaks were used for structure determination and validation. For the final structures, approximately 20 NOEs were violated out of a total of 180 NOEs. Because of the low R-factor, we know that the distance restraints were set to better reflect the NOE intensities and the structures are in better agreement with the data.

By analyzing the percent differences of the major and minor orientations of binding of ActD and by identifying the specific orientation of the intercalated chromophore, some trends were observed. In the three sequences with the strongest preference of one orientation above the other, a guanine is on the 3′ side of the benzenoid d(GC) intercalation site. When this guanine is replaced by an inosine, a significant reduction in preference of orientation is observed, going from 67% major and 33% minor to 58% major and 42% minor. It is hypothesized

that for the sequences with a guanine on the 3' side of the intercalation site, the O5' on the DNA backbone of the intercalation cytosine is not accessible to the chromophore's N2 group upon initial binding of ActD. When this O5' of the backbone comes to the surface of the minor groove, it is possible for the chromophore's N2 group to hydrogen bond with this O5' (Kamitori & Takusagawa, 1994). Perhaps the 5'-GCG-3' sequence is less likely to adopt the *trans-trans-trans* backbone conformation between the guanine and cytosine, making the backbone O5' of cytosine less available for hydrogen bond formation with the N2 group of ActD. Thus the quinoid side of the chromophore is less likely to approach and be able to bind to this sequence. This may be the basis of why ActD prefers one orientation above another, though further experimentation is needed to test this hypothesis.

More research needs to be performed to answer the question of what drives the asymmetric binding of ActD. The work accomplished in this thesis shows how ActD binds in the two orientations in a non-self-complementary DNA sequence and sets the framework and basis for future research. The solution structures of ActD bound in two orientations are presented and the structural details of how ActD binds in both orientations are addressed. Although the two orientations are formed in different ratios, their structures are found to be similar, indicating that the basis of orientational preference is likely not structural in origin. Intermolecular NOEs are observed between ActD and the four bases surrounding the intercalation site, with similar NOE intensities on both sides of the intercalation site. The binding of ActD affects the local environment of the terminal base pairs as well, evidenced by the different resonance values for these nucleotides in each complex. The most interesting result is that the two orientations of ActD binding are not present in equal concentrations. My studies show that the asymmetric binding is greatly

influenced by the asymmetric DNA sequence flanking the intercalation site, although the evidence for why this is so is not discovered in the detailed molecular structure. The complete analysis of every possible non-self-complementary base surrounding the d(GC) intercalation site led to the identification of the types of sequences that one side of the chromophore prefers above the other side. This information is essential for the complete understanding of Actinomycin D's preference of orientation.

Bibliography

Aboul-ela, F. & Varani, G. (1995). *Curr. Opin. Biotech.* **6**, 89-95.

Addess, K. J., Gilbert, D. E., Olsen, R. K. & Feigon, J. (1992). *Biochemistry* **31**, 339-50.

Addess, K. J., Sinsheimer, J. S. & Feigon, J. (1993). *Biochemistry* **32**, 2498-2508.

Aivasashvilli, V. A. & Beabealashvilli, R. S. (1983). *FEBS Lett.* **160**(1,2), 124-128.

Angerman, N. S., Victor, T. A., Bell, C. L., Danyluk, S. S. (1972). *Biochemistry* **11**, 2402-2411

Arnott, S., Chandrasekaran, R., Hukins, D. W. L., Smith, P. J. C., Watts, L. (1974). *J. Mol. Biol.* **88**, 523-533

Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A. G. W., Ratcliff, R. L. (1980). *Nature* **283**, 743-745

Barsukov, I. L. & Lian, L.-Y. (1993). . In *NMR of Macromolecules: A Practical Approach* (Roberts, G. C. K., ed.), pp. 342-351. Oxford University Press, New York.

Bax, A., Sklenar, V. & Summers, M. F. (1986). *J. Magn. Reson.* **70**, 327-331.

Bax, A., Clore, G. M. & Gronenborn, A. M. (1990). *J. Magn. Reson.* **88**, 425-431.

Behe, M. & Felsenfeld, G. (1981). *Proc. Natl. Acad. Sci. USA* **78**, 1619-1623.

Boelens, R., Konig, T. M. G., Kaptein, R. (1988). *J. Mol. Struct.* **173**, 299

Borgias, B. A. & James, T. L. (1990). *J. Magn. Reson.* **87**, 475-487.

Bothner-By, A. A., Stephens, R. L. & Lee, J. (1984). *J. Am. Chem. Soc.* **106**(3), 811-813.

Brown, L. R. & Farmer II, B. T. (1989). *Methods in Enzymology* **176**, 199-216.

Brown, D. R., Kurz, M., Kearns, D. R. & Hsu, V. L. (1994). *Biochemistry* **33**(3), 651-664.

Brown, S. C., K., M., Levenson, C. & Shafer, R. H. (1984). *Biochemistry* **23**, 403-408.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W., L. & Warren, G., L. (1998). *Acta Crystall.* **D54**, 905-921.

Cerami, A., Reich, E., Ward, D. C., Goldberg, I. H. (1967). *Proc. Natl. Acad. Sci. USA* **57**, 1036-1042.

Cheatham, T. E. & Kollman, P. A. (1996). *J. Mol. Biol.* **259**, 434-444.

Chen, F.-M. (1988). *Biochemistry* **27**(17), 6393-6397.

Chen, F.-M. (1998). *Biochemistry* **37**, 3955-3964.

Chen, H., Liu, X. & Patel, D. J. (1996). *J. Mol. Biol.* **258**, 457-479.

Chou, S.-H., Cheng, J.-W., Reid, B. R. (1992). *J. Mol. Biol.* **228**, 138-155.

Davey, C., Pennings, S. & Allan, J. (1997). *J. Mol. Biol.* **267**, 276-288.

Delepierre, M., Van Heijenoort, C., Igolen, J., Pothier, J., Le Bret, M., Roques, B. P. (1989). *J. Biomol. Struct. Dyn.* **7**(3), 557-589.

Dervan, P. (1986). *Science* **232**, 464-471.

Farber, S. J. (1966). *J. Amer. Med. Assoc.* **198**, 826-836

Farmer II, B. T., Macura, S. & Brown, L. R. (1987). *J. Magn. Reson.* **72**, 347-352.

Farmer II, B. T., Macura, S. & Brown, L. R. (1988). *J. Magn. Reson.* **80**, 1-22.

Farmer, B. T. I., Muller, L., Nikonowicz, E. P. & Pardi, A. (1993). *J. Am. Chem. Soc.* **115**, 11040-11041.

Feigon, J. (1993). *Curr. Biol.* **3**, 611-613.

Feigon, J., Koshlap, K. M. & Smith, F. W. (1995). . In *Methods in Enzymology* (James, T. L., ed.), pp. 225-255. Academic Press, San Diego.

Felsenfeld, G. & Miles, H. T. (1967). *Annu. Rev. Biochem.* **31**, 407-448.

Fletcher, M. C., Fox, K. R. (1996). *Eur. J. Biochem.* **237**, 164-170.

Fox K. R. & Waring, M. J. (1984). *Nucl. Acid. Res.* **12**, 9271-9285.

Fox, K. R. & Howarth, N. R. (1985). *Nucleic Acids Res.* **13**, 8695-8714.

Frank-Kamenetskii, M. D. & Mirkin, S. M. (1995). *Annu. Rev. Biochem.* **64**, 65-95.

Frederick, C., Saal, D., van der Marel, G., van Boom, J., Wang, A. & Rich, A. (1987). *Biopolymers* **26**, S145-S160.

Fujii, S., Wang, A., van der Marel, G., van Boom, J. & Rich, A. (1982). *Nucleic Acids Res.* **10**, 7879-7892.

Fuller, W., Wilkins, M. H. F., Wilson, H. R., Hamilton, L. D. (1965) *J. Mol. Biol.* **12**, 60-80

Gao, X., Mirau, P. & Patel, D. J. (1992). *J. Mol. Biol.* **223**, 259-79.

Gao, X. & Patel, D. J. (1989). *Biochemistry* **28**, 751-62.

Gao, X. & Patel, D. J. (1990). *Biochemistry* **29**, 10940-56.

Gee, J. E. & Miller, D. M. (1992). *Am. J. Med. Sci.* **304**, 366-372.

Geierstanger, B. H. & Wemmer, D. H. (1995). *Annu. Rev. Biophys. Biomol. Struct.* **24**, 463-93.

Giessner-Prettre, C., Pullman, B., Tran-Dinh, S., Neumann, J.-M., Huynh-Dinh, T., Igolen, J. (1984). *Nucl. Acid. Res.* **12**, 3271-3281.

Gilbert, D. E. & Feigon, J. (1999). *Curr. Opin. Struc. Biol.* **9**, 305-314.

Ginell, S., Lessinger, L., Berman, H. M. (1988). *Biopolymers* **27**, 843-864.

Goldberg, I. H., Rabinowitz, M. & Reich, E. (1962). *Proc. Natl. Acad. Sci. USA* **48**, 2094.
Goodisman, J. & Dabrowiak, J. C. (1992). *Biochemistry* **31**, 1058-1064.

Goodisman, J., Rehfuss, R., Ward, B. & Dabrowiak, J. C. (1992). *Biochemistry* **31**, 1046-1058.

Gorenstein, D. G., Lai, K. & Shah, D. O. (1984). *Biochemistry* **23**, 6717-6723.

Gronenborn, A. M. & Clore, G. M. (1990). *Prog. NMR Spectrosc.* **17**.

Grosschedl, R. (1995) *Curr. Opin. Cell Biol.* **7**, 362-370

Hall, K., Cruz, P. & Chamberlin, M. J. (1985). *Arch. Biochem. Biophys.* **236**, 47-51.

Hare, D. R. & Reid, B. R. (1986). *Biochemistry* **25**, 5341-5350.

Havel, T. (1990). *Biopolymers* **29**, 1565.

Heus, H. A., Wijmenga, S. S., van de Ven, F. J. M. & Hilbers, C. W. (1994). *J. Am. Chem. Soc.* **116**, 4983-4984.

Hilbers, C. W., Heus, H. A., Van Dongen, M. J. P. & Wijmenga, S. S. (1994).. In *Nucleic Acids and Molecular Biology*, p. 56. Springer, Berlin.

Hosur, R. V., Govil, G. & Miles, H. T. (1988). *Mag. Res. in Chem.* **26**, 927-944.

Ikuta, S., Chattopadhyaya, R., Ito, H., Dickerson, R. E. & Kearns, D. R. (1986). *Biochemistry* **25**, 4840-4849.

Ivancic, M. & Hsu, V. L. (2000). *Biopolymers* **54**, 35-43.

Jones, G. P. (1966). *Phys. Rev.* **148**, 332.

Kamitori, S. & Takusagawa, F. (1992). *J. Mol. Biol.* **225**, 445-456.

Kamitori, S. & Takusagawa, F. (1994). *J. Am. Chem. Soc.* **116**(10), 4154-4165.

Kay, L. E., Keifer, P. & Saarinen, T. (1992). *J. Am. Chem. Soc.* **114**, 10663-10665.

Kettani, A., Kumar, R. A. & Patel, D. J. (1995). *J. Mol. Biol.* **254**, 638-656.

Kirk, J. M. (1960). *Biochem. Biophys. Acta* **42**, 167-169.

Kojima, C., Ono, A. & Kainosho, M. (2000). *J. Biomol. NMR* **18**, 269-277.

Krugh, T. R. (1972). *Proc. Natl. Acad. Sci. USA* **69**(7), 1911-1914.

Krugh, T. R. & Neely, J. W. (1973). *Biochemistry* **12**(9), 1775-1782.

Kumar, V. D., Harrison, R. W., Andrews, L. C. & Weber, I. T. (1992). *Biochemistry* **31**, 1541-1550.

Kumar, V. D. & Weber, I. T. (1993). *Nucleci Acids Res.* **21**(9), 2201-8.

Kyogoku, Y., Lord, R. C. & Rich, A. (1967). *Proc. Natl. Acad. Sci. USA* **57**, 250-257.

Lam, S. L. & Au-Yeung, S. C. F. (1997). *J. Mol. Biol.* **266**, 745-760.

Lavery, R. & Sklenar, H. (1988). *J. Biomol. Struct. Dynam.* **6**, 63-91.

Lefebvre, A., Mauffret, O., El Antri, S., Monnot, M., Lescot, E. & Fermandjian, S. (1995). *Eur. J. Biochem.* **229**, 445-454.

Legault, P., Farmer, B. T. I., Mueller, L. & Pardi, A. (1994). *J. Am. Chem. Soc.* **116**, 2203-2204.

Lewis, J. & Bird, A. (1991). *FEBS Letters* **285**(2), 155-159.

Lian, C., Robinson, H. & Wang, A. H.-J. (1996). *J. Am. Chem. Soc.* **118**, 8791-8801.

Liu, H., Spielmann, P., Ulyanov, N. B., Wemmer, D. E. & James, T. L. (1995). *J. Biomol. NMR* **6**, 390-402.

Liu, X., Chen, H. & Patel, D. (1991). *J. Biomol. NMR* **1**, 323-347.

Low, C. M. L., Drew, H. R. & Waring, M. J. (1984). *Nucleic Acids Res.* **12**, 4865-79.

Low, C. M. L., Fox, K. R., Olsen, R. K. & Waring, M. J. (1986). *Nucleic Acids Res.* **14**, 2015-33.

Luisi, B. (1995). In *DNA-protein Structural Interactions*, (D. Lilley, ed.) pp. 1-48, IRL Press, New York

Macura, S. & Ernst, R. R. (1980). *Mol. Phys.* **41**, 95.

MacDonald, D., Herbert, K., Zhang, X., Polgruto, T. & Ponzy, L. (2001). *J. Mol. Biol.* **306**, 1081-1098.

Marcourt, L., Cordier, C., Couesnon, T. & Dodin, G. (1999). *Eur. J. Biochem.* **265**, 1032-1042.

Marino, J. P., Schwalbe, H., Anklin, C., Bermel, W., Crothers, D. M. & Greisinger, C. (1994). *J. Am. Chem. Soc.* **116**, 6472-6473.

Marion, D. & Wüthrich, K. (1983). *Biochem. Biophys. Res. Commun.* **113**, 967-974.

Masse, J. E., Bortmann, P., Dieckmann, T. & Feigon, J. (1998). *Nuc. Acid. Res.* **26**(11), 2618-2624.

Mauger, A. B. (1980). The Actinomycins. In *Topics in Antibiotic Chemistry* (Sammes, P. G., ed.), Vol. 5, pp. 229-305. Ellis Horwood, Chichester, England.

Meienhofer, J. & Atherton, E. (1977) *Structure-Afinity Relationship among the Semisynthetic Antibiotics* (Pearlman, D., ed.), pp. 427-529; Academic Press, Inc., New York

Mooers, B., Schroth, R., Baxter, W. & Ho, P. (1995). *J. Mol. Biol.* **249**, 772-784.

Muller, W. & Crothers, D. M. (1968). *J. Mol. Biol.* **35**, 251-290.

Nikonowicz, E. P. & Gorenstein, D. G. (1990). *Biochemistry* **29**, 8845-8858.

Nikonowicz, E. P. & Pardi, A. (1992). *J. Am. Chem. Soc.* **114**, 9202-9203.

Nilges, M. (1993). *Proteins: Struct. Funct. Genet.* **17**, 297-309.

Nilges, M. (1995). *J. Mol. Biol.* **245**, 645-660.

Neuhaus, D. & Williamson, M. (1989) *The Nuclear Orevhauser Effect in Structural and Conformational Analysis*, pp. 170-173, VCH Publishers, Inc., New York

Nonin, S., Leroy, J.-L., Gueron, M. (1995). *Biochemistry* **34**, 10652-10659

Oda, Y., Uesugi, S., Ikehara, M., Kawase, Y. & Ohtsuka, E. (1991). *Nucleic Acids Res.* **19**(19), 5263-7.

Ono, A., Tate, S., Ishido, Y. & Kainosho, M. (1994). *J. Biomol. NMR* **4**, 581-586.

Orbons, L. P. M. & Altone, C. (1986). *Eur. J. Biochem.* **160**, 141-148.

Orbons, L. P. M., van der Marel, G. S., van Boom, J. H. & Altona, C. (1986). *Eur. J. Biochem.* **160**, 131-139.

Palmer III, A. G., Cavanagh, J., Wright, P. E. & Rance, M. (1991). *J. Magn. Reson.* **93**, 151-170.

Patel, D. J. (1974). *Biochemistry* **13**(11), 2396-2402.

Patel, D. J. (1976). *Biopolymers* **15**, 533-558.

Patel, D. J., Kozlowski, S. A., Rice, J. A., Broka, C., Itakura, K. (1981). *Proc. Natl. Acad. Sci. USA* **78**(12), 7281-7284.

Petersheim, M., Mehdi, S. & Gerlt, J. A. (1984). *J. Am. Chem. Soc.* **106**, 439-440.

Piotto, M., Saudek, V. & Sklenar, V. (1992). *J. Biomol. NMR* **2**, 661-665.

Rajagopal, P., Gilbert, D. E., van der Marel, G. A., van Boom, J. H. & Feigon, J. (1988). *J. Magn. Reson.* **78**, 526-537.

Razin, A. & Cedar, H. (1991). *Microbiol. Rev.* **55**, 451-458.

Reich, E., Franklin, R. M., Shatkin, A. J. & Tatum, E. L. (1961). *Science* **134**, 556-557.

Reid, B. (1986). . In *NMR in the Life Sciences* (Bradbury, E. M. & Nicolini, C., eds.). Plenum Press, New York.

Reid, D. G., Salisbury, S. A. & Williams, D. H. (1983). *Biochemistry* **22**, 1377-1385.

Sarma, M. H., Gupta, G. & Sarma, R. H. (1986). *Biochemistry* **25**, 3659-3665.

Sastry, M. & Patel, D. J. (1993). *Biochemistry* **32**, 6588-6604.

Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H. & Kaptein, R. (1984). *Biochemistry* **23**, 1371-1376.

Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S. J., Sorensen, O. W. & Greisinger, C. (1994). *J. Biomol. NMR* **4**, 301-306.

Schmieder, P., Ippel, J. H., van der Elst, H., van der Marel, G. A., van Boom, J. H., Altona, C. & Kessler, H. (1992). *Nucleic Acids Res.* **20**, 4747-4751.

Shinomiya, M., Chu, W., Carlson, R. G., Weaver, R. F. & Takusagawa, F. (1995). *Biochemistry* **34**, 8481-8491.

Sklenar, V. & Bax, A. (1987). *J. Am. Chem. Soc.* **109**, 7525-7526.

Sklenar, V., Miyashiro, H., Zon, G., Miles, T. & Bax, A. (1986). *FEBS Letters* **203**(1), 94-98.

Sklenar, V., Peterson, R. D., Rejante, M. R. & Feigon, J. (1994). *J. Biomol. NMR* **4**, 117-122.

Sklenar, V., Peterson, R. D., Rejante, M. R., Wang, E. & Feigon, J. (1993). *J. Am. Chem. Soc.* **115**, 12181-12182.

Sobell, H. M. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 5328-5331.

Sobell, H. M. & Jain, S. C. (1972). *J. Mol. Biol.* **68**, 21-34.

Sobell, H. M., Jain, S. C., Sakore, T. D. & Nordman, C. E. (1971). *Nature* **231**, 200-205.

Solomon, I. (1955). *Phys. Rev.* **99**, 559

Stefl, R., Trantirek, L., Vorlickova, M., Koca, J., Sklenar, V. & Kypr, J. (2001). *J. Mol. Biol.* **307**, 513-524.

Sun, J.-S., Garestier, T. & Helene, C. (1996). *Curr. Opin. Struct. Biol.* **6**, 327-333.

Takusagawa, F., Dabrow, M., S., N. & Berman, H. M. (1982). *Nature* **296**, 466-469.

Tate, S., Ono, A. & Kainosho, M. (1994). *J. Am. Chem. Soc.* **116**, 5977-5978.

Tippin, D., Ramakrishnan, B. & Sundaralingam, M. (1997). *J. Mol. Biol.* **270**, 247-258.

Tjandra, N., Omchinski, J. G., Gronenborn, A. M., Clore, G. M. & Bax, A. (1997). *Nature Struct. Biol.* **4**(9), 732-738.

Tjandra, N., Tate, S., Ono, A., Kainosho, M. & Bax, A. (2000). *J. Am. Chem. Soc.* **122**, 6190-6200.

Tolman, J. R., Flanagan, J. M., Kennedy, M. A. & Prestegard, J. H. (1995). *Proc. Natl. Acad. Sci. USA* **92**, 9279-9283.

Tonelli, M., James, T. L. (1998). *Biochemistry* **37**, 11478-11487.

Trantirek, L., Stefl, R., Vorlickova, M., Koca, J., Sklenar, V. & Kypr, J. (2000). *J. Mol. Biol.* **297**, 907-922.

Uesugi, S., Oda, Y., Ikehara, M., Kawase, Y. & Ohtsuka, E. (1987). *J. Biol. Chem.* **262**(15), 6965-8.

Ulrich, E. L., John, E.-M. M., Gough, G. R., Brunden, M. J., Gilham, P. T., Westler, W. M., Markley, J. L. (1983). *Biochemistry* **22**, 4362-4365.

Van de Ven, F. J. & Hilbers, C. W. (1988). *Eur. J. Biochem.* **178**, 1.

Van Dongen, M. J. P., Wijmenga, G. A., Van der Marel, J. H., Van Boom, J. H. & Hilbers, C. W. (1996). *J. Mol. Biol.* **263**, 715.

Van Dyke, M. W. & Dervan, P. B. (1983). *Biochemistry* **22**, 2372-77.

Vermeulen, A., Zhou, H. & Pardi, A. (2000). *J. Am Chem. Soc.* **122**, 9638-9647.

Voet, D. & Voet, J. G. (1990) In *Biochemistry*, pp. 795-805, John Wiley & Sons, Inc. , USA

Watson, J.D. & Crick F.H. C. (1953). *Nature* **171**, 737-738.

Wang, E. & Feigon, J. (1999). . In *Oxford Handbook of Nucleic Acid Structures* (S., Neidle, ed.), pp. 355-388. Oxford University Press, New York.

Waring, M. J. (1970). *J. Mol. Biol.* **54**, 247-279.

Waterloh, K. & Fox, K. R. (1991). *Journal of Biological Chemistry* **266**(10), 6381-6388.

Wells, R. D. & Larson, J. E. (1970). *J. Mol. Biol.* **49**, 319-342.

Westernik, H. P., van der Marel, G. A., van Boom, J. H. & Haasnoot, C. A. G. (1984). *Nuc. Acid. Res.* **12**, 4323-4338.

Wijmenga, S. S., Mooren, M. M. & Hilbers, C. W. (1993). In *NMR of Macromolecules: A Practical Approach* (Roberts, G. C. K., ed.), p.217-288. Oxford University Press, New York.

Wijmenga, S. S. & van Buuren, B. N. M. (1998). *Prog. in NMR Spect.* **32**, 287-387.

Williamson, J. R., Raghuraman, M. K. & Cech, T. R. (1989). *Cell* **59**, 871-880.

Wüthrich, K., Billeter, M. & Braun, W. (1983). *J. Mol. Biol.* **169**, 949-961.

Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., USA

Xuan, J. C. & Weber, I. T. (1992). *Nucleic Acids Res.* **20**(20), 5457-64.

Zhu, L. & Reid, B. R. (1995). *J. Magn. Reson.*, **Series B 106**, 227-235

Appendix 1

Protocol for DNA hexamer purification

The DNA hexamers were synthesized via the phosphoramidite method at the Center for Gene Research and Biotechnology at Oregon State University. Each DNA sequence was synthesized in 3 μmol quantities and purified using polyacrylamide gel electrophoresis in the following manner. The sequences were hexamers, thus a 30% polyacrylamide and 8M urea concentration were used for running the preparatory gel. Two gels of 10 cm by 20 cm by 0.5 cm each were run concurrently using a Hoefer SE600 series gel apparatus (Amersham Pharmacia Biotech). Each gel was loaded with 0.5 μmol of DNA and poly acrylamide gel electrophoresis (PAGE) was performed for approximately 4 to 5 hours. At the end of the run, the DNA band was detected via UV shadowing and excised. Extraction was performed using electroelution in a 1X tris-borate ethylenediamine-tetraacetic acid (TBE) buffer (89 mM tris, 89 mM boric acid, 2 mM EDTA). Electroelution was run at 150 V for 3 hours and 80 V overnight, while the DNA was trapped between a BT2 membrane (Schelicher & Schuell) and two M1000 membranes (Millipore regenerated cellulose). Residual acrylamide and TBE were eliminated via a diethylaminoethyl (DEAE) Sephacel anion exchange column. After loading the sample, contaminants were rinsed off of the column by use of a low salt buffer (10 mM sodium phosphate, pH 7.6, 1 mM ethylenediamine-tetraacetic acid (EDTA), and 0.2 M NaCl), and the sample

was eluted with a high salt buffer (10 mM sodium phosphate buffer, pH7.6 1 mM EDTA, and 1.5 M NaCl). The last step in the purification process was the use of a C18 cartridge to desalt the sample. The C18 Sep Pak cartridges (Waters) were activated by 10 ml methanol followed by 10 ml ddH_2O. The sample was then loaded, the column rinsed with 5 ml H_2O, and the DNA eluted in 3 times 1 ml aliquots of 60% methanol. The sample was lyophilized two or three times to purge the methanol and suspended in 500 µl ^2H_2O. The purity of the sample was checked using a ^1H 1D spectrum.

Appendix 2

^{1}H and ^{31}P assignments of free DNA

The following seven tables are ^{1}H and ^{31}P assignments of the free DNA hexamers for which the structures were not solved. For these sequences the orientations of ActD binding were about 57% to 43% with the exception of the d(CCGCCG)•d(CGGCGG) and d(CTGCGG)•d(C^{5me}CGCAG) sequences, with the exact values listed in Table III.6 of Chapter III. These DNA hexamers were assigned prior to complexation with ActD.

Table A2.1. ^{1}H and ^{31}P assignments of d(CAGCAG)•d(CTGCTG) reported in ppm*.

nucleotide	H8/H6	H5/H2/Me	H1'	H2'/H2"	H3'	H4'	amino	imino	^{31}P
C1	7.66	5.94	5.60	1.88/2.34	4.68	4.02	7.10/8.23		-3.50
A2	8.29	7.80	5.96	2.78/2.88	5.01	4.36	6.56/7.98		-3.65
G3	7.73		5.77	2.51/2.63	4.95	4.37		12.82 (H1)	-3.54
C4	7.32	5.34	5.49	1.88/2.24	4.78	4.11	6.45/8.38		-3.49
A5	8.18	7.79	6.04	2.68/2.85	5.00	4.38	6.49/8.03		-3.58
G6	7.72		6.02	2.43/2.23	4.61	4.16		13.18 (H1)	n/a
C7	7.84	5.93	5.91	2.13/2.56	4.67	4.10	7.20/7.88		-3.73
T8	7.51	1.68	5.79	2.21/2.54	4.89	4.18		14.04 (H3)	-3.56
G9	7.96		5.94	2.70/2.70	4.98	4.40		12.82 (H1)	-3.49
C10	7.46	5.37	5.97	2.02/2.45	4.75	4.22	6.69/8.28		-3.77
T11	7.35	1.68	5.82	1.94/2.33	4.83	4.10		14.23 (H3)	-3.52
G12	7.94		6.16	2.61/2.34	4.66	4.15		12.94 (H1)	n/a

*ppm values are referenced to the solvent resonance signal at the appropriate temperature

Table A2.2 ^{1}H and ^{31}P assignments of d(CAGCCG)•d(CGGCTG) reported in ppm.

nucleotide	H8/H6	H5/H2/Me	H1'	H2'/H2''	H3'	H4'	aminos	iminos	^{31}P
C1	7.62	5.90	5.53	1.82/2.30	4.65	4.01			-3.85
A2	8.27	8.40	5.94	2.77/2.86	5.00	4.36			-4.01
G3	7.75		5.78	2.57/2.61	4.97	4.40	7.39/7.80	12.82	-3.91
C4	7.35	5.28	5.95	1.99/2.41	4.77	4.19	6.32/8.15		-4.00
C5	7.45	5.63	5.57	1.97/2.29	4.78	4.07	7.00/8.67		-3.60
G6	7.93		6.15	2.62/2.35	4.66	4.17			
C7	7.56	5.84	5.68	1.83/2.33	4.64	4.03			-3.85
G8	7.92		5.51	2.69	4.95	4.30	7.44/8.15	13.21	-3.74
G9	7.83		5.97	2.64/2.74	4.98	4.44	7.00/7.35	13.00	-4.02
C10	7.40	5.32	5.95	1.99/2.45	4.77	4.25	6.59/8.22		-4.16
T11	7.32	1.67	5.80	1.94/2.33	4.82	4.10			-3.88
G12	7.93		6.15	2.62/2.35	4.66	4.17			

*ppm values are referenced to the solvent resonance signal at the appropriate temperature

Table A2.3 ^{1}H and ^{31}P assignments of d(CAGCGG)•d(CCGCTG) reported in ppm*.

nucleotide	H8/H6	H5/H2/Me	H1'	H2'/H2"	H3'	H4'	aminos	iminos	^{31}P
C1	7.66	5.94	5.58	1.87/2.33	4.67	4.02			-3.21
A2	8.29	8.41	5.94	2.79/2.88	5.01	4.36			-3.36
G3	7.75		5.78	2.54/2.64	4.97	4.38		12.82	-3.25
C4	7.25	5.30	5.66	1.78/2.24	4.79	4.13	8.35/6.41		-3.26
G5	7.84		5.60	2.63/2.69	4.94	4.31		13.24	-3.05
G6	7.79		6.15	2.51/2.32	4.62	4.20			
C7	7.74	5.92	5.66	2.05/2.51	4.65	4.10			-3.25
C8	7.54	5.66	5.57	2.15/2.44	4.85	4.12	8.65/6.99		-3.03
G9	7.95		5.96	2.71/2.71	4.99	4.39		13.04	-3.32
C10	7.44	5.39	5.98	2.02/2.46	4.76	4.21	8.30/6.64		-3.49
T11	7.34	1.70	5.82	1.94/2.34	4.83	4.10			-3.23
G12	7.94		6.16	2.62/2.35	4.67	4.17			

*ppm values are referenced to the solvent resonance signal at the appropriate temperature

Table A2.4 ^{1}H and ^{31}P assignments of d(CTGCCG)•d(CGGCAG) reported in ppm*.

nucleotide	H8/H6	H5/H2/Me	H1'	H2'/H2"	H3'	H4'	iminos	aminos	^{31}P
C1	7.84	5.94	5.91	2.13/2.56	4.67	4.09			-3.07
T2	7.50	1.69	5.77	2.20/2.52	4.88	4.16			-3.55
G3	7.95		5.94	2.68/2.72	4.99	4.39		12.87	-3.59
C4	7.41	5.39	5.98	2.05/2.42	4.80	4.21	6.45/8.31		-3.60
C5	7.47	5.66	5.58	1.99/2.31	4.80	4.07	7.06/8.76		-3.27
G6	7.95		6.17	2.62/2.34	4.66	4.16			
C7	7.58	5.89	5.72	1.87/2.36	4.67	4.03			-3.53
G8	7.93		5.50	2.73	4.96	4.28		13.26	-3.38
G9	7.81		5.92	2.60/2.72	4.99	4.41		13.01	-3.71
C10	7.34	5.40	5.46	1.94/2.26	4.81	4.09	6.47/8.46		-3.45
A11	8.18	7.81	6.03	2.70/2.87	5.00	4.38			-3.60
G12	7.71		6.02	2.43/2.24	4.61	4.16			

*ppm values are referenced to the solvent resonance signal at the appropriate temperature

Table A2.5 ^{1}H and ^{31}P assignments of d(CCGCCG)•d(CGGCGG) reported in ppm*.

nucleotide	H8/H6	H5/H2/Me	H1'	H2'/H2"	H3'	H4'	aminos	iminos	^{31}P
C1	7.74	5.93	5.97	2.05/2.51	4.65	4.10			-3.33
C2	7.53	5.66	5.56	2.14/2.43	4.85	4.11	7.05/8.74		-3.10
G3	7.94		5.93	2.72	4.99	4.38		13.08	-3.42
C4	7.40	5.41	5.99	2.04/2.43	4.81	4.20	6.41/8.31		-3.37
C5	7.47	5.68	5.57	1.99/2.31	4.80	4.08	6.97/8.66		-3.00
G6	7.95		6.17	2.63/2.35	4.66	4.17			n/a
C7	7.58	5.88	5.71	1.86/2.35	4.67	4.03			-3.29
G8	7.94		5.49	2.73	4.96	4.28		13.26	-3.11
G9	7.84		5.95	2.61/2.73	5.00	4.41		13.05	-3.48
C10	7.27	5.36	5.65	1.84/2.26	4.81	4.11	6.46/8.44		-3.28
G11	7.84		5.62	2.64/2.70	4.95	4.32		13.27	-3.13
G12	7.79		6.15	2.51/2.32	4.63	4.20			n/a

*ppm values are referenced to the solvent resonance signal at the appropriate temperature

Table A2.6 ^{1}H and ^{31}P assignments of free d(CTGCGG)•d(C^{5me}CGCAG) obtained from NMR spectra recorded at 298K and reported in ppm*.

nucleotide	H8/H6	H5/H2/Me	H1'	H2'/H2"	H3'	H4'	amino	imino	^{31}P
C1	7.83	5.96	5.94	2.14/2.57	4.69	4.11			-3.56
T2	7.51	1.72	5.79	2.22/2.52	4.89	4.18			-3.40
G3	7.94		5.93	2.67/2.74	4.99	4.38		12.88	-3.46
C4	7.32	5.43	5.67	1.86/2.30	4.81	4.15	6.45/8.48		-3.29
G5	7.83		5.68	2.63/2.71	4.96	4.32		13.28	-3.26
G6	7.79		6.15	2.50/2.33	4.63	4.19			
C7	7.81	5.96	5.95	2.03/2.55	4.67	4.11			-3.47
mC8	7.41	1.79	5.62	2.15/2.43	4.86	4.13	6.61/8.85		-3.22
G9	7.91		5.89	2.68	4.99	4.37		13.02	-3.46
C10	7.37	5.47	5.52	1.92/2.26	4.81	4.11	6.45/8.46		-3.32
A11	8.18	*	6.04	2.70/2.86	5.01	4.38		13.98	-3.44
G12	7.72		6.02	2.44/2.27	4.62	4.16			

*ppm values are referenced to the solvent resonance signal at the appropriate temperature

Table A2.7 ^{1}H and ^{31}P assignments of free d(CTGCIG)•d(CCGCAG) obtained from NMR spectra recorded at 298K reported in ppm. The amino and imino proton resonances were obtained from NOESY spectra recorded in H_2O at 283K reported in ppm*.

nucleotide	H8/H6	H5/H2/Me	H1'	H2'/H2"	H3'	H4'	amino	imino	^{31}P
C1	7.82	5.95	5.96	2.15/2.54	4.68	4.10	7.17/7.90		-3.49
T2	7.50	1.74	5.76	1.97/2.39	4.83	4.12			-3.39
G3	7.98		5.88	2.69	4.97	4.39		12.97	-3.33
C4	7.38	5.48	5.54	1.88/2.25	4.76	4.17	6.55/8.46		-3.41
I5	8.17	*	6.07	2.71/2.86	4.98	4.36			-3.50
G6	7.74		6.01	2.46/2.29	4.63	4.15			
C7	7.72	5.92	5.99	2.08/2.51	4.65	4.11			-3.32
C8	7.51	5.71	5.82	2.20/2.50	4.87	4.18			-3.41
G9	7.95		5.92	2.68	4.98	4.37		12.86	-3.34
C10	7.36	5.48	5.54	1.86/2.21	4.78	4.10	6.67/8.45		-3.29
A11	8.19	*	5.97	2.69/2.78	4.97	4.35			-3.33
G12	7.80		6.07	2.53/2.32	4.65	4.14		13.18	

*ppm values are referenced to the solvent resonance signal at the appropriate temperature

Appendix 3

Assignments of the complexes with modified DNA bases

The following tables are ^{1}H and ^{31}P assignments of the DNA hexamers with modified bases in complex with ActD, this includes the d(CTGCGG)•d(C^{5me}CGCAG) and d(CTGCIG)•d(CCGCAG) sequences. The last two tables include the few key assignments of ActD in complex with the above mentioned sequences.

Table A3.1 The DNA ^{1}H and ^{31}P assignments of the ActD:d(CTGCGG)•d(C^{5me}CGCAG) complex obtained from spectra recorded at 298K and reported in ppm.

nucleotide	complex	H8/H6	H5/H2/Me	H1'	H2'/H2''	H3'	H4'	amino	imino	^{31}P
C1	MAJOR	7.73	5.92	5.99	1.91/2.43	4.63	4.07			-3.84
	minor	n/a								
T2	MAJOR	7.22	1.76	5.64	1.61/1.76	4.74	3.98			-3.91
	minor	n/a								
G3	MAJOR	7.65		5.67	2.63/2.77	4.95	3.99	7.24/7.93	12.79	-1.82
	minor	n/a								-0.88
C4	MAJOR	7.28	5.90	6.03	1.25/1.73	4.41	4.17	6.83/7.89		-3.91
	minor	7.31	5.77	5.81	1.79/2.08	4.51	3.97	6.79/8.01		-3.56
G5	MAJOR	7.85		5.25	2.55/2.66	4.89	4.28	7.07/7.90	13.26	-3.25
	minor	7.87		5.46	2.66	4.91	3.99			
G6	MAJOR	7.82		6.14	2.56/2.35	4.63	4.15			
	minor	7.89		5.90	2.62/2.56	4.62	4.07			
C7	MAJOR	7.73	5.92	5.98	1.85/2.43	4.63	4.07			-3.83
	minor	7.72	5.99	5.99	1.91/2.41	4.61	4.06			
mC8	MAJOR	7.03	1.84	5.74	1.31/1.67	4.72	3.96	6.72/8.99		-3.91
	minor	7.11	1.83	5.64	1.43/1.73	4.70	3.97			
G9	MAJOR	7.66		5.61	2.55/2.62	4.90	3.99	7.30/8.18	12.45	-1.02
	minor	7.61		5.63	2.60/2.70	4.91	3.99	7.29/8.06	12.67	-2.18
C10	MAJOR	7.45	5.80	5.80	2.02/2.20	4.55	4.04	6.72/8.08		-3.67
	minor	7.36	5.97	5.86	1.36/1.73	4.42	4.17			
A11	MAJOR	8.20	7.87	5.87	2.70	4.98	4.22			
	minor	8.21		5.67	2.69	4.93	4.22			
G12	MAJOR	7.71		5.67	2.70/2.59	4.91	3.99			
	minor	n/a		n/a	n/a		n/a			

Table A3.2 The DNA ^{1}H and ^{31}P assignments of the ActD:d(CTGCIG)•d(CCGCAG) complex obtained from spectra recorded at 308K and reported in ppm.

nucleotide	complex	H8/H6	H5/H2/Me	H1'	H2'/H2''	H3'	H4'	amino	imino	^{31}P
C1	MAJOR	7.76	5.94	6.02	2.02/2.46	4.68	4.11			-3.76
	minor	7.72	5.92	6.03	1.97/2.44	4.65	4.12			-3.75
T2	MAJOR	7.22	1.79	5.71	1.59/1.79	4.77	4.01			-3.83
	minor	7.14	1.84	5.76	1.69/1.84	4.77	4.01			-3.90
G3	MAJOR	7.65		5.67	2.63/2.75	4.98	3.93	7.28/8.00	12.80	-2.01
	minor	7.72		5.68	2.62/2.69	4.94	3.96	7.38/7.75	12.61	-0.89
C4	MAJOR	7.35	5.98	5.96	1.39/1.73	4.44	4.22	6.87/7.92		-3.67
	minor	7.38	5.99	5.90	1.43/1.79	4.45	3.98	7.08/8.06		-3.74
I5	MAJOR	8.23		5.95	2.69	4.94	4.02			-3.12
	minor	8.21		5.76	2.68	4.97	4.03			-3.42
G6	MAJOR	n/a								
	minor	n/a								
C7	MAJOR	7.72		6.03	2.05/2.38	4.66				-3.54
	minor	7.73	5.92	6.04	1.97/2.34	4.66	4.12			-3.55
C8	MAJOR	7.19	5.83	5.80	1.25/1.71	4.70	3.97			-3.93
	minor	7.21	5.77	5.79	1.27/1.79	4.68	4.01			-3.90
G9	MAJOR	7.72		5.68	2.61/2.69	4.95	3.89	7.29/7.79	12.40	-0.99
	minor	7.67		5.67	2.64/2.75	4.97	3.87	7.41/7.74	12.29	-1.98
C10	MAJOR	7.46	5.83	5.44	2.04/2.21	4.56	4.07	8.13/6.71		-2.90
	minor	7.41	5.82	5.89	1.95/2.14	4.57	4.08	8.14/6.93		-2.86
A11	MAJOR	8.22		5.94	2.69	5.01	4.24			-3.62
	minor	8.24		5.80	2.67	4.98	4.26			-3.63
G12	MAJOR	7.78		6.03	2.59/2.35	4.66	4.05			
	minor	7.75		5.97	2.55/2.34	4.65	4.06			

Table A3.3 A few key ^{1}H assignments of ActD in the ActD:d(CTGCGG)•d(C^{5me}CGCAG) complex.

complex11	6CH3	4CH3	H7	H8	Thr-γH
MAJOR	2.04	1.68	6.64	7.12	1.39
minor	2.02	1.75	6.61	7.04	1.36

Table A3.4 Several ^{1}H assignments of ActD in the ActD:d(CTGCIG)•d(CCGCAG) complex.

complex12	6CH3	4CH3	H7	H8	Thr-γH
MAJOR	2.05	1.71	6.64	7.11	1.40
minor	2.04	1.78	6.62	7.09	1.39

Appendix 4

Details of the assignments of the cyclic pentapeptide rings of ActD in the complex

Sequential connectivities within the individual cyclic pentapeptide lactone rings were identified from the NOE crosspeak patterns in the NOESY spectrum of the complex. As mentioned in chapter 3 of this thesis, differences between the chemical shifts of the chromophore H8/H7 protons in the two complexes (Figure 8) make it possible to distinguish between the two complexes formed. The major complex exhibits several correlations between the chromophore H8 proton (7.10 ppm) and the G3/C4 residues of the DNA, while the minor complex exhibits connectivities between the chromophore H8 proton (7.03 ppm) and the G9/C10 residues of the DNA. The benzenoid and quinoid pentapeptide lactone rings can be assigned via the Thr methyl groups on the benzenoid and quinoid side, which are distinguished by observing the NOE between the chromophore's H8 and H7 protons and the methyl group of Thr on the benzenoid ring only. The benzenoid Thr γ-methyl group resonates at 1.36 ppm, while the quinoid Thr methyl group resonates at 1.33 ppm. The distinction between these two chemical shifts can be followed to the Thr Hα chemical shifts, locating the benzenoid Thr Hα at 4.87 ppm (major) and 4.71 ppm (minor), and the quinoid Thr Hα at 4.60 ppm (major) and 4.70 ppm (minor). This in turn can be followed to the crosspeaks of Thr Hβ, finding the benzenoid resonances at 5.17/5.13 ppm and both quinoid resonances at 5.10 ppm.

Other NOEs exist between the Thr γ-methyl groups and other protons of the peptide chain, these being the DVal γ-methyl groups and the N-methylvaline (NMV) γ-methyl groups as well as the NMV Hα and Hβ protons. These NOEs are observed for both the pentapeptide rings. Additional NOEs are observed across the benzenoid and quinoid pentapeptide lactone rings, from the benzenoid side Thr γ-methyl to the quinoid side Pro Hδ, Hγ and Hα, and vice versa. NOEs are also observed from the downfield shifted Pro Hβ resonances to the C4H1′ and C10H1′ in both complexes. However C4H1′ in the major complex is spatially proximal to a Hβ proton of the Pro in the quinoid chain, while C10H1′ is near the Hβ proton of the benzenoid chain Pro. For the minor complex, the situation is reversed, yet the severe overlap in the H1′ region of ActD:d(CTGCGG)•d(CCGCAG), makes these crosspeaks difficult to distinguish from several other crosspeaks. In this manner it is possible to identify four different downfield shifted Pro Hβ protons, each of which belongs to a different complex and to a different peptide chain. NOEs between the Pro Hα protons and Pro Hβ protons, clearly identify the benzenoid Pro Hα at 6.17 ppm, a higher field than the quinoid Pro Hα at 6.23 ppm. This is in agreement with several previous investigations (Brown et al., 1994; Chen et al., 1996; Delepierre, 1989; Liu et al., 1991).

The NOE between the Pro Hα proton and the DVal Hα proton, which is very intense, and the NOEs between Pro Hβ and both of DVal γ-methyl protons, permit the assignment of the four DVal spin systems. NOEs between the Pro Hα protons and both Sar Hα protons allow the assignments of the corresponding Sar resonances. The Sar N-methyl protons are also identified via NOEs to Pro Hα,

Hγ and one of the Pro Hβ protons. The Pro Hα protons show NOEs to the N-methyl groups of Sar and N-methyl groups of NMV. Yet because of spectral overlap in this region, the carbon chemical shifts obtained from a ^{1}H-^{13}C HSQC experiment were used to distinguish between the different N-methyl groups. The NMV N-methyl group carbon resonance occurs around 40 ppm, while the Sar N-methyl group carbon resonates around 35 ppm (Brown et al., 1994). Using this $^{1}H^{13}C$ HSQC experiment, the chemical shifts of the protons in the N-methyl groups were assigned. Knowing the resonances of the N-methyl NMV groups, the rest of the NMV spin system was possible to assign. Here additional NOEs between the NMV N-methyl group and the Pro Hα proton as well as the Sar N-methyl groups proved to be helpful. The weak NOEs between the Thr γ-methyl protons and the N-methyl groups of NMV confirm the assignments of the different Thr spin systems. In the manner described here, all the nonexchageable protons in the cyclic pentapeptide lactone moieties of both complexes were identified and are listed in Table III.7 of Chapter III.

The exchangeable protons of ActD in the complex are a pair of amide protons associated with each of the LThr and DVal residues on the benzenoid and quinoid rings. The NH protons of DVal on the benzenoid and quinoid rings were assigned to resonances at 8.07 and 8.13 ppm, respectively, based on NOE crosspeaks to their own Hα protons, Hβ protons, and γ-methyl protons. The NH protons of L-Thr on the benzenoid and quinoid rings were assigned to resonances at 7.91 and 7.99 ppm based on their NOEs to the corresponding Hα protons, Hβ protons, γ -methyl protons and some intermolecular NOEs discussed in Chapter 3.