PS04.04.11 DAUNOMYCIN INDUCES DEGLYCOSYLATION OF DNA: STRUCTURE OF d(CG[glucose-T]ACG)-DAUNOMYCIN COMPLEX. YiGui Gao, Howard Robinson, +Jacques H. van Boom, Andrew H.-J. Wang, Biophysics Division & Dept. of Cell & Structural Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA, +Leiden Institute of Chemistry, Gorlaeus Laboratory, 2300 RA Leiden, The Netherlands

DNA from certain organisms contains unusually-modified DNA bases. For example, glycosylated-T or C at C5 position of the base is found in *Trypanosoma Brucei*. The role of those modified bases remains largely unknown, although in trypanosome they are believed to be involved in the regulation of gene expression. Ethidium bromide, an intercalator, has been shown to have an cytotoxic effect towards trypanosomes.

We have undertaken a structural study in which intercalator anticancer drug daunomycin was added to the modified DNA hexamer d(CG[glucose-T]ACG). The solution structure of the hexamer was determined by NMR. In addition, the three-dimensional molecular and crystal structure of the complex of daunomycin and the hexamer was determined at 1.7Å by X-ray diffraction analyses. Crystal data: *P1*, $a = 18.63$ Å, $b = 20.01$ Å, $c = 26.54$ Å, $\alpha = 69.30°$, $\beta = 90.31°$, $\gamma = 108.16°$, R=17.9%, 3385 reflections at 2σ. Two daunomycin molecules bind to the DNA double helix. Unexpectedly we found that in the crystal structure one the glucose moieties is missing from the DNA. The biological implication of this will be addressed.

This work is supported by American Cancer Society DHP-114 to AHJW.

PS04.04.12 SURVEYING AND CLASSIFYING NUCLEIC ACID STRUCTURES. Anke Gelbin, Les Clowney, Shu-Hsin Hsieh, Christine Zardecki, John Westbrook and Helen M. Berman, Department of Chemistry, Rutgers University, Piscataway NJ 08855.

Nucleic acids are highly flexible molecules that can assume a wide variety of shapes. A major challenge is to be able to relate the sequence and experimental conditions to the conformations. There are now over 400 structures of nucleic acid-containing crystals stored in the Nucleic Acid Database archive. The variety of ways in which these molecules can be classified yield different types of information that can be used to help relate the structures to their sequences and other properties. The tools that have been developed for classification and analysis, as well as some structure survey results, will be presented.

Support for this work has been provided by the NSF BIR 9510703.

PS04.04.13 POLYPURINE TRACT DNA SEQUENCE, CAAAGAAAAG CONTAINING 2 MOLECULES IN ASYMMETRIC UNIT. IS THE 2ND HELIX DISORDERED? Gye Won Han, Mary L. Kopka, David S. Goodsell and Richard E. Dickerson, Molecular Biology Institute, Department of Chemistry and Biochemistry, University of California at Los Angeles, CA 90095, USA

The DNA duplex decamer CAAAGAAAAG : CTTTTCTTTG was synthesized and crystallized to explore the structure of the polypurine tract of HIV-1. Although the polypurine tract consists of 16 purines, initially a 10-mer was used because of stacking considerations. A 15-mer also was synthesized for crystallization. This sequence is highly conserved in HIV, thus making an attractive target for drug design.

The native crystal is monoclinic in space group C2 with 2 helices (A and B) in asymmetric unit. Cell dimensions are a=59.82 Å, b=28.28 Å, c=72.49 Å and β=103.9°. X-ray data were collected using RAXIS-II at -180°C up to 1.85 Å resolution.

Using single isomorphous replacement methods with CAAAGAAAAG : $CTU_{Br}TTCTTTG$ gives only one Br site in difference Patterson map, and this is consistent with the molecular replacement result where a clear hydration spine appeared in helix A, but not in the 2nd helix.

Diffusion of cisplatin (cis-dichlorodiamino platinum (II)) into the native crystal once again gives only one Pt site in helix A in the Difference Fourier map using SIR_{Br} phases. We believe the 2nd helix may be disordered and in order to improve the phases, a second derivative CAAAGAAAAG : $CTTTTC_{Br}TTTG$ for MIR method will be used.

PS04.04.14 POTENTIALLY RIGHT HANDED SEQUENCE CRYSTALLIZES AS LEFT HANDED DNA: THE CRYSTAL STRUCTURE OF d(CCCGGG). P. Karthe[1], S. Krishnaswamy[2] & N. Gautham[1], [1]Department of Crystallography & Biophysics, University of Madras, Guindy Campus, Madras 600 025, India; [2]Bioinformatics Centre, School of Biotechnology, Madurai Kamaraj University, Madurai - 625 021, India

The DNA duplex d(CCCGGG).d(CCCGGG) has only one alternating pyrimidine- purine base step. Despite this, it crystallizes as a left handed helix and packs into a orthorhombic unit cell with a=17.76, b=30.92, c=43.92 Å, similar to the one observed previously for Z DNA hexamers. Moreover, the structure exhibits several remarkable features that are not hitherto observed in left handed Z DNA. The most stricking of these is that the successive base pairs with in the central tetranucleotide show uniform values of twist and rise, resulting in a novel uniform left handed DNA double helix. This structure thus demonstrates that DNA can take up a left handed conformation in the absence of stretches of alternating pyrimidine-purine sequences, and also that like right handed DNA, left handed DNA too can exist in polymorphic forms.

PS04.04.15 A NOVEL [G-(G.C)] BASE-TRIPLET: MODEL FOR BASE PAIR RECOGNITION DURING HOMOLOGOUS RECOMBINATION. Blaine H. M. Mooers & P. Shing Ho, Dept. of Biochemistry & Biophysics, Oregon State Univ., Corvallis, OR 97331

In the crystal structure of the nonamer d(GCGTACGCG), the 3'-terminal guanine forms a base-triplet in the minor groove of B-DNA that is consistent with recent biochemical evidence about the structure of RecA-DNA triplex complexes. During homologous recombination, RecA protein polymerizes on single-stranded DNA, and the resulting nucleoprotein filament incorporates a homologous region of double-stranded DNA into a RecA protein coated DNA triplex, otherwise known as "R-DNA". R-DNA has an extended and unwound conformation. The third strand in this enzymatically formed triplex is parallel to the homologous strand in the Watson-Crick duplex, and it has been generally thought to lie in the major groove of the DNA duplex. However, a recent study by Baliga et al., Proc. Natl. Acad. Sci., USA, 92, 10393-10397 (1995), indicates that the third DNA strand lies in the minor groove of the parent duplex. None of the available structures for DNA triplets, however, provide a model for how the third strand can recognize the homologous duplex in the minor groove. We present the 2.5 Angstrom structure of a base-triplet which has features common to R-DNA in that the third base sits in the minor groove parallel to the homologous strand in the Watson-Crick duplex. In the crystal lattice, the first eight nucleotides of the nonamer sequence form a standard B-DNA duplex with a complimentary strand. These duplexes stack end-to-end but fail to form continuous helices because the terminal base pairs of adjacent stacks are underwound with respect to each other. This underwinding places

the orphaned 3U-terminal end in the minor groove of the adjacent duplex where it symmetrically pairs with the 5'-terminal guanine to form a d[G*(G.C)] base-triplet. Our findings extend to the minor groove a DNA hydrogen bonding pattern which permits base-pair recognition during homologous recombination.

PS04.04.16 PHASING IN DRUG-DNA SEQUENCE RECOGNITION: STRUCTURE OF A TRIS (BENZIMIDAZOLE) - DNA COMPLEX. Stephen Neidle, George R Clark+, Emily J Gray, Yu-Hua Li# and Werner Leupin#, The CRC Biomolecular Structure Unit, The Institute of Cancer Research, Sutton, Surrey, SM2 5NG, UK, #F. Hoffman-la-Roche Ltd, Preclinical Research Pharma Gene Technologies, CH-4002 Basel, Switzerland, +Permanent address: Chemistry Department, University of Auckland, New Zealand

Effective recognition of a DNA sequence longer than ca 3-4 base pairs in length requires drug and base pairs to be in register along the complete length of the drug. We have studied this problem in the context of the crystal structure of a complex between a tris(benzimidazole) drug and the oligonucleotide duplex $d(CGCAAATTTGCG)_2$. This has been determined to 2.2 Å resolution and refined to an R of 17.4%. The drug is bound in the minor groove region and covers ca 7? base pairs. There is an extensive set of hydrogen bonds between the imidazole rings and N_3/O_2 atoms of the A:T base pairs. These have exceptionally high propeller twists, and five out of the six have lost one of their two Watson-Crick hydrogen bonds; this is compensated for by a series of major-groove three-centre hydrogen bonds.

The drug itself is highly twisted in order to achieve maximum hydrogen- bonding register with the A:T base pair edges. The DNA is deformed beyond what has been observed in other minor-groove drug crystal structures, with evidence of local helix unwinding and extension. These changes are necessary for effective DNA recognition of every benzimidazole sub-unit to take place, and thus for each to be in phase with base pairs.

PS04.04.17 CRYSTAL STRUCTURE OF A 1:1 COMPLEX BETWEEN NETROPSIN AND $D(CGCAATTGCG)_2$. Christine M. Nunn, Neil Spink, *Elspeth Garman, Stephen Neidle, CRC, Biomolecular Structure Unit, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, SM2 SNG, UK, *Laboratory of Molecular Biophysics, University of Oxford, Oxford, OX1 3QU, UK

DNA minor groove binding drugs such as netropsin, distamycin and pentamidine have been extensively studied bound to DNA dodecamer sequences which contain a central AT-rich base-pair region. In this study the natuarally occuring antibiotic netropsin has been cocrystallised with the DNA decamer $d(CGCAATTGCG)_2$ and the structure determined to 2.4Å resolution. The netropsin molecule displays AT specificity with hydrogen bonding contacts from the amide NH groups of netropsin to adenine N3 and thymine O2 atoms lying along the floor of the groove.

The crystal structure of native $d(CGCAATTGCG)_2$ has been determined in two previous studies.[1,2] Within the netropsin: DNA complex the central eight bases of each single strand form Watson-Crick base-pairs and duplex of the type $d(GCAATTGC)_2$, whilst the terminal 5'-C and G-3' bases are unpaired. The two terminal guanosines of each single strand lie within the minor groove of a symmetry-related duplex with hydrogen-bonding interactions via guanine atoms N2 and N3 and atom O4' of the deoxyribose sugar to a C G base-pair. The terminal unpaired cytosine bases lie within the major groove of adjacent duplexes to form base triplets of the type C ·G ·C with hydrogen-bonding interactions to a C ·G base-pair.

1. N. Spink et al., Proc. Natl. Acad. Sci. USA (1995) **92** 10767-10771.
2. See Poster by Wood, Nunn & Neidle.

PS04.04.18 CRYSTALLIZATION AND PRELIMINARY X-RAY STRUCTURANALYSIS OF THE STRUCTURAL DOMAIN E OF THERMUS FLAVUS 5S rRNA. Markus Perbandt*, Alexis Nolte*, Siegfried Lorenz*, Jens Peter Fuerste*, Christian Betzel**, Volker A. Erdmann*, *Institute of Biochemistry, Freie Universitaet Berlin, Thielallee 63,14195 Berlin, Germany, **Institute of Physiological Chemistry, University of Hamburg, 20246 Hamburg, Germany

The ribosomal 5S RNA is an essential constituent of the large ribosomal subunit. To overcome the difficulties of crystallizing large RNA molecules such as 5S rRNA, we decided to devide the 5S rRNA in five domains A through E to determine their structure. Recently we determined the crystal structural of the helical domain A. Here we report the preliminary structural results of the chemically synthesized domain E of the Thermus flavus 5S rRNA. The crystal form is trigonal with cell dimension: a = b = 42.80 Å and c = 162.20 Å. Diffraction-data to 2.8 Å have been recorded and the structure solution is currently underway by means of MIR and MAD techniques.

PS04.04.19 CRYSTAL STRUCTURES OF DNA TARGETS OF THE E2 PROTEIN FROM BOVINE PAPILLOMAVIRUS-1. H. Rozenberg[1], D. Rabinovich[1], R. H. Hegde[2] and Z. Shakked[1] [1]Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel [2]Skirball Inst. of Biomolecular Medicine, NYU Medical center, New York, NY 10016, USA

The E2 protein is the dominant transcriptional regulator of papillomaviruses. In bovine papillomavirus-1 (BPV-1), E2 binds sequence-specifically to a consensus sequence $ACCN_6GGT$ found 17 times in the BPV-1 genome. We crystallized several DNA oligomers incorporating the consensus motif (underlined, see Table). The structure of the DNA binding domain of the bovine protein complexed to a 17-mer target (CCGACCGACGTCGGTCG) has been determined previously (Hegde *et al.*, 1992).

The structures of three dodecamers: (1), (2) and (3), were solved by a novel modification of the classical molecular Fourier transform method ("MFT" available at http://www.weizmann.ac.il/~csrabin1/MFT/). The structures of (1) and (2) were refined to R-factors of 21 and 17% respectively. The refinement of (3) is in progress. The free targets adopt B-DNA structures with 10.5-10.7 bp/turn comparable to the complexed target. Sequence (2) is identical to the central 12 base-pair region of the bound sequence. The three unique duplexes of (2) adopt similar helical conformations. The conserved ACC/GGT motifs are relatively straight as in the complexed target whereas the central six base-pair region is bent by nearly 14° toward the major groove which is comparable in magnitude (20°) but in opposite direction to that of the complexed DNA. This observation indicates that the central GACGTC region is inherently flexible. Thus, the present system provides an example where the flexibility of the DNA region which is not contacted by the protein is an important determinant of sequence-specific recognition.

Crystal data: sequence, space group, unit-cell dimensions, no. of duplexes in the asymmetric unit and resolution (measured at 120K)
(1) ACCGGTACCGGT $P4_3$, (40.2, 40.2, 57.6 Å), 1, 2.8 Å
(2) ACCGACGTCGGT $P1$,(40.5,40.1,40.5Å, 82.6,116.2,80.7°),3,2.6Å
(3) ACCGACGTCGGT $R3$, (63.1, 63.1, 44.5 Å), 1, 2.1 Å
(4) ACCGACGIUCGGT $P1$, (40.3,40.3,40.2 Å, 87.1,87.3,68.0°),3,3.2Å
(5) CCGACCGACGTCGGTCG, $R3$, (100.7, 100.7, 79.8 Å), 3, 8.0Å