

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<sup>+</sup>, Emily J Gray, Yu-Hua Li<sup>#</sup> and Werner Leupin<sup>#</sup>, The CRC Biomolecular Structure Unit, The Institute of Cancer Research, Sutton, Surrey, SM2 5NG, UK, <sup>#</sup>F. Hoffman-la-Roche Ltd, Preclinical Research Pharma Gene Technologies, CH-4002 Basel, Switzerland, <sup>+</sup>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)<sub>2</sub>. 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<sub>3</sub>/O<sub>2</sub> 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)<sub>2</sub>.** Christine M. Nunn, Neil Spink, <sup>\*</sup>Elsbeth Garman, Stephen Neidle, CRC, Biomolecular Structure Unit, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, SM2 5NG, UK, <sup>\*</sup>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 naturally occurring antibiotic netropsin has been cocrystallised with the DNA decamer d(CGCAATTGCG)<sub>2</sub> and the structure determined to 2.4 Å 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)<sub>2</sub> has been determined in two previous studies.<sup>1,2</sup> 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)<sub>2</sub>, whilst the terminal 5'-C and G-3' bases are unpaired. The two terminal guanines 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 STRUCTURAL ANALYSIS OF THE STRUCTURAL DOMAIN E OF THERMUS FLAVUS 5S rRNA.** Markus Perbandt\*, Alexis Nolte\*, Siegfried Lorenz\*, Jens Peter Fuerste\*, Christian Betzel<sup>1\*\*</sup>, Volker A. Erdmann\*, <sup>\*</sup>Institute of Biochemistry, Freie Universitaet Berlin, Thielallee 63,14195 Berlin, Germany, <sup>\*\*</sup>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 divide the 5S rRNA in five domains A through E to determine their structure. Recently we determined the crystal structure 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<sup>1</sup>, D. Rabinovich<sup>1</sup>, R. H. Hegde<sup>2</sup> and Z. Shakked<sup>1</sup> <sup>1</sup>Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel <sup>2</sup>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<sub>6</sub>GGT 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<sub>3</sub>, (40.2, 40.2, 57.6 Å), 1, 2.8 Å
- (2) ACCGACGTCGGT P1, (40.5, 40.1, 40.5 Å, 82.6, 116.2, 80.7°), 3, 2.6 Å
- (3) ACCGACGTCGGT R3, (63.1, 63.1, 44.5 Å), 1, 2.1 Å
- (4) ACCGACGTCGGT P1, (40.3, 40.3, 40.2 Å, 87.1, 87.3, 68.0°), 3, 3.2 Å
- (5) CCGACCGACGTCGGTCG, R3, (100.7, 100.7, 79.8 Å), 3, 8.0 Å