ELSEVIER

Nucleic acid crystallography: current progress
Martin Egli

Fifty years after the publication of the DNA double helix model by Watson and Crick, new nucleic acid structures keep emerging at an ever-increasing rate. The past three years have brought a flurry of new oligonucleotide structures, including those of a Hoogsteen-paired DNA duplex, Holliday junctions, DNA–drug complexes, quadruplexes, a host of RNA motifs and various nucleic acid analogues. Major advances were also made in terms of the structure and function of catalytic RNAs. These range from improved models of the phosphodiester cleavage reactions catalyzed by the hairpin and hepatitis delta virus ribozymes to the visualization of a complete active site of a group I self-splicing intron with bound 5′- and 3′-exons. These triumphs are complemented by a refined understanding of cation–nucleic-acid interactions and new routes to the generation of derivatives for phasing of DNA and RNA structures.

Addresses
Department of Biochemistry, Vanderbilt University, School of Medicine, Nashville, TN 37232, USA
e-mail: martin.egli@vanderbilt.edu

Abbreviations
LPTL   lonepair triloop
MAD   multiple-wavelength anomalous dispersion
MTE   2′-O-[2-(methoxythio)ethyl]-modified
PNA   peptide nucleic acid
PPT   polypurine tract
RT   reverse transcriptase
SAD   single-wavelength anomalous dispersion
TNA   L-α-threofuranosyl nucleic acid

Introduction
During the past five years, we have witnessed stunning successes of X-ray crystallographic structure determination that reached a pinnacle with the 3D-structures at atomic resolution of the ribosome and its large and small subunits. Considering the wealth of information on RNA–RNA and RNA–protein interactions now emerging from these structures, is there still a need for more structural studies of nucleic acids 'after the ribosome' [1,2]? The answer is a resounding Yes! Firstly, a look at newly determined structures of DNA, RNA and their analogues over the past two years makes clear the incredibly versatile nature of nucleic acid tertiary structure. Long gone are the days when it was possible to group nucleic acid crystal structures simply into A-, B- and Z-form DNA double helices [3]. But even in the case of double helical fragments, it is noteworthy that the available data are still very limited. For example, about 85% of all crystallized B-form duplex sequences start with cytosine and only two sequences consisting only of AT base pairs have been crystallized [4]. Thus, no structure of a single molecular species alone, even when it concerns a molecule as complicated as the ribosome, can possibly account for the complexity of nucleic acid structure overall. Secondly, the devil is in the details. Despite improvements in the average resolution of crystal structures, there remains a need to visualize structural details for better understanding of structure–function and structure–stability relationships. And thirdly, despite the identification of recurring motifs important for RNA tertiary structure formation (i.e. the mediation of side-by-side packing of helices by unpaired adenosines [5]), there are numerous highly complex RNAs and protein–RNA assemblies that have not yet disclosed their structural secrets [6]. Among them are the spliceosome, the signal recognition particle and telomerase. We can indeed look forward to a steady stream of new structures of nucleic acid and protein–nucleic-acid complexes in the years to come.

Here, I briefly review selected crystallographic studies of nucleic acids that have been published since 2002. A search in the Nucleic Acid Database [7] reveals that the number of released non-protein containing nucleic acid structures solved by X-ray crystallography has increased by 140 between January 2002 and June 2004 to a current total of 805. This review does not consider complexes between nucleic acids and proteins. A breakdown into numbers and types of structural motifs is provided in Table 1. The following sections will highlight structural studies of single-, double- and four-stranded DNAs, modified DNAs, DNA–drug complexes, RNAs, RNA–aptamer complexes, catalytic RNAs, and chemically modified nucleic acids. The role of metal cations in nucleic acid tertiary structure will also be discussed, as will improvements in phasing strategies for DNA and RNA structures. Because of space limitations, many structural studies could not be considered, but their omission is by no means a reflection of lower quality or importance.

Duplex DNA
Practically all double-helical nucleic acids studied in the past 20 years are composed either of only G•C base-pairs
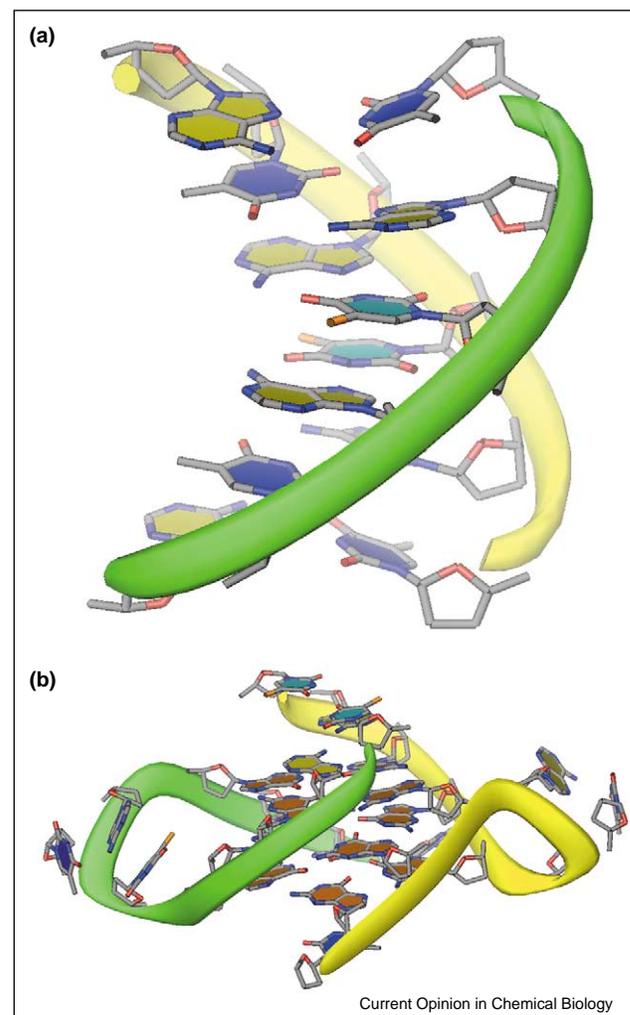
Nucleic acid crystal structures 2002–2004.

Category	No. structures released Jan 2002–June 2004	All structures
A-DNA	5	121
B-DNA	17	153
Z-DNA	5	59
DNA quadruplexes	16	28
DNA junctions	10	16
DNA drug	19	186
DNA other double-helical	13	16
RNA double helices	17	76
RNA quadruplexes	3	4
RNA junctions	1	3
Transfer RNA	0	13
Ribozyme	3	15
RNA other single-stranded	10	19
RNA drug	5	26
RNA–DNA hybrid	8	49
Small; non-double-helical	3	55
Nucleic acid mimetics	4	8
Ribosome (RNA)	3	6

[a]Breakdown into categories of nucleic-acid-only containing crystal structures released by the Nucleic Acid Database (NDB [7]). Of a total of 805 such structures currently available through the NDB, 140 were released between January of 2002 and June of 2004. Note: if the numbers per category are added together, the resulting numbers will be greater than 140 (past 2.5 years) and 805 (total). This is because structures are grouped within more than one category in some cases.

Figure 1

Current Opinion in Chemical Biology

Structures of DNAs with unusual pairing motifs. **(a)** The sequence d(ATATAT) forms an antiparallel duplex with Hoogsteen base-pairing [8•]. Protein Data Bank (PDB) code 1GQU. **(b)** DNA quadruplex formed by the dimeric form of the human telomeric sequence TAGGGTTAGGGT [18•] (PDB code 1K8P).

or A•T flanked by G•C pairs. Thus, only a little 3D-structural work has been directed at AT-rich sequences, although their functional relevance is well known. Fibre diffraction studies of AT-rich sequences provided evidence for considerable structural polymorphism. Therefore, it is possible that the mostly canonical geometries observed for A•T paired regions in the structures of oligonucleotide duplexes are partly a consequence of the constraints exerted by the G•C stretches at both ends. Interestingly, the crystal structure of d(ATATAT) showed antiparallel orientation of strands with formation of Hoogsteen pairs between A and T [8•]. The double helix closely resembles B-DNA, but the orientations of hydrogen bond acceptors and donors of bases on the floors of the major and minor grooves differ from B-form DNA (Figure 1a). For example, the minor groove is narrower and more hydrophobic in the Hoogsteen duplex. The same structure was recently observed in a second crystal form, whereas an NMR investigation of the hexamer in solution revealed adoption of a standard Watson–Crick paired B-form structure [9]. A non-canonical arrangement of bases was also observed in the structure of d(GCGAAAGCT) [10]. The four nucleotides CGAA engage in homo base-pairs (underlined; C•C+, G•G and A•A) with the equivalent nucleotides from a second strand, the two tetramers adopting parallel orientation. The second halves of the decamers then split away to form antiparallel duplexes with two further strands.

While the tetramer d(CGAA) alone adopts a right-handed parallel duplex in the crystal [11], the fragments d(GCGAAGC) and d(GCGAAAGC) form antiparallel duplexes with sheared G•A pairs in the centre and switching of G and A partners between adjacent duplexes [12]. In another study, a recent crystal structure at 2.1 Å resolution has shed light on the 3D lattice formed by a single-stranded oligodeoxynucleotide GGACAGATGG-GAG [13•]. One often wonders what structures a single-stranded DNA might adopt. We have hundreds of examples of structures of double-stranded DNAs in the crystal and in solution but very little is known regarding the conformational variations in a single strand. The 13-mer (the 3′-terminal G is disordered in the crystal) snakes through 3D space in a rather complicated manner,

forming numerous non-canonical interactions with symmetry-related strands. Remarkably, only two among the observed base pairs are of the standard Watson–Crick type and the backbone between the third and the fourth nucleotide assumes a sharp kink. The structure probably provides just a glimpse at the potential complexity of DNA structure when non-canonical pairing modes are considered. Although DNA cannot possibly match RNA in terms of structural complexity and numbers of folding motifs, its chemistry offers an ideal basis for the formation of predictable 2D and 3D arrays. The work by Paukstelis and Seeman and co-workers is an excellent demonstration that DNA 'wears many hats' and, used correctly, can serve as a building material for supramolecular assemblies and nanoscale structures. An unusual conformational behaviour was also observed in the structure of the dodecamer duplex formed by d(CATGGGCCCATG) [14]. It lies on a structural continuum along the transition between A- and B-DNA. The double helix exhibits base pair inclinations that are representative of B-DNA but the values of other parameters such as twist, groove width, sugar pucker as well as slide and x-displacement resemble those of A-DNA. The structure of the intermediate supports a 'slide first - roll later' mechanism for the A→B helix transition.
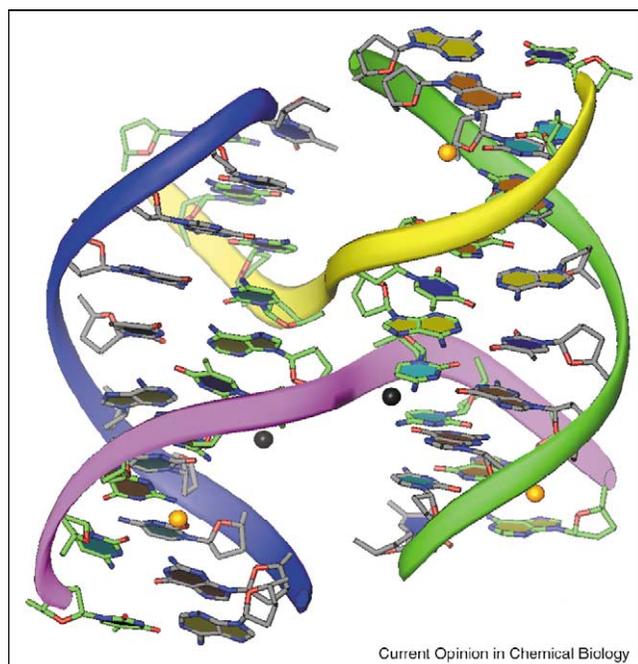
DNA quadruplexes

After a burst of research regarding the structure of G-rich four-stranded motifs in the early 1990s, followed by a relatively quiet period, the number of structures involving guanine tetrads has increased substantially in the past couple of years. Two new structures of $[d(TGGGGT)]_4$ have demonstrated formation of thymine tetrads, suggesting that yeast telomere sequences $d(TG_{1–3})$ might form continuous quadruplex structures [15]. A complex and intriguing new motif was found in the crystal structure of d(gcGA[G]Agc). Two octamers pair into a stack of intercalated G and A residues framed by G•A and g•c base pairs, the core constituted by a tandem of stacked Gs [16•]. Four of these dimers then assemble into a DNA octaplex in which the eight central Gs arrange into a stacked double G-quartet. Replacing the fifth residue [G] in the above octamer sequence by a T still produces a stable intercalated duplex ([16•] and cited references). This remarkable structure gives rise to a potential folding mechanism for eight tandem repeats of $d(ccGA[G]_4Agg)$ (such repetitive regions are encountered throughout the genomes of all vertebrates), whereby two quadruplexes associate into an octaplex under formation of four central G-tetrad layers. An unusual form of self-association is also seen in the structure of d(GCATGCT). Two strands fold into a complementary quadruplex, whereby pairs of Gs and Cs arrange into two layers of base-tetrads with A and T residues forming short loops. This crystal structure has recently been re-determined at atomic resolution ([17] and cited references). Crystal structures of human telomeric DNA were assessed using the dimeric form

of TAGGGTTAGGGT and the 22-mer $AG_3(T_2AG_3)_3$. Both revealed quadruplexes with parallel orientations of strands [18•] (Figure 1b). Most of the deoxyguanosine sugars exhibit the C2′-endo pucker and the glycosylic bonds in the two four-stranded motifs adopt exclusively the anti conformation. The G-rich quadruplexes are extensively stabilized by π stacking interactions between layers of guanines. Moreover, potassium ions are trapped in the core between stacked G-quartets, spaced at ca. 2.7 Å from each of a total of eight O6 carbonyl groups. The relative orientation of strands in the human telomere quadruplexes differs from that in the potassium-containing quadruplex adopted by two $d(G_4T_4G_4)$ strands. The structure of this sequence was recently revisited in two crystal forms and found to be consistent with that previously observed for the quadruplex in solution, where two dodecamers form an antiparallel diagonal arrangement ([19] and cited references) The basis of stabilization of G-quadruplex structures by quadruplex-selective ligands has also been assessed based on the $d(G_4T_4G_4)$ sequence from *Oxytricha nova*. In the crystal structure of the complex between the DNA quadruplex and a disubstituted aminoalkylamido acridine compound, the acridine moiety sits at one end of the stack of G tetrads and forms specific hydrogen bonds to thymine bases in the loops [20].

Holliday junctions

Several X-ray crystallographic studies have provided a more detailed picture of the 3D structure of the four-way or Holliday junction formed during DNA-strand exchange in recombination (reviewed in [21] and [22]). Remarkably, DNA decamers with sequences CCGGG<u>ACC</u>GG, CCGGT<u>ACC</u>GG, and TCGGT<u>ACC</u>GA fold into four-stranded complexes instead of forming double helices with mismatched base pairs in the centre (Figure 2). The ACC trinucleotide (underlined) forms the core of the junction and its 3′-C•G base pairs help to stabilize the junction by engaging in direct and water-mediated hydrogen bonds to phosphate groups at the strand crossover. The four strands adopt an X conformation whereby an angle of about 40° relates the two interconnected duplexes across the junction. Site-selective ion binding and its role in the conformation of the Holliday junction were probed with Na^+, Ca^{2+} and Sr^{2+} in the case of the last of the three above sequences [23]. The effects of sequence on the conformation of the Holliday junction have also been investigated in some detail [24]. Thus, the DNA sequences CCGGC<u>GCC</u>GG and CCAGT<u>AC-(Br5U)</u>GG were shown to also crystallize as Holliday junctions. Therefore, the central trinucleotide motif can be extended from ACC to PuCPy, where Pu is either G or A and Py is C, Me5C or Br5U, but not T. The stability of the four-way junction appears to be influenced by the electronic properties of substituents at the 5-positions of Py bases. Lastly, the structure in a second crystal form of the junction formed by the above decamer
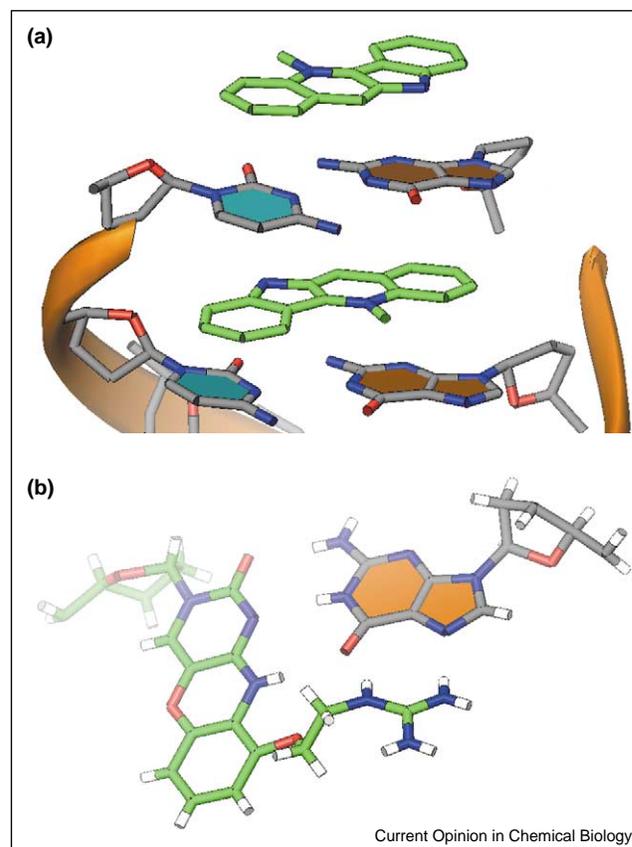
Holliday junction formed by four strands d(TCGGTACCGA) [23] (PDB code 1L4J).

Drug–DNA interactions and artificial pairing systems. **(a)** The drug cryptolepine intercalates at CC sites [35•] (PDB code 1K9G). **(b)** Five Watson–Crick and Hoogsteen hydrogen bonds are formed in the G-clamp•G pair [39•] (PDB code 1KGK).

CCGGTACCGG has demonstrated the subtle interplay of sequence and lattice in shaping the conformation of Holliday junctions [25].

DNA–drug interactions

Several new crystal structures of DNA–drug complexes, involving both groove binders and intercalators have been analysed during the past two years. The origins of specificity of the anticancer antibiotic chromomycin for the GGCC sequence were revealed in the structure of the drug complexed to the DNA octamer duplex [d(TTGGCCAA)]$_2$ [26]. Dimers of the drug bind at and widen the minor groove in addition to kinking the DNA. No fewer than six G-specific hydrogen bonds between the drug and DNA form the basis for sequence specificity. Preferred binding of the anticancer drug actinomycin D to DNA GpC sequences flanked by T:T mismatches in (CTG)$_n$ trinucleotide repeats whose expansion is correlated with certain neurological disorders was characterized in the crystal structure of the drug bound to d(ATGCTGCAT) [27]. Several structures at medium resolution were published of 1:1 complexes between the minor groove binders netropsin and distamycin and DNA decamer duplexes [28,29]. The hexamers d(CGTACG) and d(CGATCG) in complex with intercalating agents have been analysed crystallographically in quite a few studies. Several new DNA-intercalator structures involving the two hexamers have been published recently, among them complexes with the inactive

derivative of an antitumor acridinecarboxamide compound [30], a bisacridine that was shown to fuse DNA duplexes [31], a disaccharide anthracycline that interacts with the DNA in the minor groove, but also has one of its sugar rings protruding out from the DNA helix and contacting an adjacent DNA duplex [32], the trioxatriangulenium ion that when irradiated in the intercalated form leads to DNA strand cleavage primarily at GpG steps [33], and, finally, the complex between daunomycin and the DNA–RNA chimeric hexamer d(C)r(G)d(ATCG) [34]. The crystal structure of the antimalarial and cytotoxic drug cryptolepine bound to d(CCTAGG) revealed for the first time a planar compound intercalating at a nonalternating pyrimidine-pyrimidine step [35•] (Figure 3a).

RNA–DNA hybrids

Two recent papers have shed light on the rather complex structural polymorphism exhibited by RNA–DNA hybrids. A high-resolution structure of viral polypurine tract (PPT) RNA — the PPT serves as a primer and is essential in retroviruses for replication by reverse transcriptase (RT) — in complex with DNA has uncovered

several unusual features [36]. The investigated sequence r(CAAAGAAAAG) constitutes two-thirds of HIV-1 PPT and contains a highly deformable AGA motif that is used by the RT enzyme for unpairing of bases when it binds PPT. The sugar of the second adenine (underlined in the sequence) adopts a C2′-endo pucker, characteristic of B-form DNA, whereas all other riboses in the structure exhibit the expected C3′-endo conformation. This conformational switch at the 5′-end of the PPT could adversely affect binding of the RNA–DNA hybrid at the active site of the RNase H domain of RT, thus causing it to pause and leaving the PPT intact. The PPT-RNA–DNA duplex was crystallized in two forms and for one of them a novel type of intermolecular intercalation involving dimerization of duplexes by 'base-pair swapping' was reported [37].

Nucleic acid analogs and adducts

Structural investigations of chemically modified nucleic acids and nucleic acid analogues have been spurred by various motives such as the development of nucleic acid therapeutics and molecular probes and a better understanding of issues relating to DNA recognition and repair as well as the etiological and biophysical properties of nucleic acids, among others. Three crystallographic studies have provided insight into the conformational and biochemical properties of chemically modified nucleic acids and analogues. The structure of a DNA duplex with incorporated 2′-O-[2-(methoxythio)ethyl]-modified (MTE) thymidines helped rationalize the improved RNA affinity and protein-binding properties of oligonucleotides containing MTE residues relative to DNA and DNA phosphorothioates [38]. Consistent with the dramatic enhancements of the thermodynamic stability of DNAs bearing guanidino G-clamps, a cytosine analogue that was designed to form Hoogsteen-type hydrogen bonds with G in addition to the standard Watson–Crick ones, the crystal structure at atomic resolution of a modified DNA decamer duplex established the existence of five hydrogen bonds in the G-clamp•G pair [39•] (Figure 3b). The third structure concerns the geometry of a duplex formed between a chiral peptide nucleic acid (PNA) with D-lysine modifications in the backbone of the central portion and the cDNA [40]. The heteroduplex has close similarity to the P-helix conformation seen in other PNA structures, indicating PNA's intrinsic preference for this geometry and its relatively restricted conformational flexibility. DNAs with base lesions continue to be the focus of structural investigations, and have recently included the *cis-syn* thymidine dimer [41], 3,N(4)-etheno-2′-deoxycytidine [42] and N(4)-methoxy-2′-deoxycytidine [43]. Two analogues, formyluridine and thymidine without the exocyclic 2-keto oxygen and N1 substituted by carbon were incorporated into the sequence d(CGCGAATTCGCG) to analyse the roles of Mg^{2+} in crystal packing [44] and of minor groove functional groups in DNA hydration [45], respectively.

The structure of the DNA hairpin capped by a bis-(2-hydroxyethyl)stilbene-4,4′-diether moiety (Sd2) d(GT$_4$G)-Sd2-d(CA$_4$C) was determined at 1.5 Å resolution [46]. The structure revealed conformational variations of the Sd2 moieties between the two independent molecules in the crystal. It also provided evidence for narrowing of the A-tract minor groove as a result of Mg^{2+} coordination and provided a better understanding of the origins of hairpin stability and the photochemical behaviour of Sd2-conjugated DNAs in solution. Finally, two structural studies were directed at the conformational properties of L-α-threofuranosyl nucleic acid (TNA) [47,48]. TNA features a shorter backbone than DNA but exhibits stable self-pairing and cross-pairs with DNA and RNA. In the crystal structures of both B-form and A-form DNA duplexes with incorporated TNA thymidines, the threose sugars adopted the C4′-exo pucker. These structures also demonstrated a close conformational resemblance between TNA and RNA, a finding that should be of particular interest in view of the fact that TNA is the focus of intense research directed at an etiology of nucleic acid structure.

RNA oligonucleotides

X-ray crystallographic structure determination of RNA continues to be a very active field of research. At least as far as fragments with lengths of up to ca. 40 nucleotides are concerned, the structural efforts have greatly profited from recent advances in the solid phase synthesis of RNA [49]. The crystal structure of a 27-nucleotide fragment of *Escherichia coli* 23S rRNA in which the conserved GAGA tetraloop was replaced by GUAA revealed the detailed geometry and hydration of the loop and the so-called A-minor motif [50]. The comparison of the GYRA (Y is C or U; R is A or G) loop conformation with those observed for other tetraloops of the GNRA type (N is A, G, C or U) also pointed to distinct differences in the geometry of these important RNA building blocks and protein as well as RNA recognition motifs. Short RNAs that form hairpins in solution but crystallize as duplexes with mismatched base pairs have contributed considerably to our understanding of the complexity of RNA structure. In a further example, the sequence r(GGUCACAGCCC) crystallized as a double helix with six non-canonical pairings [51]. The crystal structure of an 18-mer self-complementary duplex that contains most of the target stem recognized by the Com zinc-finger protein from bacteriophage Mu also featured mismatched base pairs [52]. In the structure of a 26-nucleotide RNA comprising the so-called loop E motif, an arrangement termed the hook-turn was observed [53]. It consists of an A-form helix that splits into two separate strands after a G•A base pair, whereby the strand containing the G describes a 180° turn in the space of two nucleotides. Similar structures also occur in 16S and 23S rRNAs. Another rRNA structural motif, the lonepair triloop (LPTL), was first postulated on the basis of phylogenetic covariation analyses [54]. Its hallmarks

are a single ('lone') base pair capped by a hairpin loop with three nucleotides. The two nucleotides immediately preceding and following the lonepair nucleotides are not paired, thus restricting the length of the helix to just one base pair. LPTLs are found throughout rRNAs and in one instance in the T loop of tRNA, and a few of them appear to be associated with ribosomal function. The structure of the hepatitis C virus internal ribosome entry site (HCV IRES) RNA determined at 2.8 Å resolution showed an overall arrangement that involved two helical junctions [55]. One of the junctions features virtually seamless stacking between the two participating helices, whereas the stacking between helices in the other is interrupted. The distorted geometry of the junction may provide a crucial recognition motif for the translational machinery. The relative orientation of helices at a junction is probably also important for function in the case of RNA pseudoknots that are stimulating ribosomal frameshifting. The crystal structure at atomic resolution of a −1 type frameshifting RNA pseudoknot from a luteovirus offered insight into the fine structure of the pseudoknot motif, Mg^{2+} binding and hydration [56].
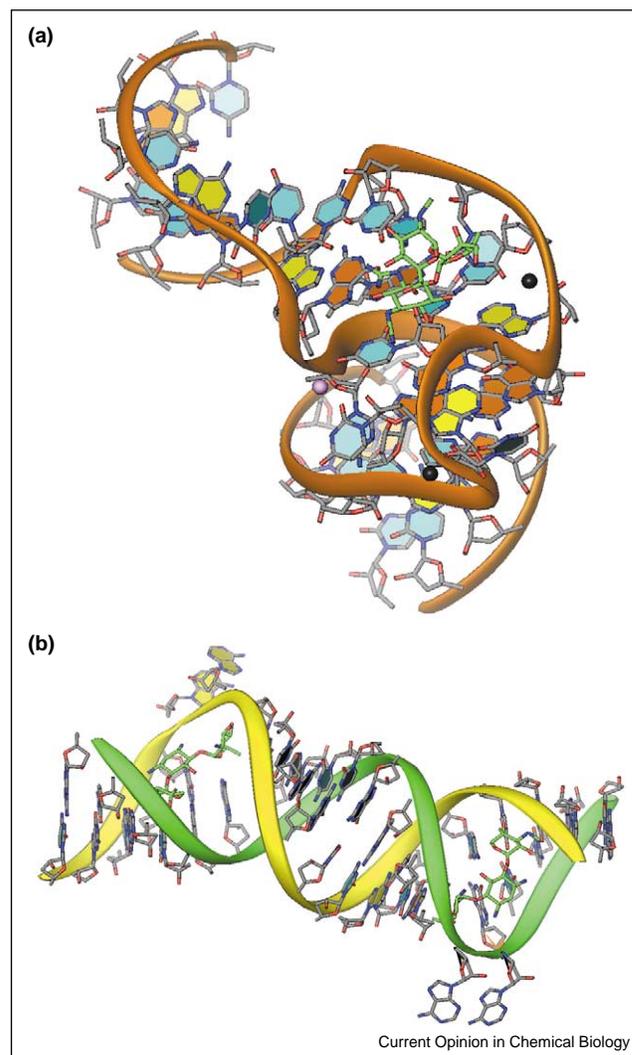
RNA–drug interactions

Structures of antibiotics bound to ribosomal subunits and rRNA oligonucleotide fragments as well as *in vitro* selected RNA aptamers provide for an interesting comparison between the various interaction modes underlying specificity. In the structure of the complex between streptomycin and a 40-mer RNA aptamer, specific interactions comprise direct hydrogen bonds between functional groups of the streptose and RNA base edges in the drug-binding pocket [57•] (Figure 4a). On the other hand, the majority of contacts between RNA and streptomycin in the structure of the drug bound to the 16S rRNA decoding aminoacyl-tRNA site (A site) involve backbone phosphate groups. Structures of complexes between an oligoribonucleotide rRNA fragment and aminoglycoside antibiotics that decrease accuracy of translation by binding to the ribosomal A site have permitted a visualization of the detailed drug–RNA interactions (i.e. [58]) (Figure 4b). They have also helped rationalize the role of conserved RNA residues and the sensitiveness exhibited by the drugs to common A-site point mutations. The binding sites on the ribosome of another class of drugs, the macrocyclic antibiotics (including carbomycin A, spiramycin, tylosin and azithromycin) have been probed on the basis of their complexes with the large ribosomal subunit of a halophilic organism [59]. Their locations in the polypeptide tunnel adjacent to the peptidyl transfer centre suggest that they inhibit protein synthesis by blocking the path by which nascent polypeptides exit the ribosome.

Catalytic RNAs

Research targeted at the tertiary structure elucidation of ribozymes and mechanistic aspects of RNA-catalyzed

Figure 4

Aptamer–drug interactions. **(a)** Complex between the antibiotic streptomycin and an *in vitro* selected 40-mer RNA aptamer [57•] (PDB code 1NTB). **(b)** Geneticin bound to a bacterial 16S ribosomal RNA A-site oligonucleotide [58] (PDB code 1MWL).

reactions has been very productive in the past three years. Ribonuclease P was one of the first catalytic RNAs to be discovered; it is composed of an RNA and a protein component and is present in all organisms. RNase P is an endonuclease and processes the 5′-end of tRNA by cleaving a precursor, thus leading to tRNA maturation. The RNA component is catalytically active in the absence of the protein component and contains a specificity domain as well as a catalytic domain. The crystal structure of the 154-nucleotide specificity domain of RNase P from *Bascillus subtilis* has recently been determined and has provided a first model of the overall architecture of this key ribozyme [60•]. The catalytically essential domains of another ribozyme, the group II self-splicing intron that catalyzes auto-excision from precursor

RNA transcripts, have been visualized at 3 Å resolution [61]. The crystal structure of the 70-nucleotide fragment containing domains 5 and 6 of the intron with the mechanistically important branch-point adenosine revealed a surprising two-nucleotide bulge around the branch point. Several recent publications have reported new insights regarding the catalytic mechanisms of ribozymes. Some of them, especially the hammerhead ribozyme, have been the target of intense structural and functional studies during the past decade. New structural data suggest that switching between the hammerhead RNA's ligase and nuclease activities is determined by the relative orientations of stem I and stem II [62]. Earlier work had demonstrated the rate-limiting step for the hammerhead ribozyme-catalyzed cleavage reaction to be a pH-dependent conformational change that is driven by the deprotonation of the 2'-OH moiety at the cleavage site [63]. A conformational switch also controls hepatitis delta virus ribozyme catalysis [64]. Comparisons of crystal structures of the ribozyme before and after cleavage showed a sizable conformational change in the RNA after cleavage along with ejection of the Mg^{2+} ion that is required for strand cleavage (Figure 5). Metal ion binding is also at the core of another small ribozyme motif, the *in vitro* selected leadzyme. New crystallographic data are consistent with a model whereby binding of the Pb^{2+} ion to a pre-catalytic conformation of the ribozyme in conjunction with a flexible trinucleotide bulge appears to bring about configurations of the ribose $2'$-O^- nucleophile and the scissile bond conducive to cleavage [65]. On the contrary, the hairpin ribozyme does not appear to employ metal cations as catalytic cofactors and the RNA was believed to either rely on binding energy to catalyse the cleavage reaction or employ another mechanism, i.e. general acid-base catalysis; (reviewed in [66,67]). The crystal structure of a vanadate-hairpin ribozyme has now provided evidence that transition state stabilization is probably the key strategy used by this catalytic RNA motif to cleave site-specifically [68]. Finally, the crystal structure of a complete self-splicing group I intron with bound exons and metal ion cofactors has been determined very recently [69•]. The complex represents the splicing intermediate before exon ligation and has yielded a detailed mechanism for the second step of splicing that probably involves two Mg^{2+} ions (Figure 6).

## DNA– and RNA–metal ion interactions

Metal cations are an integral part of nucleic acid structure and are known to affect DNA and RNA stability, conformation and packing. Ions also act as cofactors at the active sites of most catalytic RNAs. The potential role of alkali metal cations (Na^+, K^+) or ammonium ions in the control of the particular conformational features displayed by A-tract DNA has received special attention by nucleic acid crystallographers in recent years (reviewed in [70]). It has also been pointed out that Mg^{2+} only rarely interacts directly with DNA and that solvated ions can

Figure 5
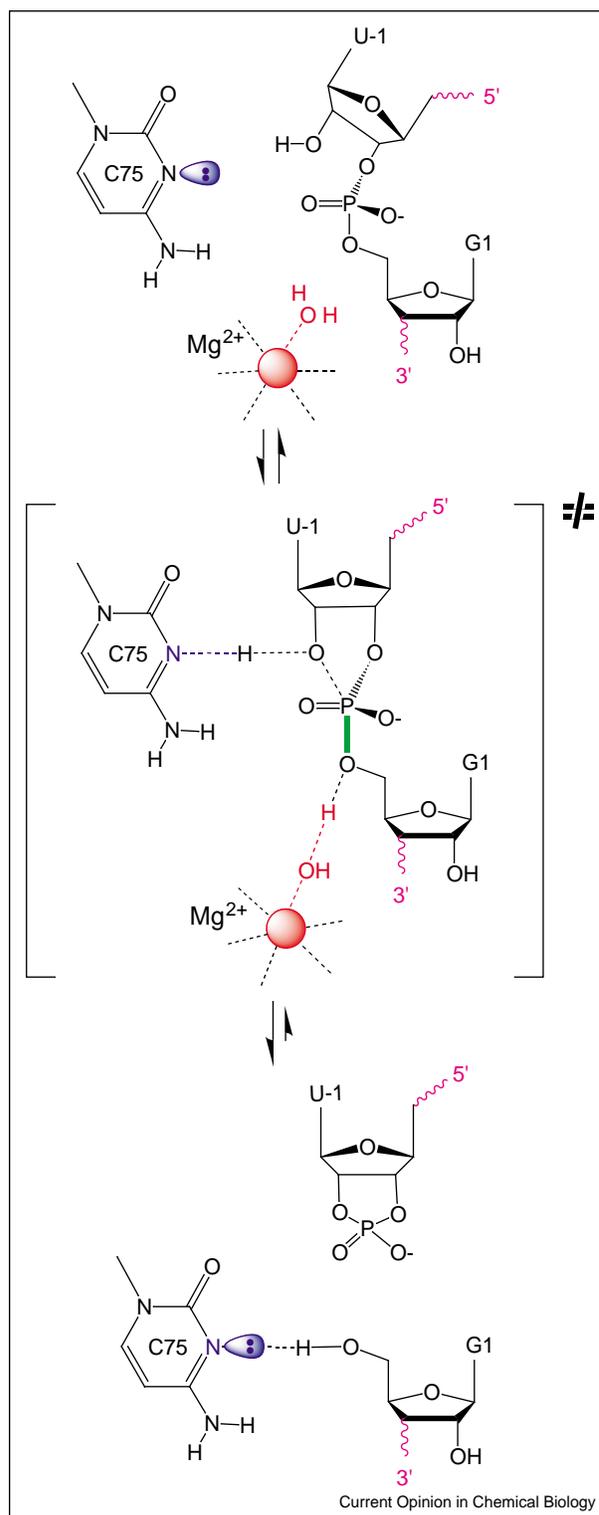
Proposed mechanism of the RNA cleavage reaction catalyzed by the hepatitis delta virus ribozyme based on the recent crystal structure of the ribozyme [64] (PDB code 1VBX).
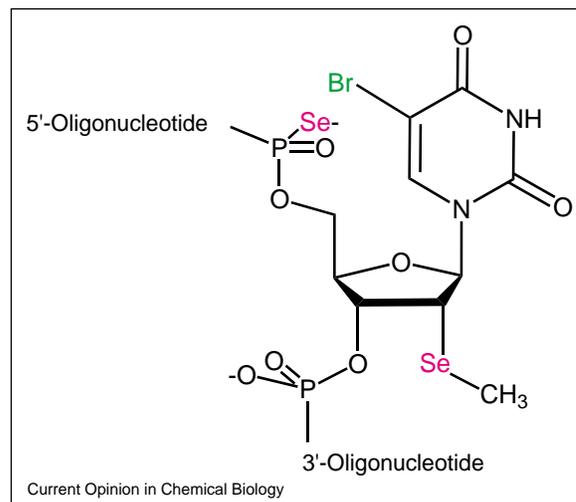
Figure 6
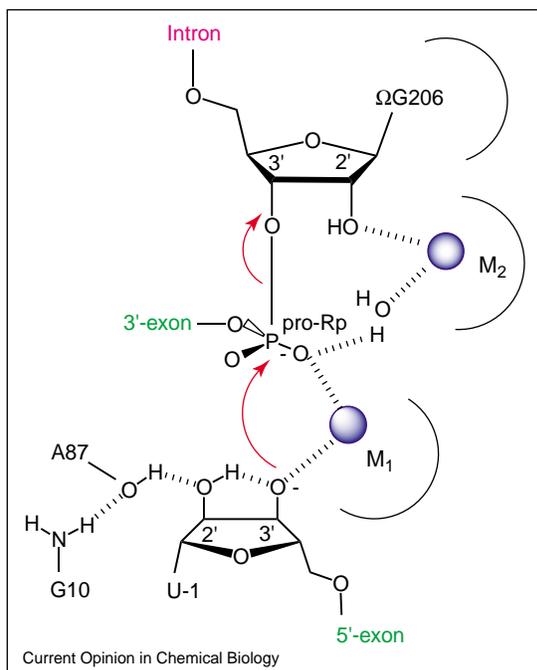
Proposed reaction mechanism for the second step of splicing by a bacterial intron based on the crystal structure of the self-splicing group I intron with both exons from bacterium *Azoarcus* sp. [69•] (PDB code1T42).

Figure 7

Covalent nucleotide modifications for phasing of nucleic acid crystal structures [80,81•].

thus be viewed as hydrogen bond donors in addition to just point charges [71]. A detailed analysis of Mg^{2+}–DNA interactions in the crystal structure of the Dickerson–Drew B-form DNA dodecamer at atomic resolution demonstrated that at least one Mg^{2+} ion is present at every site where phosphate groups from neighbouring duplexes are closely spaced in the crystal lattice [72]. Several crystal structures of DNA duplexes were reported with an emphasis on the specific binding modes of Co^{2+} hexammine [73,74], Zn^{2+} [73,75] and Ni^{2+} [73,76]. The structure of the complex between a DNA hexamer and the intercalating drug adriamycin crystallized in the presence of Tl^+ showed binding of metal ions exclusively in the major groove [77]. Localized inorganic ions appeared to be displaced from the minor groove by the cationic functional group of adriamycin. A systematic crystallographic analysis of the coordination of 13 different metal ions to two related RNA duplexes was carried out and has enhanced our knowledge of the various binding modes of middle- and heavy-weight metal cations to RNA [78].

Phasing methods

Multi- or single-wavelength anomalous dispersion (MAD or SAD, respectively) in combination with halogenated derivatives is considered the most common and convenient technique for determination of new nucleic acid crystal structures that do not yield to molecular replacement. However, both iodine and bromine at the 5-position of pyrimidines (Figure 7) are photolabile, and debromination has been shown to occur at a relatively moderate X-ray dose [79]. Moreover, it is not uncommon that halogen derivatives produce a crystal form that differs from that of wild-type DNA or RNA, or that the presence of bromine or iodine prevents crystallization altogether. Recently, synthetic methods have been developed to introduce selenium at alternative positions within the nucleic acid framework to facilitate crystallographic phasing (Figure 7). As proof of principle, the crystal structure of a DNA duplex with single 2'-MeSe-uracil residues per strand was determined by MAD [80]. Phosphoroselenoates are stable to oxidation on a crystallographic time-scale and are suitable for structure determination by MAD or SAD [81•]. This approach was recently used successfully to solve the crystal structure of an oligo-2',3'-dideoxyglucopyranose (homo-DNA) duplex that had resisted all attempts to phase it for a decade (M Egli, unpublished results).

Conclusions

For all the excitement about the successes in nucleic acid crystallography and the growing number of structures of protein–nucleic acid complexes, we should not forget that many questions remain unanswered or that focusing on a limited set of sequences or motifs may have led us to not fully appreciate the potential structural complexity of the nucleic acids. The latter applies particularly to DNA. Recent examples including the structures of four-stranded motifs (G-rich quadruplexes; i.e. [16•,18•]), duplexes that lack Watson–Crick-type base pairs (the so-called Hoogsteen duplex [8•]), and the 3D-array con-

stituted by a single-stranded oligodeoxynucleotide [13•] all demonstrate that the conformational space accessible to DNA is larger than we may commonly assume. G-rich motifs are instructive in this respect. Although the majority of these quadruplexes feature similar G-tetrad stacks in their cores, variations in the relative polarities of strands and conformational and sequence diversity in the loops connecting the G-stretches give rise to a myriad of different arrangements that may correlate with a host of functions. Concerning the sequence dependence of the geometry of double helices, it is clear that all except two sequences investigated to date with X-ray crystallographic methods encompassed GC base pairs. Therefore, there is room for further investigations into the particular geometries of DNAs with unusual sequences, such as AT-rich DNAs, and those containing tandem or triplet repeats.

Another area where knowledge gaps exist at the structural level concerns the interactions between metal ions and nucleic acids and the potential consequences of ions for modulating DNA and RNA structure. Further, the role of cations in the phosphodiester cleavage reaction catalyzed by small ribozymes still deserves our attention. It is noteworthy that the recent success in the visualization of the second step of the splicing reaction catalyzed by a group I intron RNA [69•] is not paralleled by a similar breakthrough in the structure-based dissection of the mechanistic details of the hammerhead RNA-mediated phosphodiester cleavage reaction. In addition, compared with the analysis of the complexes between DNA and small molecules that has furnished a multitude of structures of DNAs with bound intercalators and groove binders, the interactions between RNAs and small molecules have only been studied to a limited degree. Given RNA's potential importance as a drug target, this is certainly a field that might profit enormously from a more extensive number of structures that, in combination with computational simulations, could drive the discovery of new drug leads including antibiotic agents. In addition to a host of other functions of RNA molecules, the phenomenon that oligoribonucleotides act as gene silencers (RNA interference or RNAi) has further increased the interest in RNA and its interactions with proteins that eventually lead to the suppression of gene expression. Structural studies of RNA and RNA–protein complexes will play an important role in understanding the principles underlying RNA interference.

Update

Recent work by Subirana and colleagues has established, on the basis of a crystal structure of a DNA heptanucleotide, that branching points produced by Ni^{2+} ions attached to N7 of guanine can lead to the generation of three-dimensional arrays [82]. These arrays are stabilized by phosphate–Ni^{2+}–guanine interactions, and large solvent cavities among DNA duplexes are reminiscent of those observed in the 3D-array formed by a single-stranded DNA 13-mer [13•]. A new analogue of the minor groove binding ligand Hoechst 33342, methylproamine, was found in *in vitro* studies to be a potent radioprotector [83]. DNA-bound methylproamine presumably acts as a reducing agent through electron transfer, thereby repairing transient radiation-induced oxidizing species on DNA. The crystal structure of the complex between methylproamine and a DNA with sequence CGCGAA-TTCGCG confirmed binding of the ligand in the central minor groove of the dodecamer. Finally, using a combination of fluorescence measurements and X-ray crystal structures of RNA oligonucleotides containing the ribosomal decoding site, Hermann and colleagues have created an elegant assay to affinity-screen the ribosomal target against RNA binders for the discovery of novel antibiotics [84].

Acknowledgements

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Moore PB, Steitz TA: **After the ribosome structures: how does peptidyl transferase work?** *RNA* 2003, **9**:155-159.

2. Wilson DN, Nierhaus KH: **The ribosome through the looking glass**. *Angew Chem Int Ed Engl* 2003, **42**:3464-3486.

3. Rich A: **The double helix: a tale of two puckers**. *Nat Struct Biol* 2003, **10**:247-249.

4. Subirana JA: **DNA discoveries through crystallography**. *Nature* 2003, **423**:683.

5. Battle DJ, Doudna JA: **Specificity of RNA-RNA helix recognition**. *Proc Natl Acad Sci USA* 2002, **99**:11676-11681.

6. Doudna JA: **Structural genomics of RNA**. *Nat Struct Biol* 2000, **7**:954-956.

7. Berman HM, Olson WK, Beveridge DL, Westbrook J, Gelbin A, Demeny T, Hsieh S-H, Srinivasan AR, Schneider B: **The nucleic acid database: a comprehensive relational database of three-dimensional structures of nucleic acids**. *Biophys J* 1992, **63**:751-759. (http://www.ndbserver.rutgers.edu).

8. Abrescia NGA, Thompson A, Huynh-Dinh T, Subirana JA:
• **Crystal structure of an antiparallel DNA fragment with Hoogsteen base pairing**. *Proc Natl Acad Sci USA* 2002, **99**:2806-2811.
The double helical structure formed by an oligonucleotide containing A•T base pairs only reveals absence of Watson–Crick hydrogen bonds. An intriguing finding, although the biological relevance of this pairing mode remains to be determined.

9. Abrescia NGA, Gonzalez C, Gouyette C, Subirana JA: **X-ray and NMR studies of the DNA oligomer d(ATATAT): Hoogsteen base pairing in duplex DNA**. *Biochemistry* 2004, **43**:4092-4100.

10. Sunami T, Kondo J, Kobuna T, Hirao I, Watanabe K, Miura K, Takenaka A: **Crystal structure of d(GCGAAAGCT) containing a parallel-stranded duplex with homo base pairs and an anti-parallel duplex with Watson-Crick base pairs**. *Nucleic Acids Res* 2002, **30**:5253-5260.

11. Kobuna T, Sunami T, Kondo J, Takenaka A: **X-ray structure of d(CGAA), parallel-stranded DNA duplex with homo base pairs**. *Nucleic Acids Res Suppl* 2002:179-180.

12. Sunami T, Kondo J, Hirao I, Watanabe K, Miura K, Takenaka A: **Structure of d(GCGAAGC) and d(GCGAAAGC) (tetragonal form); a switching of partners of the sheared G:A pairs to form a functional G:AxA:G crossing**. *Acta Crystallogr D* 2004, **60**:422-431.

13. Paukstelis PJ, Nowakwoski J, Birktoft JJ, Seeman NC:
● **Crystal structure of a continuous three-dimensional DNA lattice**. *Chem Biol* 2004, **11**:1119-1126.
This 3D structure of a single-stranded DNA allows a glimpse at the multitude of conformations and interactions accessible to DNA when non-canonical pairing modes are involved.

14. Ng HL, Dickerson RE: **Mediation of the A/B-DNA helix transition by G-tracts in the crystal structure of duplex CATGGGCCCATG**. *Nucleic Acids Res* 2002, **30**:4061-4067.

15. Caceres C, Wright G, Gouyette C, Parkinson G, Subirana JA: **A thymine tetrad in d(TGGGGT) quadruplexes stabilized with Tl⁺/Na⁺ ions**. *Nucleic Acids Res* 2004, **32**:1097-1102.

16. Kondo J, Adachi W, Umeda S, Sunami T, Takenaka A: **Crystal**
● **structures of a DNA octaplex with I-motif of G-quartets and its splitting into two quadruplexes suggest a folding mechanism of eight tandem repeats**. *Nucleic Acids Res* 2004, **32**:2541-2549.
A fascinating tertiary and quaternary structural motif that could explain the involvement of tandem repeats in genetic diseases. The observed packaging of repeats may facilitate slippage during replication.

17. Thorpe JH, Teixeira SCM, Gale BC, Cardin CJ: **Crystal structure of the complementary quadruplex formed by d(GCATGCT) at atomic resolution**. *Nucleic Acids Res* 2003, **31**:844-849.

18. Parkinson GN, Lee MPH, Neidle S: **Crystal structure of parallel**
● **quadruplexes from human telomeric DNA**. *Nature* 2002, **417**:876-880.
New structures of G-rich sequences related to human telomeric DNA that demonstrate the versatility of four-stranded motifs and the role of metal ions in promoting and stabilizing a particular topology.

19. Haider S, Parkinson GN, Neidle S: **Crystal structure of the potassium form of an Oxytricha nova G-quadruplex**. *J Mol Biol* 2002, **320**:189-200.

20. Haider SM, Parkinson GN, Neidle S: **Structure of a G-quadruplex-ligand complex**. *J Mol Biol* 2003, **326**:117-125.

21. Eichman BF, Ortiz-Lombardia M, Aymami J, Coll M, Ho PS: **The inherent properties of DNA four-way junctions: comparing the crystal structures of holliday junctions**. *J Mol Biol* 2002, **320**:1037-1051.

22. Hays FA, Watson J, Ho PS: **Caution! DNA crossing: crystal structures of Holliday junctions**. *J Biol Chem* 2003, **278**:49663-49666.

23. Thorpe JH, Gale BC, Teixeira SCM, Cardin CJ: **Conformational and hydration effects of site-selective sodium, calcium and strontium ion binding to the DNA Holliday junction structure d(TCGGTACCGA)(4)**. *J Mol Biol* 2003, **327**:97-109.

24. Hays FA, Vargason JM, Ho PS: **Effect of sequence on the conformation of DNA holliday junctions**. *Biochemistry* 2003, **42**:9586-9597.

25. Thorpe JH, Teixeira SCM, Gale BC, Cardin CJ: **Structural characterization of a new crystal form of the four-way Holliday junction formed by the DNA sequence d(CCGGTACCGG)₂: sequence versus lattice?**. *Acta Crystallogr D Biol Crystallogr* 2002, **58**:567-569.

26. Hou MH, Robinson H, Gao Y-G, Wang AHJ: **Crystal structure of the [Mg²⁺-(chromomycin A3)₂]-d(TTGGCCAA)₂ complex reveals GGCC binding specificity of the drug dimer chelated by a metal ion**. *Nucleic Acids Res* 2004, **32**:2214-2222.

27. Hou MH, Robinson H, Gao YG, Wang AHJ: **Crystal structure of actinomycin D bound to the CTG triplet repeat sequences linked to neurological diseases**. *Nucleic Acids Res* 2002, **30**:4910-4917.

28. Shi K, Mitra SN, Sundaralingam M: **Structure of the 1:1 netropsin-decamer d(CCIICICCII)₂ complex with a single bound netropsin**. *Acta Crystallogr D* 2002, **58**:601-606.

29. Uytterhoeven K, Sponer J, Van Meervelt L: **Two 1: 1 binding modes for distamycin in the minor groove of d(GGCCAATTGG)**. *Eur J Biochem* 2002, **269**:2868-2877.

30. Adams A, Guss JM, Denny WA, Wakelin LPG: **Crystal structure of 9-amino-N-[2-(4-morpholinyl)ethyl]-4-acridinecarboxamide bound to d(CGTACG)₂: implications for structure-activity relationships of acridinecarboxamide topoisomerase poisons**. *Nucleic Acids Res* 2002, **30**:719-725.

31. Teixeira SCM, Thorpe JH, Todd AK, Powell HR, Adams A, Wakelin LPG, Denny WA, Cardin CJ: **Structural characterisation of bisintercalation in higher-order DNA at a junction-like quadruplex**. *J Mol Biol* 2002, **323**:167-171.

32. Temperini C, Messori L, Orioli P, Di Bugno C, Animati F, Ughetto G: **The crystal structure of the complex between a disaccharide anthracycline and the DNA hexamer d(CGATCG) reveals two different binding sites involving two DNA duplexes**. *Nucleic Acids Res* 2003, **31**:1464-1469.

33. Reynisson J, Schuster GB, Howerton SB, Williams LD, Barnett RN, Cleveland CL, Landman U, Harrit N, Chaires JB: **Intercalation of trioxatriangulenium ion in DNA: binding, electron transfer, X-ray crystallography, and electronic structure**. *J Am Chem Soc* 2003, **125**:2072-2083.

34. Shi K, Pan B, Sundaralingam M: **Structure of a B-form DNA/RNA chimera (dC)(rG)d(ATCG) complexed with daunomycin at 1.5 Å resolution**. *Acta Crystallogr D Biol Crystallogr* 2003, **59**:1377-1383.

35. Lisgarten JN, Coll M, Portugal J, Wright CW, Aymami J:
● **The antimalarial and cytotoxic drug cryptolepine intercalates into DNA at cytosine-cytosine sites**. *Nat Struct Biol* 2002, **9**:57-60.
Unlike all other intercalating agents studied by crystallographic means to date that typically insert at a CpG step, cryptolepine is bound at a CpC step.

36. Kopka ML, Lavelle L, Han GW, Ng HL, Dickerson RE: **An unusual sugar conformation in the structure of an RNA/DNA decamer of the polypurine tract may affect recognition by RNase H**. *J Mol Biol* 2003, **334**:653-665.

37. Han GW, Kopka ML, Langs D, Sawaya MR, Dickerson RE: **Crystal structure of an RNA.DNA hybrid reveals intermolecular intercalation: dimer formation by base-pair swapping**. *Proc Natl Acad Sci USA* 2003, **100**:9214-9219.

38. Prakash TP, Manoharan M, Kawasaki AM, Fraser AS, Lesnik EA, Sioufi N, Leeds JM, Teplova M, Egli M: **2′-O-[2-(methylthio)ethyl]-modified oligonucleotide: an analogue of 2′-O-[2-(methoxy)-ethyl]-modified oligonucleotide with improved protein binding properties and high binding affinity to target RNA**. *Biochemistry* 2002, **41**:11642-11648.

39. Wilds CJ, Maier MA, Tereshko V, Manoharan M, Egli M: **Direct**
● **observation of a cytosine analogue that forms five hydrogen bonds to guanosine: guanidino G-clamp**. *Angew Chem Int Ed Engl* 2002, **41**:115-157.
Designed to mimic the arginine–guanine pair commonly observed in complexes between protein and DNA, the G-clamp is shown to engage in Hoogsteen and Watson–Crick hydrogen bonds to G.

40. Menchise V, De Simone G, Tedeschi T, Corradini R, Sforza S, Marchelli R, Capasso D, Saviano M, Pedone C: **Insights into peptide nucleic acid (PNA) structural features: the crystal structure of a D-lysine-based chiral PNA-DNA duplex**. *Proc Natl Acad Sci USA* 2003, **100**:12021-12026.

41. Park H, Zhang K, Ren Y, Nadji S, Sinha N, Taylor JS, Kang C: **Crystal structure of a DNA decamer containing a cis-syn thymine dimer**. *Proc Natl Acad Sci USA* 2002, **99**:15965-15970.

42. Freisinger E, Fernandes A, Grollman AP, Kisker C: **Crystallographic characterization of an exocyclic DNA adduct: 3,N4-etheno-2′-deoxycytidine in the dodecamer 5′-CGCGAATTepsilonCGCG-3′**. *J Mol Biol* 2003, **329**:685-697.

43. Hossain MT, Kondo J, Ueno Y, Matsuda A, Takenaka A: **X-Ray analysis of d(CGCGAATTXGCG)(2) containing a 2′ or**

minute-deoxy-N(4)-methoxycytosine residue at X: a characteristic pattern of sugar puckers in the crystalline state of the Dickerson-Drew type DNA dodecamers. *Biophys Chem* 2002, **95**:69-77.

44. Tsunoda M, Kondo J, Karino N, Ueno Y, Matsuda A, Takenaka A: **Water mediated Dickerson-Drew-type crystal of DNA dodecamer containing 2′-deoxy-5-formyluridine**. *Biophys Chem* 2002, **95**:227-233.

45. Woods KK, Lan T, McLaughlin LW, Williams LD: **The role of minor groove functional groups in DNA hydration**. *Nucleic Acids Res* 2003, **31**:1536-1540.

46. Egli M, Tereshko V, Mushudov GN, Sanishvili R, Liu X, Lewis FD: **Face-to-face and edge-to-face π-π interactions in a synthetic DNA hairpin with a stilbenediether linker**. *J Am Chem Soc* 2003, **125**:10842-10849.

47. Wilds CJ, Wawrzak Z, Krishnamurthy R, Eschenmoser A, Egli M: **Crystal structure of a B-form DNA duplex containing (L)-alpha-threofuranosyl (3′→2′) nucleosides: a four-carbon sugar is easily accommodated into the backbone of DNA**. *J Am Chem Soc* 2002, **124**:13716-13721.

48. Pallan PS, Wilds CJ, Wawrzak Z, Krishnamurthy R, Eschenmoser A, Egli M: **Why does TNA pair more strongly with RNA than with DNA? – an answer from X-ray analysis.** *Angew Chem Int Ed Engl* 2003, **115**:5893-5895.

49. Marshall WS, Kaiser RJ: **Recent advances in the high-speed solid phase synthesis of RNA**. *Curr Opin Chem Biol* 2004, **8**:222-229.

50. Correll CC, Swinger K: **Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 Å resolution**. *RNA* 2003, **9**:355-363.

51. Kacer V, Scaringe SA, Scarsdale JN, Rife JP: **Crystal structures of r(GGUCACAGCCC)₂**. *Acta Crystallogr D* 2003, **59**:423-432.

52. Lima S, Hildenbrand J, Korostelev A, Hattman S, Li H: **Crystal structure of an RNA helix recognized by a zinc-finger protein: an 18-bp duplex at 1.6 Å resolution**. *RNA* 2002, **8**:924-932.

53. Szep S, Wang J, Moore P: **The crystal structure of a 26-nucleotide RNA containing a hook-turn**. *RNA* 2003, **9**:44-51.

54. Lee JC, Cannone JJ, Gutell RR: **The lonepair triloop: a new motif in RNA structure**. *J Mol Biol* 2003, **325**:65-83.

55. Kieft JS, Zhou K, Grech A, Jubin R, Doudna JA: **Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation**. *Nat Struct Biol* 2002, **9**:370-374.

56. Egli M, Minasov G, Su L, Rich A: **Metal ion coordination and conformational flexibility in a viral RNA pseudoknot at atomic resolution**. *Proc Natl Acad Sci USA* 2002, **99**:4302-4307.

57. Tereshko V, Skripkin E, Patel DJ: **Encapsulating streptomycin**
• **within a small 40-mer RNA**. *Chem Biol* 2003, **10**:175-187.
A detailed look at the versatility of substrate recognition by RNA. The binding mode discussed here contrasts with that seen for streptomycin bound to the A site of the ribosome.

58. Vicens Q, Westhof E: **Crystal structure of geneticin bound to a bacterial 16S ribosomal RNA A site oligonucleotide**. *J Mol Biol* 2003, **326**:1175-1188.

59. Hansen JL, Ippolito JA, Ban N, Nissen P, Moore PB, Steitz TA: **The structures of four macrolide antibiotics bound to the large ribosomal subunit**. *Mol Cell* 2002, **10**:117-128.

60. Krasilnikov AS, Yang X, Pan T, Mondragon A: **Crystal structure**
• **of the specificity domain of ribonuclease P**. *Nature* 2003, **421**:760-764.
The first three-dimensional model for an important domain of a universally conserved catalytic RNA is reported.

61. Zhang L, Doudna JA: **Structural insights into group II intron catalysis and branch-site selection**. *Science* 2002, **295**:2084-2088.

62. Dunham CM, Murray JB, Scott WG: **A helical twist-induced conformational switch activates cleavage in the hammerhead ribozyme**. *J Mol Biol* 2003, **332**:327-336.

63. Murray JB, Dunham CM, Scott WG: **A pH-dependent conformational change, rather than the chemical step, appears to be rate-limiting in the hammerhead ribozyme cleavage reaction**. *J Mol Biol* 2002, **315**:121-130.

64. Ke A, Zhou K, Ding F, Cate JHD, Doudna JA: **A conformational switch controls hepatitis delta virus ribozyme catalysis**. *Nature* 2004, **429**:201-205.

65. Wedekind JE, McKay DB: **Crystal structure of the leadzyme at 1.8 Å resolution: metal ion binding and the implications for catalytic mechanism and allo site ion regulation**. *Biochemistry* 2003, **42**:9554-9563.

66. Ryder SP, Strobel SA: **Comparative analysis of hairpin ribozyme structures and interference data**. *Nucleic Acids Res* 2002, **30**:1287-1291.

67. Ferré-D'Amaré AR, Rupert PB: **The hairpin ribozyme: from crystal structure to function**. *Biochem Soc Trans* 2002, **30**:1105-1109.

68. Rupert PB, Massey AP, Sigurdsson ST, Ferré-D'Amaré AR: **Transition state stabilization by a catalytic RNA**. *Science* 2002, **298**:1421-1424.

69. Adams PL, Stahley MR, Kosek AB, Wang J, Strobel SA: **Crystal**
• **structure of a self-splicing group I intron with both exons**. *Nature* 2004, **430**:45-50.
An outstanding structure of the RNA molecule whose discovery provided the initial demonstration that not all enzymes are protein. For the first time, all players, including intron, exons and metal ion cofactors are visualized and a double-metal ion mechanism is proposed for the second step of splicing.

70. Egli M: **DNA-cation interactions: quo vadis?** *Chem Biol* 2002, **9**:277-286.

71. Subirana JA, Soler-Lopez M: **Cations as hydrogen bond donors: a view of electrostatic interactions in DNA**. *Annu Rev Biophys Biomol Struct* 2003, **32**:27-45.

72. Egli M, Tereshko V: **Lattice- and sequence-dependent binding of Mg²⁺ in the crystal structure of a B-DNA dodecamer**. In *Curvature and Deformation of Nucleic Acids: Recent Advances, New Paradigms*. Edited by Stellwagen N, Mohanty U: ACS Symposium Series 2004. Vol. 884:87-109.

73. Labiuk SL, Delbaere LTJ, Lee JS: **Cobalt(II), nickel(II) and zinc(II) do not bind to intra-helical N(7) guanine positions in the B-form crystal structure of d(GGCGCC)**. *J Biol Inorg Chem* 2003, **8**:715-720.

74. Ramakrishnan B, Sekharudu C, Pan B, Sundaralingam M: **Near-atomic resolution crystal structure of an A-DNA decamer d(CCCGATCGGG): cobalt hexammine interaction with A-DNA**. *Acta Crystallogr D Biol Crystallogr* 2003, **59**:67-72.

75. Soler-Lopez M, Malinina L, Tereshko V, Zarytova V, Subirana JA: **Interaction of zinc ions with d(CGCAATTGCG) in a 2.9 Å resolution X-ray structure**. *J Biol Inorg Chem* 2002, **7**:533-538.

76. Abrescia NA, Huynh-Dinh T, Subirana JA: **Nickel-guanine interactions in DNA: crystal structure of nickel-d[CGTGTACACG]₂**. *J Biol Inorg Chem* 2002, **7**:195-199.

77. Howerton SB, Nagpal A, Williams LD: **Surprising roles of electrostatic interactions in DNA-ligand complexes**. *Biopolymers* 2003, **69**:87-99.

78. Ennifar E, Walter P, Dumas P: **A crystallographic study of the binding of 13 metal ions to two related RNA duplexes**. *Nucleic Acids Res* 2003, **31**:2671-2682.

79. Ennifar E, Carpentier P, Ferrer JL, Walter P, Dumas P: **X-ray-induced debromination of nucleic acids at the Br K absorption edge and implications for MAD phasing**. *Acta Crystallogr D Biol Crystallogr* 2002, **58**:1262-1268.

80. Teplova M, Wilds CJ, Wawrzak Z, Tereshko V, Du Q, Carrasco N, Huang Z, Egli M: **Covalent incorporation of selenium into oligonucleotides for X-ray crystal structure determination via MAD: proof of principle**. *Biochimie* 2002, **84**:849-858.

81. Wilds CJ, Pattanayek R, Pan C, Wawrzak Z, Egli M:
• **Selenium-assisted nucleic acid crystallography: use of**

phosphoroselenoates for MAD phasing of a DNA structure. *J Am Chem Soc* 2002, **124**:14910-14916.
A 'last resort' approach to phasing of nucleic acid crystal structures that resist all other attempts to solve them. Allows screening of more potential derivatives than is possible by replacing T and C with Br5U and Br5C, respectively.

82. Valls N, Gouyette C, Subirana JA: **A cubic arrangement of DNA double helices based on nickel-guanine interactions**. *J Am Chem Soc* 2004, **126**:7812-7816.

83. Martin RF, Broadhurst S, Reum ME, Squire CJ, Clark GR, Lobachevsky PN, White JM, Clark C, Sy D, Spotheim-Maurizot M, Kelly DP: ***In vitro* studies with methylproamine: a potent new radioprotector**. *Cancer Res* 2004, **64**:1067-1070.

84. Shandrick S, Zhao Q, Han Q, Ayida BK, Takahashi M, Winters G, Simonsen KB, Vourloumis D, Hermann T: **Monitoring molecular recognition of the ribosomal decoding site**. *Angew Chem Int Ed* 2004, **43**:3177-3182.