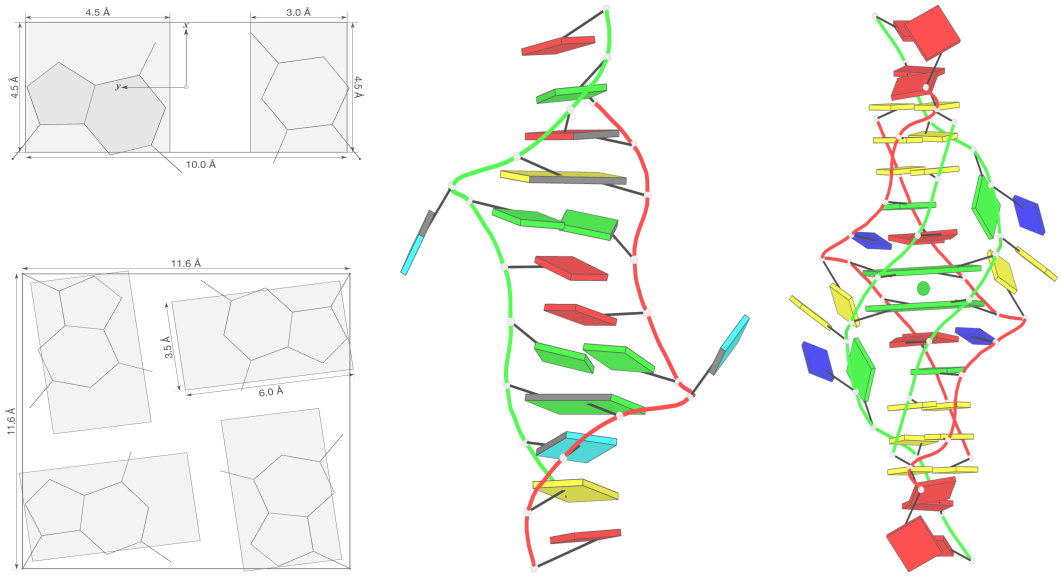
DSSR: an integrated software tool for Dissecting the Spatial Structure of RNA

An Overview (User Manual for DSSR 2.3)



modified nucleotide, non-canonical base pair, helix, stem, coaxial stacking
hairpin/internal/junction loop, kink turn, G-quadraplex, i-motif, pseudoknot
comprehensive characterization of DNA-protein and RNA-protein complexes
cartoon-block innovative schematics in PyMOL, SQL-like feature queries in Jmol
in silico base mutations, regular models, customized rebuilding, template-based modeling

Dr. Xiang-Jun Lu (律祥俊)

Department of Biological Sciences
Columbia University, New York

Contents

List of Figures

1 Overview of DSSR

As the number of experimentally solved RNA-containing structures deposited in the PDB (Burley *et al.*, 2018) grows, it is becoming increasingly important to characterize the geometric features of these molecules consistently and efficiently. Existing RNA bioinformatics tools (Lemieux and Major, 2002; Yang *et al.*, 2003; Sarver *et al.*, 2008) are fragmented, and suffer in either scope or usability. DSSR (Lu *et al.*, 2015) is an integrated software tool for Dissecting the Spatial Structure of RNA. It has been designed from ground up to streamline the analyses and annotations of 3D RNA structures. Despite its acronym, DSSR is actually not limited to RNA; it works on any nucleic-acid-containing structures, including DNA-protein or RNA-protein complexes. While PDB entries represent a typical use case, 3D nucleic acids structures derived from other sources, including molecular dynamic (MD) simulations (Tan *et al.*, 2018), can also be handled by DSSR.

DSSR has been integrated into Jmol and PyMOL, bringing unmatched searching capabilities and innovative visualization styles into 3D nucleic acid structures (Hanson and Lu, 2017; Lu, 2020). It includes modules for homology modeling via *in silico* base mutations, easy generation of regular helical models, and the creation of customized structures with user-specific base sequences and rigid-body parameters. These modeling features not only cover a plethora of common use cases, but also enable many novel applications.

Over the past few years, DSSR has been widely cited in literature (Bayrak *et al.*, 2017; Desai *et al.*, 2017; Meier *et al.*, 2018; Tan *et al.*, 2018; Berger *et al.*, 2019; Giambaşu *et al.*, 2019; Cai *et al.*, 2020; Chan *et al.*, 2020; Cuturello *et al.*, 2020; Kribelbauer *et al.*, 2020; Baulin *et al.*, 2020; Zhao *et al.*, 2020; Miskiewicz *et al.*, 2020; Afek *et al.*, 2020; Shi *et al.*, 2020) [15 items] and adopted into many other structural bioinformatics resources (Lu *et al.*, 2018; Zok *et al.*, 2018; Antczak *et al.*, 2019; Gallego *et al.*, 2019; Thiel *et al.*, 2019; Yesselman *et al.*, 2019; Sagendorf *et al.*, 2020; Zok *et al.*, 2020; Chen *et al.*, 2020; Saaidi *et al.*, 2020; Becquey *et al.*, 2020) [11 items]. These 26 sample, yet representative, DSSR citations are from the following 15 leading scientific journals: (i) *Biochemistry*, (ii) *Bioinformatics* (4 times), (iii) *BMC Bioinformatics*, (iv) *Brief. Bioinformatics*, (v) *F1000Research*, (vi) *J. Am. Chem. Soc.*, (vii) *J. Chem. Theory Comput.*, (viii) *J. Mol. Biol.*, (ix) *Nature* (2 times), (x) *Nat. Commun.*, (xi) *Nat. Nanotechnol.*, (xii) *Nucleic Acids Res.* (7 times), (xiii) *Proc. Natl. Acad. Sci.*, (xiv) *RNA* (2 times), and (xv) *Science*. They clearly demonstrate the huge impact DSSR has already achieved in the broad field of DNA/RNA structural bioinformatics.

DSSR has much to offer. It integrates an unprecedented set of features into a single tool, including analysis/annotation, visualization, and modeling. It delivers a great user experience by solving problems and saving time. It is a 'solid software product' in structural bioinformatics of nucleic acids.

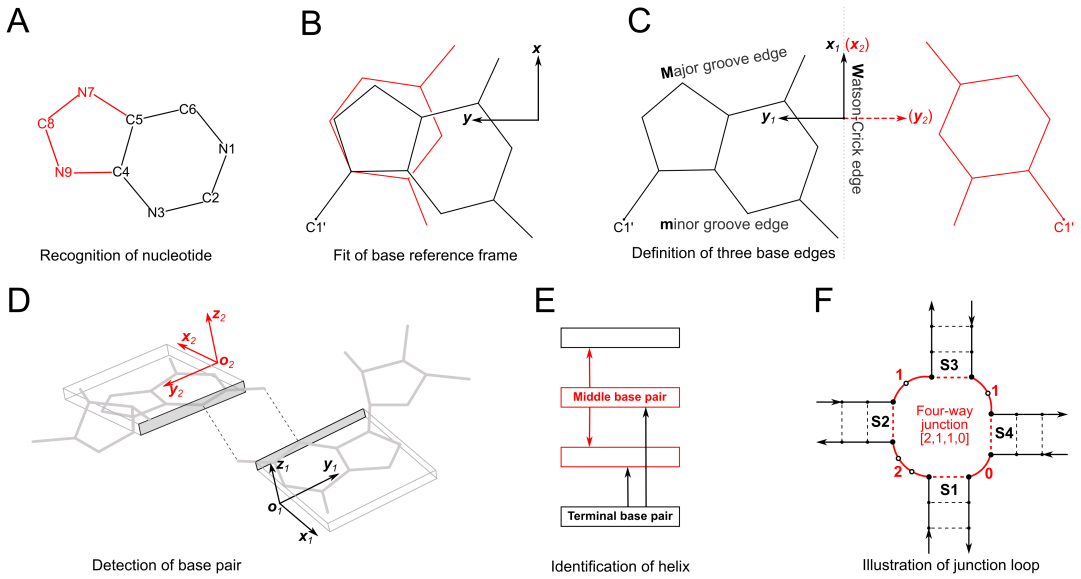


Figure 1: Definitions of key nucleic acid structural components in DSSR [Reproduced from Figure 1 of the DSSR paper (Lu *et al.*, 2015)]. (A) Nucleotides are recognized using standard atom names and base planarity. This method works for both the standard (A, C, G, T and U) and modified nucleotides, regardless of their tautomeric or protonation states. (B) Bases are assigned a standard reference frame (Olson *et al.*, 2001) that is independent of sequence identity: purines and pyrimidines are symmetrically placed with respect to the sugar. (C) The standard base frame is derived from an idealized Watson-Crick base pair, and defines three base edges (Watson-Crick, minor groove, and Major groove) that are used to classify pairing interactions. (D) Base pairs are identified from the coplanarity of base rings and the occurrence of H-bonds. This geometric algorithm can find canonical (Watson-Crick and G–U wobble) as well as non-canonical pairs. Higher-order (three or more) coplanar base associations, termed multiplets, are also detected. (E) Helices are defined by stacking interactions of base pairs, regardless of pairing type (canonical or otherwise) or backbone connectivity (covalently connected or broken). A helix consists of at least two base pairs. A stem is defined as a special type of helix, made up of canonical pairs with a continuous backbone along each strand. Coaxial stacking is defined by the presence of two or more stems within one helix. (F) 'Closed' loops of various types (hairpin, bulge, internal, and junction loops) are delineated by stems, and specified by the lengths of the intervening, consecutive nucleotide segments.

1.1 Structural analysis and annotation

Figure 1 outlines some key algorithms underlying the DSSR program. Starting from a 3D nucleic-acid-containing structure in PDB or mmCIF format, DSSR uses standard atom names and base planarity to detect nucleotides, including modified ones (Figure 1A). It employs the standard base reference frame (Olson *et al.*, 2001, Figure 1B,C) and a set of simple geometric criteria (Figure 1D) to identify all existent base pairs: either canonical Watson-Crick (WC) and wobble pairs or non-canonical pairs with at least one hydrogen bond (H-bond). The latter pairs may include normal or modified bases, regardless of tautomeric or protonation state. DSSR uses the six standard rigid-body base-pair parameters (Shear, Stretch, Stagger, Propeller, Buckle, and Opening) to rigorously quantify the spatial disposition of any two interacting bases. See worked examples (Appendix A) on base-pair parameters. Where applicable, the program also denotes a base-pair by common names (including WC, wobble G–U, reverse WC, Hoogsteen, reverse Hoogsteen, sheared G–A, Calcutta U–U, dinucleotide platform, etc.), the Saenger (1984) classification scheme of 28 H-bonding types, and the Leontis and Westhof (2001, LW) nomenclature of 12 basic geometric classes.

DSSR detects multiplets (triplets or higher-order base associations) by searching horizontally in the plane of the associated base-pair for further H-bonding interactions. The program determines double-helical regions (Figure 1E) by exploring vertically in the neighborhood of selected base-pairs for stacking interactions, regardless of pairing type or backbone connection (e.g., coaxial stacks). DSSR then identifies hairpin loops, bulges, internal loops, and multi-branch (junction) loops (Figure 1F). The program outputs RNA secondary structure in three commonly used formats—dot-bracket notation (dbn), connectivity table (`.ct`), and base-pair sequence (`.bpseq`)—that can be fed directly into visualization tools such as VARNA (Darty *et al.*, 2009). DSSR derives proper dbn for RNA with higher-order pseudoknots, and it can also produce pseudoknot-free secondary structures.

In DSSR, each helix/stem is characterized by a least-squares fitted helical axis, and dinucleotide steps are classified into the most common A-, B-, or Z-form double helical forms (where appropriate) and quantified by local step and helical parameters. See worked examples (Appendix A) on these parameters. DSSR calculates commonly used backbone torsion angles, including three sets of virtual η/θ torsions (Li *et al.*, 2019), classifies the backbone into BI/BII conformations and the sugar into C2′/C3′-endo puckers, and assigns the consensus RNA backbone suite names (Richardson *et al.*, 2008). The program automatically identifies A-minor interactions, splayed-apart dinucleotide conformations, atom-base capping interactions,

ribose zippers, G-quadruplexes, i-motifs, kissing loops, U-turns, k-turns, etc. DSSR also reports non-pairing interactions (H-bonding or base-stacking) between two nucleotides and contacts involving phosphate groups.

1.2 Integrations into Jmol and PyMOL

Working with Robert Hanson and Thomas Holder respectively, I initiated the integrations of DSSR into Jmol and PyMOL, two of the most popular molecular viewers. The DSSR-Jmol and DSSR-PyMOL integrations lead to unparalleled search capabilities and innovative visualization styles of 3D nucleic acid structures.

The DSSR-Jmol paper (Hanson and Lu, 2017), "DSSR-enhanced visualization of nucleic acid structures in Jmol", introduces a flexible, powerful SQL-like query language that employs the standard JSON interface between the two programs. Users can now select DSSR-derived structural features (such as base pairs, double helices, and various loops) as easily as they can select protein α-helices and β-sheets. Moreover, fine-grained characteristics of RNA structural features can be queried. See the supplemental PDF (Appendix E) of Hanson and Lu (2017), which serves as the user manual, for details. Visit `http://jmol.x3dna.org/` to have a try. Here are some simple examples:

```
SELECT WITHIN(dssr, "nts WHERE is_modified = true") # modified nucleotides
SELECT pairs # all pairs
Select WITHIN(dssr, "pairs WHERE name = 'Hoogsteen'") # Hoogsteen pairs
SELECT WITHIN(dssr, "pairs WHERE name != 'WC'") # non-Watson-Crick pairs
SELECT junctions # all junctions loops
select within(dssr, "junctions WHERE num_stems = 4") # four-way junction loops
```

The DSSR-PyMOL paper (Lu, 2020), "DSSR-enabled innovative schematics of 3D nucleic acid structures with PyMOL", presents schematic block representations in diverse styles (see Figure 2). These DSSR blocks can be seamlessly integrated into PyMOL and complement its other popular visualization options. In addition to portraying individual base blocks, DSSR can draw WC pairs as long blocks and highlight the minor-groove edges. Notably, DSSR can dramatically simplify the depiction of G-quadruplexes by automatically detecting G-tetrads and treating them as large square blocks. The DSSR-enabled innovative schematics with PyMOL are aesthetically pleasing and highly informative; the base identity, pairing geometry, stacking interactions, double-helical stems, and G-quadruplexes are immediately obvious (Figure 2). These features can be accessed via four interfaces: the command-line interface, the DSSR plugin for PyMOL, the web application, and the web application programming

interface. The supplemental PDF (Appendix G) of Lu (2020) serves as a practical guide, with complete and reproducible examples. The web application interface (`http://skmatic.x3dna.org/`) also provides pre-calculated schematics and meta information of nucleic-acid-containing structures in the PDB. It is intuitive, allowing even novices or occasional users to get started quickly. Here are some examples:

```
http://skmatic.x3dna.org/pdb/2lx1  # with internal loop 5'GAGU-3'UGAG
http://skmatic.x3dna.org/pdb/2grb  # an RNA quadruplex containing inosine-tetrad
http://skmatic.x3dna.org/pdb/6vu1  # an RNA hairpin
http://skmatic.x3dna.org/pdb_entries  # 12 randomly picked entries
http://skmatic.x3dna.org/pdb_entries/recent-week
http://skmatic.x3dna.org/pdb_entries/recent-month
```

The `http://skmatic.x3dna.org/` website has been recommended in Faculty Opinions as "simple and effective". It is classified as "Good for Teaching".

1.3 Versatile modeling capabilities

DSSR has been augmented with a module for *in silico* base mutations that are context sensitive. Powered by the analysis engine already in DSSR, this modeling module allows users to perform base mutations with unprecedented flexibility and convenience. By default, the mutation preserves both the geometry of the sugar-phosphate backbone and the base reference frame (position and orientation). As a result, re-analyzing the mutated model gives the same base-pair and step parameters as those of the original structure. Here are some potential applications:

```
mutate all bases in hairpin loops to a specific base (e.g., G)
mutate all non-stem bases to a specific base (e.g., U)
mutate bases 2-12 to a specific base (e.g., A) regardless of context
mutate selected bases in a given structure to a new sequence (e.g, 3-6 to AUAU)
mutate all bases of the same type to another (e.g., A to G)
mutate all bases of the same type to another (e.g., C to U) except for some
mutate all G-C WC pairs to C-G pairs, and A-U to U-A
mutate all G-tetrads in G-quadruplexes to non-G-tetrads (e.g., U-tetrads)
```

From early on, 3DNA (Lu and Olson, 2003) contains the `fiber` program that can be used to build models of over fifty types of uniform helical structures. These regular models are based primarily on the fiber diffraction work of Arnott (1999). The fiber module in DSSR has obsoleted the 3DNA `fiber` program, with improved usability.

The 3DNA `rebuild` program complements the `analyze` routine. These two programs are a defining feature of 3DNA (Lu and Olson, 2003, 2008; Li *et al.*, 2019). The DSSR analyze

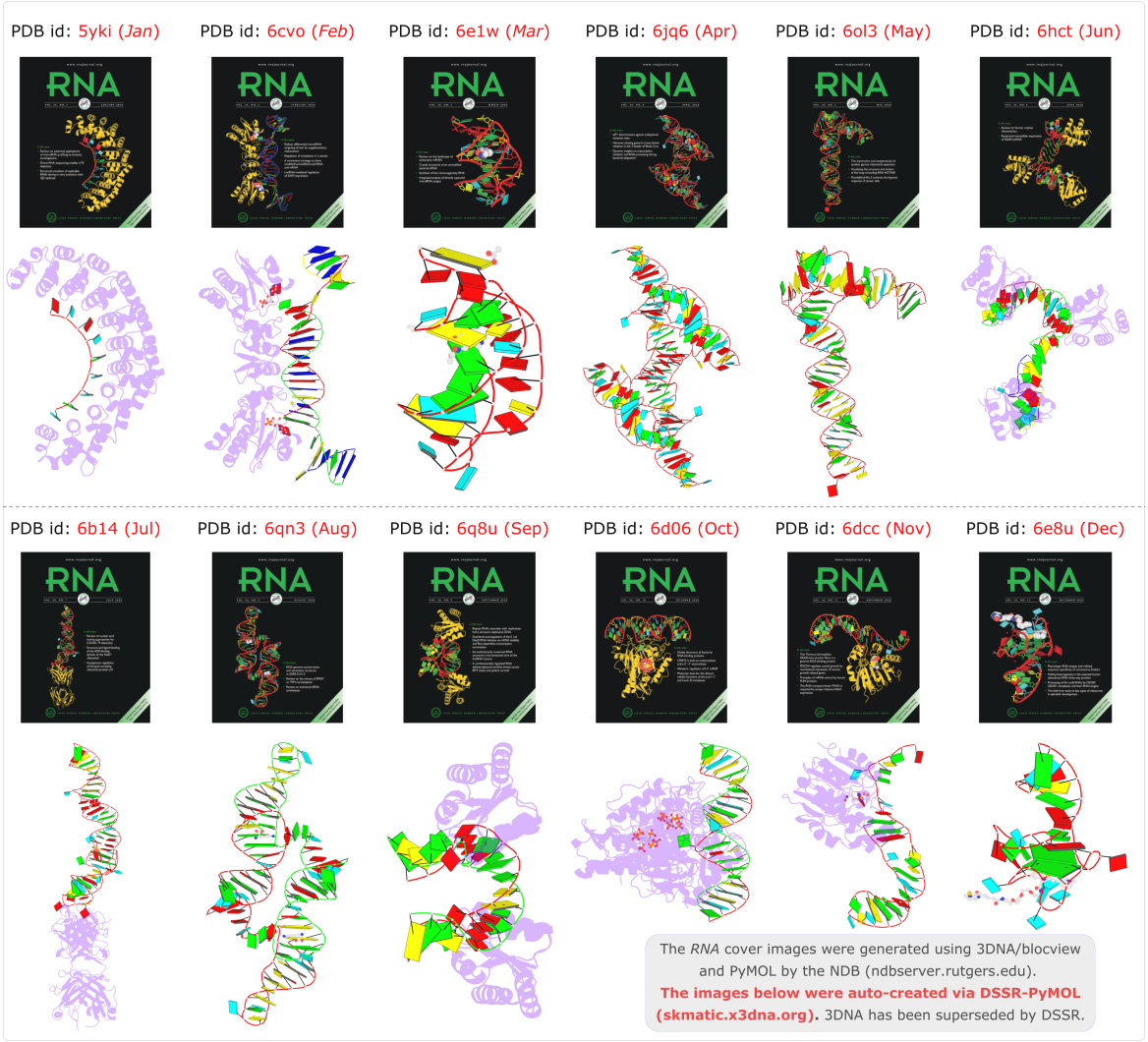


Figure 2: Cover images of all the 12 issues of the *RNA Journal* in 2020. The cover images were generated using 3DNA/blocview and PyMOL by the Nucleic Acid Database (NDB, Narayanan *et al.*, 2014). The corresponding images below the covers were created automatically via DSSR-PyMOL (Lu, 2020). The DSSR-PyMOL approach is easier to use, has more features, and produces better graphics. 3DNA has be superseded by DSSR. The 3DNA/blocview images will be replaced by the DSSR-PyMOL schematics.

module has completely surpassed the 3DNA `analyze` program. The rebuild module in DSSR replaces the 3DNA `rebuild` program, with enriched functionality.

1.4 DSSR vs. 3DNA vs. SCHNAaP/SCHNArP

From a historical perspective, DSSR is built upon the 3DNA suite of software programs for the analysis, rebuilding, and visualization of 3D nucleic acid structures (Lu and Olson, 2003, 2008; Li *et al.*, 2019). 3DNA was derived from the SCHNAaP and SCHNArP pair of programs for rigorous analysis and reversible rebuilding of double-helical nucleic acid structures (Lu *et al.*, 1997*a*,*b*). Specifically, the algorithms that underpinned SCHNAaP/SCHNArP laid the foundation of `analyze`/`rebuild`, two core components of the 3DNA suite. The idea of representing bases and WC-pairs as rectangular blocks came from the pioneering work of Calladine *et al.* (2004). The block schematics were first implemented in SCHNAaP/SCHNArP, adapted into 3DNA, and greatly expanded in DSSR (Lu, 2020). See Figure 2.

DSSR and 3DNA also take advantage of the standard base reference frame, which was approved by the Nomenclature Committee of IUBMB (NC-IUBMB)/IUPAC-IUBMB Joint Commission on Biochemical Nomenclature (JCBN) (Olson *et al.*, 2001). This international standard came about, in large part, due to the foundation laid by our contribution on resolving the discrepancies among nucleic acid conformational analyses (Lu and Olson, 1999; Lu *et al.*, 1999).

Just as 3DNA has replaced SCHNAaP/SCHNArP in functionality and real-world applications, DSSR has superseded 3DNA. Key 3DNA features for analysis, visualization, modeling, and utilities have been integrated into DSSR, with vastly enhanced functionality and significantly improved usability. Notably, the mutate/fiber/rebuild modules in DSSR completely supersedes the `mutate_bases`, `fiber`, and `rebuild` programs distributed with 3DNA v2.4. While 3DNA is still maintained, no additional features other than bug fixes are scheduled. In short, 3DNA is becoming the past; DSSR is the future.

1.5 Quality control of DSSR

DSSR is written in strict ANSI C, as a single command-line program. It is self-contained, with zero runtime dependencies on third-party libraries. The program has been extensively tested using all nucleic-acid-containing structures in the PDB and is continuously developed following user feedback. It is also regularly checked using Valgrind to avoid memory leaks.

DSSR is efficient and robust, with sensible defaults for input and intuitive outputs, making it accessible to a broad audience.

Overall, DSSR possesses an unmatched set of features, far more than a typical user would normally need. Some new features in DSSR or variations of well-known ones may come as a surprise even to experts in the field. By connecting dots in DNA/RNA structural bioinformatics, DSSR makes common tasks simple and sophisticated applications feasible.

The design and implementation of DSSR have benefited greatly from my extensive user-support experience, improved programming skills, and accumulated domain knowledge. Considering usability, interoperability, features, and support, DSSR easily stands out among its 'competitors'. I strive to make DSSR a pragmatic tool that the DNA/RNA structural bioinformatics community can count on.

References

Afek,A., Shi,H., Rangadurai,A., Sahay,H., Senitzki,A., Xhani,S., Fang,M., Salinas,R., Mielko,Z., Pufall,M.A., Poon,G.M.K., Haran,T.E., Schumacher,M.A., Al-Hashimi,H.M. and Gordân,R. (2020) DNA mismatches reveal conformational penalties in protein–DNA recognition. *Nature,* **587** (7833), 291–296.

Antczak,M., Zablocki,M., Zok,T., Rybarczyk,A., Blazewicz,J. and Szachniuk,M. (2019) RNAvista: a webserver to assess RNA secondary structures with non-canonical base pairs. *Bioinformatics,* **35** (1), 152–155.

Arnott,S. (1999) Polynucleotide secondary structures: an historical perspective. In *Oxford Handbook of Nucleic Acid Structure*, (Neidle,S., ed.). Oxford University Press, first edition, pp. 1–38.

Baulin,E., Metelev,V. and Bogdanov,A. (2020) Base-intercalated and base-wedged stacking elements in 3D-structure of RNA and RNA–protein complexes. *Nucleic Acids Res.,* **10.1093/nar/gkaa610**.

Bayrak,C.S., Kim,N. and Schlick,T. (2017) Using sequence signatures and kink-turn motifs in knowledge-based statistical potentials for RNA structure prediction. *Nucleic Acids Res.,* **45** (9), 5414–5422.

Becquey,L., Angel,E. and Tahi,F. (2020) RNANet: an automatically built dual-source dataset integrating homologous sequences and RNA structures. *Bioinformatics,* **10.1093/bioinformatics/btaa944**.

Berger,K.D., Kennedy,S.D. and Turner,D.H. (2019) Nuclear magnetic resonance reveals that GU base pairs flanking internal loops can adopt diverse structures. *Biochemistry,* **58** (8), 1094–1108.

Burley,S.K., Berman,H.M., Christie,C., Duarte,J.M., Feng,Z., Westbrook,J., Young,J. and Zardecki,C. (2018) RCSB Protein Data Bank: Sustaining a living digital data resource that enables breakthroughs in scientific research and biomedical education. *Protein Sci.,* **27** (1), 316–330.

Cai,Z., Cao,C., Ji,L., Ye,R., Wang,D., Xia,C., Wang,S., Du,Z., Hu,N., Yu,X., Chen,J., Wang,L., Yang,X., He,S. and Xue,Y. (2020) RIC-seq for global in situ profiling of RNA–RNA spatial interactions. *Nature,* **582** (7812), 432–437.

Calladine,C.R., Drew,H.R., Luisi,B.F. and Travers,A.A. (2004) *Understanding DNA: The molecule & how it works*. Third edition, Elsevier, Amsterdam.

Chan,C.W., Badong,D., Rajan,R. and Mondragón,A. (2020) Crystal structures of an unmodified bacterial tRNA reveal intrinsic structural flexibility and plasticity as general properties of unbound tRNAs. *RNA,* **26** (3), 278–289.

Chen,X., Khan,N.S. and Zhang,S. (2020) LocalSTAR3D: a local stack-based RNA 3D structural alignment tool. *Nucleic Acids Res.,* **48** (13), e77.

Cuturello,F., Tiana,G. and Bussi,G. (2020) Assessing the accuracy of direct-coupling analysis for RNA contact prediction. *RNA,* **26** (5), 637–647.

Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics,* **25** (15), 1974–5.

Desai,N., Brown,A., Amunts,A. and Ramakrishnan,V. (2017) The structure of the yeast mitochondrial ribosome. *Science,* **355** (6324), 528–531.

Gallego,D., Darré,L., Dans,P.D. and Orozco,M. (2019) VeriNA3d: An R package for nucleic acids data mining. *Bioinformatics,* **35** (24), 5334–5336.

Giambaşu,G.M., Case,D.A. and York,D.M. (2019) Predicting site-binding modes of ions and water to nucleic acids using molecular solvation theory. *J. Am. Chem. Soc.,* **141** (6), 2435–2445.

Hanson,R.M. and Lu,X.J. (2017) DSSR-enhanced visualization of nucleic acid structures in Jmol. *Nucleic Acids Res.,* **45** (W1), W528–W533.

Kribelbauer,J.F., Lu,X.J., Rohs,R., Mann,R.S. and Bussemaker,H.J. (2020) Toward a mechanistic understanding of DNA methylation readout by transcription factors. *J. Mol. Biol.,* **432** (6), 1801–1815.

Lemieux,S. and Major,F. (2002) RNA canonical and non-canonical base pairing types: a recognition method and complete repertoire. *Nucleic Acids Res.,* **30** (19), 4250–63.

Leontis,N.B. and Westhof,E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA,* **7** (4), 499–512.

Li,S., Olson,W.K. and Lu,X.J. (2019) Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures. *Nucleic Acids Res.,* **47** (W1), W26–W34.

Lu,J.S., Bindewald,E., Kasprzak,W.K. and Shapiro,B.A. (2018) RiboSketch: versatile visualization of multi-stranded RNA and DNA secondary structure. *Bioinformatics,* **34** (24), 4297–4299.

Lu,X.J. (2020) DSSR-enabled innovative schematics of 3D nucleic acid structures with PyMOL. *Nucleic Acids Res.,* **48** (13), e74.

Lu,X.J., Babcock,M.S. and Olson,W.K. (1999) Overview of Nucleic Acid Analysis Programs. *J. Biomol. Struct. Dyn.,* **16** (4), 833–843.

Lu,X.J., Bussemaker,H.J. and Olson,W.K. (2015) DSSR: An integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res.,* **43** (21), e142.

Lu,X.J., El Hassan,M. and Hunter,C. (1997*a*) Structure and conformation of helical nucleic acids: Analysis program (SCHNAaP). *J. Mol. Biol.,* **273** (3), 668–680.

Lu,X.J., El Hassan,M. and Hunter,C. (1997*b*) Structure and conformation of helical nucleic acids: Rebuilding program (SCHNArP). *J. Mol. Biol.,* **273** (3), 681–691.

Lu,X.J. and Olson,W.K. (1999) Resolving the discrepancies among nucleic acid conformational analyses. *J. Mol. Biol.,* **285** (4), 1563–75.

Lu,X.J. and Olson,W.K. (2003) 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.,* **31** (17), 5108–5121.

Lu,X.J. and Olson,W.K. (2008) 3DNA: A versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.,* **3** (7), 1213–1227.

Meier,M., Moya-Torres,A., Krahn,N.J., McDougall,M.D., Orriss,G.L., McRae,E.K.S., Booy,E.P., McEleney,K., Patel,T.R., McKenna,S.A. and Stetefeld,J. (2018) Structure and hydrodynamics of a DNA G-quadruplex with a cytosine bulge. *Nucleic Acids Res.,* **46** (10), 5319–5331.

Miskiewicz,J., Sarzynska,J. and Szachniuk,M. (2020) How bioinformatics resources work with G4 RNAs. *Brief. Bioinformatics,* **10.1093/bib/bbaa201**.

Narayanan,B.C., Westbrook,J., Ghosh,S., Petrov,A.I., Sweeney,B., Zirbel,C.L., Leontis,N.B. and Berman,H.M. (2014) The Nucleic Acid Database: New features and capabilities. *Nucleic Acids Res.,* **42** (D1), D114–D122.

Olson,W.K., Bansal,M., Burley,S.K., Dickerson,R.E., Gerstein,M., Harvey,S.C., Heinemann,U., Lu,X.J., Neidle,S., Shakked,Z., Sklenar,H., Suzuki,M., Tung,C.S., Westhof,E., Wolberger,C. and Berman,H.M. (2001) A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.,* **313** (1), 229–237.

Richardson,J.S., Schneider,B., Murray,L.W., Kapral,G.J., Immormino,R.M., Headd,J.J., Richardson,D.C., Ham,D., Hershkovits,E., Williams,L.D., Keating,K.S., Pyle,A.M., Micallef,D., Westbrook,J. and Berman,H.M. (2008) RNA backbone: consensus all-angle conformers and modular string nomenclature. *RNA,* **14** (3), 465–81.

Saaidi,A., Allouche,D., Regnier,M., Sargueil,B. and Ponty,Y. (2020) IPANEMAP: integrative probing analysis of nucleic acids empowered by multiple accessibility profiles. *Nucleic Acids Res.,* **48** (15), 8276–8289.

Saenger,W. (1984) *Principles of nucleic acid structure.* Springer advanced texts in chemistry, first edition, Springer-Verlag, New York.

Sagendorf,J.M., Markarian,N., Berman,H.M. and Rohs,R. (2020) DNAproDB: An expanded database and web-based tool for structural analysis of DNA-protein complexes. *Nucleic Acids Res.,* **48** (D1), D277–D287.

Sarver,M., Zirbel,C.L., Stombaugh,J., Mokdad,A. and Leontis,N.B. (2008) FR3D: finding local and composite recurrent structural motifs in RNA 3D structures. *J. Math. Biol.,* **56** (1-2), 215–52.

Shi,H., Rangadurai,A., Abou Assi,H., Roy,R., Case,D.A., Herschlag,D., Yesselman,J.D. and Al-Hashimi,H.M. (2020) Rapid and accurate determination of atomistic RNA dynamic ensemble models using NMR and structure prediction. *Nat. Commun.,* **11** (1), 5531.

Tan,D., Piana,S., Dirks,R.M. and Shaw,D.E. (2018) RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proc. Natl. Acad. Sci.,* **115** (7), E1346–E1355.

Thiel,B.C., Beckmann,I.K., Kerpedjiev,P. and Hofacker,I.L. (2019) 3D based on 2D: Calculating helix angles and stacking patterns using forgi 2.0, an RNA Python library centered on secondary structure elements. *F1000Research,* **8**, 287.

Yang,H., Jossinet,F., Leontis,N., Chen,L., Westbrook,J., Berman,H. and Westhof,E. (2003) Tools for the automatic identification and classification of RNA base pairs. *Nucleic Acids Res.,* **31** (13), 3450–3460.

Yesselman,J.D., Eiler,D., Carlson,E.D., Gotrik,M.R., d'Aquino,A.E., Ooms,A.N., Kladwang,W., Carlson,P.D., Shi,X., Costantino,D.A., Herschlag,D., Lucks,J.B., Jewett,M.C., Kieft,J.S. and Das,R. (2019) Computational design of three-dimensional RNA structure and function. *Nat. Nanotechnol.,* **14** (9), 866–873.

Zhao,J., Kennedy,S.D., Berger,K.D. and Turner,D.H. (2020) Nuclear Magnetic Resonance of Single-Stranded RNAs and DNAs of CAAU and UCAAUC as Benchmarks for Molecular Dynamics Simulations. *J. Chem. Theory Comput.,* **16** (3), 1968–1984.

Zok,T., Antczak,M., Zurkowski,M., Popenda,M., Blazewicz,J., Adamiak,R.W. and Szachniuk,M. (2018) RNApdbee 2.0: multifunctional tool for RNA structure annotation. *Nucleic Acids Res.,* **46** (W1), W30–W35.

Zok,T., Popenda,M. and Szachniuk,M. (2020) ElTetrado: A tool for identification and classification of tetrads and quadruplexes. *BMC Bioinformatics,* **21** (1), 40.