X3DNA-DSSR, a resource for structural bioinformatics of nucleic acids

Xiang-Jun Lu, Ph.D.

Department of Biological Sciences
Columbia University

December 9, 2021 (BioExcel Webinar)

DSSR: Dissecting the Spatial Structure of RNA

It excels in RNA/DNA structural bioinformatics.

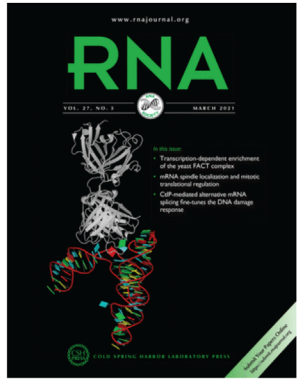
Structural analysis

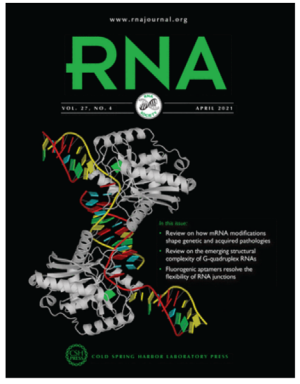
"Figures featuring cryo-EM maps were generated using **Chimera** (38). Maps colored by local resolution were generated using estimations of resolution by **ResMap** (50). Figures featuring only models were generated using **PyMOL** (51).

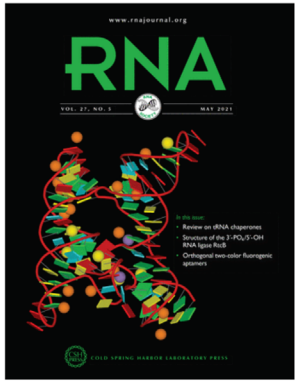
The secondary structure diagram for *S. cerevisiae* 15S rRNA was created by modifying the diagram from the **Comparative RNA Website** (46) with base pair information extracted from the final model using **DSSR** (52). ..."

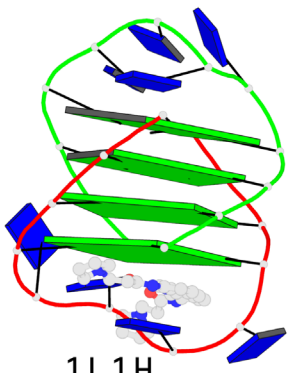
— Desai, ..., Ramakrishnan (2017), *Science*
"The structure of the yeast mitochondrial ribosome"

Block-view schematics

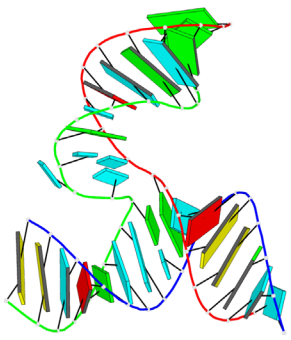




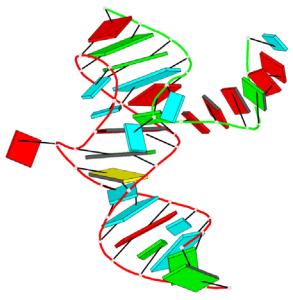




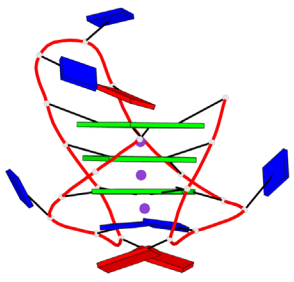
1L1H



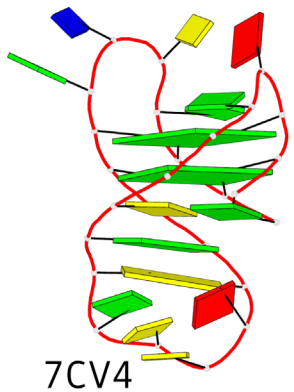
4KZ2



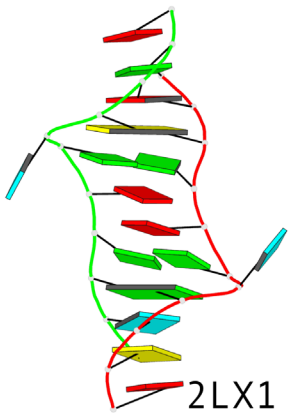
3B31



6JKN

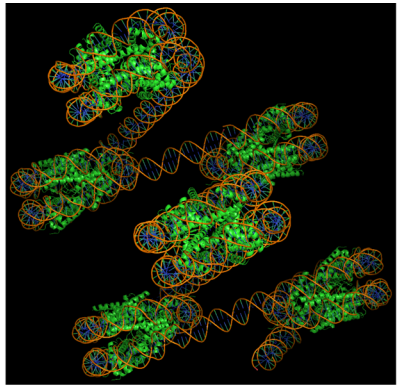


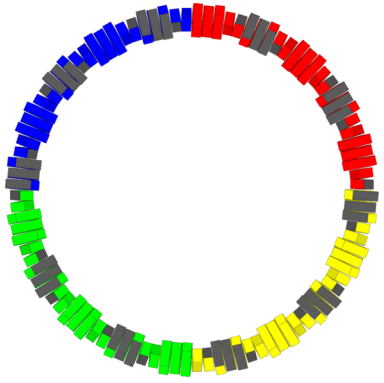
7CV4

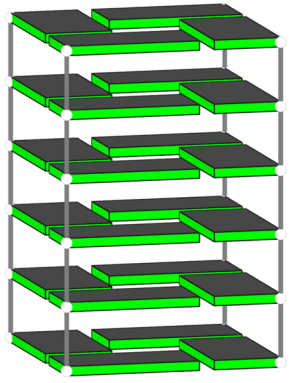


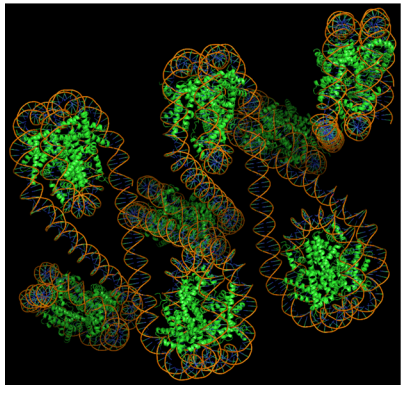
2LX1

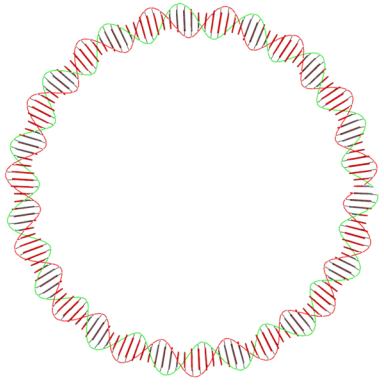
Model building

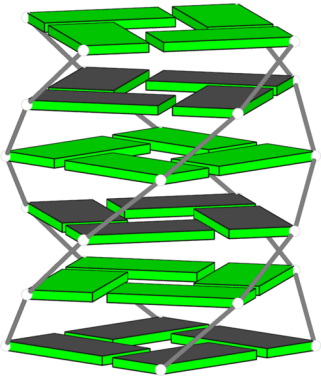








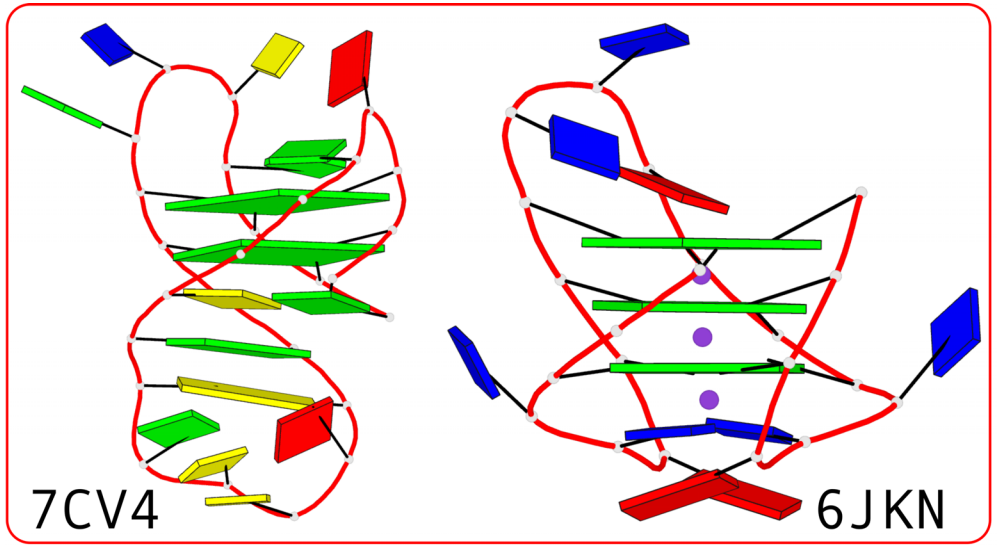


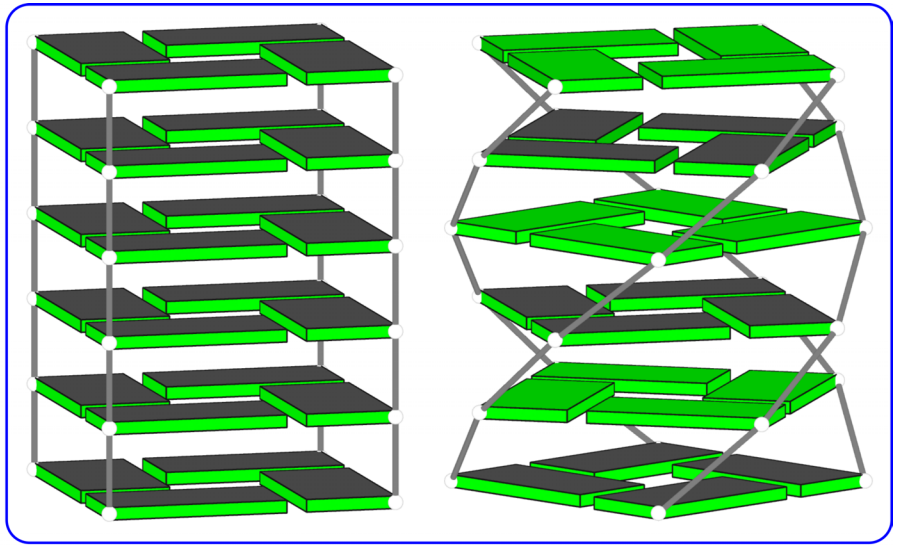


Easy integration

URS	RiboSketch	RNApdbee	RNAMotifContrast
Forgi	RNAvista	VeriNA3d	IsRNA2
RNAMake	ElTetrado	DNAproDB	SHAPELoop
R2DT	IPANEMAP	RNANet	3DSSR

G4-unique features





Identification and annotations **Model building**

Schematic representations

27 MARCH 2020• VOL 367 ISSUE 6485 1405

Underpromise, overdeliver

The majority of crises that most of us have lived through have not looked to science for immediate answers. In many cases, much of the scientific analysis came after the fact—the effects of climate change on extreme weather events; the causes of nuclear accidents; and the virology of outbreaks that were contained such as severe acute respiratory syndrome (SARS) in 2002–2003 or Middle East respiratory syndrome (MERS) in 2012. Now, science is being asked to provide a rapid solution to a problem that is not completely described.

I am worried that science may end up overpromising on what can be delivered in response to coronavirus disease 2019 (COVID-19). This isn't because I think the scientific community has bad intentions or will purposefully overhype anything, but because of what science can report in real time. It is difficult to share progress with adequate caveats about how long things might take or whether they will work at all. The scientific

> "...engendering false hope

remdesivir, novel antivirals, and numerous antibodies. These are exciting possibilities, but also extremely speculative. Political overhyping of such approaches is extremely dangerous—it risks creating false expectations and depleting drugs needed to treat diseases for which they are approved. And it sets science up to overpromise and underdeliver.

As for vaccines, we know so little about SARS-CoV-2. Developing a vaccine could take at least a year and a half—as many experts have suggested—or maybe won't happen at all. Fortunately, a clinical trial for a vaccine is already underway in the United States, but the public must be told that these early vaccines may not work or be safe—that this vaccine is only being tested for safety, not efficacy, at this point.

Scientists involved in COVID-19 research know these caveats. But the general public—who are agonizing over how long this pandemic will last, how it will affect the economy, and whether they and their loved ones will be safe—are looking for

H. Holden Thorp
Editor-in-Chief,
Science journals.
hthorp@aaas.org;
@hholdenthorp

Published results of DSSR are reproducible.

In fact, DSSR has a lot more to offer.

It enables innovative, cutting-edge applications.

DSSR is built upon 3DNA.

3DNA: 3-Dimensional Nucleic Acids (1999-2002)

5108–5121 *Nucleic Acids Research, 2003, Vol. 31, No. 17*
DOI: 10.1093/nar/gkg680

2003

Cited 1,740 times

3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures

Xiang-Jun Lu and Wilma K. Olson*

2008

Cited 588 times

3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures

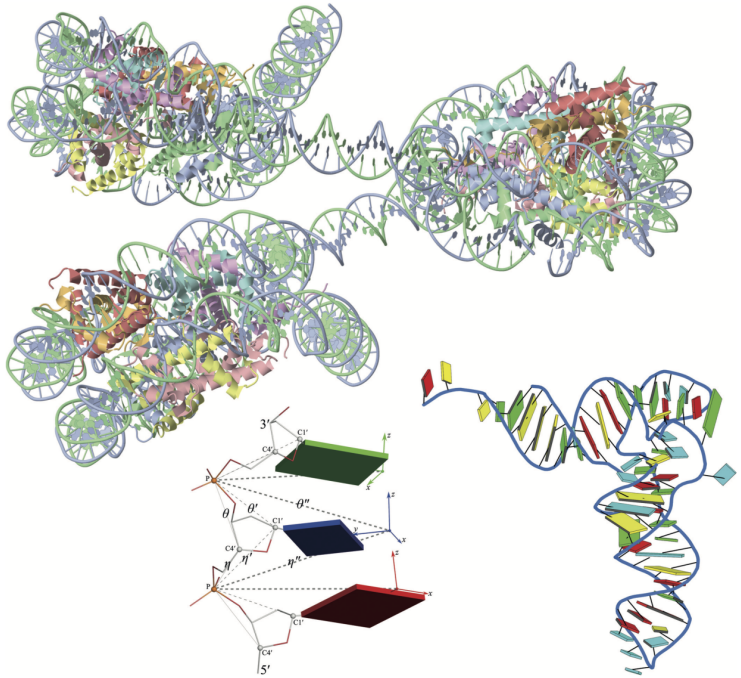
Xiang-Jun Lu[1,2] & Wilma K Olson[1]

NATURE PROTOCOLS | VOL.3 NO.7 | 2008 | 1213

PRINT ISSN: 0305-1048
ONLINE ISSN: 1362-4962

Nucleic Acids Research

VOLUME 47 WEB SERVER ISSUE JULY 2, 2019
https://academic.oup.com/nar



OXFORD UNIVERSITY PRESS

Open Access
No barriers to access – all articles freely available online

MIX Paper from responsible sources FSC® C007785

W26–W34 *Nucleic Acids Research, 2019, Vol. 47, Web Server issue*
doi: 10.1093/nar/gkz394

2019

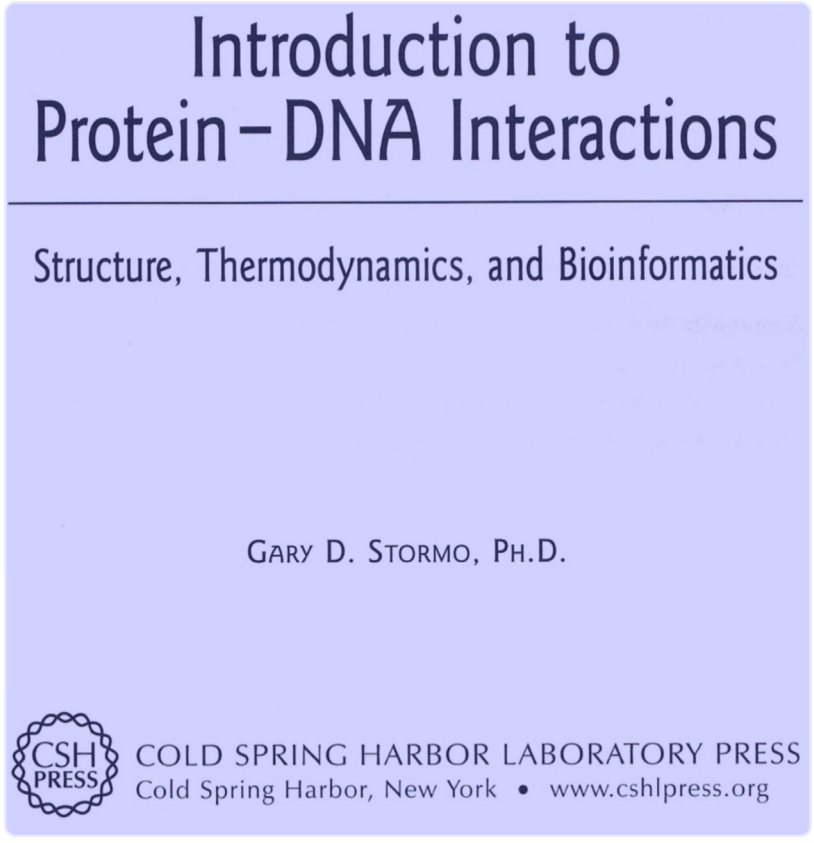
Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures

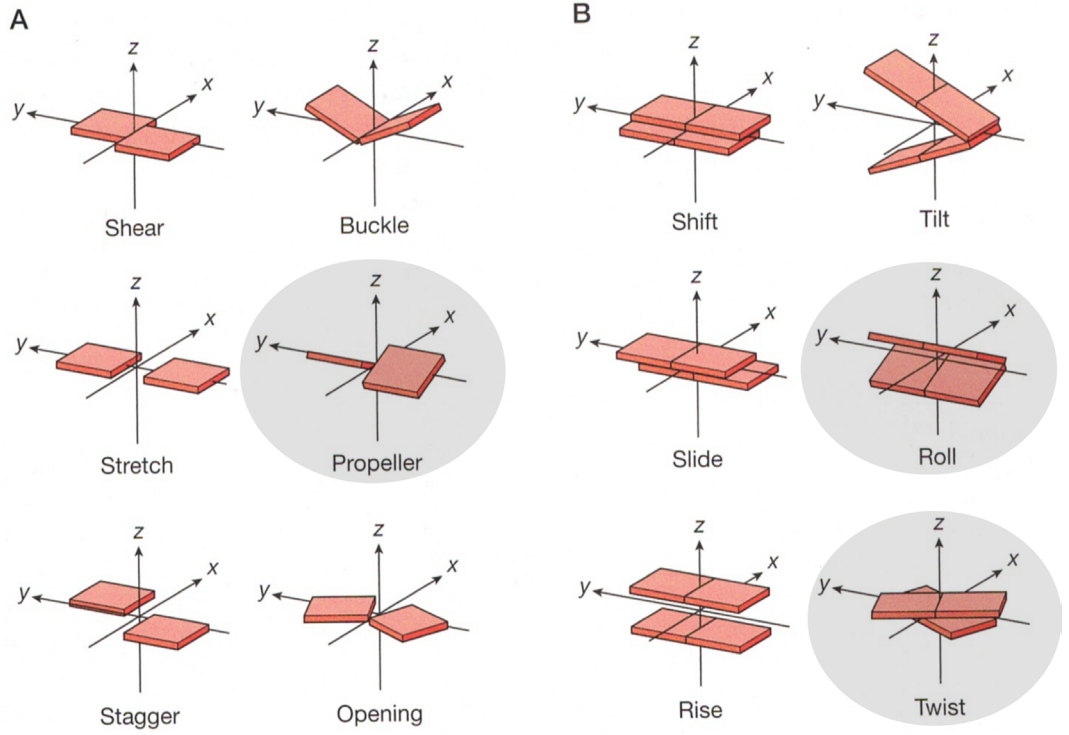
Shuxiang Li[1], Wilma K. Olson[1,*] and Xiang-Jun Lu[2,*]

http://web.x3dna.org/

The cover image is reproducible.

3DNA calculates DNA shape parameters





A, B redrawn from Lu & Olson (2003)

3DNA parameters in the Nucleic Acid Database (NDB)

NUCLEIC ACID DATABASE

A Portal for Three-dimensional Structural Information about Nucleic Acids
As of 1-Dec-2021 number of released structures: 11785

Search DNA | Search RNA | Advanced Search

Enter an NDB ID or PDB ID
Search for released structures

NDB ID: BDL084 PDB ID: 355D

Base Pair Morphology Step Parameters
CSV Format

Model Number	Step Number	Step Name	Shift	Slide	Rise	Tilt	Roll	Twist	X-Displacement	Y-Displacement	Helical Rise	Inclination	Tip	Helical Twist
1	5	AA_DA5DA6:DT19DT20_BB	0.171	-0.325	3.298	-0.33	0.459	37.52	-0.566	-0.31	3.293	0.713	0.514	37.524
1	6	AA_DA6DT7:DA18DT19_BB	-0.011	-0.601	3.219	-0.311	-2.675	32.403	-0.61	-0.034	3.256	-4.783	0.556	32.512
1	1	AA_DC1DG2:DC23DG24_BB	0.087	0.039	3.2	-3.216	8.52	32.731	-1.273	-0.654	3.09	14.766	5.573	33.94
1	3	AA_DC3DG4:DC21DG22_BB	-0.138	0.593	3.0	0.967	11.3	25.114	-1.405	0.518	2.977	24.457	-2.092	27.519
1	9	AA_DC9DG10:DC15DG16_BB	0.7	0.776	3.068	-3.656	4.18	26.581	0.66	-2.36	3.031	8.967	7.844	27.145
1	11	AA_DC11DG12:DC13DG14_BB	-0.309	0.211	3.174	-0.679	6.692	33.31	-0.686	0.423	3.161	11.528	1.169	33.964
1	2	AA_DG2DC3:DG22DC23_BB	0.496	0.668	3.691	2.847	-9.055	43.879	1.792	-0.363	3.517	-11.952	-3.757	44.844
1	4	AA_DG4DA5:DT20DC21_BB	-0.453	-0.139	3.388	-1.585	1.373	37.5	-0.402	0.489	3.396	2.133	2.463	37.556
1	10	AA_DG10DC11:DG14DC15_BB	-1.311	0.36	3.371	-2.853	-9.368	41.601	1.467	1.504	3.297	-12.975	3.951	42.688
1	7	AA_DT7DT8:DA17DA18_BB	-0.082	-0.397	3.216	1.681	-0.974	33.744	-0.529	0.408	3.218	-1.676	-2.893	33.798
1	8	AA_DT8DC9:DG16DA17_BB	-0.267	-0.226	3.465	0.684	-1.686	42.136	-0.13	0.446	3.467	-2.344	-0.951	42.174

The above values were obtained using first alternate conformation only and calculated by 3DNA program.

Xiang-Jun Lu & Wilma K. Olson (2003). '3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures', *Nucleic Acids Res.***31(17)**, 5108-21.
Xiang-Jun Lu & Wilma K. Olson (2008). '3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures', *Nat Protoc.***3(7)**, 1213-27.

3D-DART (Netherlands) and do_x3dna (Germany)

Published online 5 May 2009 *Nucleic Acids Research, 2009, Vol. 37, Web Server issue* ***W235–W239***
doi:10.1093/nar/gkp287

3D-DART: a DNA structure modelling server

Marc van Dijk and Alexandre M. J. J. Bonvin*

Bijvoet Center for Biomolecular Research, Science Faculty, Utrecht University, The Netherlands

3DNA-driven DNA Analysis and Rebuilding Tool

"3D-DART uses the DNA rebuild functionality of the well-known software package 3DNA Lu *et al.* and extends its functionally ..."

Structural bioinformatics *Bioinformatics, 31(15), 2015, 2583–2585*

do_x3dna: a tool to analyze structural fluctuations of dsDNA or dsRNA from molecular dynamics simulations

Rajendra Kumar* and Helmut Grubmüller

Department of Theoretical and Computational Biophysics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, Göttingen 37077, Germany

"It extends the capability of the 3DNA package to GROMACS MD trajectories and includes new ..."

DSSR is the next generation of 3DNA.

Continuously developed for 10 years (NIH R01 grant)

Expert domain knowledge of nucleic acid structures

Detail-oriented software engineering skills

Extensive user-support experience (3DNA Forum)

3DNA+more much=DSSR

Three DSSR papers, all in *Nucleic Acids Research* (NAR)

DSSR: an integrated software tool for dissecting the spatial structure of RNA

Xiang-Jun Lu ✉, Harmen J. Bussemaker, Wilma K. Olson

Nucleic Acids Research, Volume 43, Issue 21, 2 December 2015, Page e142,
https://doi.org/10.1093/nar/gkv716
Published: 15 July 2015 **Article history** ▼

2015

DSSR-enhanced visualization of nucleic acid structures in Jmol

Robert M. Hanson, Xiang-Jun Lu ✉

Nucleic Acids Research, Volume 45, Issue W1, 3 July 2017, Pages W528–W533,
https://doi.org/10.1093/nar/gkx365
Published: 03 May 2017 **Article history** ▼

2017

DSSR-enabled innovative schematics of 3D nucleic acid structures with PyMOL

Xiang-Jun Lu ✉

Nucleic Acids Research, Volume 48, Issue 13, 27 July 2020, Page e74,
https://doi.org/10.1093/nar/gkaa426
Published: 22 May 2020 **Article history** ▼

2020

Identification and analysis

Block-view schematics

Advanced model building

Integration into other resources

Features tailored to G-quadruplexes

Identification and analysis

Block-view schematics

Advanced model building

some typical examples, based on literature citations

How does DSSR actually work, *in command line*?

using the classic yeast tRNA^{Phe} as an example (PDB id: 1EHZ)

```
ls -lh   # in a demo folder with these 3 files
    -rw-r--r--  1 xiangjunlu  staff   245K Dec  8 08:26 1ehz.cif
    -rw-r--r--  1 xiangjunlu  staff   223K Dec  8 08:26 1ehz.pdb
    -rwxr-xr-x  1 xiangjunlu  staff   1.8M Dec  8 08:26 x3dna-dssr*

./x3dna-dssr -h              # to get started right away

./x3dna-dssr -i=1ehz.pdb     # PDB input file
./x3dna-dssr -i=1ehz.cif     # mmCIF input file

./x3dna-dssr -i=1ehz.pdb --json  # to get JSON output for easy parsing
```

Running DSSR on the classic yeast tRNA^{Phe} (PDB id: 1EHZ)

```
./x3dna-dssr -i=1ehz.pdb -o=1ehz.out

    total number of nucleotides: 76
           modified nucleotides: 14
    total number of base pairs: 34
    total number of helices: 2
    total number of stems: 4
    total number of isolated WC/wobble pairs: 1
    total number of multiplets: 4
    total number of atom-base capping interactions: 4
    total number of splayed-apart dinucleotides: 9
                        consolidated into units: 6
    total number of hairpin loops: 3
    total number of junctions: 1
    total number of non-loop single-stranded segments: 1
    total number of kissing loops: 1
```

Running DSSR on the classic yeast tRNA^{Phe} (PDB id: 1EHZ)

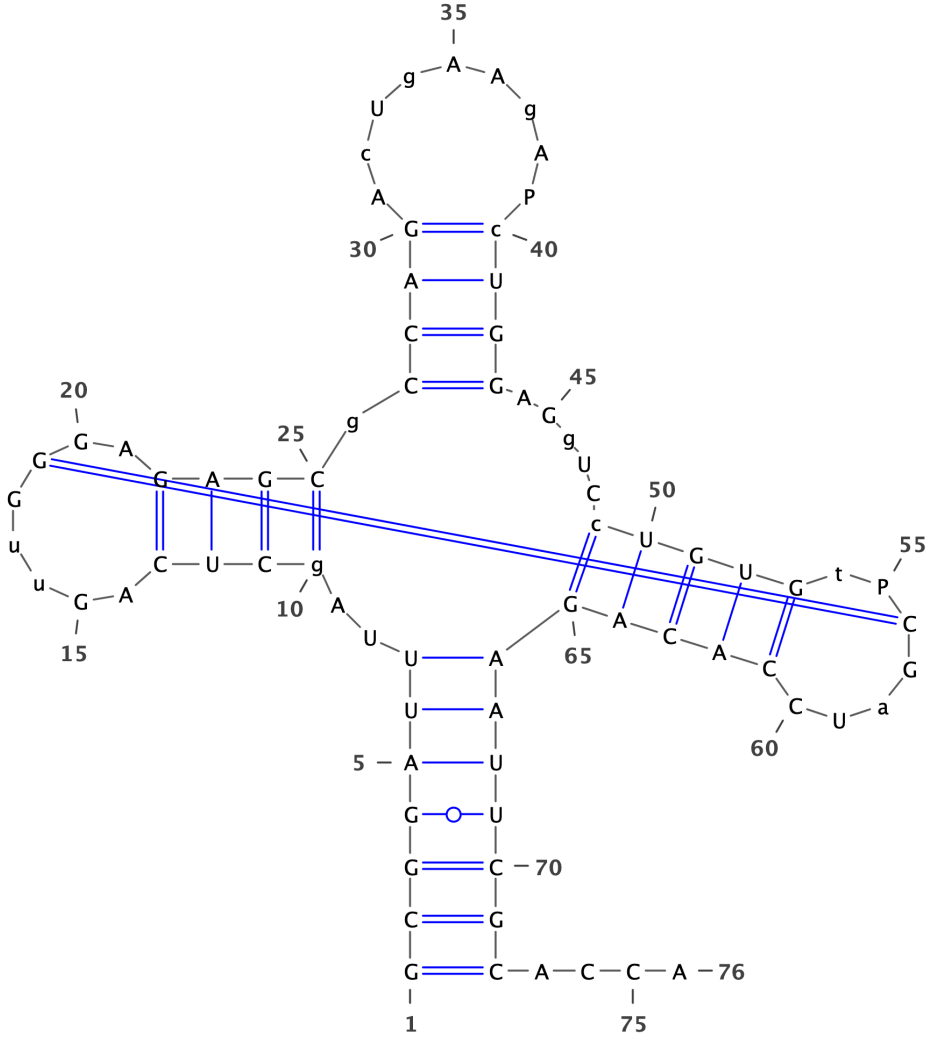
```
List of 34 base pairs
     nt1             nt2             bp  name         Saenger    LW   DSSR
   1 A.G1            A.C72           G-C WC           19-XIX     cWW  cW-W
   2 A.C2            A.G71           C-G WC           19-XIX     cWW  cW-W
   3 A.G3            A.C70           G-C WC           19-XIX     cWW  cW-W
   4 A.G4            A.U69           G-U Wobble       28-XXVIII  cWW  cW-W
   5 A.A5            A.U68           A-U WC           20-XX      cWW  cW-W
   6 A.U6            A.A67           U-A WC           20-XX      cWW  cW-W
   7 A.U7            A.A66           U-A WC           20-XX      cWW  cW-W
   8 A.U8            A.A14           U-A rHoogsteen   24-XXIV    tWH  tW-M
   9 A.U8            A.A21           U+A --           --         tSW  tm+W
  ..................................................................
  16 A.G15           A.C48           G+C rWC          22-XXII    tWW  tW+W
  ..................................................................
  33 A.G53           A.C61           G-C WC           19-XIX     cWW  cW-W
  34 A.5MU54         A.1MA58         t-a rHoogsteen   24-XXIV    tWH  tW-M
```

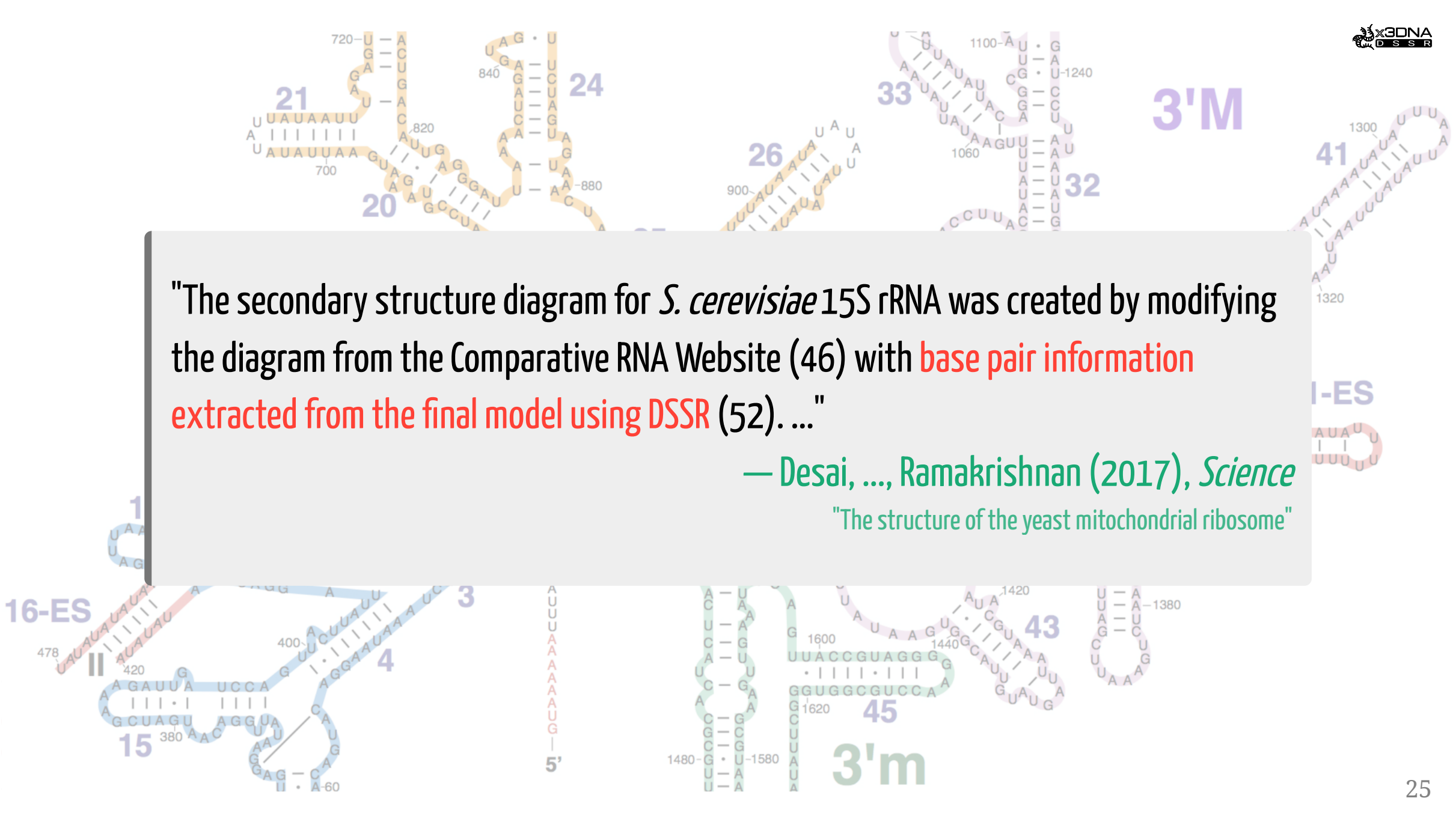
Running DSSR on the classic yeast tRNA^{Phe} (PDB id: 1EHZ)

DSSR derives 2D structure in:

- .bpseq (Comparative RNA Web)
- .ct (Connectivity Table)
- .dbn (Dot-Bracket Notation)

which can be directly fed into VARNA, for example, for visualization.





"The secondary structure diagram for *S. cerevisiae* 15S rRNA was created by modifying the diagram from the Comparative RNA Website (46) with base pair information extracted from the final model using DSSR (52). ..."

— Desai, ..., Ramakrishnan (2017), *Science*

"The structure of the yeast mitochondrial ribosome"

> "Conformations of the ssRNA tetramers were categorized based on their base-stacking patterns, which were analyzed by the program DSSR."
>
> — Tan, ..., DE Shaw (2018), *Proc. Natl. Acad. Sci.*
>
> "RNA force field with accuracy comparable to state-of-the-art protein force fields"

> "..., we utilized the Dissecting the Spatial Structure of RNA (DSSR) tool to identify the 2D structure from the 3D structures. This important addition to the procedure not only automates the process, it also removes human error ..."
>
> — Hurst and Chen (2021), *J. Phys. Chem. B*
>
> "Sieving RNA 3D structures with SHAPE and evaluating mechanisms driving sequence-dependent reactivity bias"

"Watson–Crick and non-Watson–Crick base pairs were identified using the DSSR software."

— Cai, ... Xue (2020), *Nature*

"RIC-seq for global *in situ* profiling of RNA-RNA spatial interactions"

"The overall shape of the DNA was characterized by analysing the following shape parameters: minor groove width, major groove width, local helical bending, bending direction and local helical twisting, ... using X3DNA-DSSR."

— Afek, ... Al-Hashimi, and Raluca Gordân (2020), *Nature*

"DNA mismatches reveal conformational penalties in protein–DNA recognition"

Identification and analysis

Block-view schematics

Advanced model building

simple, effective, and aesthetically pleasing

rendered with PyMOL (web-interface, plugin, CLI, API)

Base blocks make base identity, pairing/stacking obvious

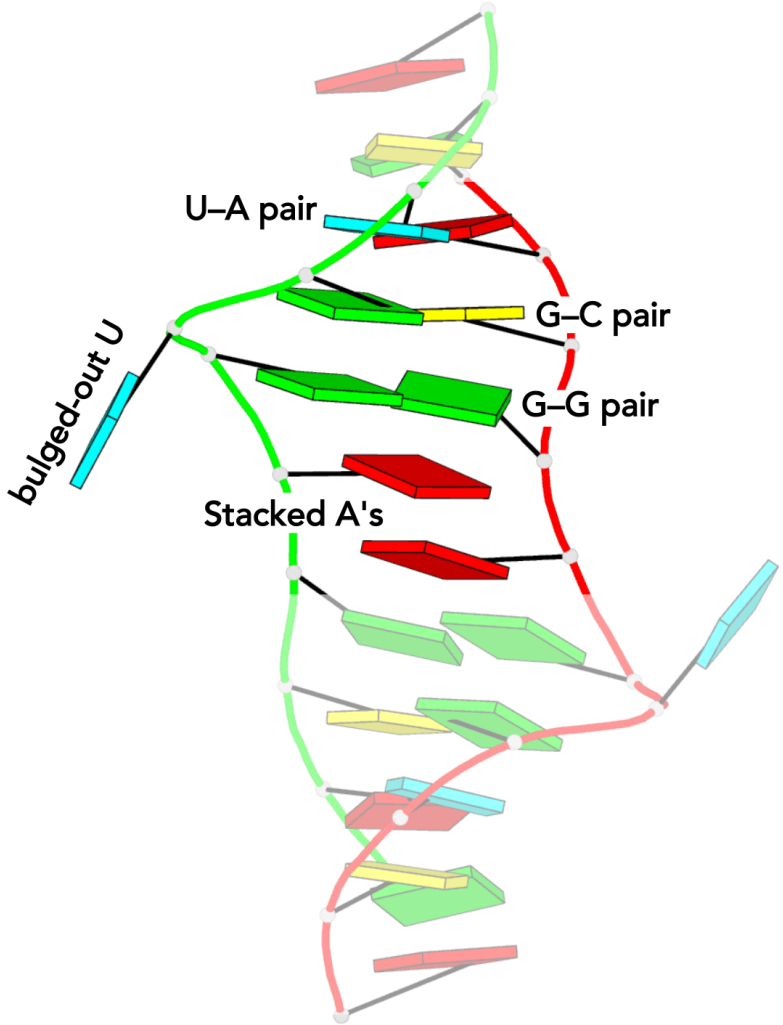
G: green

C: yellow

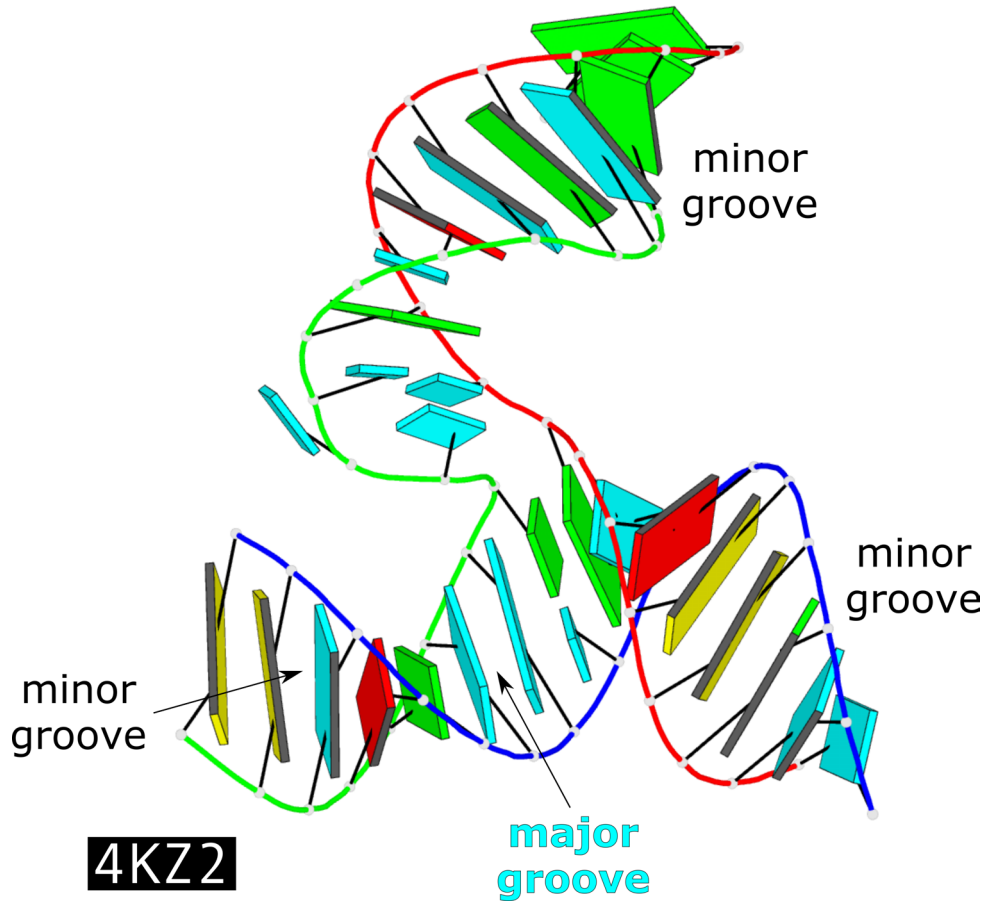
A: red

U: cyan

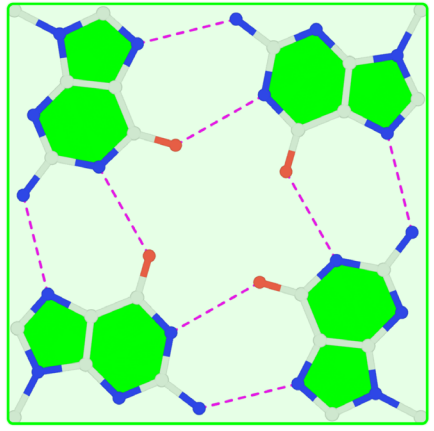
T: blue

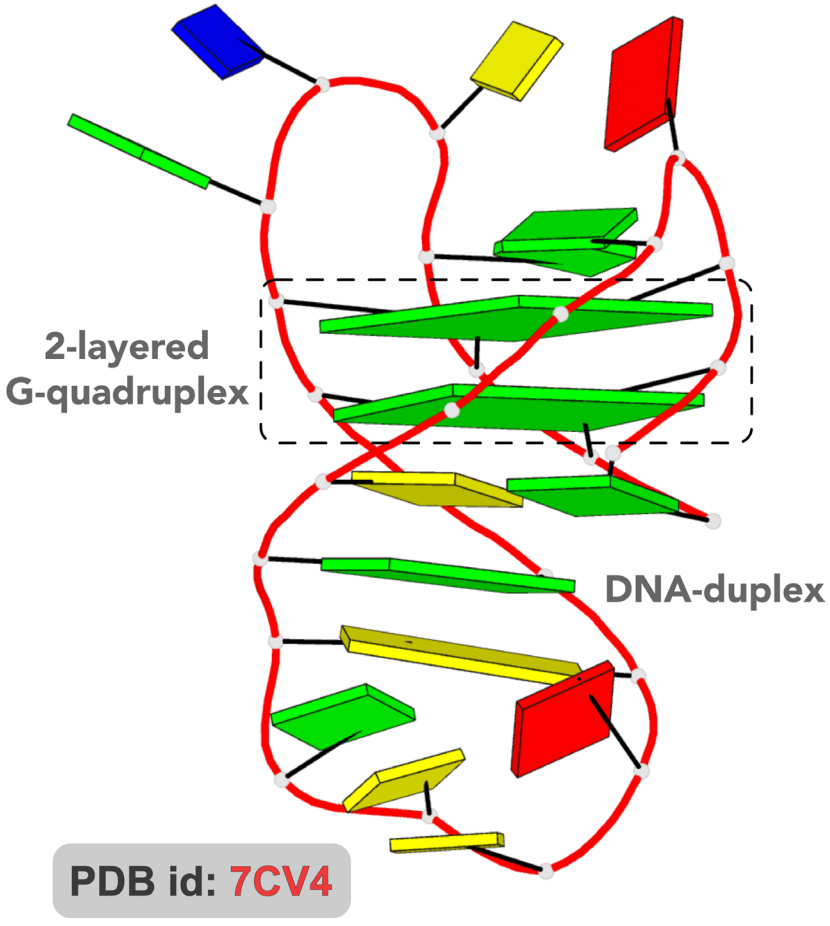


WC-pair blocks reveal double-stranded regions and grooves

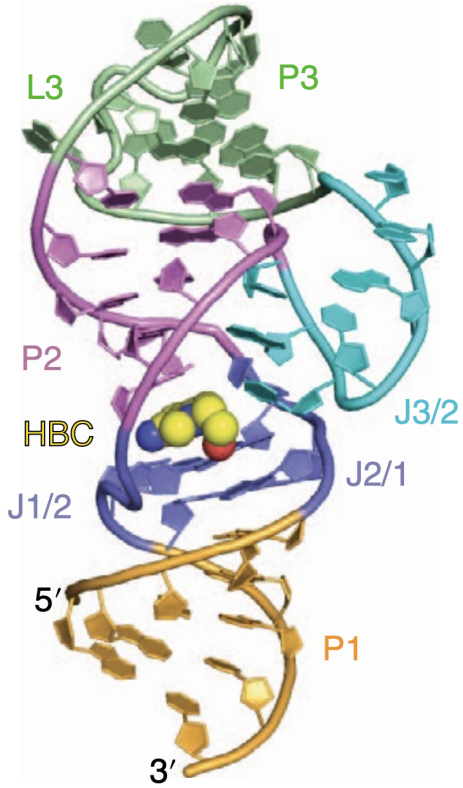


G-tetrad square blocks simplify G-quadruplexes

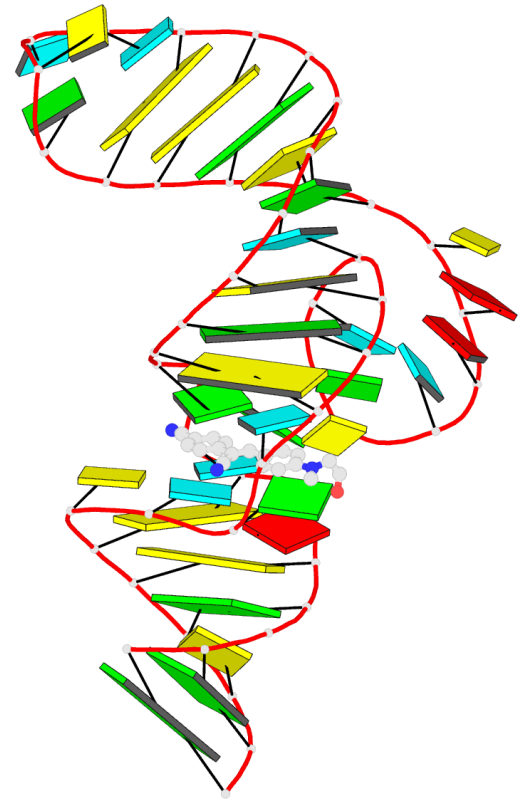




cf.#1 Pepper aptamer in complex with HBC (PDB id: 7EOH)

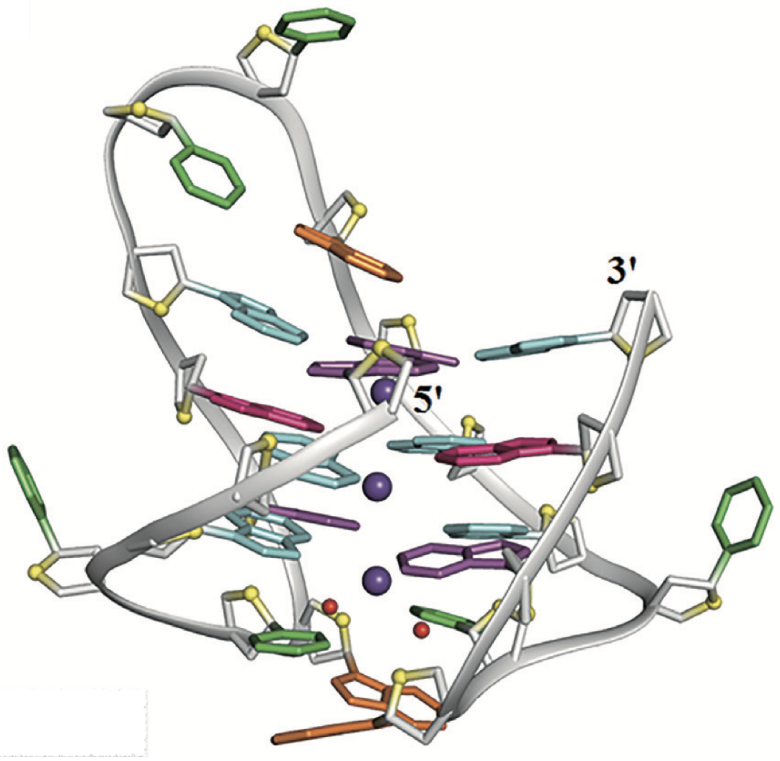


Huang *et al.* (2021) *Nat. Chem. Biol.*

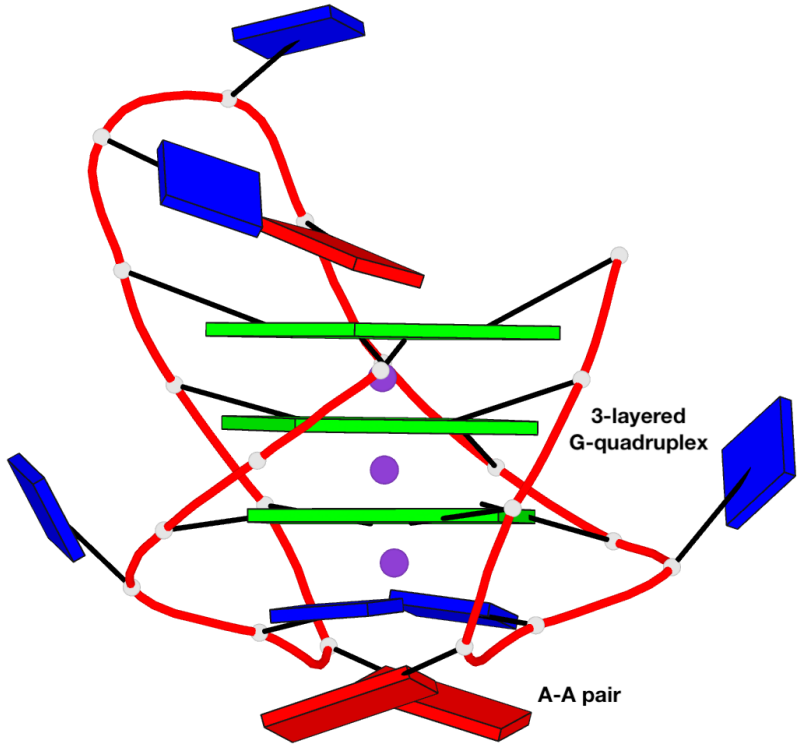


DSSR-PyMOL schematic

cf.#2 Chair-type telomeric DNA G-quadruplex (PDB id: 6JKN)



Geng *et al.* (2019) *Nucleic Acids Res.*



DSSR-PyMOL schematic

The block-view schematics are highly acclaimed.

** Very Good
(Good for Teaching)

"simple and effective"

"I highly recommend it!"

Rated ★★ **Very Good**

Faculty Opinions

14 Aug 2020

Quentin Vicens | **Faculty Member**

Chemical Biology / Macromolecular Assembly & Chemistry

University of Colorado Denver - School of Medicine
Aurora, CO
USA

Classified as

Good for Teaching

I really enjoyed “playing” with the revised and expanded version of Dissecting the Spatial Structure of RNA (DSSR) described by Xiang-Jun Lu in this July issue of NAR. The software is known to generate ‘block view’ representations of nucleic acids that make many parameters more immediately visible, such as base composition, stacking, and groove depth. This new version includes Watson-Crick pairs shown as single rectangles, and G quadruplexes as large squares, making such regions more quickly distinguishable from other regions within an overall tertiary structure. I was amazed at how simple and effective the web interface was, and I liked the possibility to download a PyMOL session to look at molecules under different angles. If need be, blocks can be further edited in PyMOL using the provided plugin (see on page 35). I highly recommend it!

Cover images of the *RNA Journal* (2021) provided by the NDB

"generated using DSSR and PyMOL (Lu XJ. 2020. *Nucleic Acids Res* **48**: e74)."

January

February

March

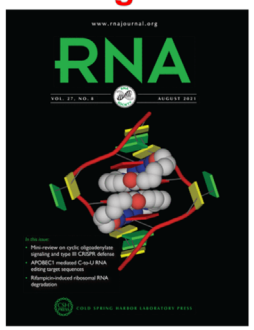
April

May

June

July

August



September

October

November

December

http://skmatic.x3dna.org/ – it is **easy** to use!

Specify a PDB id or a number for pre-calculated results

PDB id (4-char: e.g., 1ehz · 2grb · 2hoj)
1ehz

Random sample (3 to 99: e.g., 12 · 30 · 60)
12

Submit | Reset

Specify a PDB or mmCIF coordinate file and set options

Input a URL (e.g., PDB · mmCIF)
https://files.rcsb.org/download/1EHZ.pdb.gz

Upload a file (e.g., PDB · mmCIF)
Choose File no file selected

☐ With six orthogonal views ☐ Viewed in raw coordinates ◉ As for PDB entries

Styles (e.g.#1 · e.g.#2)
face [list of keywords]

Colors (e.g.#1 · e.g.#2)
A:red, C:yellow, G:green, T:blue, U:cyan

Depth (e.g., thin · thick)
0.5

Submit | Reset

Identification and analysis

Block-view schematics

Advanced model building

RNA modeling, DNA-protein complexes, G-quadruplexes

regular fiber-based models, customized models by sequence/parameters

3DNA/DSSR for RNA modeling by Merck scientists

"3DNA program for RNA modeling"

Merck.com (Apr 21, 2015)

"I am writing to see if I can use 3DNA for my RNA modeling research. There are very limited number of commercially available modeling tools for DNA/RNA. I am very interested in testing your programs."

3DNA/DSSR for RNA modeling by Merck scientists

"3DNA program for RNA modeling"

Merck.com (Apr 21, 2015)

"I am writing to see if I can use 3DNA for my RNA modeling research. There are very limited number of commercially available modeling tools for DNA/RNA. I am very interested in testing your programs."

ARTICLE *Nature* (2015)

doi:10.1038/nature15542

Selective small-molecule inhibition of an RNA structural element

John A. Howe[1]*, Hao Wang[1]*, Thierry O. Fischmann[1]*, Carl J. Balibar[1], Li Xiao[1], Andrew M. Galgoci[1], Juliana C. Malinverni[1], Todd Mayhood[1], Artjohn Villafania[1], Ali Nahvi[2], Nicholas Murgolo[1], Christopher M. Barbieri[1], Paul A. Mann[1], Donna Carr[1], Ellen Xia[1], Paul Zuck[3], Dan Riley[3], Ronald E.Painter[1], Scott S. Walker[1], Brad Sherborne[1], Reynalda de Jesus[1], Weidong Pan[1], Michael A. Plotkin[1], Jin Wu[1], Diane Rindgen[1], John Cummings[1], Charles G. Garlisi[1], Rumin Zhang[1], Payal R. Sheth[1], Charles J. Gill[1], Haifeng Tang[1] & Terry Roemer[1]

3DNA/DSSR for RNA modeling by Merck scientists

"3DNA program for RNA modeling"

Merck.com (Apr 21, 2015)

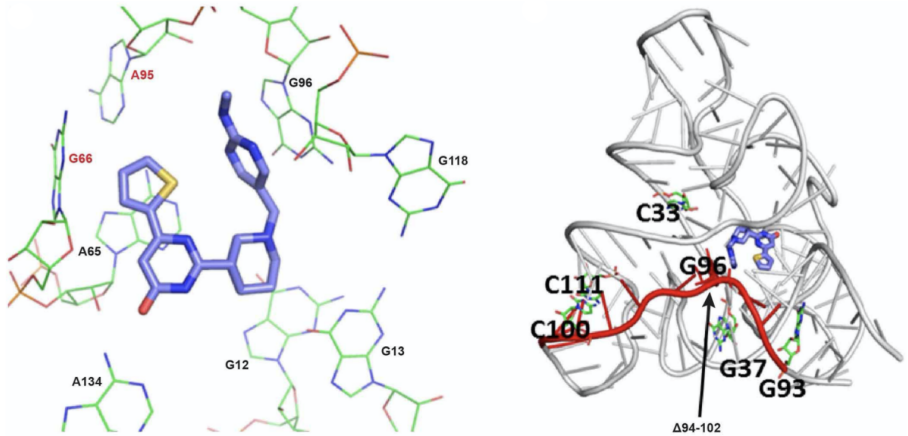
"I am writing to see if I can use 3DNA for my RNA modeling research. There are very limited number of commercially available modeling tools for DNA/RNA. I am very interested in testing your programs."

ARTICLE Nature (2015)

doi:10.1038/nature15542

Selective small-molecule inhibition of an RNA structural element

John A. Howe[1]*, Hao Wang[1]*, Thierry O. Fischmann[1]*, Carl J. Balibar[1], Li Xiao[1], Andrew M. Galgoci[1], Juliana C. Malinverni[1], Todd Mayhood[1], Artjohn Villafania[1], Ali Nahvi[2], Nicholas Murgolo[1], Christopher M. Barbieri[1], Paul A. Mann[1], Donna Carr[1], Ellen Xia[1], Paul Zuck[3], Dan Riley[3], Ronald E.Painter[1], Scott S. Walker[1], Brad Sherborne[1], Reynalda de Jesus[1], Weidong Pan[1], Michael A. Plotkin[1], Jin Wu[1], Diane Rindgen[1], John Cummings[1], Charles G. Garlisi[1], Rumin Zhang[1], Payal R. Sheth[1], Charles J. Gill[1], Haifeng Tang[1] & Terry Roemer[1]



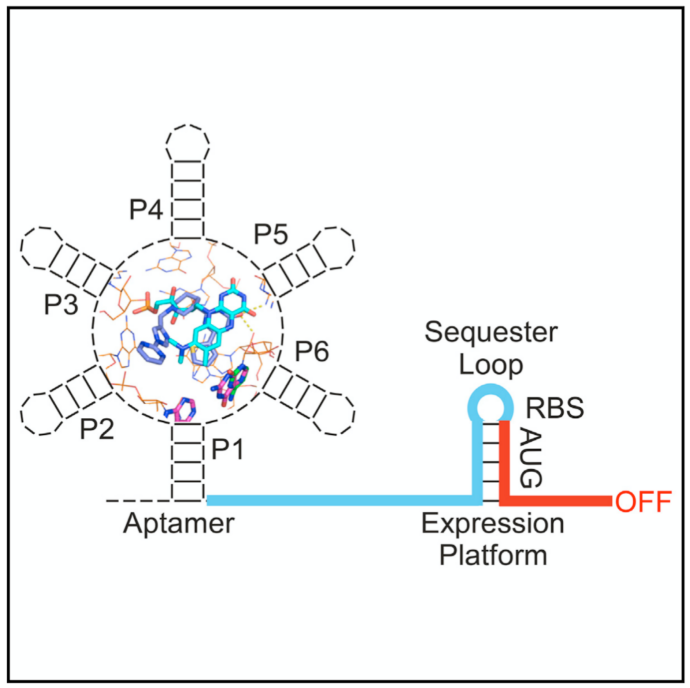
"**The homology model** was constructed using program mutate_bases of the 3DNA package using the *F. nucleatum* impX riboswitch aptamer X-ray structure as the template and the FMN aptamer alignment of *E. coli*, *F. nucleatum*, *P. aeruginosa* and *A. baumannii*."

3DNA/DSSR for RNA modeling by Merck scientists

Cell Chemical Biology (2017)

Dual-Targeting Small-Molecule Inhibitors of the *Staphylococcus aureus* FMN Riboswitch Disrupt Riboflavin Homeostasis in an Infectious Setting

Graphical Abstract



Authors

Hao Wang, Paul A. Mann, Li Xiao, ..., Amy Flattery, Matthias Mack, Terry Roemer

Correspondence

terry_roemer@merck.com

In Brief

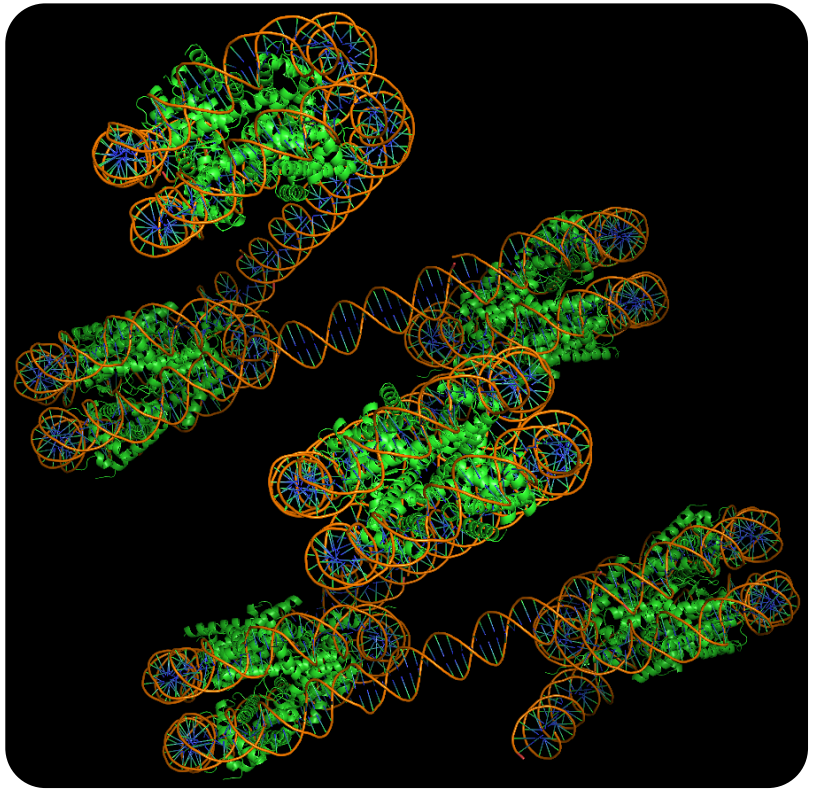
Wang et al. demonstrate that ribocil-C and roseoflavin selectively target functionally related non-coding RNA structural elements termed FMN riboswitches controlling gene expression of riboflavin biosynthesis and uptake. Such targets and cognate inhibitors offer new opportunities and challenges to antibiotic discovery.

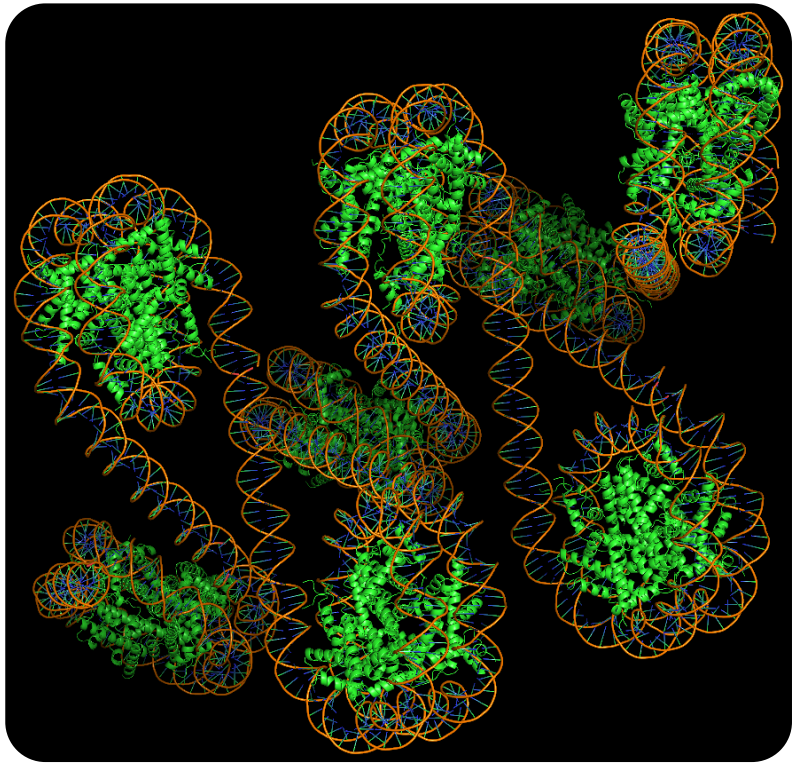
"**The homology models** of both S. aureus SA1 and SA2 FMN aptamers **were constructed using program mutate_bases (Lu and Olson, 2003) of the 3DNA package**, with the F. nucleatum riboswitch aptamer X-ray structure as the template."

The 3DNA **mutate_bases** program has been **integrated into DSSR and *substantially improved*.**

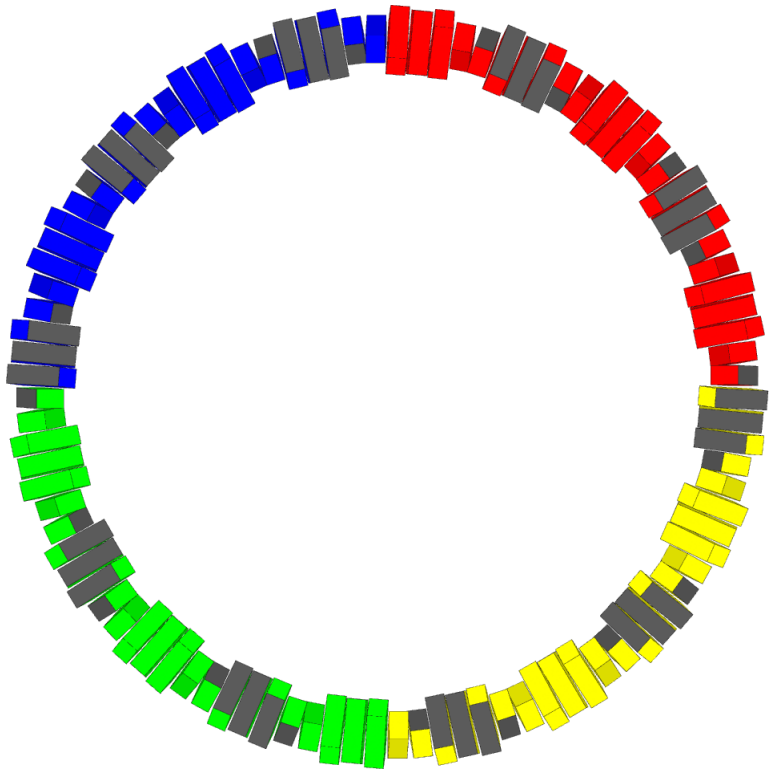
Template-based assembly of DNA-protein complexes

two chromatin-like models, derived from PDB id: 4XZQ

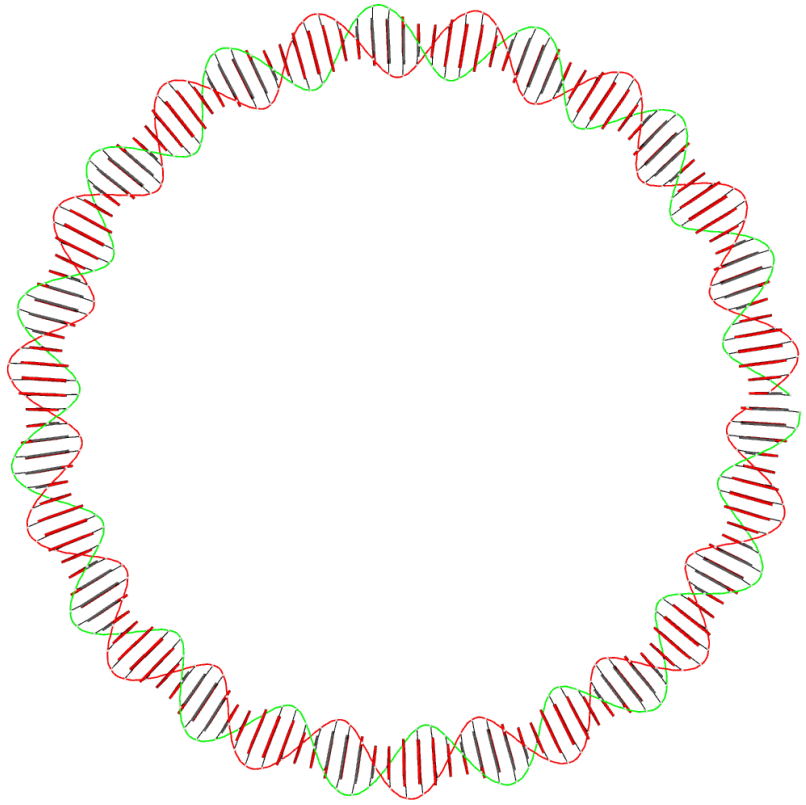




Circular DNA in perfectly planar geometry

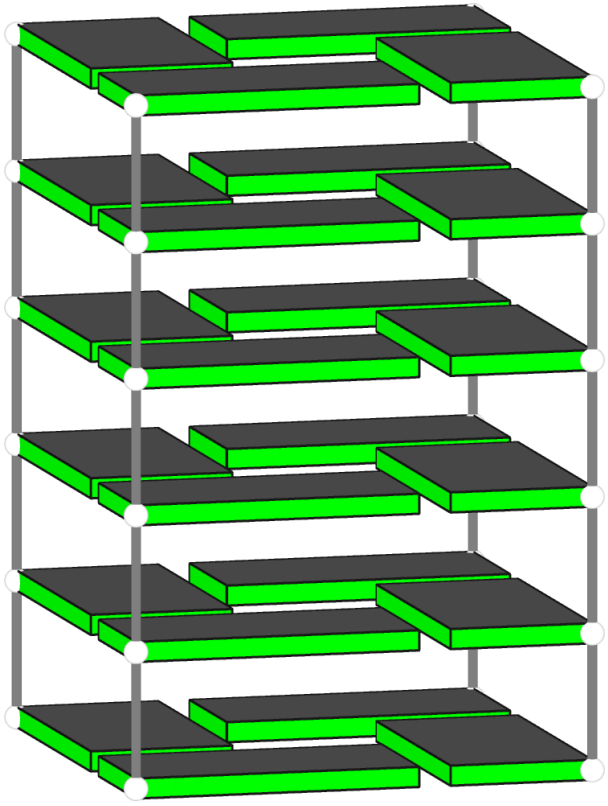


100 base pairs

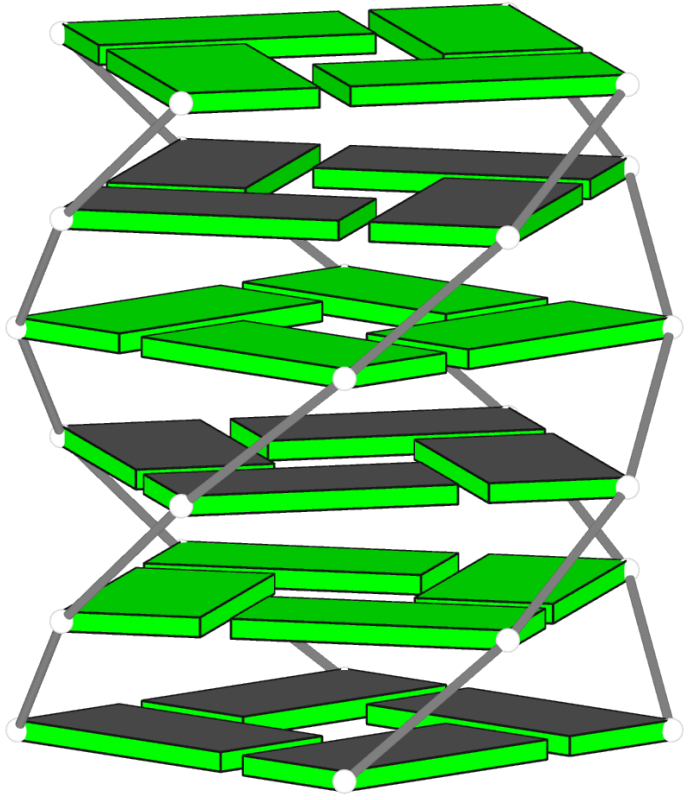


150 base pairs

Unique capabilities for modeling G-quadruplexes



no twist, similar faces



30° twist, alternating faces

Identification and analysis

Block-view schematics

Advanced model building

Integration into other resources

DSSR-Jmol integration: SQL-like queries

```
select junctions
```

to select all (multi-brunch) junction loops

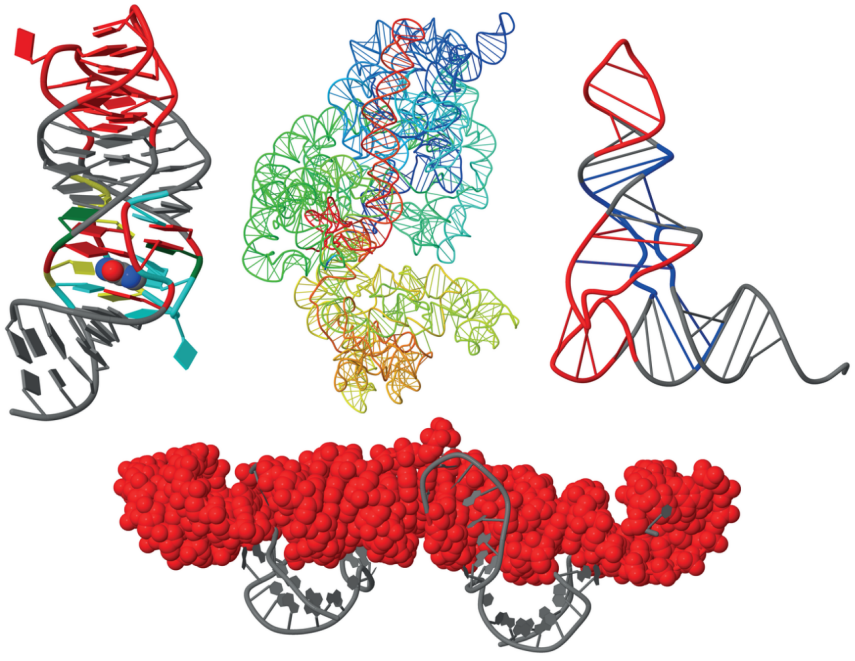
```
SELECT WITHIN(dssr,
"pairs WHERE name != 'WC'")
```

to select all non-Watson-Crick pairs

PRINT ISSN: 0305-1048
ONLINE ISSN: 1362-4962

Nucleic Acids Research

VOLUME 45 WEB SERVER ISSUE JULY 3, 2017
https://academic.oup.com/nar



OXFORD UNIVERSITY PRESS

Open Access
No barriers to access – all articles freely available online

DSSR-enhanced visualization of nucleic acid structures in Jmol

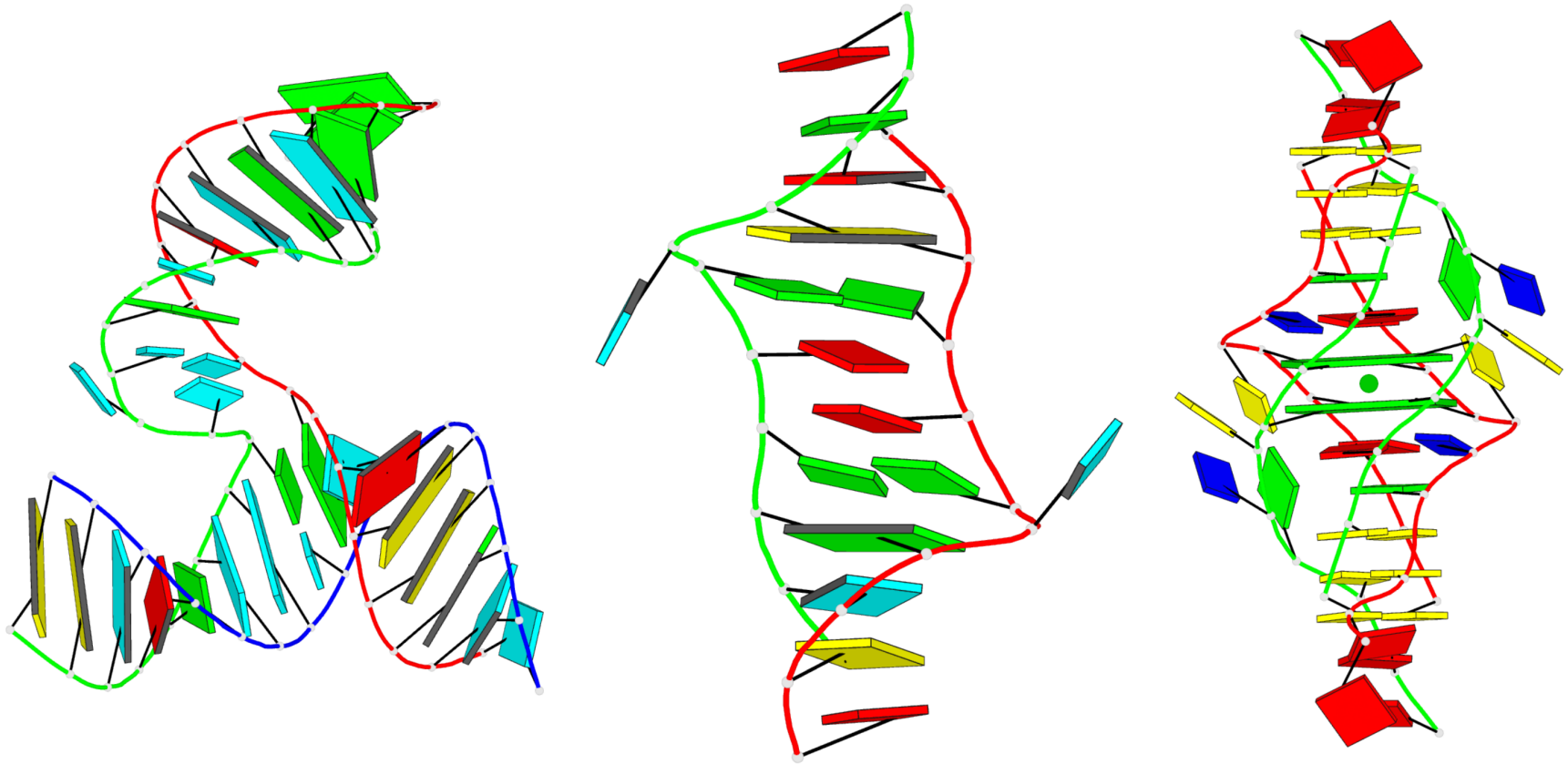
Robert M. Hanson, Xiang-Jun Lu

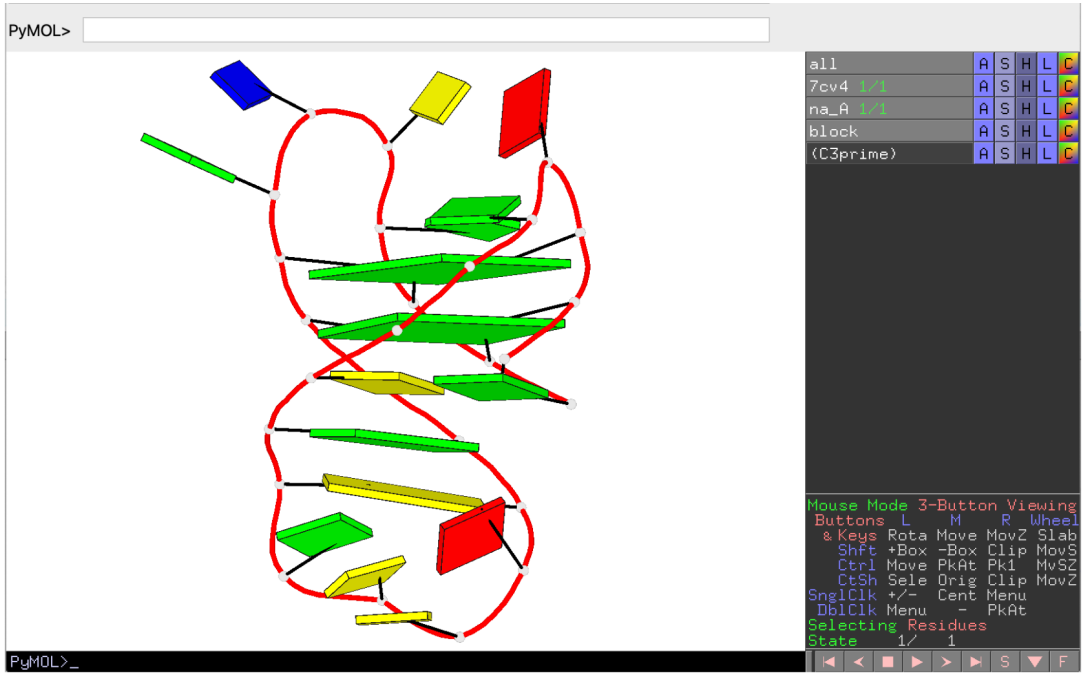
Nucleic Acids Research, Volume 45, Issue W1, 3 July 2017, Pages W528–W533,
https://doi.org/10.1093/nar/gkx365
Published: 03 May 2017 **Article history** ▼

2017

Dr. Robert Hanson
Principal Jmol Developer
St. Olaf College, Minnesota

DSSR-PyMOL integration: innovative schematics





dssr_block plugin

DSSR-enabled innovative schematics of 3D nucleic acid structures with PyMOL

Xiang-Jun Lu

Nucleic Acids Research, Volume 48, Issue 13, 27 July 2020, Page e74,
https://doi.org/10.1093/nar/gkaa426
Published: 22 May 2020 **Article history** ▼ 2020

"Finally, all results reported here are completely reproducible."

Thomas Holder
Principal PyMOL Developer
Schrödinger, Inc.

Sixteen published resources/pipelines employing DSSR

URS	RiboSketch	RNApdbee	RNAMotifContrast
Forgi	RNAvista	VeriNA3d	IsRNA2
RNAMake	ElTetrado	DNAproDB	SHAPELoop
R2DT	IPANEMAP	RNANet	3DSSR

PDB-REDO makes use of DSSR

Recent joint-publication with PDB-REDO

research papers

STRUCTURAL BIOLOGY

ISSN 2059-7983

New restraints and validation approaches for nucleic acid structures in *PDB-REDO*

Ida de Vries,[a] Tim Kwakman,[a] Xiang-Jun Lu,[b] Maarten L. Hekkelman,[a] Mandar Deshpande,[c] Sameer Velankar,[c] Anastassis Perrakis[a]* and Robbie P. Joosten[a]*

[a]Oncode Institute and Division of Biochemistry, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands, [b]Department of Biological Sciences, Columbia University, New York, NY 10027, USA, and [c]Protein Data Bank in Europe (PDBe), European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL–EBI), Wellcome Genome Campus, Hinxton CB10 1SD, United Kingdom. *Correspondence e-mail: a.perrakis@nki.nl, r.joosten@nki.nl

Received 26 May 2021
Accepted 26 July 2021

Edited by A. G. Cook, University of Edinburgh, United Kingdom

Keywords: nucleic acid restraints; Watson–Crick base pairs; validation; *PDB-REDO*; x3*DNA-DSSR*.

The quality of macromolecular structure models crucially depends on refinement and validation targets, which optimally describe the expected chemistry. Commonly used software for these two procedures has been designed and developed in a protein-centric manner, resulting in relatively few established features for the refinement and validation of nucleic acid-containing

DSSR's systematic, integrated approach enables novel applications to be developed **rapidly**.

http://snap-5mc.x3dna.org/ (SNAP is now part of DSSR)

- Transcription factor-DNA complexes containing 5-methyl-cytosines in the PDB.
- Kribelbauer *et al. Journal of Molecular Biology* (2020) **432**, 1801--1815

Annotated list of i-motifs (C+C pairs) in the PDB (*unpublished*)

Identification and analysis

Block-view schematics

Advanced model building

Features tailored to G-quadruplexes

schematics, modeling, identification and annotations

How many G-quadruplexes (G4) in the PDB?

Counts from leading experts:

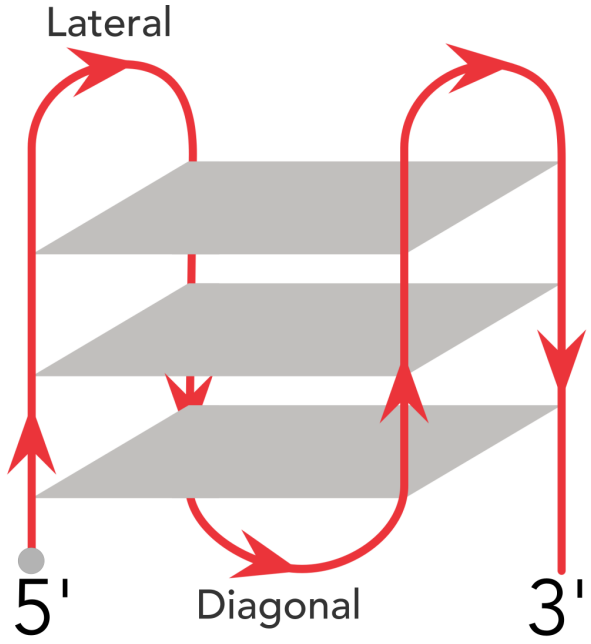
- Mergny (*Biochimie*, 2020): 200+
- Ferré-D'Amaré (*RNA*, 2021): 246
- Neidle (*JBC*, 2021): 520

The actual counts:

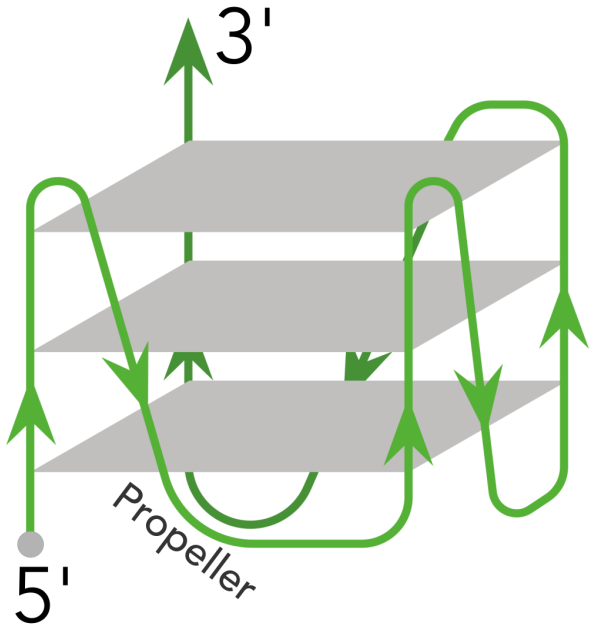
- 372, end of 2020
- 415 (Dec. 2, 2021)

DSSR identifies G4s automatically, using atomic coordinates

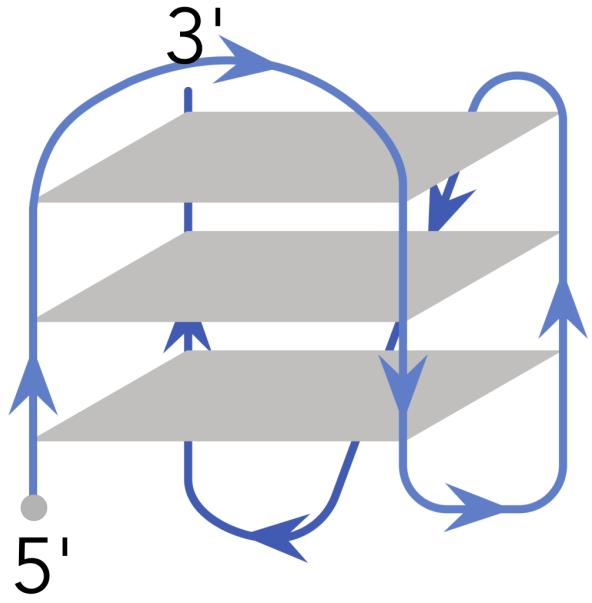
Gn-L1-Gn-L2-Gn-L3-Gn: different loops, highly polymorphic



Anti-parallel (2+2)



Parallel (4+0)



Mixed (3+1)

How to characterize G-quadruplexes systematically?

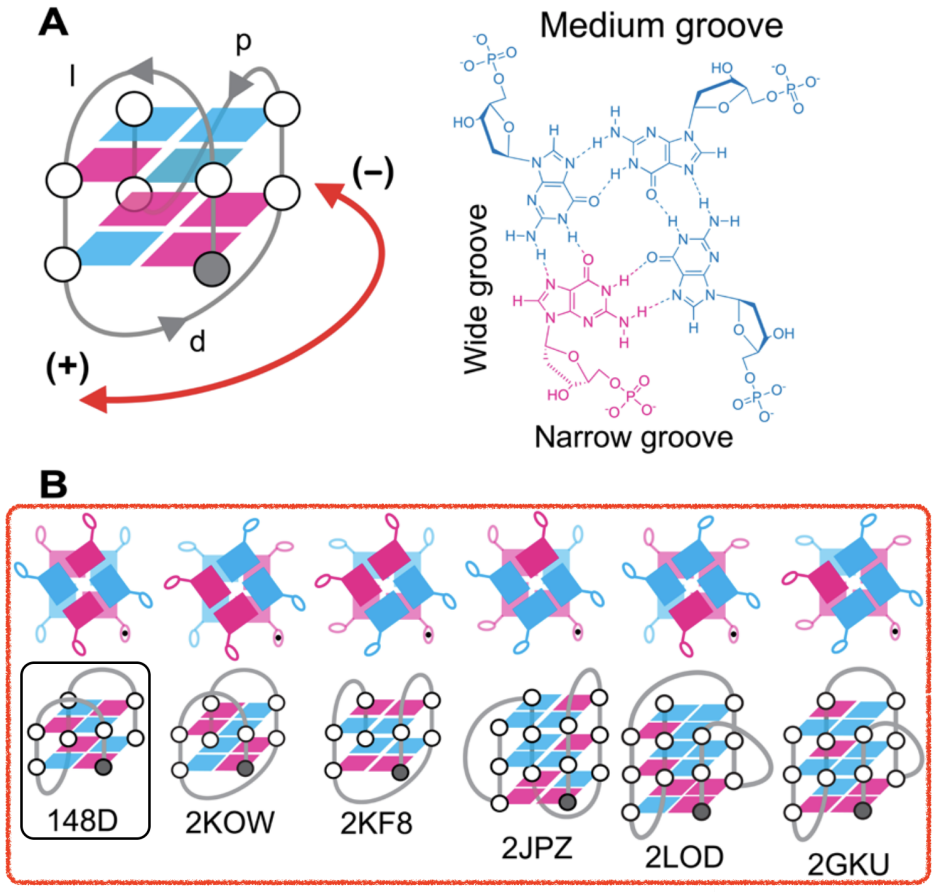


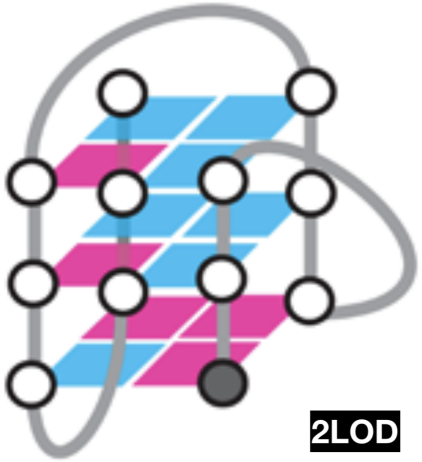
Fig. 1. Structural descriptors of canonical quadruplexes.

Webba da Silva (2018)
"Encoding canonical DNA quadruplex structure",
Sci. Adv.

e.g., 148D (chair-type)

2(+Ln+Lw+Ln)

Two/six example descriptors were assigned incorrectly



3(–P–Ln–Lw) ❌
3(-P-Lw-Ln) ✅ (DSSR)

3(–PD+Ln) ❌
3(-PD+Lw) ✅ (DSSR)

> "2GKU: you are correct- our mistake. There can not be a –Ln after the –P."
>
> — Dr. Webba da Silva

G4DB: A curated list of G-quadruplexes in the PDB

dynamic table, flexible search: 64 hits (out of 415) with term aptamer

Show All entries

aptamer

Showing 1 to 64 of 64 entries (filtered from 415 total entries)

Previous 1 Next

PDB ID	CLASS	METHOD	AUTHORS	REFERENCE	ANNOTATION
7ntu	hydrolase	X-ray (3.1 Å)	Troisi R, Balasco N, Santamaria A, Vitagliano L, Sica F	(2021) "Structural and functional analysis of the simultaneous binding of two duplex/quadruplex aptamers to human alpha-thrombin." *Int.J.Biol.Macromol.*, **181**, 858-867. doi: 10.1016/j.ijbiomac.2021.04.076.	X-ray structure of the complex between human alpha thrombin and two duplex-quadruplex aptamers: nu172 and hd22_27mer. ➥ 6 G-tetrads, 2 G4 helices, 2 G4 stems, 2(+Ln+Lw+Ln), chair(2+2), UDUD
7oax	RNA	X-ray (2.24 Å)	Mieczkowski M, Steinmetzger C, Bessi I, Lenz AK, Schmiedel A, Holzapfel M, Lambert C, Pena V, Hobartner C	(2021) "Large Stokes shift fluorescence activation in an RNA aptamer by intermolecular proton transfer to guanine." *Nat Commun*, **12**, 3549. doi: 10.1038/s41467-021-23932-0.	Crystal structure of the chili RNA aptamer in complex with dmhbo+. ➥ 8 G-tetrads, 4 G4 helices

G4DB: A curated list of G-quadruplexes in the PDB

22 hits left with combined search terms aptamer chair

Show All entries

aptamer chair

Showing 1 to 22 of 22 entries (filtered from 415 total entries)

Previous 1 Next

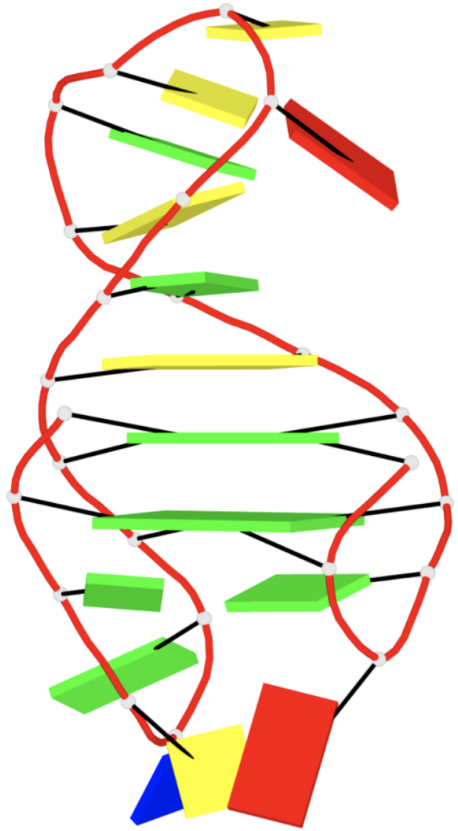
PDB ID	CLASS	METHOD	AUTHORS	REFERENCE	ANNOTATION
7ntu	hydrolase	X-ray (3.1 Å)	Troisi R, Balasco N, Santamaria A, Vitagliano L, Sica F	(2021) "Structural and functional analysis of the simultaneous binding of two duplex/quadruplex aptamers to human alpha-thrombin." *Int.J.Biol.Macromol.*, **181**, 858-867. doi: 10.1016/j.ijbiomac.2021.04.076.	X-ray structure of the complex between human alpha thrombin and two duplex-quadruplex aptamers: nu172 and hd22_27mer. ➥ 6 G-tetrads, 2 G4 helices, 2 G4 stems, 2(+Ln+Lw+Ln), chair(2+2), UDUD
6z8w	hydrolase	X-ray (1.73 Å)	Smirnov I, Kolganova N, Troisi R, Sica F, Timofeev E	(2021) "Expanding the recognition interface of the thrombin-binding aptamer HD1 through modification of residues T3 and T12." *Mol Ther Nucleic Acids*, **23**, 863-871. doi: 10.1016/j.omtn.2021.01.004.	X-ray structure of the complex between human alpha thrombin and a thrombin binding aptamer variant (tba-3g), which contains 1-beta-d-glucopyranosyl residue in the side chain of thy3 at n3.. ➥ 2 G-tetrads, 1 G4 helix, 1 G4 stem, 2(+Ln+Lw+Ln), chair(2+2), UDUD

Comprehensive annotations of G4 (DSSR-G4DB)

Stem#1, 2 G-tetrad layers, 3 loops, INTRA-molecular, UDUD, anti-parallel, 2(+Ln+Lw+Ln), chair(2+2)

Webba da Silva nomenclature

common name



```
1  glyco-bond=s-s- sugar=---- groove=wnwn Major-->WC nts=4 GGGG A.DG2,A.DG25,A.DG18,A.DG7
2  glyco-bond=-s-s sugar=3--- groove=wnwn WC-->Major nts=4 GGGG A.DG3,A.DG24,A.DG19,A.DG6
 step#1  mm(<>,outward)  area=12.77 rise=3.77 twist=20.4
 strand#1  U DNA glyco-bond=s- sugar=-3 nts=2 GG A.DG2,A.DG3
 strand#2  D DNA glyco-bond=-s sugar=-- nts=2 GG A.DG25,A.DG24
 strand#3  U DNA glyco-bond=s- sugar=-- nts=2 GG A.DG18,A.DG19
 strand#4  D DNA glyco-bond=-s sugar=-- nts=2 GG A.DG7,A.DG6
loop#1 type=lateral   strands=[#1,#4] nts=2 AG A.DA4,A.DG5
loop#2 type=lateral   strands=[#4,#3] nts=10 CGCGCCAGCG A.DC8,A.DG9,A.DC10,A.DG11,A.DC12,
loop#3 type=lateral   strands=[#3,#2] nts=4 GTCG A.DG20,A.DT21,A.DC22,A.DG23
```

rigid-body parameters

Download PDB file

Interactive view in 3Dmol.js

loop type and identity

PDB id: 7CV4 (in chair shape)

Identification and analysis

Block-view schematics

Advanced model building

Integration into other resources

Features tailored to G-quadruplexes

The details are **ESSENTIAL**

- Strict ANSI C (80,000+ lines of code)

```
-ansi -Wextra ... -Wunused -Wshadow -Werror
```

- Valgrind `--leak-check=full`

```
HEAP SUMMARY:
    in use at exit: 0 bytes in 0 blocks
  total heap usage: 51,432 allocs, 51,432 frees, 672,550,630 bytes allocated

All heap blocks were freed -- no leaks are possible
```

- Tested using all RNA/DNA-containing structures in the PDB

DSSR is an integrated software tool
with unmatched capabilities
in RNA/DNA structural bioinformatics.

Three **web** resources on x3dna.org

http://web.x3dna.org – Updated web-interface to 3DNA

http://skmatic.x3dna.org – Schematics, JSON + human-readble output

http://G4.x3dna.org – Annotated G-quadruplexes in the PDB

The command-line program: x3dna-dssr

- Tiny size (<2MB), no dependencies
- No setup, or configuration needed

`x3dna-dssr -h` to get started right away

Professional user manual (236 pages)

Harmen Bussemaker and Wilma Olson

Brady Butterfield and **Beth Kauderer**

The 3DNA/DSSR user community

Thank YOU!