INTEGRATIVE STRUCTURAL BIOLOGY OF HIGHER-ORDER DNA G-QUADRUPLEXES

Robert C. Monsen, PhD
Trent Lab, University of Louisville, KY

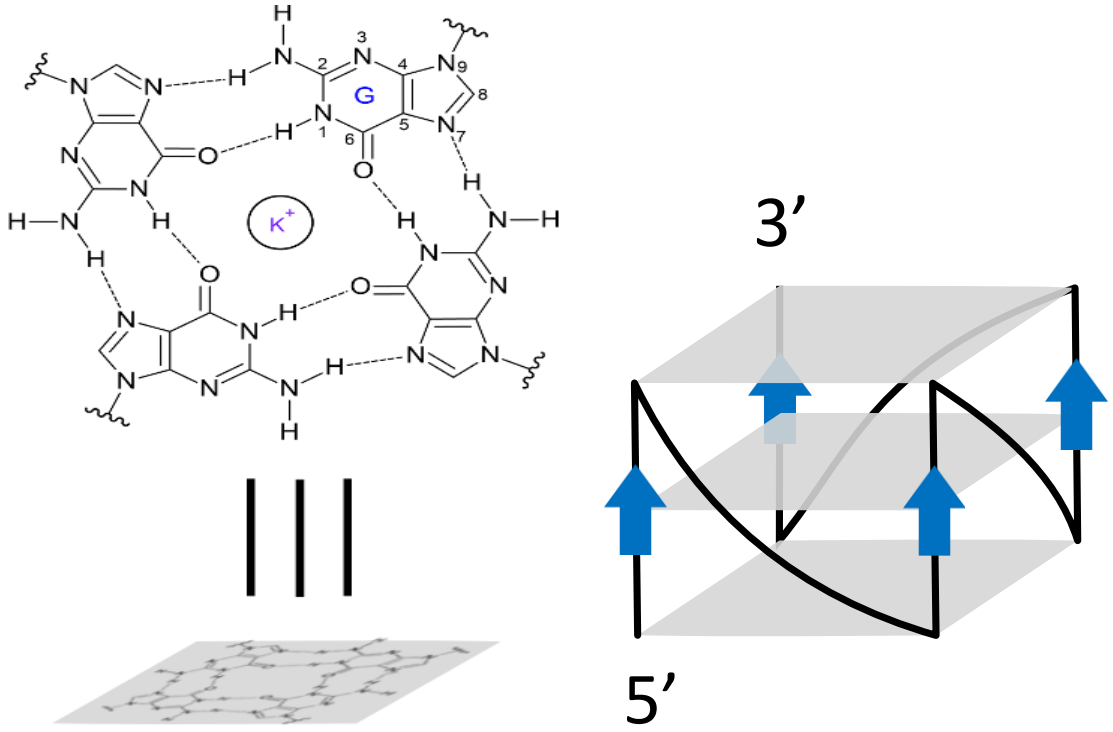
HOS2022

WHAT ARE G-QUADRUPLEXES?

Tetraplex DNA structures composed of square planar guanine tetrads

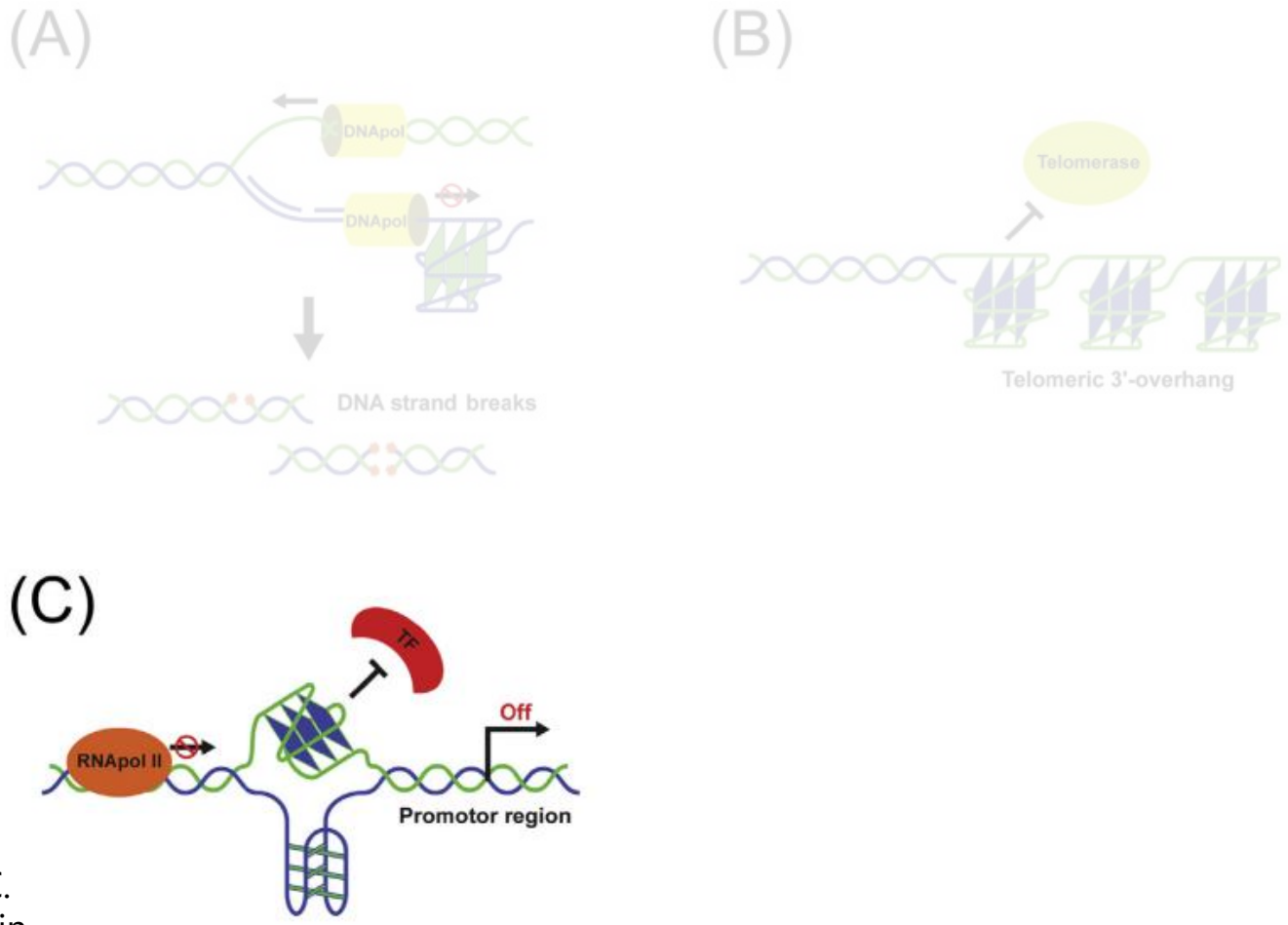
Putative G4 motif:

G_x-L_y-G_x-L_y-G_x-L_y-G_x



WHERE DO DNA G-QUADRUPLEXES FORM?

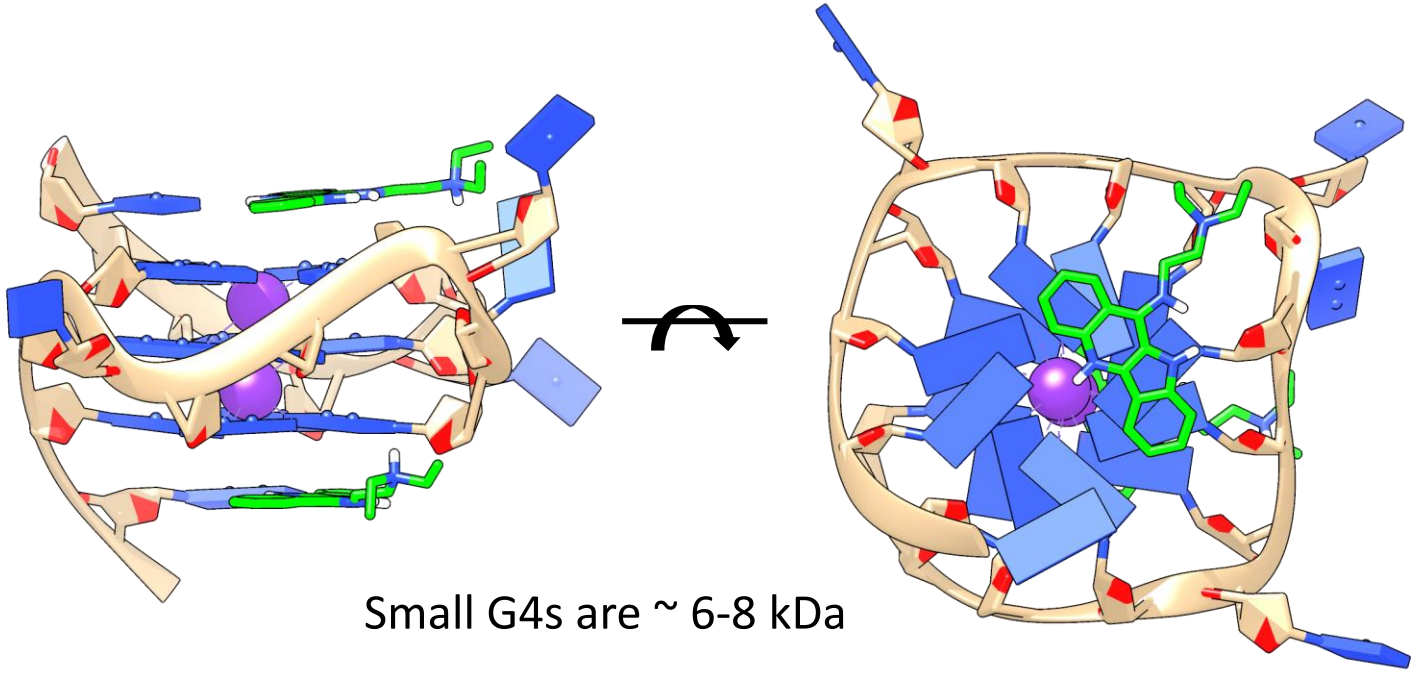
A) Arrest of DNA polymerase
B) Telomere elongation interference
C) Promoter activity control



Carvalho, J., Mergny, J.L., Salgado, G.F., Queiroz, J.A. and Cruz, C. (2020) G-quadruplex, Friend or Foe: The Role of the G-quartet in Anticancer Strategies. *Trends Mol Med*.

G-QUADRUPLEXES AS DRUG TARGETS: SMALL, SIMPLE "IDEALIZED" G-QUADRUPLEXES ARE POOR TARGETS

1. Only simple quadruplex sequences are typically amenable to NMR/XRD
2. Seldomly have 'druggable' pockets
3. G-tetrad faces enrich for non-specific scaffolds



Small G4s are ~ 6-8 kDa

Receptors with common features

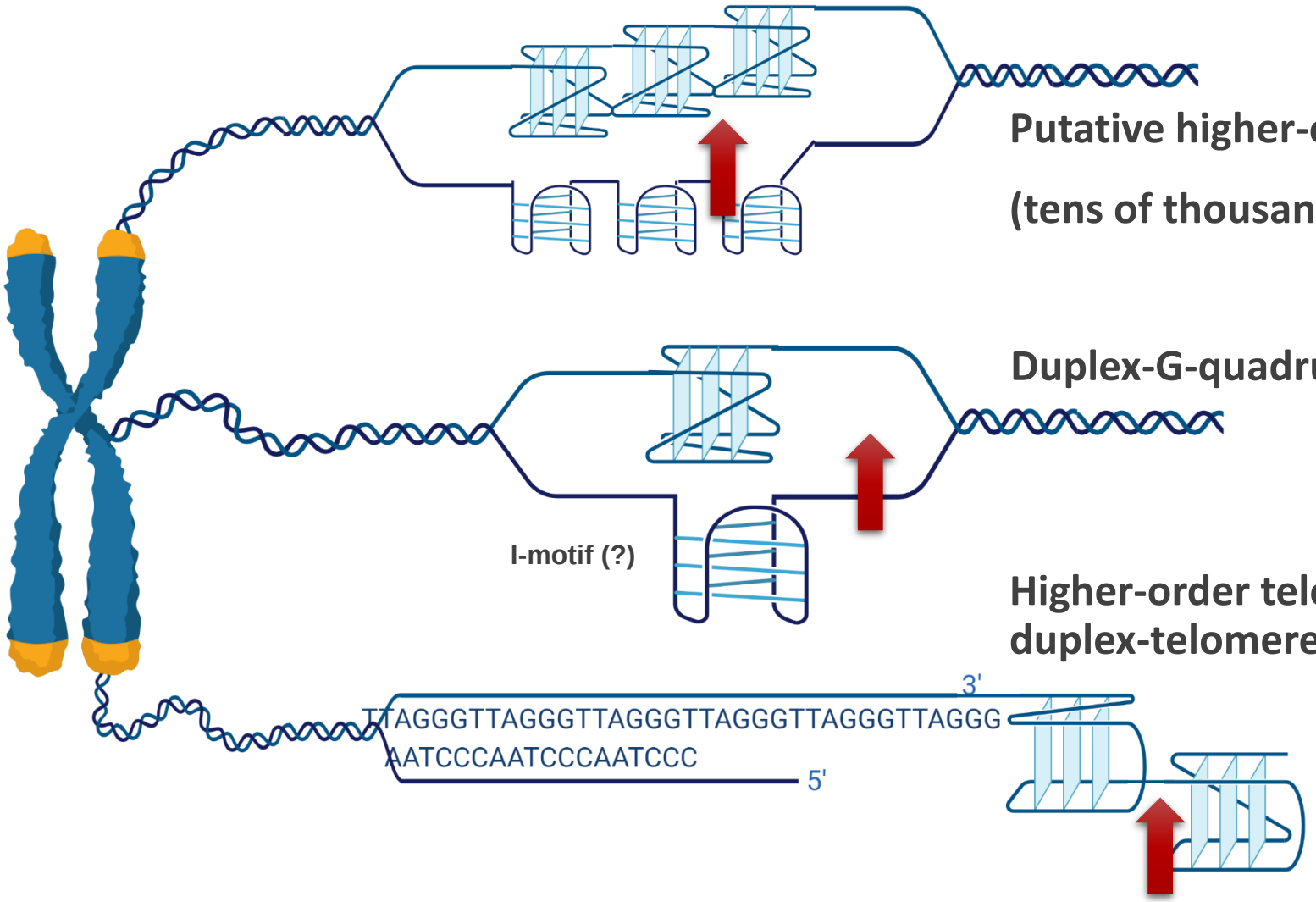
↓

Non-selective drugs

(PDB 2L7V, sequence from modified c-Myc promoter: TGA**GGG**T**GGG**TA**GGG**T**GGG**TAA)

Dai J, Carver M, Hurley LH, Yang D. Solution structure of a 2:1 quindoline-c-MYC G-quadruplex: insights into G-quadruplex-interactive small molecule drug design. J Am Chem Soc. 2011 Nov 9;133(44)

HIGHER-ORDER G-QUADRUPLEX SYSTEMS COULD OFFER RICHER DRUG TARGETING LANDSCAPE



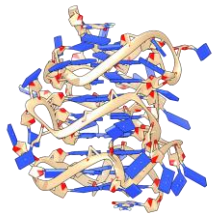
Putative higher-order promoter G-quadruplexes

(tens of thousands identified!)

Duplex-G-quadruplex junction

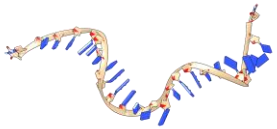
Higher-order telomere or duplex-telomere junctions

INTEGRATIVE STRUCTURAL BIOLOGY (ISB) APPROACH

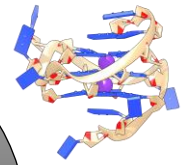


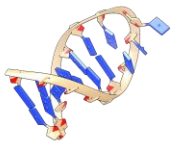
Optimized ISB model for drug targeting

(1) Identify putative G-quadruplex sequence



(2) Secondary Structure analysis

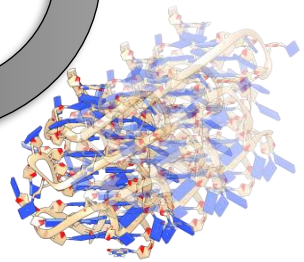




CD, 1D NMR, DNase

(3) Modeling and Simulation

Refine



(4) Tertiary Structure analysis

SAXS, AUC

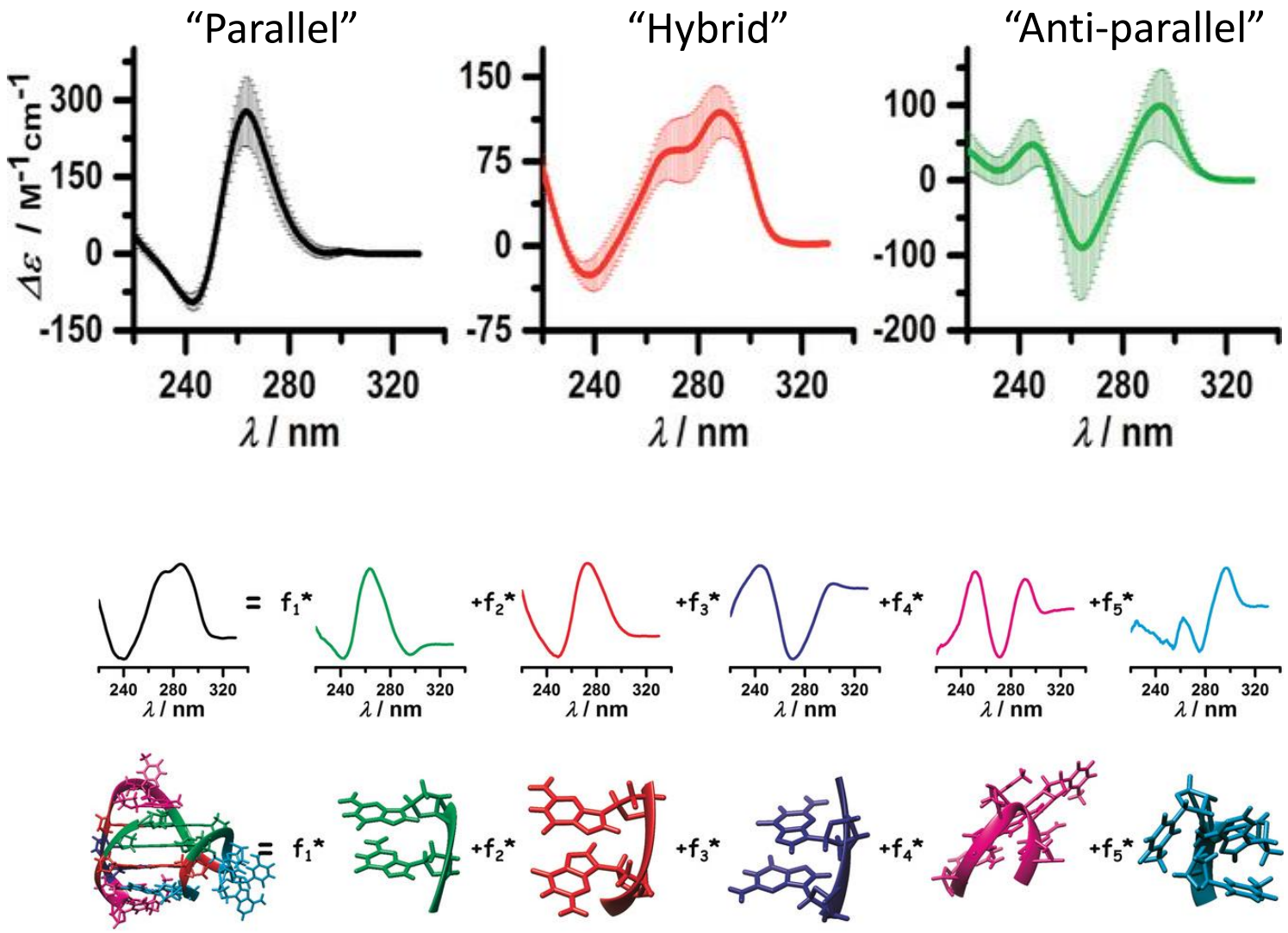
(5) Model validation or reformulate hypothesis

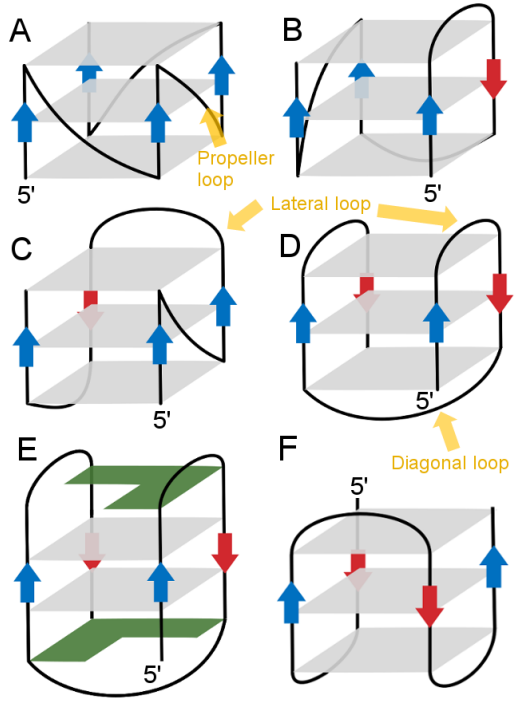
Compare to experimental:

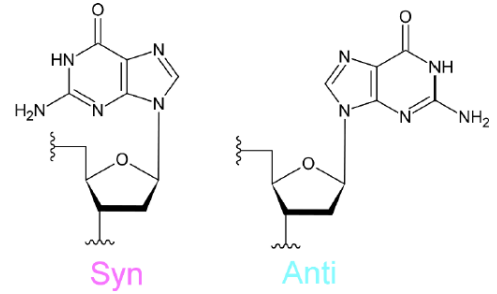
$S_{exp.} = S_{model}$

$\chi^2_{SAXS} = \sim 1.0$

CIRCULAR DICHROISM FOR QUANTITATION OF G4 2° STRUCTURE



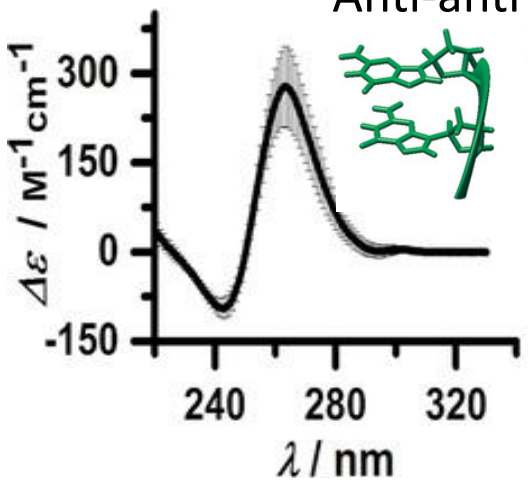


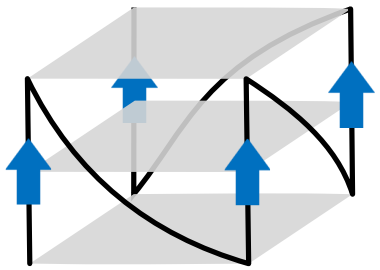


Del Villar-Guerra R, Trent JO, Chaires JB. G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. Angew Chem Int Ed Engl. 2018 Jun 11;57(24):7171-7175.

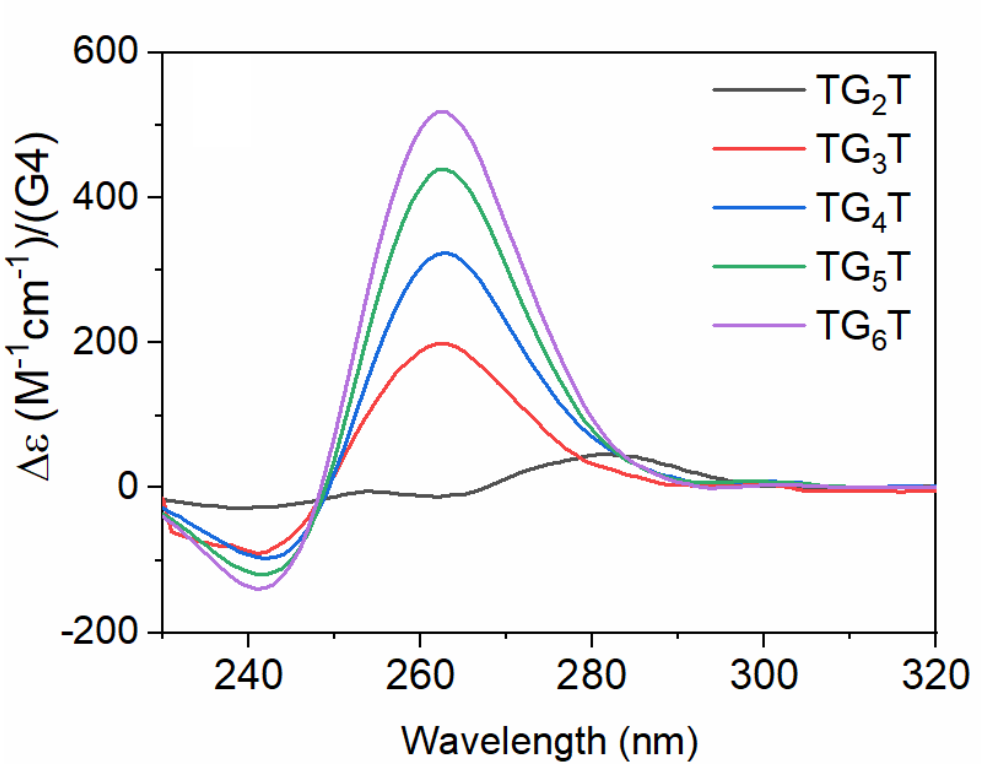
CD 264-NM CORRELATES WITH # G-TETRAD STACKS IN PARALLEL QUADRUPLEXES

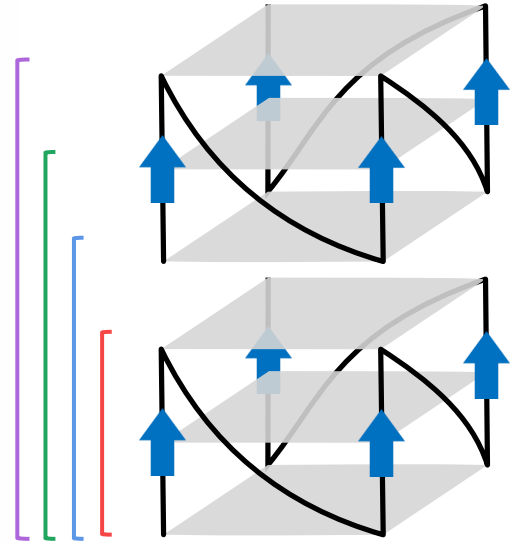
Anti-anti



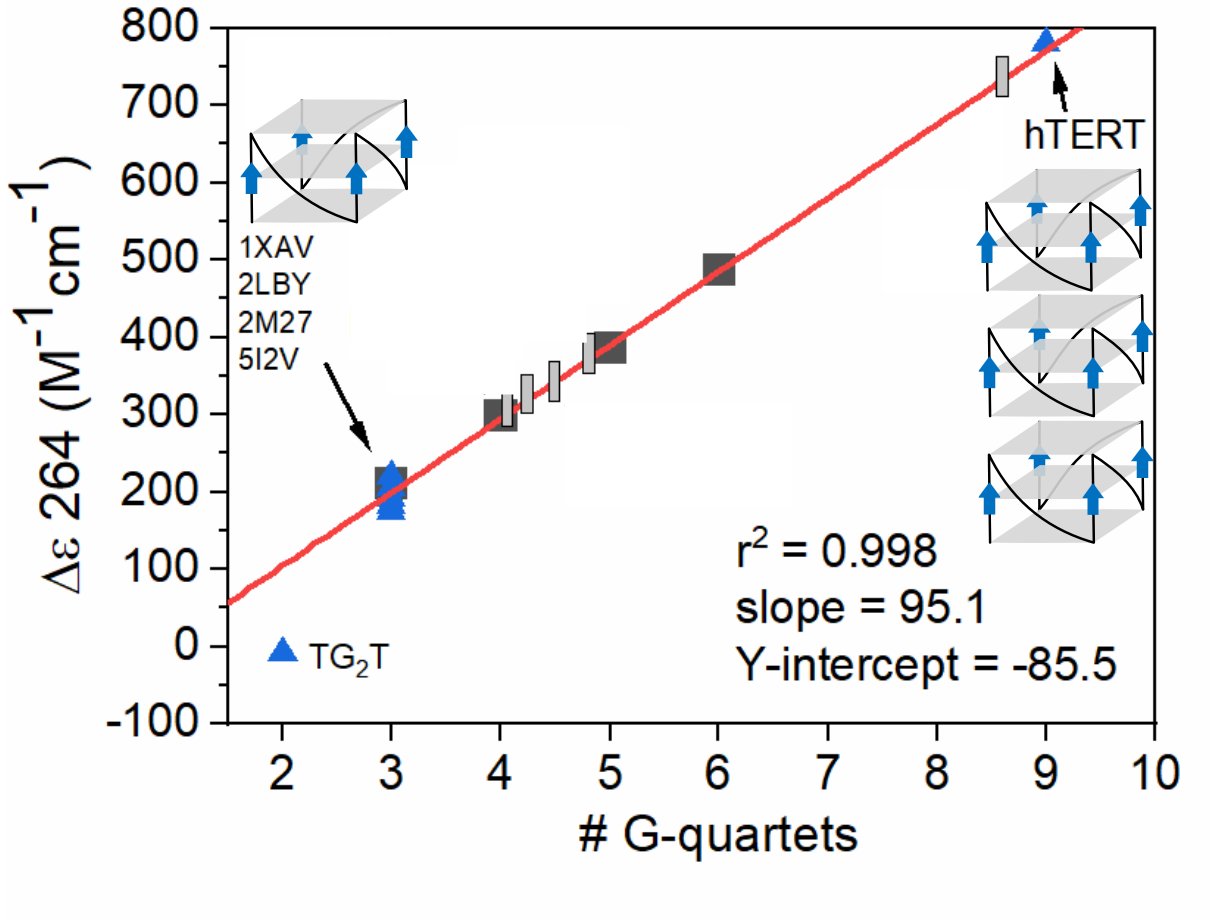


Parallel





KNOWN PARALLEL PROMOTER G-QUADRUPLEXES FALL ON REGRESSION LINE



Nucleic Acids Research, 2022 **1–21**
https://doi.org/10.1093/nar/gkac182

Long promoter sequences form higher-order G-quadruplexes: an integrative structural biology study of *c-Myc*, *k-Ras* and *c-Kit* promoter sequences

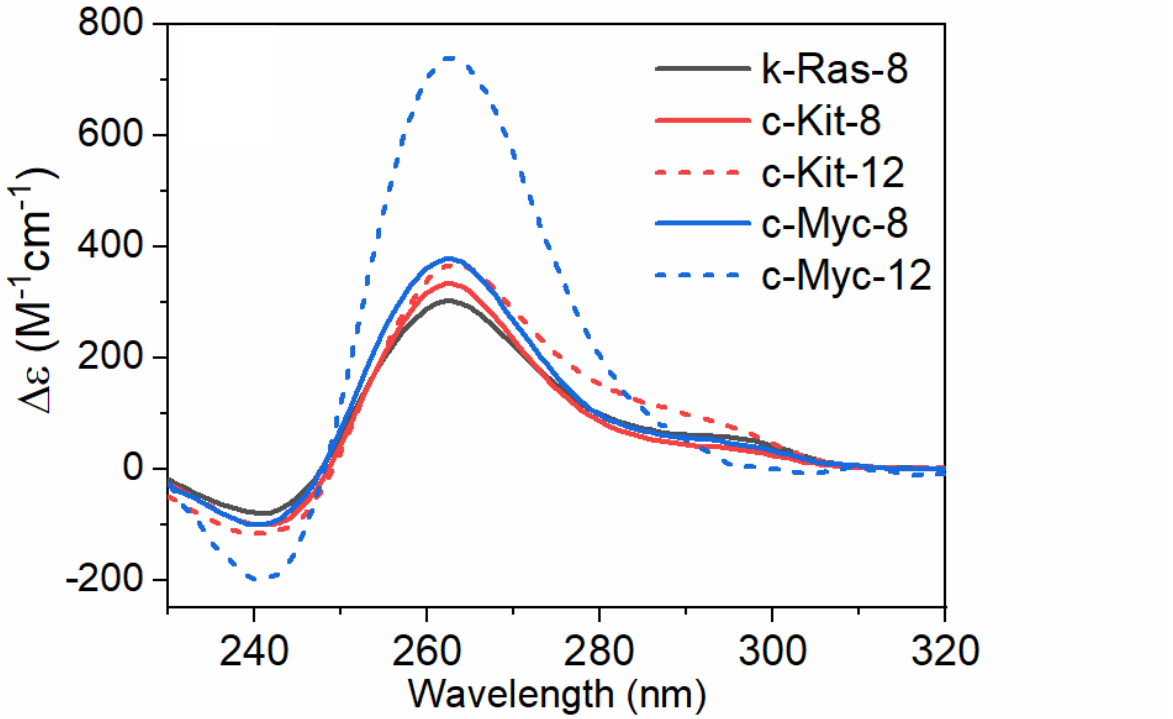
Robert C. Monsen [1], Lynn W. DeLeeuw[1], William L. Dean[1], Robert D. Gray[1], Srinivas Chakravarthy[2], Jesse B. Hopkins [2], Jonathan B. Chaires [1,3,4,*] and John O. Trent [1,3,4,*]

[1]UofL Health Brown Cancer Center, University of Louisville, Louisville, KY 40202, USA, [2]The Biophysics Collaborative Access Team (BioCAT), Department of Biological, Chemical, and Physical Sciences, Illinois Institute of Technology, Chicago, IL 60616, USA, [3]Department of Medicine, University of Louisville, Louisville, KY 40202, USA and [4]Department of Biochemistry and Molecular Genetics, University of Louisville, Louisville, KY 40202, USA

Received January 31, 2022; Revised March 03, 2022; Editorial Decision March 04, 2022; Accepted March 21, 2022

HIGHER-ORDER PROMOTER G-QUADRUPLEXES ARE ALL PARALLEL

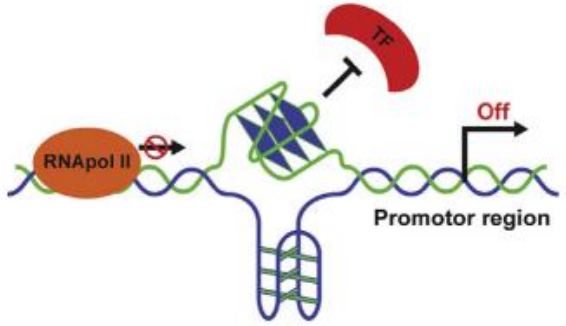
Gene name-(# of G runs)	Sequence	Length (nt)
c-Myc-8	GGG GAG GGT GGG GAG GGT GGG GAA GGT GGG GAG G	34
c-Myc-12	GGG AAC CCG GGA GGG GCG CTT ATG GGG AGG GTG GGG AGG GTG GGG AAG GTG GGG AGG AGA CTC AGC CGG G	70
c-Kit-8	GGG CGG GCG CGA GGG AGG GGA GGC GAG GGG CGT GG	35
c-Kit-12	GGG CGG GCG CGA GGG AGG GGA GGC GAG GGG CGT GGC CGG CGC GCA GAG GGA GGG CGC TGG G	64
k-Ras-8	GGG AGC GGC TGA GGG CGG TGT GGG AAG AGG GAA GAG GGG GAG G	43

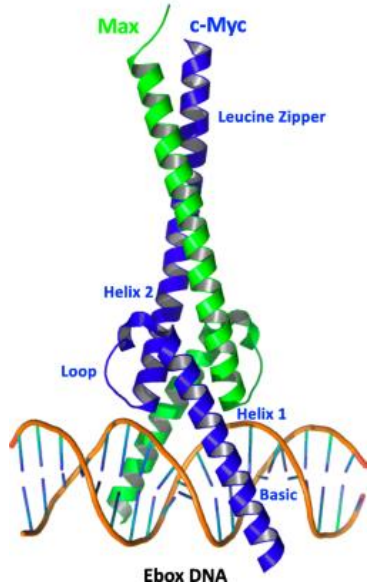


Monsen RC, DeLeeuw LW, Dean WL, Gray RD, Chakravarthy S, Hopkins JB, Chaires JB, Trent JO. Long promoter sequences form higher-order G-quadruplexes: an integrative structural biology study of c-Myc, k-Ras and c-Kit promoter sequences. Nucleic Acids Res. 2022.

THE C-MYC PROTEIN IS AN 'UNDRUGGABLE' CANCER TARGET

- Overexpressed in ~80% of solid tumors
- Regulates cell growth, differentiation, and metabolism
- c-Myc is designated as an "undruggable" transcription factor because of short half life and no persistent structure to target



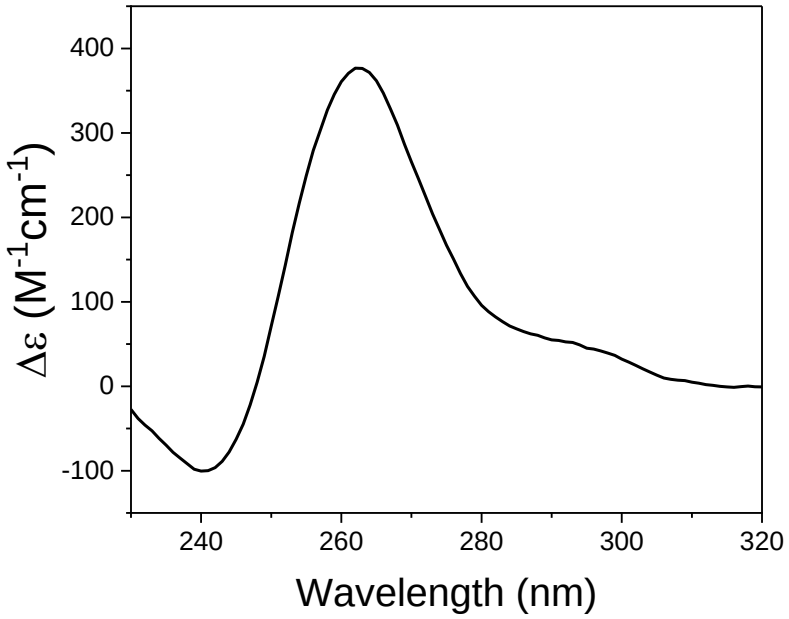


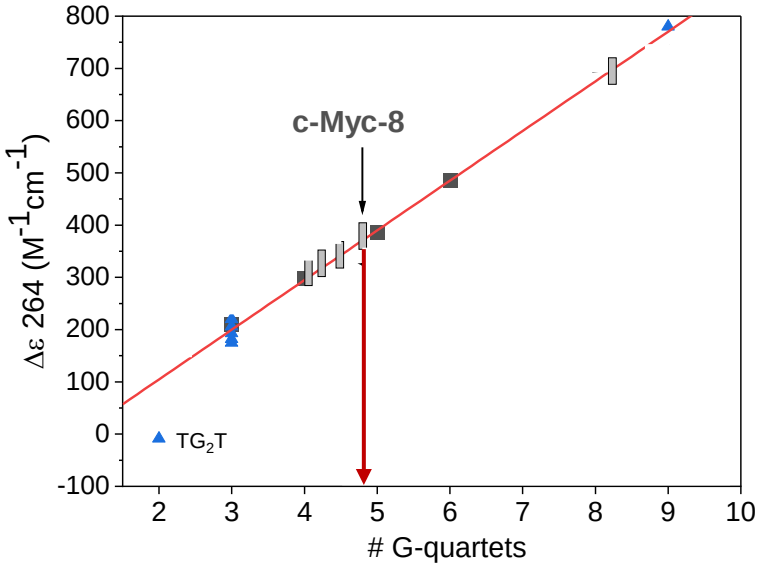
Madden, S.K., de Araujo, A.D., Gerhardt, M. et al. Taking the Myc out of cancer: toward therapeutic strategies to directly inhibit c-Myc. Mol Cancer 20, 3 (2021)

EXTENDED "C-MYC-8" SEQUENCE PREDICTED TO BE A 5-STACK PARALLEL QUADRUPLEX

GGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGAGG

1 2 3 4 5 6 7 8



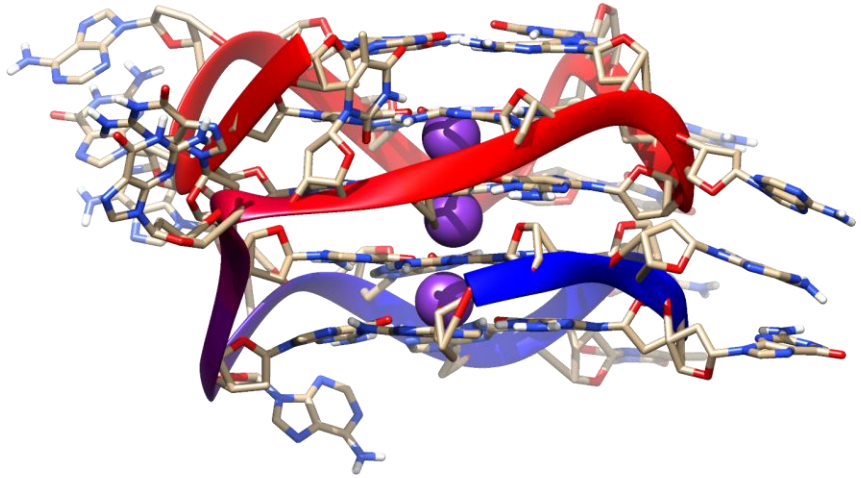


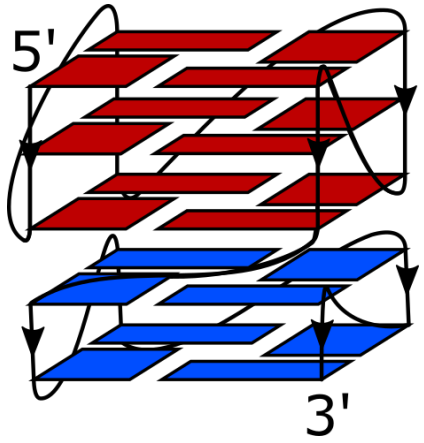
Annealed in K^+-phosphate buffer (185 mM KCl, pH 7.2)

5-TETRAD PARALLEL G4 MODEL IS CONSISTENT WITH MEASURED SEDIMENTATION COEFFICIENT

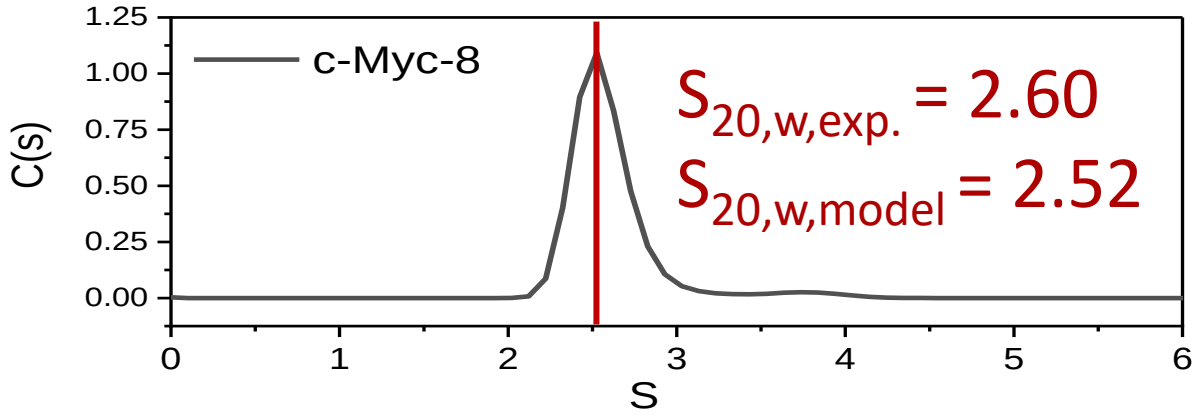
10 20 30

5'–GGGGAGGGTG GGGAGGGTGG GGAAGGTGGG GAGG





3-tetrad
+
2-tetrad

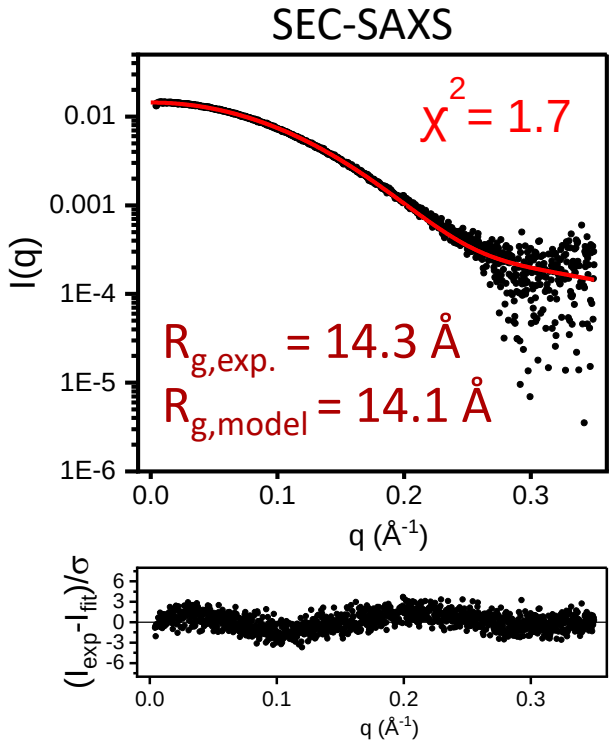


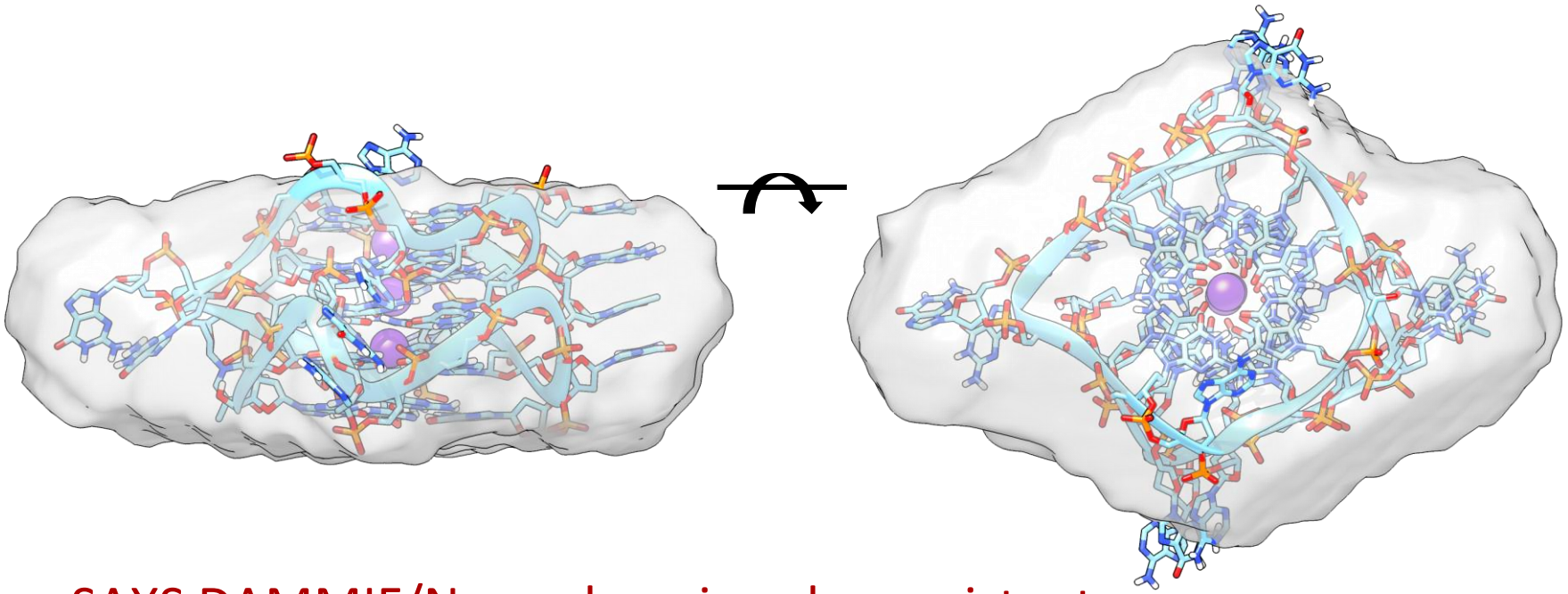
*vbar = 0.55 cm^3/g

*Hydrodynamic calculations with HYDROPRO10

Ortega, A., Amoros, D. and Garcia de la Torre, J. (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. Biophys J, 101, 892-898.

5-TETRAD MODEL HAS EXCELLENT AGREEMENT WITH SMALL ANGLE X-RAY SCATTERING DATA





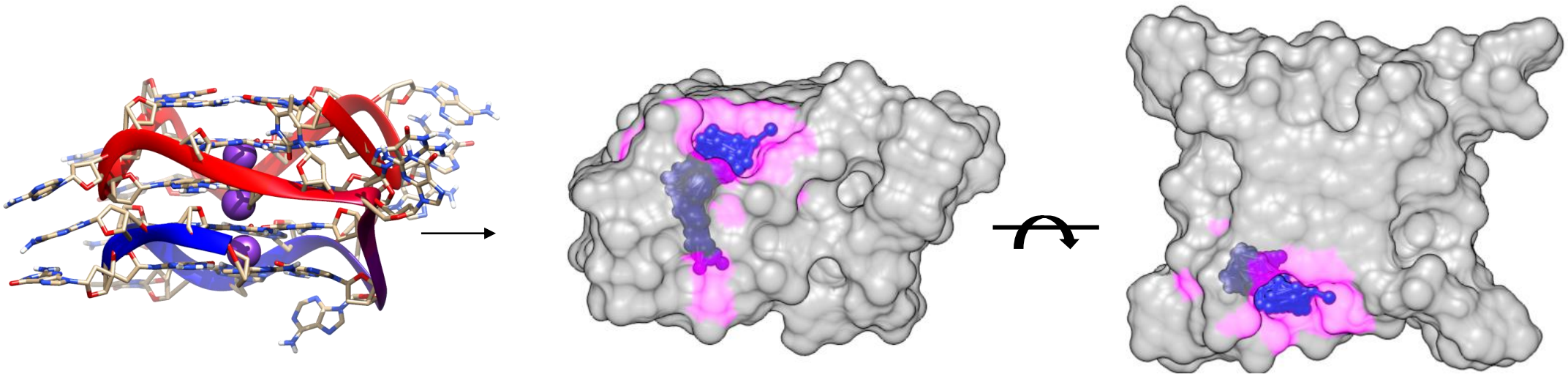
SAXS DAMMIF/N envelope is only consistent with a 3+2 stacked parallel G4 model!

CRYOSL reduced χ^2:

$$\chi^2(r_0, \delta\rho) = \frac{1}{N_p}\sum_{i=1}^{N_p}\left[\frac{I_e(s_i) - cI(s_i, r_0, \delta\rho)}{\sigma(s_i)}\right]^2,$$

Svergun D.I., Barberato C. & Koch M.H.J. (1995) J. Appl. Cryst. 28, 768-773.

C-MYC-8: STACKED G4 MODEL HAS "PROTEIN-LIKE" LOOP POCKETS

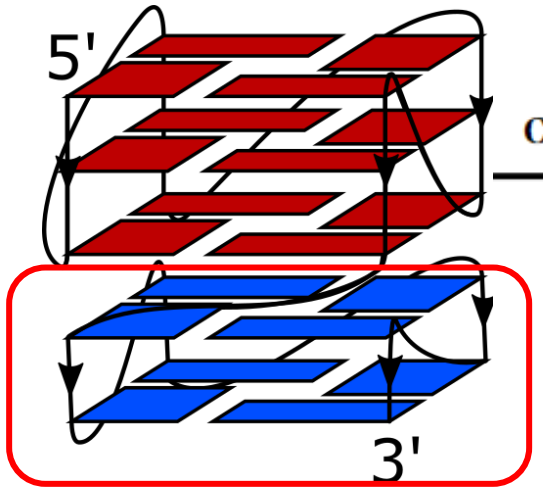


SiteMap scoring:
0.63 (undruggable)
0.87 (difficult)
1.10 (druggable)

SiteMap score = 0.89 (i.e. druggable pocket)
(compared to 0.55 to 0.67 for smaller systems)

Halgren, T. (2007), New Method for Fast and Accurate Binding-site Identification and Analysis. Chemical Biology & Drug Design, 69: 146-148.

THE C-MYC-8 HOS IS UNPRECEDENTED



Chemistry–A European Journal
Full Paper
doi.org/10.1002/chem.202100895
Chemistry Europe
European Chemical Societies Publishing

G-Quadruplex Formation by DNA Sequences Deficient in Guanines: Two Tetrad Parallel Quadruplexes Do Not Fold Intramolecularly

Iva Kejnovská[+],[a] Petr Stadlbauer[+],[a] Lukáš Trantírek,[b] Daniel Renčiuk,[a] Martin Gajarský,[b] Daniel Krafčík,[b] Jan Palacký,[a] Klára Bednářová,[a] Jiří Šponer,[a] Jean-Louis Mergny,[a] and Michaela Vorlíčková*[a]

June, 2021

ACKNOWLEDGEMENTS

The Trent Lab group

- John Trent, Ph.D. (PI)
- Lynn DeLeeuw, M.S.
- Jon Maguire, B.S.

The Chaires Lab group

- J. Brad Chaires, Ph.D.
- Bob Gray, Ph.D.
- Bill Dean, Ph.D.

BioCAT Beamline

- Srinivas Chakravarthy, Ph.D.
- Jesse Hopkins, Ph.D.

Argonne National Laboratory

National Center for CryoEM Access and Training

UNIVERSITY OF LOUISVILLE

Funding source: National Institutes of Health (NIH) [GM077422]