



HAL
open science

**Recognition of different base tetrads by RHAU
(DHX36): X-ray crystal structure of the G4 recognition
motif bound to the 3'-end tetrad of a DNA
G-quadruplex**

Brahim Heddi, Vee Vee Cheong, Emmanuelle Schmitt, Yves Mechulam, Anh
Tuân Phan

► **To cite this version:**

Brahim Heddi, Vee Vee Cheong, Emmanuelle Schmitt, Yves Mechulam, Anh Tuân Phan. Recognition of different base tetrads by RHAU (DHX36): X-ray crystal structure of the G4 recognition motif bound to the 3'-end tetrad of a DNA G-quadruplex. *Journal of Structural Biology*, 2020, 209 (1), pp.S1047-8477(19)30210-2. 10.1016/j.jsb.2019.10.001 . hal-02355168

HAL Id: hal-02355168

<https://polytechnique.hal.science/hal-02355168v1>

Submitted on 2 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2
3
4 **Recognition of different base tetrads by RHAU: X-ray crystal structure of the**
5 **G4 recognition motif bound to the 3'-end tetrad of a DNA G-quadruplex**
6
7
8
9

10
11 Brahim Heddi^{1,2*}, Vee Vee Cheong¹, Emmanuelle Schmitt³,
12 Yves Mechulam³ and Anh Tuân Phan^{1*}
13
14
15

16
17 ¹School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang
18 Link, Singapore 637371, Singapore
19

20
21 ²Laboratoire de Biologie et de Pharmacologie Appliquée, CNRS UMR 8113, Ecole Normale
22 Supérieure Paris-Saclay, 61 Avenue du président Wilson, Cachan 94235, France
23

24
25 ³Laboratoire de Biochimie, UMR 7654, Ecole Polytechnique, Centre National de la Recherche
26 Scientifique, Palaiseau 91128, France
27
28
29
30
31
32
33
34
35

36 *Corresponding author: phantuan@ntu.edu.sg and brahim.heddi@ens-cachan.fr
37
38
39
40
41
42
43

44 Running title: Protein recognition of different base tetrads
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **ABSTRACT**
5

6 G-quadruplexes (G4) are secondary structures of nucleic acids that can form in cells and have
7
8 diverse biological functions. Several biologically important proteins interact with G-
9 quadruplexes, of which RHAU – a helicase from the DEAH-box superfamily, was shown to bind
10
11 and unwind G-quadruplexes efficiently. We report a high-resolution (1.5 Å) X-ray co-crystal
12
13 structure of an N-terminal fragment of RHAU bound to the exposed tetrad of a parallel-stranded
14
15 G-quadruplex. The RHAU peptide folds into an L-shaped α -helix, and binds to the G-
16
17 quadruplex through π -stacking and electrostatic interactions. X-ray crystal structure of our
18
19 complex identified key amino-acid residues important for G-quadruplex-peptide binding
20
21 interaction at the 3'-end G•G•G•G tetrad. Together with previous solution and crystal structures
22
23 of RHAU bound to the 5'-end G•G•G•G and A•T•G•G tetrads, our crystal structure highlights
24
25 the occurrence of a robust G-quadruplex recognition motif within RHAU that can adapt to
26
27 different accessible tetrads.
28
29
30
31
32
33
34
35
36

37 Keywords: G-quadruplex; X-ray crystallography; RHAU helicase; DHX36; G4R1; DEAH-box
38 family.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

INTRODUCTION

G-quadruplexes are four-stranded nucleic acid secondary structures formed by the stacking of G-tetrads (planar G•G•G•G structure formed by Hoogsteen base pairing) and stabilized by cations [1]. G-quadruplexes are polymorphic and can adopt diverse topologies [2, 3] as exemplified by parallel (all four strands in the same direction) and non-parallel (one or two strands in the opposite direction) G-quadruplexes.

G-quadruplexes form in cells[4-6], and are involved in key cellular processes including transcription, replication and translation, and implicated in genomic instability and aging[7-12]. Several biologically relevant proteins interact with G-quadruplexes[6, 13] by promoting G-quadruplex formation/stabilization, or through the deformation of G-quadruplexes[12]. Among them several biologically important helicases bind and unwind G-quadruplexes with high affinity and efficiency[14]. However, molecular details on how proteins interact with G-quadruplex remain very scarce [15].

RHAU (also called DHX36), a helicase of the DEAH-box superfamily which binds and unwinds RNA and DNA G-quadruplexes with high affinity and efficiency[16-19], is reported to exhibit a 100-fold binding preference for parallel G-quadruplexes as compared to non-parallel G-quadruplexes[20]. RHAU is composed of an N-terminal domain containing the RHAU-specific motif (RSM), a helicase core and a C-terminal domain, and is conserved across numerous organisms[17]. RHAU is involved in spermatogenesis[21-23], heart development during embryogenesis[24], hematopoiesis[25], and is shown to regulate several RNA[26, 27] and mRNA[24, 28, 29], likely through interaction with important G-quadruplex forming sequences present in the 5'- or 3'-UTR.

1
2
3
4
5
6
7 In 2015, we reported a high-resolution NMR solution structure of the N-terminal RSM-
8 containing 18-aa fragment of RHAU (*Rhau18*) in complex with a parallel-stranded G-
9 quadruplex[20]. The solution structure identified *Rhau18* to form an α -helix G-quadruplex
10 recognition motif, with the peptide bound to the 5'-end terminal parallel G-quadruplex via
11 hydrophobic/stacking and electrostatic interactions. As this peptide-quadruplex binding mode
12 resembles the binding mode of most G-quadruplex ligands[30], our NMR solution structure
13 suggested the preference of RHAU for parallel-stranded G-quadruplex to be mainly due to their
14 G-tetrad accessibility and steric hindrance of the protein[20].
15
16
17
18
19
20
21
22
23
24
25
26
27

28 The recognition mode of RHAU for G-quadruplex identified in *Rhau18* was similarly observed
29 in a recent X-ray crystal structure of a truncated bovine RHAU co-crystallized with a parallel G-
30 quadruplex[31]. The structure identified a bovine RHAU α -helix fragment to utilize
31 hydrophobic/stacking and electrostatic interactions to bind to a 5'-end A•T•G•G tetrad (Figure
32 S1A). Structural comparison between the bovine RHAU-quadruplex complex and human
33 RHAU-quadruplex complex finds key residues participating in the binding interaction between
34 peptide and DNA to be highly conserved across the two species, with very similar N-terminal
35 amino acids identified to interact with an accessible 5'-end tetrad (Figure S1B). This finding
36 reinforced our hypothesis of RHAU preferentiality to parallel G-quadruplex to be dependent on
37 the availability of accessible tetrads[20]. Binding affinity between the full-length RHAU and a
38 G-quadruplex can be further enhanced by electrostatic interactions between the negatively-
39 charged backbone of the G-quadruplex and the positively-charged cavity of RHAU[31, 32].
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **Rhau18** was also shown to bind to a 3'-end G-tetrad when the 5'-end tetrad binding interface is
5 occupied by another **Rhau18** peptide [20]. It was recently shown that RHAU was pulling one
6 residue at a time with 3' to 5' directionality to destabilize and resolve the G-quadruplex
7 structure[31]. With the 3'-end G-tetrad possessing opposite polarity to a 5'-end tetrad, the
8 binding of RHAU to a 3'-end G-tetrad may imply a different unfolding mechanism to be
9 involved. Through structural comparison of RHAU-quadruplex complexes with different
10 peptide-quadruplex binding interface, better understanding of how RHAU adapts to different
11 tetrad polarity could be achieved, and possibly provide clues on its unwinding mechanism.
12
13
14
15
16
17
18
19
20
21
22
23
24
25

26 In this study, we report the co-crystallization of a dimeric parallel-stranded DNA G-quadruplex
27 with exposed 3'-end G-tetrad in complex with a 29-amino acid peptide extracted from the N-
28 terminal human helicase RHAU. Our X-ray crystal structure shows the DNA to assemble as a
29 dimeric parallel-stranded G-quadruplex, with the peptide interacting at the 3'-end terminal G-
30 tetrad. The peptide conformation in the co-crystal structure highly resembles the N-terminal
31 domain of RHAU in earlier reported structures[20, 31], adopting a predominantly α -helical
32 structure. Analysis of the X-ray co-crystal structure finds the peptide to interact with the G-
33 quadruplex through extensive CH/ π -CH₃/ π -stacking and electrostatic interactions using
34 conserved positively charged residues. Our structure suggests the N-terminal binding domain of
35 RHAU to possess a robust motif capable of binding different G-tetrad polarity.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52 **RESULTS AND DISCUSSION**

53 **Structure determination of *Rhau29-T95* co-crystal complex**

54 Crystallization trials were conducted on a 29-aa RSM-containing N-terminal RHAU fragment
55 named **Rhau29** (N_{term}-HPGHLKGREIGMWYAKKQGQKNKEAERQE-C_{term}) and the G-
56
57
58
59
60
61
62
63
64
65

1
2
3
4 quadruplex forming sequence d(GGGT)₄ named **T95**. The **Rhau29-T95** complex was
5
6 successfully crystallized (see Methods), with crystals of P2₁ space group diffracted X-rays to 1.5
7
8 Å resolution. The structure was solved using molecular replacement with a parallel-stranded G-
9
10 quadruplex (coordinates from the PDB code 2LK7 without terminal thymine residues) [33] as a
11
12 starting model and refined using REFMAC [34] to 1.5 Å resolution (Table 1 and Figure 1).
13
14
15
16
17

18
19 Figure 1A illustrates a representative portion of the final electron density of the **Rhau29-T95**
20
21 complex, showing two **Rhau29** peptides bound to a dimeric **T95** G-quadruplex on the exposed
22
23 3'-end tetrads. The asymmetric unit of the X-ray crystal structure is composed of two **Rhau29-**
24
25 **T95** complexes (Figure S2) which are nearly identical (all-atom RMSD = 0.78 Å, excluding the
26
27 terminal thymine residues, Figure S3A). Analysis of the asymmetrical unit finds the four
28
29 individual peptide-quadruplex subunits (each consisting of one DNA molecule and one **Rhau29**
30
31 molecule) to be very similar (pairwise RMSD = 0.74±0.1 Å, Figure S3B), with notable structural
32
33 differences found only in the terminal thymine residues as well as side chain of lysine residues.
34
35
36
37 Five K⁺ ions are found between the G-tetrads, within the subunits and at the stacking interface
38
39 between the two subunits (Figure 1A). A total of 262 water molecules are distributed along the
40
41 medium grooves and around the phosphate groups. In our electron density map, 48 amino acids
42
43 corresponding to the 12 C-terminal residues in each of the four **Rhau29** molecules in the
44
45 asymmetric unit could not be seen. The final refinement of the complex led to an R_{free} equal to
46
47 0.2448 (R_{work} = 0.1853), in which a long mobile segment within the four **Rhau29** molecules
48
49 likely contributed to the relatively high R_{free} value.
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Two *Rhau29* peptides encase a dimeric G-quadruplex

In the *Rhau29-T95* complex, the dimeric G-quadruplex in our X-ray crystal structure consists of two stacked parallel-stranded G-quadruplex subunits. Within each G-quadruplex subunit, the G-tetrads adopt the same hydrogen-bond directionality (G1•G5•G9•G13, G2•G6•G10•G14 and G3•G7•G11•G15), and the two subunits are stacked via their 5'-5'-end interface (Figure 1B) with partial overlap between the 5/6-membered ring of the guanine bases of each subunit (*i.e.* G1/G9*, G5/G5*, G9/G1*, G13/G13* – asterisks denote the second subunit). Such stacking mode has been observed in NMR-derived structures of similar G-quadruplex-forming sequences which also adopts a dimeric parallel-stranded G-quadruplex conformation (Figure S4)[35]. The DNA is encased between the two peptide molecules (Figure 1). Our structure shows the two *Rhau29* molecules to bind onto the 3'-end interfaces of the dimeric G-quadruplex, with each peptide covering the exposed G-tetrad of the dimeric G-quadruplex. Analysis of free dimeric parallel-stranded G-quadruplex solution structure and the *Rhau29-T95* complex X-ray crystal structure finds the 5'-5' stacking interface between the G-quadruplex subunits to remain unperturbed by the binding of *Rhau29* (Figure S4).

Rhau29 peptides folds into an L-shaped α -helix

The peptide presents an L-shaped structure with an α -helix spanning residues 7 to 15, with Leu5, Ile10, Trp13 and Tyr14 forming the hydrophobic core (Figure 2, Figure S5). The missing 12 residues at C-terminal of each *Rhau29* molecule could indicate that these residues are structured only in the presence of the full-length protein when bound to a G-quadruplex, with mutagenesis data of bovine RHAU suggesting a partial ligand-induced protein folding for the N-terminal RHAU fragment[31].

1
2
3
4 Previously, we have solved the solution structure of an 18-aa peptide *Rhau18* (N_{term}-
5 SMHPGHLKGREIGMWYAKKQ-C_{term}) bound to the 5'-end G-tetrad of a parallel-stranded G-
6 quadruplex[20], and recently the X-ray crystal structure of the bovine RHAU containing the
7 RHAU-specific motif *RhauB* (N_{term}-PGHLKGREIGLWYAKKQ-C_{term}) bound to a 5'-end
8 A•T•G•G tetrad of a parallel G-quadruplex intermediate was reported[31]. The *Rhau18* contains
9 an α -helical L-shaped motif found to be critical for G-quadruplex binding, which is preserved in
10 the C-terminal extended peptide *Rhau29* and similarly retained in bovine *RhauB*. Structure
11 comparison of *Rhau29*, *RhauB* and *Rhau18* shows the three peptide cores to be highly similar
12 (pairwise r.m.s.d. = 0.8±0.1 Å using backbone atoms from residue 3 to 16) (Figure S5).
13 However, the orientation of the N- and C-terminal residues in *Rhau18*, *RhauB* and *Rhau29*
14 structures differs. This difference could be explained by the presence of several interactions in
15 the crystal packing: interactions between His3, Pro4, Gly5, His6 and W15 of *Rhau29* and the
16 sugar-phosphate backbone of G9, G10, and G11 and T8 (Figure S6), with similar interactions
17 being observed in the X-ray crystal structure of the bovine RHAU[31]. The N-terminal region of
18 each peptide differs minutely between the structures, with *Rhau18* containing two extra residues
19 (Ser and Met) from the plasmid vector[20]. Additionally, the C-terminal tail of *Rhau29* in the X-
20 ray crystal structure is facing the 3'-end and 5'-end terminal DNA residues, while the C-terminal
21 tail of *Rhau18* in the NMR structure is facing the middle thymine loop.
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

50 **Binding of *Rhau29* to a different tetrad polarity**

51
52 In our structure, an optimal fit is found between the peptide and the 3'-end G-tetrad. The *Rhau29*
53 peptide face interacting with the G-quadruplex is remarkably planar, with C α of Gly11 pointing
54 toward the center of the G-tetrad (Figure S7). Electrostatic interactions are found between three
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 conserved residues of *Rhau29* (Lys6, Arg8 and Lys17)[23] and the phosphate groups of *T95*.
5
6 Within the asymmetric unit, the position of these residues differs slightly between the different
7
8 peptide-DNA complexes (Figure S8). This is reflected by higher temperature factors for these
9
10 peptide side chains (43.82 ± 11.96 , as compared to the mean temperature factors for all peptide
11
12 side chains, 27.00 ± 11.96). Extensive CH- π and CH₃- π interaction are also found between His1,
13
14 Gly7, Arg8, Ile10, Gly11, Met12 and Tyr14 with the guanine bases (Figure 2B, 2D).
15
16
17
18
19
20

21 Our *Rhau29-T95* structure shows *Rhau29* to bind to the 3'-end G-tetrad, in contrast to structures
22
23 of similar RHAU fragments (*Rhau18* and *RhauB*) bound to the 5'-end tetrad[20, 31].
24
25 Structurally, the 5'- and 3'-end G-tetrads in a parallel-stranded G-quadruplex are quite different:
26
27 (i) the exposed face polarity is reversed (Figure 3), which was shown to influence the stacking
28
29 energy[36]; (ii) the sugar and phosphate group orientations in a 3'-end interface are different
30
31 from those in a 5'-end interface (Figure 3). These differences can influence the recognition mode
32
33 and/or the binding affinity of G-quadruplex-interacting molecules, as recently demonstrated for
34
35 the binding of small-molecule ligands to different G-tetrad surfaces[37].
36
37
38
39
40
41

42 **RHAU contains a versatile G-quadruplex binding motif**

43
44 The *Rhau29-T95* co-crystal structure provides the first high-resolution data for the binding of a
45
46 fragment of RHAU onto the 3'-end G-tetrad of a G-quadruplex. The X-ray crystal structure
47
48 agrees with the reported NMR data of RHAU, where the 5'-end and 3'-end G-tetrads have been
49
50 identified as G-quadruplex binding sites[20]. The N-terminal 18-aa RHAU motif appears to act
51
52 as an independent domain capable of docking onto a G-quadruplex with accessible G-tetrads, in
53
54 which our data demonstrated the recognition of G-quadruplex by RHAU to be malleable and
55
56 adaptive in the binding of 5'- or 3'-end accessible tetrads despite differences in tetrad polarity (5'-
57
58
59
60
61
62
63
64
65

1
2
3
4 vs. 3'-end)[20] or composition (G•G•G•G vs. A•T•G•G)[20, 31]. Within the RHAU G-
5
6 quadruplex recognition motif, the conserved residues Arg8 and Ile10 were found to play a critical
7
8 role in resolving G-quadruplexes by the helicase[38].
9

10
11
12
13
14 Dimeric G-quadruplexes were shown to be potentially formed in TERRA[39, 40], human
15
16 telomerase RNA[41], as well as the human CEB1 minisatellites[42, 43]. In the context of a
17
18 dimeric G-quadruplex, it is speculated that the protein may be positioned on each available side
19
20 of the G-quadruplex dimer to facilitate the unwinding of each individual G-quadruplex subunit.
21
22 It is anticipated that RHAU helicase will use the same binding principles, as evidenced here, in
23
24 the recognition of different tetrads such as G•C•G•C tetrads[2, 44], or even in triads[45].
25
26
27
28
29

30 31 **MATERIALS AND METHODS**

32 33 **Sample preparation**

34
35 DNA oligonucleotides were chemically synthesized on an ABI 394 DNA/RNA synthesizer using
36
37 standard oligonucleotide synthesis cycle. Oligonucleotides were cleaved from the solid support
38
39 and deprotected with ammonium hydroxide, purified using Poly-Pak cartridges (Glen Research)
40
41 and dialyzed successively against water, KCl solution and water prior to lyophilization. Peptides
42
43 used in X-ray crystallization study were purchased from Chempeptide Ltd. DNA and peptide
44
45 samples were reconstituted in potassium cacodylate buffer.
46
47
48
49
50
51

52 53 **Crystallization, X-ray crystallography and structure refinement**

54
55 Crystallization assays were performed with sitting-drop method, using a solution of *T95* (2 mM)
56
57 and *Rhau29* (1.5 mM) in buffer containing 100 mM potassium cacodylate buffer (pH 6.5), 100
58
59
60
61
62

1
2
3
4 mM potassium chloride. Initial screening was done with Natrix commercial screens (Hampton
5
6 Research). Monoclinic crystals (space group $P2_1$) were obtained using 55% methylpentanediol as
7
8 precipitating agent, in the presence of 40 mM sodium cacodylate (pH 6.0), 80 mM potassium
9
10 chloride and 12 mM spermine tetrahydrochloride. For data collection, crystals were directly flash
11
12 frozen in liquid nitrogen. Data were collected from a single crystal at the Swiss Light Source
13
14 beamline ($\lambda = 1.00 \text{ \AA}$; Zurich, Switzerland) equipped with a Pilatus detector, and processed using
15
16 XDS [46]. Data collection statistics are summarized in Table 1. The structure was solved by
17
18 molecular replacement, using PHASER[47] and refined to 1.5 \AA resolution. The coordinates of
19
20 the DNA residues, except the terminal thymines, from the NMR structure of a parallel-stranded
21
22 G-quadruplex [33] were used as a search model. The model was refined and completed with ions
23
24 and water molecules, following several cycles of manual building using Coot[48] and energy
25
26 minimization using REFMAC [34]. The final model contains all 16 DNA nucleotides, 17 out of
27
28 29 peptide residues, 5 potassium ions, and 262 water molecules. Structures were further
29
30 visualized using PYMOL viewer program, that was also used to generate the figures[49].
31
32
33
34
35
36
37

38 **Data deposition**

39
40 The coordinates for the crystal structure of the *Rhau29-T95* complex have been deposited in the
41
42 Protein Data Bank (PDB ID: [6Q6R](#))
43
44

45 **FUNDING**

46
47 This work was supported by Singapore Ministry of Education Academic Research Fund Tier 2
48
49 (MOE2018-T2-2-029). No competing financial interests have been declared.
50
51
52

53 **ACKNOWLEDGEMENTS**

54
55 We thank Julien Lescar and Abbas El Sahili for collecting the diffraction data.
56
57
58
59
60
61
62
63
64
65

REFERENCES

- [1] Davis JT. G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed. Engl.* (2004) 43:668-98.
- [2] Patel DJ, Phan AT, Kuryavvi V. Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.* (2007) 35:7429-55.
- [3] Phan AT. Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J.* (2010) 277:1107-17.
- [4] Biffi G, Tannahill D, McCafferty J, Balasubramanian S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* (2013) 5:182-6.
- [5] Henderson A, Wu Y, Huang YC, Chavez EA, Platt J, Johnson FB, et al. Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res.* (2014) 42:860-9.
- [6] Lipps HJ, Rhodes D. G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol.* (2009) 19:414-22.
- [7] Cahoon LA, Seifert HS. An alternative DNA structure is necessary for pilin antigenic variation in *Neisseria gonorrhoeae*. *Science.* (2009) 325:764-7.
- [8] Kumari S, Bugaut A, Huppert JL, Balasubramanian S. An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.* (2007) 3:218-21.
- [9] Lopes J, Piazza A, Bermejo R, Kriegsman B, Colosio A, Teulade-Fichou MP, et al. G-quadruplex-induced instability during leading-strand replication. *EMBO J.* (2011) 30:4033-46.
- [10] Maizels N, Gray LT. The G4 genome. *PLoS Genet.* (2013) 9:e1003468.
- [11] Murat P, Balasubramanian S. Existence and consequences of G-quadruplex structures in DNA. *Curr. Opin. Genet. Dev.* (2014) 25:22-9.
- [12] Rhodes D, Lipps HJ. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* (2015) 43:8627-37.
- [13] Oganessian L, Bryan TM. Physiological relevance of telomeric G-quadruplex formation: a potential drug target. *Bioessays.* (2007) 29:155-65.
- [14] Sauer M, Paeschke K. G-quadruplex unwinding helicases and their function in vivo. *Biochem. Soc. Trans.* (2017) 45:1173-82.
- [15] Mendoza O, Bourdoncle A, Boule JB, Brosh RM, Jr., Mergny JL. G-quadruplexes and helicases. *Nucleic Acids Res.* (2016) 44:1989-2006.
- [16] Creacy SD, Routh ED, Iwamoto F, Nagamine Y, Akman SA, Vaughn JP. G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. *J. Biol. Chem.* (2008) 283:34626-34.
- [17] Lattmann S, Giri B, Vaughn JP, Akman SA, Nagamine Y. Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU. *Nucleic Acids Res.* (2010) 38:6219-33.

- 1
2
3
4 [18] Meier M, Patel TR, Booy EP, Marushchak O, Okun N, Deo S, et al. Binding of G-
5 quadruplexes to the N-terminal recognition domain of the RNA helicase associated with AU-rich
6 element (RHAU). *J. Biol. Chem.* (2013) 288:35014-27.
7
8 [19] Vaughn JP, Creacy SD, Routh ED, Joyner-Butt C, Jenkins GS, Pauli S, et al. The DEXH
9 protein product of the DHX36 gene is the major source of tetramolecular quadruplex G4-DNA
10 resolving activity in HeLa cell lysates. *J. Biol. Chem.* (2005) 280:38117-20.
11
12 [20] Heddi B, Cheong VV, Martadinata H, Phan AT. Insights into G-quadruplex specific
13 recognition by the DEAH-box helicase RHAU: Solution structure of a peptide-quadruplex
14 complex. *Proc. Natl. Acad. Sci. U.S.A.* (2015) 112:9608-13.
15
16 [21] Fu JJ, Li LY, Lu GX. Molecular cloning and characterization of human DDX36 and mouse
17 Ddx36 genes, new members of the DEAD/H box superfamily. *Acta Biochim. Biophys. Sin.*
18 (Shanghai). (2002) 34:655-61.
19
20 [22] Gao X, Ma W, Nie J, Zhang C, Zhang J, Yao G, et al. A G-quadruplex DNA structure
21 resolvase, RHAU, is essential for spermatogonia differentiation. *Cell Death Dis.* (2015) 6:e1610.
22
23 [23] Lattmann S, Giri B, Vaughn JP, Akman SA, Nagamine Y. Role of the amino terminal
24 RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the
25 DEAH-box RNA helicase RHAU. *Nucleic Acids Res.* (2010) 38:6219-33.
26
27 [24] Nie J, Jiang M, Zhang X, Tang H, Jin H, Huang X, et al. Post-transcriptional Regulation of
28 Nkx2-5 by RHAU in Heart Development. *Cell Rep.* (2015) 13:723-32.
29
30 [25] Lai JC, Ponti S, Pan D, Kohler H, Skoda RC, Matthias P, et al. The DEAH-box helicase
31 RHAU is an essential gene and critical for mouse hematopoiesis. *Blood.* (2012) 119:4291-300.
32
33 [26] Lattmann S, Stadler MB, Vaughn JP, Akman SA, Nagamine Y. The DEAH-box RNA
34 helicase RHAU binds an intramolecular RNA G-quadruplex in TERC and associates with
35 telomerase holoenzyme. *Nucleic Acids Res.* (2011) 39:9390-404.
36
37 [27] Booy EP, Meier M, Okun N, Novakowski SK, Xiong S, Stetefeld J, et al. The RNA helicase
38 RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the
39 formation of the P1 helix template boundary. *Nucleic Acids Res.* (2012) 40:4110-24.
40
41 [28] Booy EP, Howard R, Marushchak O, Ariyo EO, Meier M, Novakowski SK, et al. The RNA
42 helicase RHAU (DHX36) suppresses expression of the transcription factor PITX1. *Nucleic*
43 *Acids Res.* (2014) 42:3346-61.
44
45 [29] Huang H, Suslov NB, Li NS, Shelke SA, Evans ME, Koldobskaya Y, et al. A G-
46 quadruplex-containing RNA activates fluorescence in a GFP-like fluorophore. *Nat. Chem. Biol.*
47 (2014) 10:686-U128.
48
49 [30] Ohnmacht SA, Neidle S. Small-molecule quadruplex-targeted drug discovery. *Bioorg. Med.*
50 *Chem. Lett.* (2014) 24:2602-12.
51
52 [31] Chen MC, Tippiana R, Demeshkina NA, Murat P, Balasubramanian S, Myong S, et al.
53 Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature.* (2018)
54 558:465-9.
55
56
57
58
59
60
61
62
63
64
65

- 1
2
3
4 [32] Chen WF, Rety S, Guo HL, Dai YX, Wu WQ, Liu NN, et al. Molecular Mechanistic
5 Insights into Drosophila DHX36-Mediated G-Quadruplex Unfolding: A Structure-Based Model.
6 Structure. (2018) 26:403-15.
7
8 [33] Do NQ, Phan AT. Monomer-dimer equilibrium for the 5'-5' stacking of propeller-type
9 parallel-stranded G-quadruplexes: NMR structural study. Chemistry. (2012) 18:14752-9.
10
11 [34] Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al.
12 REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D. Biol.
13 Crystallogr. (2011) 67:355-67.
14
15 [35] Do NQ, Lim KW, Teo MH, Heddi B, Phan AT. Stacking of G-quadruplexes: NMR
16 structure of a G-rich oligonucleotide with potential anti-HIV and anticancer activity. Nucleic
17 Acids Res. (2011) 39:9448-57.
18
19 [36] Lech CJ, Heddi B, Phan AT. Guanine base stacking in G-quadruplex nucleic acids. Nucleic
20 Acids Res. (2013) 41:2034-46.
21
22 [37] Le DD, Di Antonio M, Chan LK, Balasubramanian S. G-quadruplex ligands exhibit
23 differential G-tetrad selectivity. Chem Commun (Camb). (2015) 51:8048-50.
24
25 [38] Tippana R, Chen MC, Demeshkina NA, Ferre-D'Amare AR, Myong S. RNA G-quadruplex
26 is resolved by repetitive and ATP-dependent mechanism of DHX36. Nat Commun. (2019)
27 10:1855.
28
29 [39] Martadinata H, Phan AT. Structure of human telomeric RNA (TERRA): stacking of two G-
30 quadruplex blocks in K(+) solution. Biochemistry. (2013) 52:2176-83.
31
32 [40] Martadinata H, Heddi B, Lim KW, Phan AT. Structure of long human telomeric RNA
33 (TERRA): G-quadruplexes formed by four and eight UUAGGG repeats are stable building
34 blocks. Biochemistry. (2011) 50:6455-61.
35
36 [41] Martadinata H, Phan AT. Formation of a stacked dimeric G-quadruplex containing bulges
37 by the 5'-terminal region of human telomerase RNA (hTERC). Biochemistry. (2014) 53:1595-
38 600.
39
40 [42] Piazza A, Cui X, Adrian M, Samazan F, Heddi B, Phan AT, et al. Non-Canonical G-
41 quadruplexes cause the hCEB1 minisatellite instability in *Saccharomyces cerevisiae*. Elife.
42 (2017) 6.
43
44 [43] Adrian M, Ang DJ, Lech CJ, Heddi B, Nicolas A, Phan AT. Structure and conformational
45 dynamics of a stacked dimeric G-quadruplex formed by the human CEB1 minisatellite. J. Am.
46 Chem. Soc. (2014) 136:6297-305.
47
48 [44] Zhang N, Gorin A, Majumdar A, Kettani A, Chernichenko N, Skripkin E, et al. Dimeric
49 DNA quadruplex containing major groove-aligned A-T-A-T and G-C-G-C tetrads stabilized by
50 inter-subunit Watson-Crick A-T and G-C pairs. J. Mol. Biol. (2001) 312:1073-88.
51
52 [45] Heddi B, Martin-Pintado N, Serimbetov Z, Kari TM, Phan AT. G-quadruplexes with (4n -
53 1) guanines in the G-tetrad core: formation of a G-triad.water complex and implication for small-
54 molecule binding. Nucleic Acids Res. (2016) 44:910-6.
55
56 [46] Kabsch W. Xds. Acta Crystallogr. D. Biol. Crystallogr. (2010) 66:125-32.
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

[47] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J. Appl. Crystallogr.* (2007) 40:658-74.

[48] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* (2010) 66:486-501.

[49] Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.3r1. 2010.

1
2
3
4 **FIGURE LEGENDS**
5
6
7
8
9

10 **Figure 1.** Crystal structure of *Rhau29* peptide and *T95* G-quadruplex. **(a)** Model of a dimer and
11 the corresponding electron density (2mFc0-dFc map contoured at 1.0 σ). **(b)** Schematic view of
12 the topology of the *T95* G-quadruplex dimer **(c)** ribbon view of four individual peptides/DNA
13 complexes found in the asymmetric unit viewed from the front and **(d)** the top. Proteins are in
14 green, DNA are in gray (backbone), cyan (Guanine bases), and orange (Thymine bases).
15
16
17
18
19
20
21
22
23

24 **Figure 2.** Molecular details of the DNA/peptide interaction. **(a)** and **(c)** Top and front ribbon
25 views of one complex monomer. **(b)** and **(d)** Top- and front-views of the amino-acid residues
26 interacting with the G-tetrads (only residues in direct contact with the tetrad are shown).
27
28
29
30
31
32
33

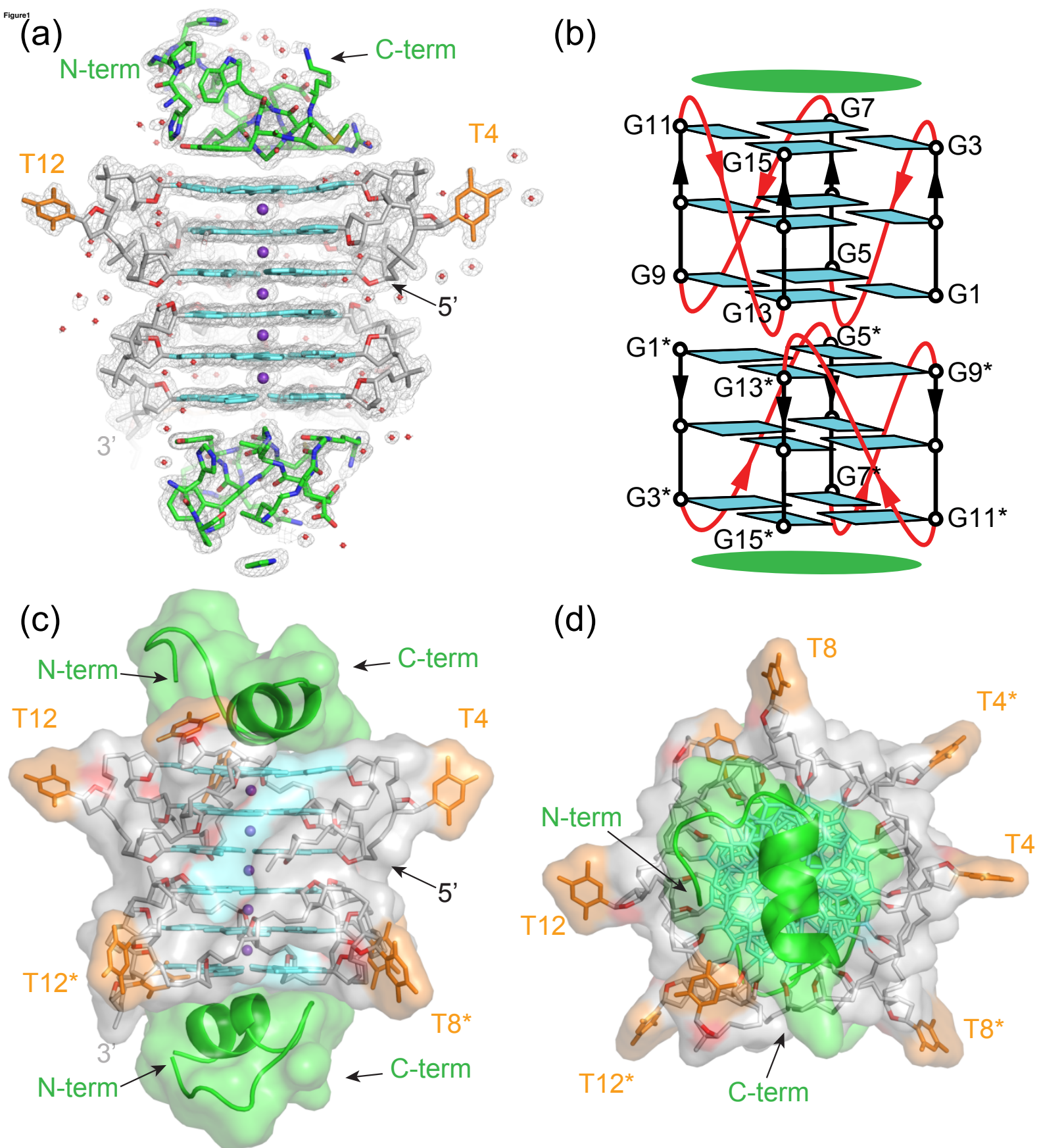
34 **Figure 3.** Comparison between X-ray (left) and NMR (right) structure. **(a)** Top view of a G-
35 tetrad oriented in 3'- and 5'-end interfaces and **(b)** front view of peptide/DNA complex. Sugars
36 are represented in ball and sticks view; O4' atoms are in red; Guanine residues are in cyan.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 1. Data collection and refinement statistics

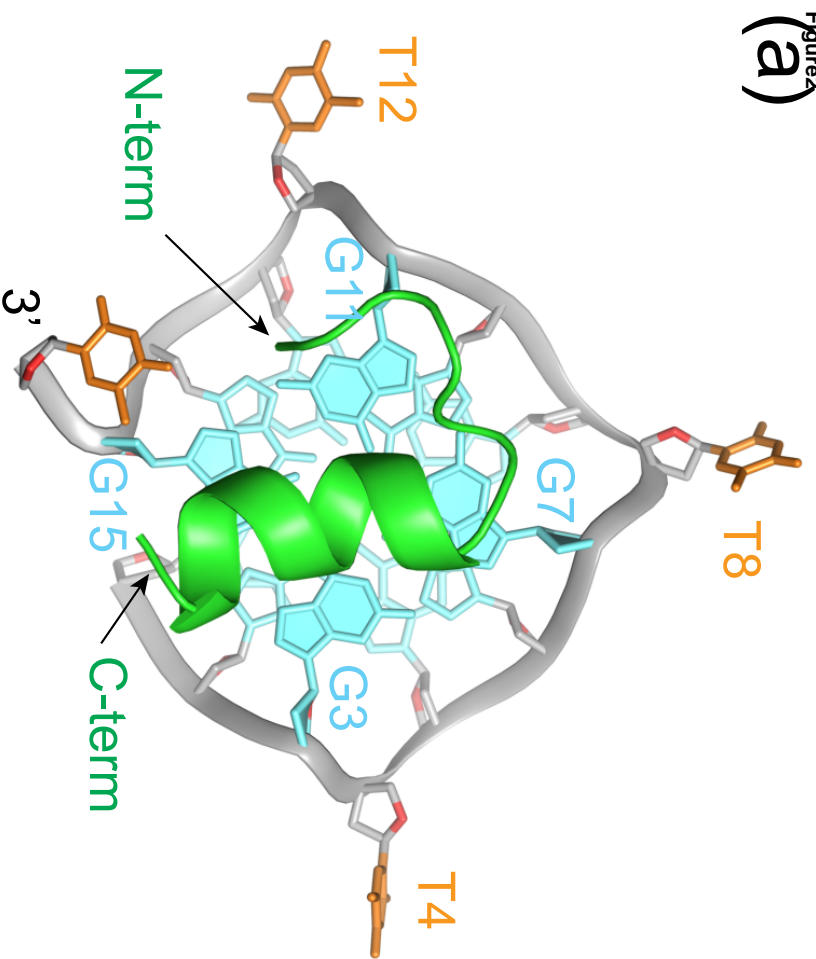
Data collection*	
Space group	P2 ₁
Cell dimensions	
a, b, c (Å)	56.27, 42.09, 61.11
α, β, γ (°)	90.0, 99.3, 90.0
Resolution (Å)	60.60 1.50
Rsym	0.083 (0.941)
I / σI	8.50 (1.34)
Completeness (%)	99.35 (96.25)
Redundancy	9.1 (5.6)
Refinement	
Resolution (Å)	44.75 - 1.50
Number of reflections	45457
Rwork / Rfree	0.1853/0.2448
Number of atoms	
DNA	1364
Protein	564
K ⁺	10
Water	265
	DNA 23.7
	Protein 28.4
B-factors (Å ²)	K ⁺ 11.1
	Water 38.4
r.m.s.d.	
Bond lengths (Å)	0.019
Bond angles (°)	1.947

*A single crystal was used for data collection. Values in parentheses are for the highest-resolution shell.

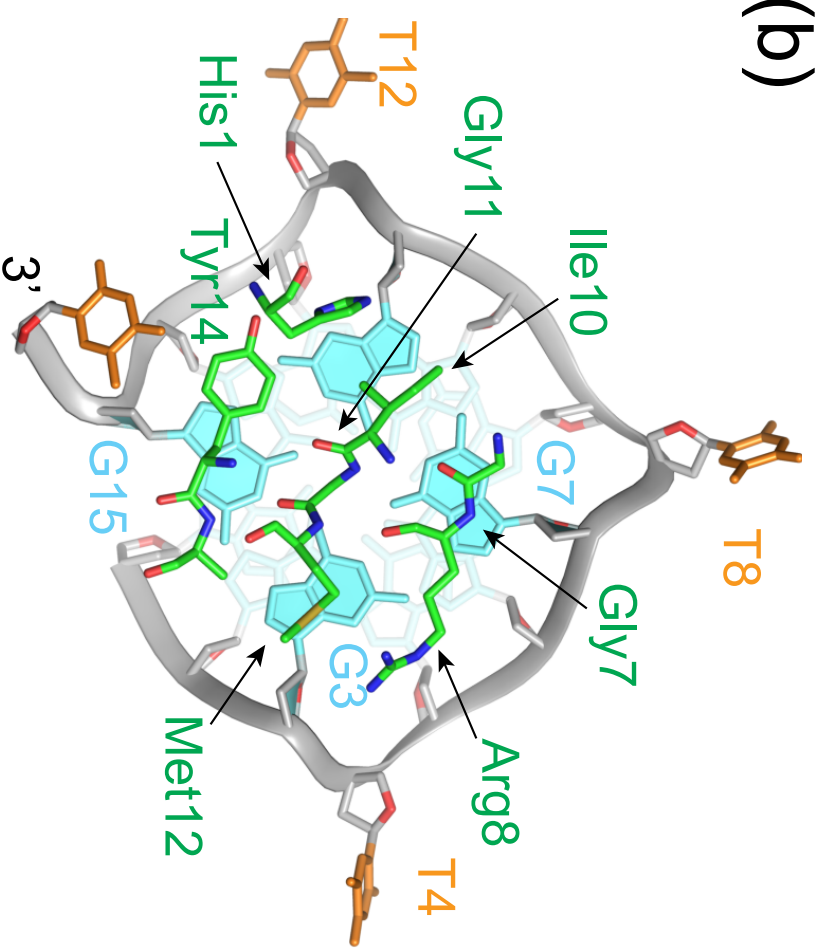
Figure 1



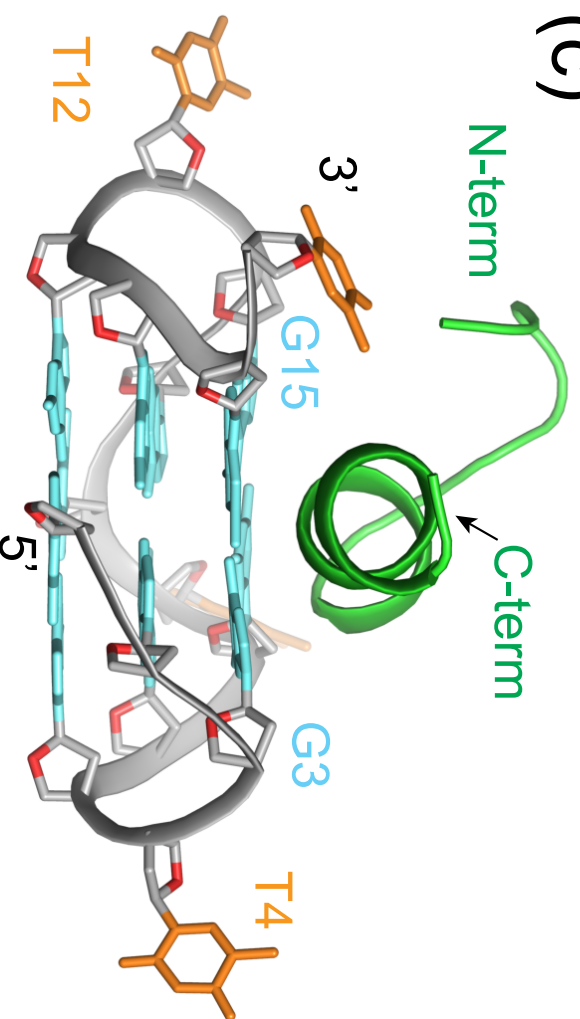
(a)



(b)



(c)



(d)

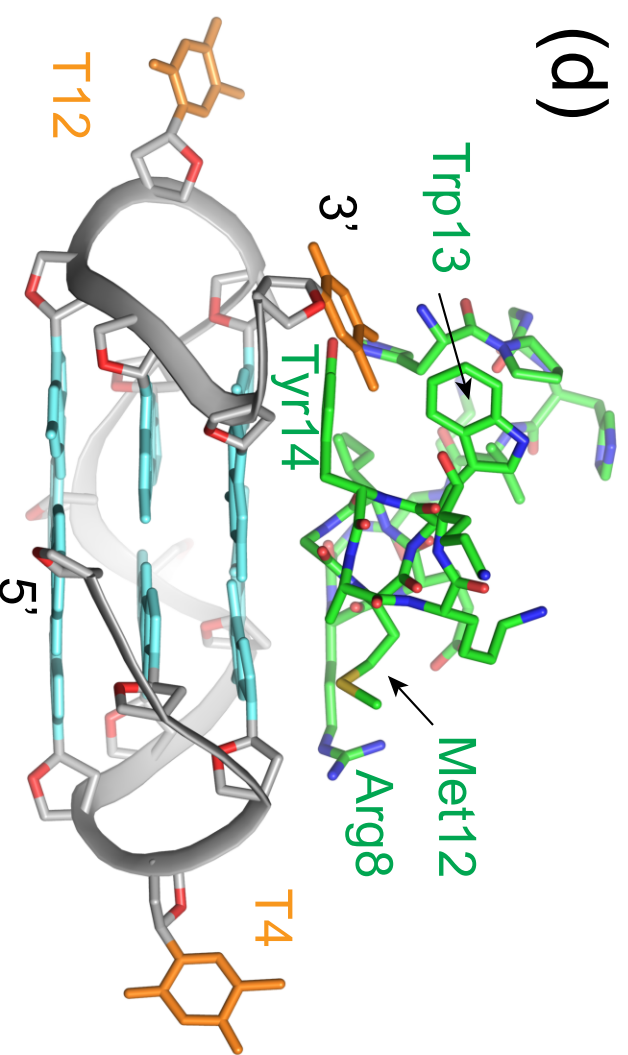
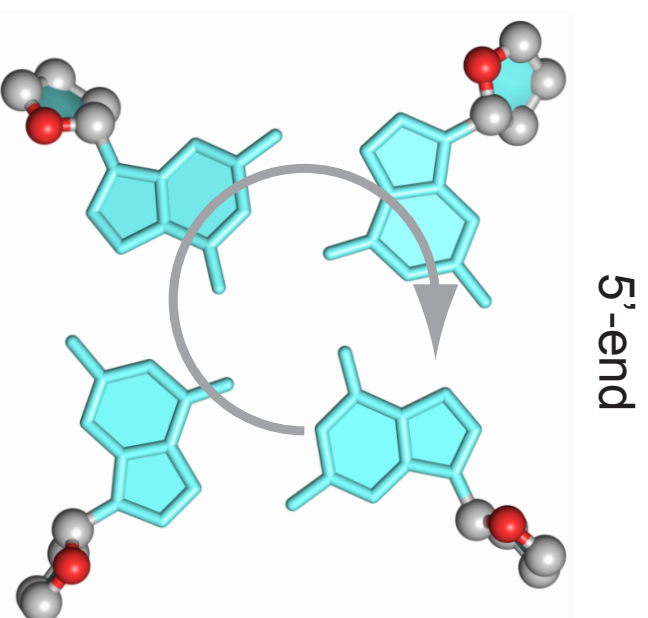
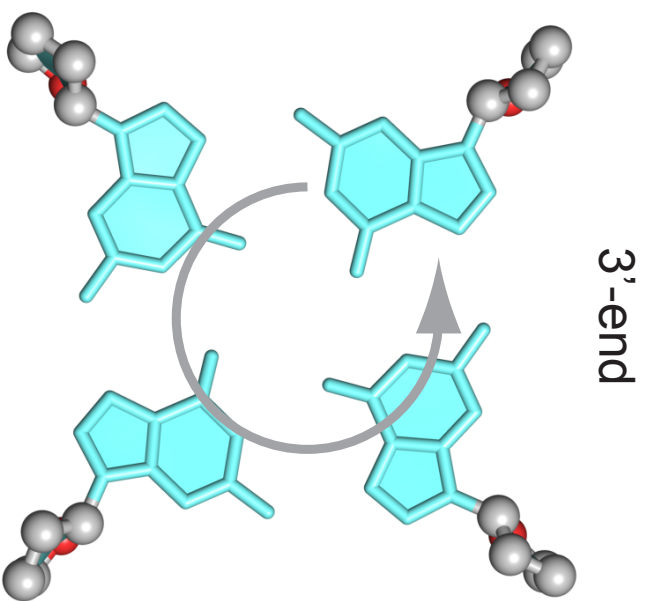


Figure 9
(a)



(b)

