Structure and conformational dynamics of a stacked dimeric G-quadruplex formed by the human CEB1 minisatellite

Michael Adrian[†], Ding Jie Ang[†], Christopher J. Lech[†], Brahim Heddi[†], Alain Nicolas[‡] and Anh Tuân Phan*[,†]

[†]School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore

and [‡]Institut Curie, Centre de Recherche, UMR3244 CNRS, Université Pierre et Marie Curie, 75248 Paris, France

SUPPORTING INFORMATION

Table S1. List of site-specific labeled *CEB1* sequences used in this work.

Name	**Sequence (5' - 3')**							
CEB1	A	GGGGGG	A	GGG	A	GGG	T	GG
G6N	A	GGGG^{N}**G**G	A	GGG	A	GGG	T	GG
G7N	A	GGGGG^{N}**G**	A	GGG	A	GGG	T	GG
G9N	A	GGGGGG	A	^{N}**G**GG	A	GGG	T	GG
G10N	A	GGGGGG	A	G^{N}**G**G	A	GGG	T	GG
G13N	A	GGGGGG	A	GGG	A	^{N}**G**GG	T	GG
G15N	A	GGGGGG	A	GGG	A	GG^{N}**G**	T	GG
G17N	A	GGGGGG	A	GGG	A	GGG	T	^{N}**G**G
G2D	A	^{D}**G**GGGGG	A	GGG	A	GGG	T	GG
G3D	A	G^{D}**G**GGGG	A	GGG	A	GGG	T	GG
G4D	A	GG^{D}**G**GGG	A	GGG	A	GGG	T	GG
G5D	A	GGG^{D}**G**GG	A	GGG	A	GGG	T	GG
G7D	A	GGGGG^{D}**G**	A	GGG	A	GGG	T	GG
G9D	A	GGGGGG	A	^{D}**G**GG	A	GGG	T	GG
G10D	A	GGGGGG	A	G^{D}**G**G	A	GGG	T	GG
G11D	A	GGGGGG	A	GG^{D}**G**	A	GGG	T	GG
G14D	A	GGGGGG	A	GGG	A	G^{D}**G**G	T	GG

Table S2. List of site-specific labeled *G(2,3)T* sequences used in this work.

Name	**Sequence (5' - 3')**							
G(2,3)T	A	TTGGGG	A	GGG	A	GGG	T	GG
^{23}*G4N*	A	TT^{N}**G**GGG	A	GGG	A	GGG	T	GG
^{23}*G5N*	A	TTG^{N}**G**GG	A	GGG	A	GGG	T	GG
^{23}*G6N*	A	TTGG^{N}**G**G	A	GGG	A	GGG	T	GG
^{23}*G7N*	A	TTGGG^{N}**G**	A	GGG	A	GGG	T	GG
^{23}*G9N*	A	TTGGGG	A	^{N}**G**GG	A	GGG	T	GG
^{23}*G10N*	A	TTGGGG	A	G^{N}**G**G	A	GGG	T	GG
^{23}*G11N*	A	TTGGGG	A	GG^{N}**G**	A	GGG	T	GG
^{23}*G13N*	A	TTGGGG	A	GGG	A	^{N}**G**GG	T	GG
^{23}*G14N*	A	TTGGGG	A	GGG	A	G^{N}**G**G	T	GG
^{23}*G15N*	A	TTGGGG	A	GGG	A	GG^{N}**G**	T	GG
^{23}*G17N*	A	TTGGGG	A	GGG	A	GGG	T	^{N}**G**G
^{23}*G10D*	A	TTGGGG	A	G^{D}**G**G	A	GGG	T	GG
^{23}*G14D*	A	TTGGGG	A	GGG	A	G^{D}**G**G	T	GG

Table S3. List of site-specific labeled *G(3,4)T* sequences used in this work.

Name	**Sequence (5' - 3')**							
G(3,4)T	A	GTTGGG	A	GGG	A	GGG	T	GG
^{34}G5N	A	GTT$^{\mathbf{N}}$**G**GG	A	GGG	A	GGG	T	GG
^{34}G6N	A	GTTG$^{\mathbf{N}}$**G**G	A	GGG	A	GGG	T	GG
^{34}G13N	A	GTTGGG	A	GGG	A	$^{\mathbf{N}}$**G**GG	T	GG
^{34}G14N	A	GTTGGG	A	GGG	A	G$^{\mathbf{N}}$**G**G	T	GG
^{34}G17N	A	GTTGGG	A	GGG	A	GGG	T	$^{\mathbf{N}}$**G**G
^{34}G2D	A	$^{\mathbf{D}}$**G**TTGGG	A	GGG	A	GGG	T	GG
^{34}G7D	A	GTTGG$^{\mathbf{D}}$**G**	A	GGG	A	GGG	T	GG
^{34}G11D	A	GTTGGG	A	GG$^{\mathbf{D}}$**G**	A	GGG	T	GG
^{34}G13D	A	GTTGGG	A	GGG	A	$^{\mathbf{D}}$**G**GG	T	GG

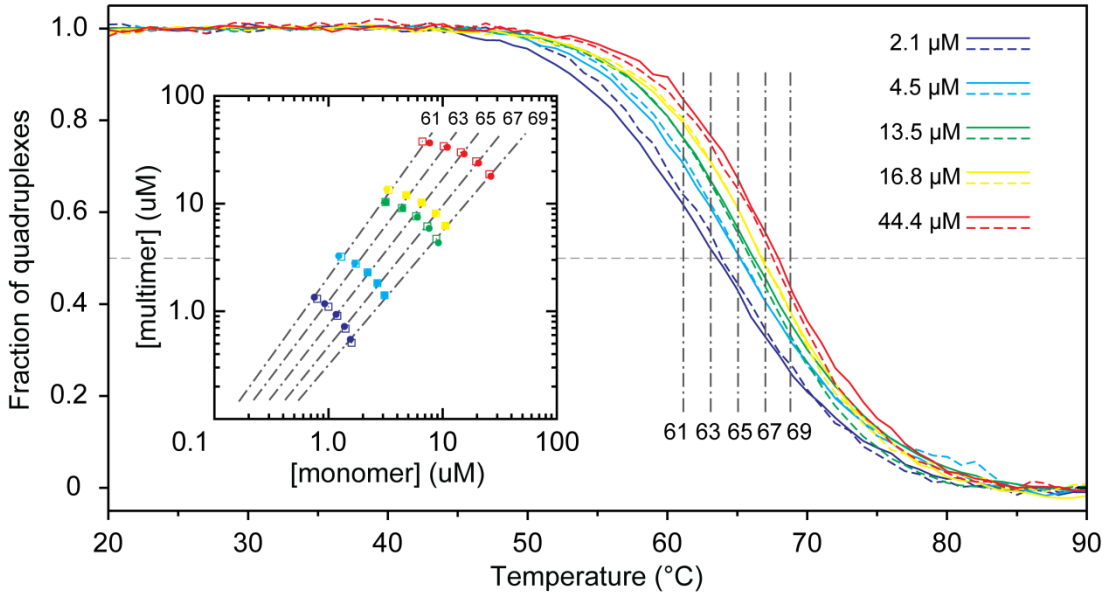

Figure S1. Melting curves of *CEB1* at different strand concentrations in 20 mM potassium phosphate buffer. Solid and dashed lines represent cooling and heating curves, respectively. Insert: Multimer vs. monomer concentrations at different representative temperatures are plotted along the regression lines of slopes ~1.35. Squares and circles correspond to cooling and heating data points, respectively.

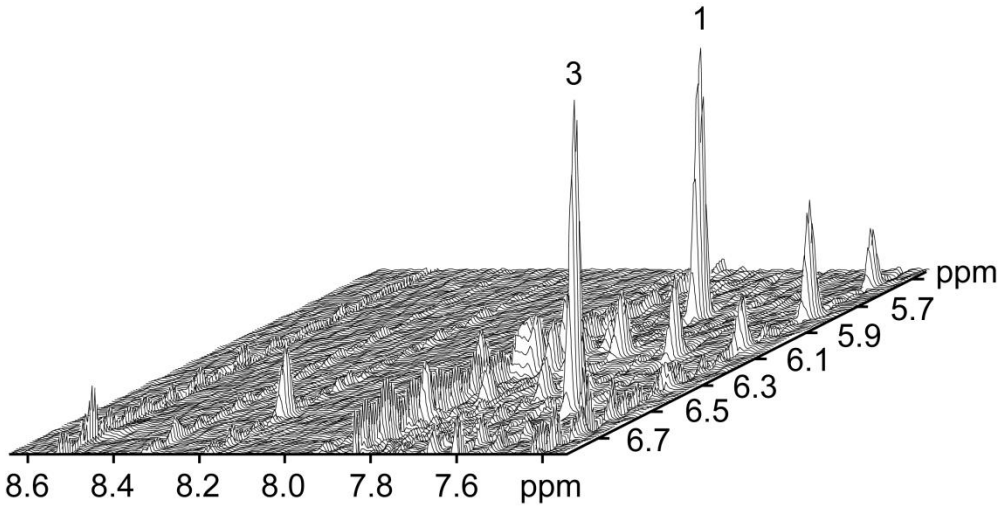

Figure S2. Stack view of NOESY spectrum (mixing time, 100 ms) showing strong intensity of intra-residue H8-H1' cross peaks from residue A1 and G3.

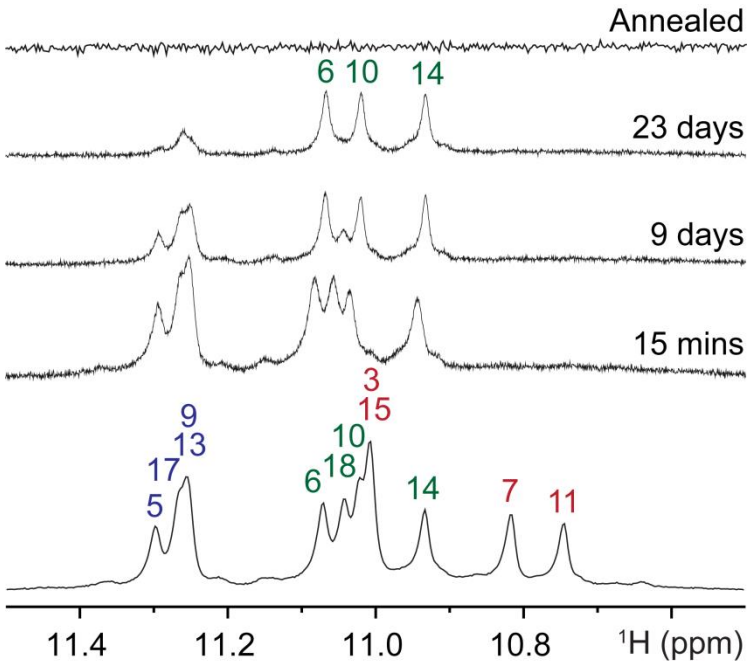

Figure S3. Real-time exchange of imino protons. Spectra were recorded at indicated time after dissolved in D_2O solvent. The top spectrum was obtained following sample annealing. Chemical shifts of outermost tetrad guanines are labeled in red; middle tetrad guanines, green; tetrad guanines at the stacking interface, blue. Minor peaks are not assigned.

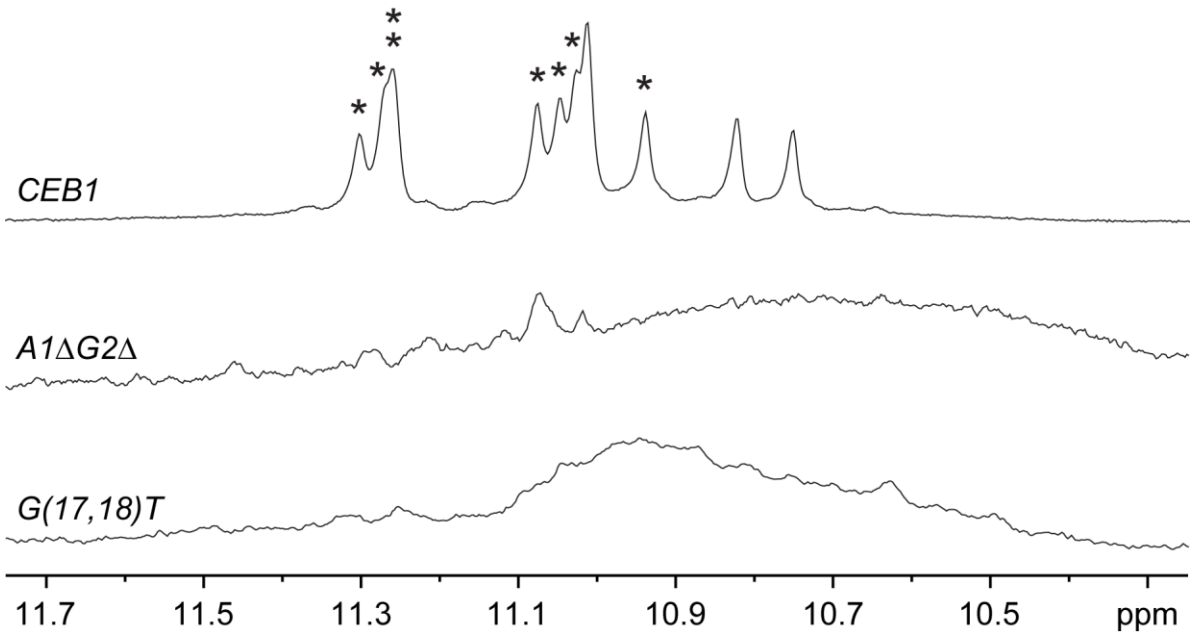

Figure S4. Imino spectra of *CEB1* mutants *A1*Δ*G2*Δ and *G(17,18)T*. Peaks marked with asterisks stayed longer than a day in D_2O at room temperature.

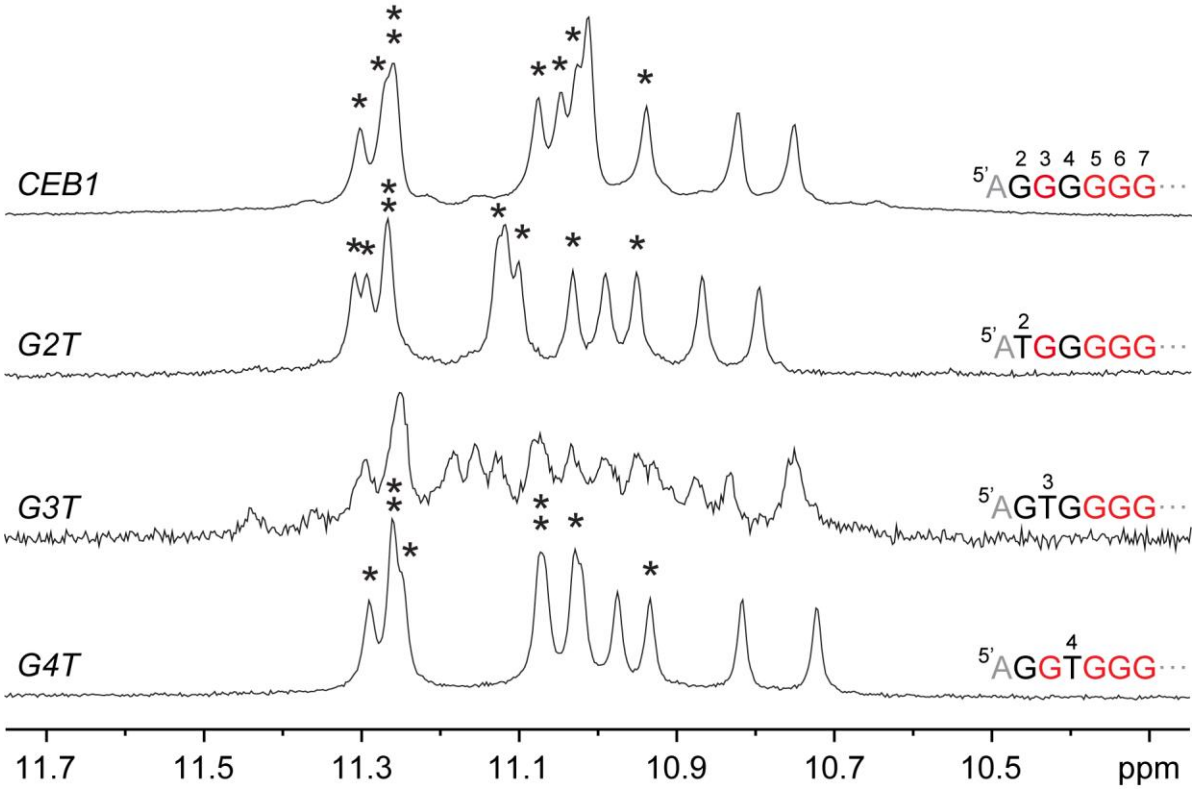

Figure S5. Imino spectra of *CEB1* mutants *G2T*, *G3T*, and *G4T*. Peaks marked with asterisks stayed longer than a day in D_2O at room temperature.

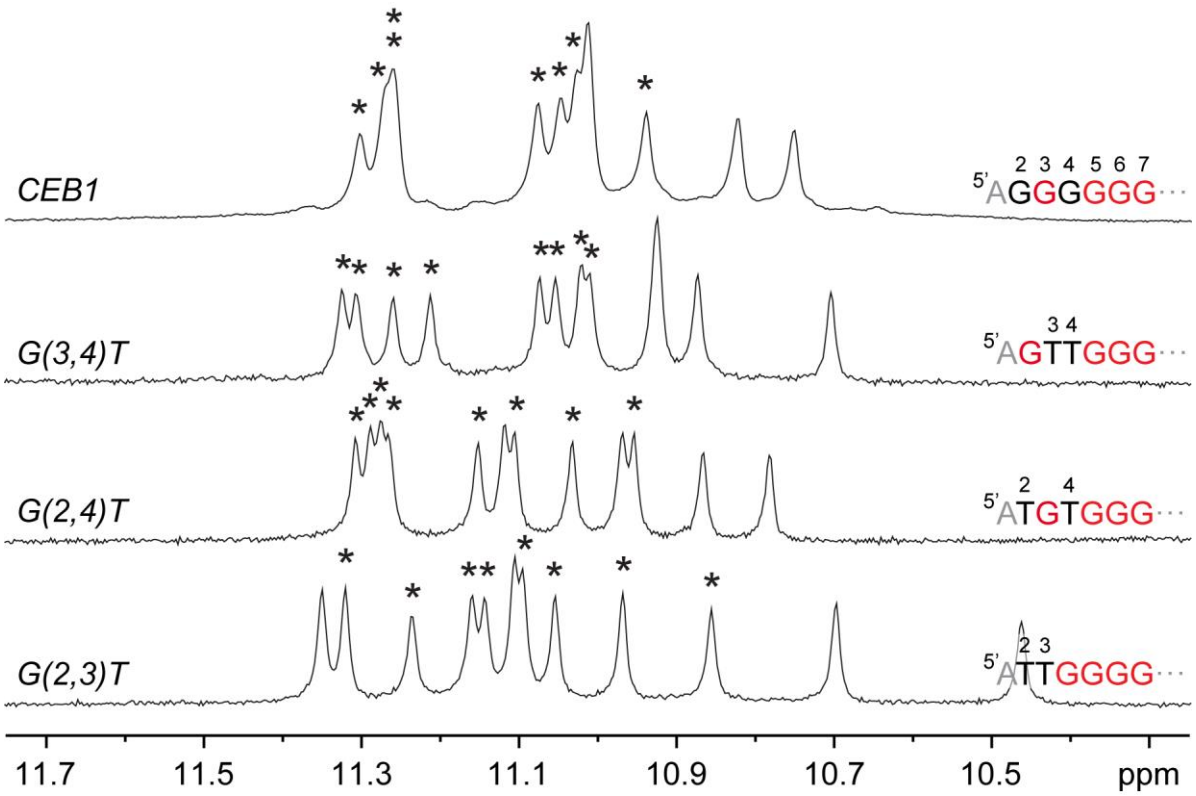

Figure S6. Imino spectra of *CEB1* mutants *G(3,4)T*, *G(2,4)T*, and *G(2,3)T*. Peaks marked with asterisks stayed longer than a day in D_2O at room temperature.

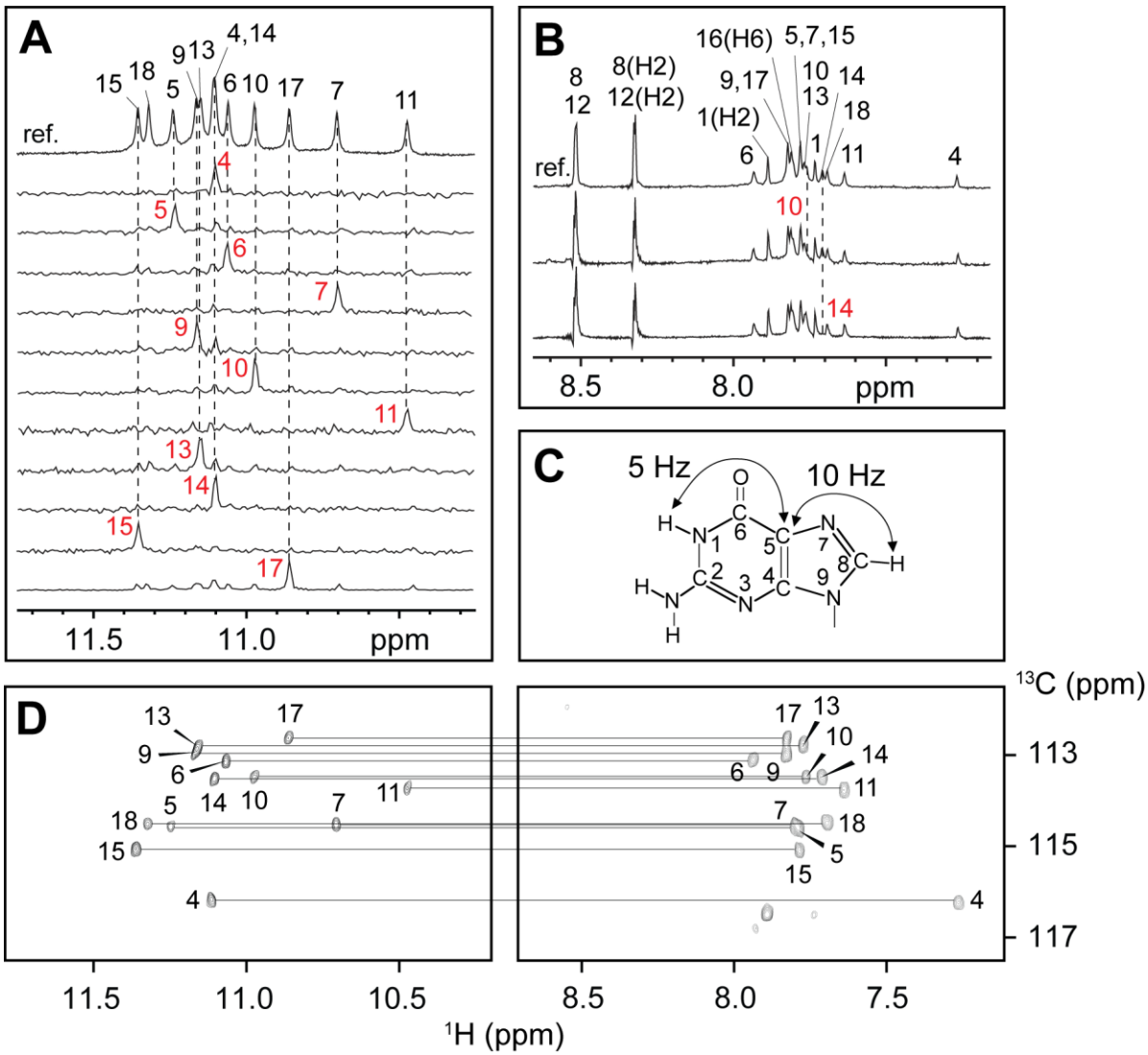

Figure S7. Imino and aromatic H8/6/2 proton assignments of *G(2,3)T*. (A) Imino proton assignments from ^{15}N-filtered spectra of samples, 2% ^{15}N-enriched at indicated positions. (B) H8 proton assignments by site-specific ^{2}H labeling at the indicated positions. The reference spectra (ref.) of imino and aromatic protons are shown at the top. (C,D) Through-bond correlations between guanine imino and H8 protons via ^{13}C5 at natural abundance, using long-range J-couplings shown in (C). Assignments are labeled with residue numbers.

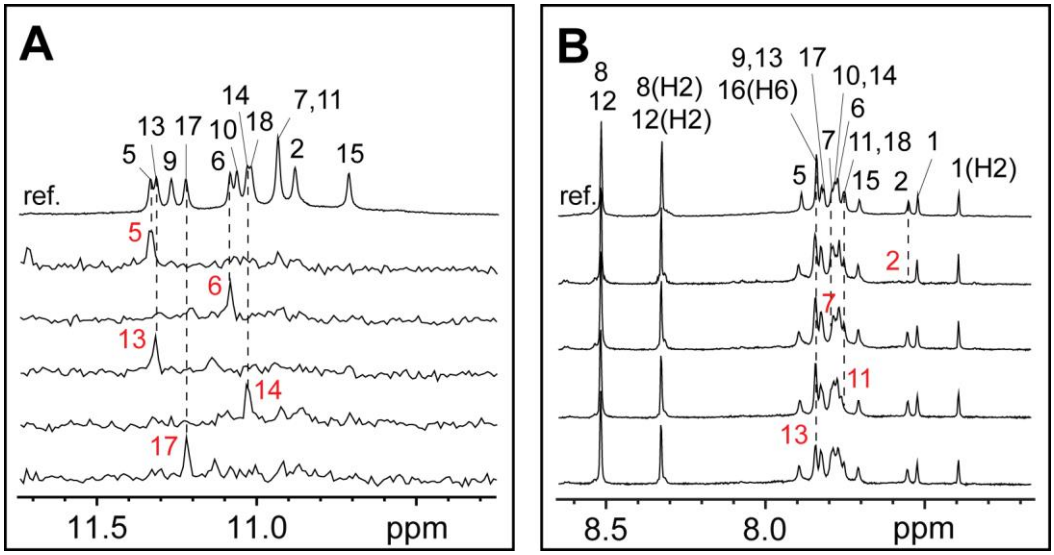

Figure S8. Imino and aromatic H8/6/2 proton assignments of *G(3,4)T*. (A) Imino proton assignments from ^{15}N-filtered spectra of samples, 2% ^{15}N-enriched at indicated positions. (B) H8 proton assignments by site-specific ^{2}H labeling at the indicated positions. The reference spectra (ref.) of imino and aromatic protons are shown at the top. Assignments are labeled with residue numbers.

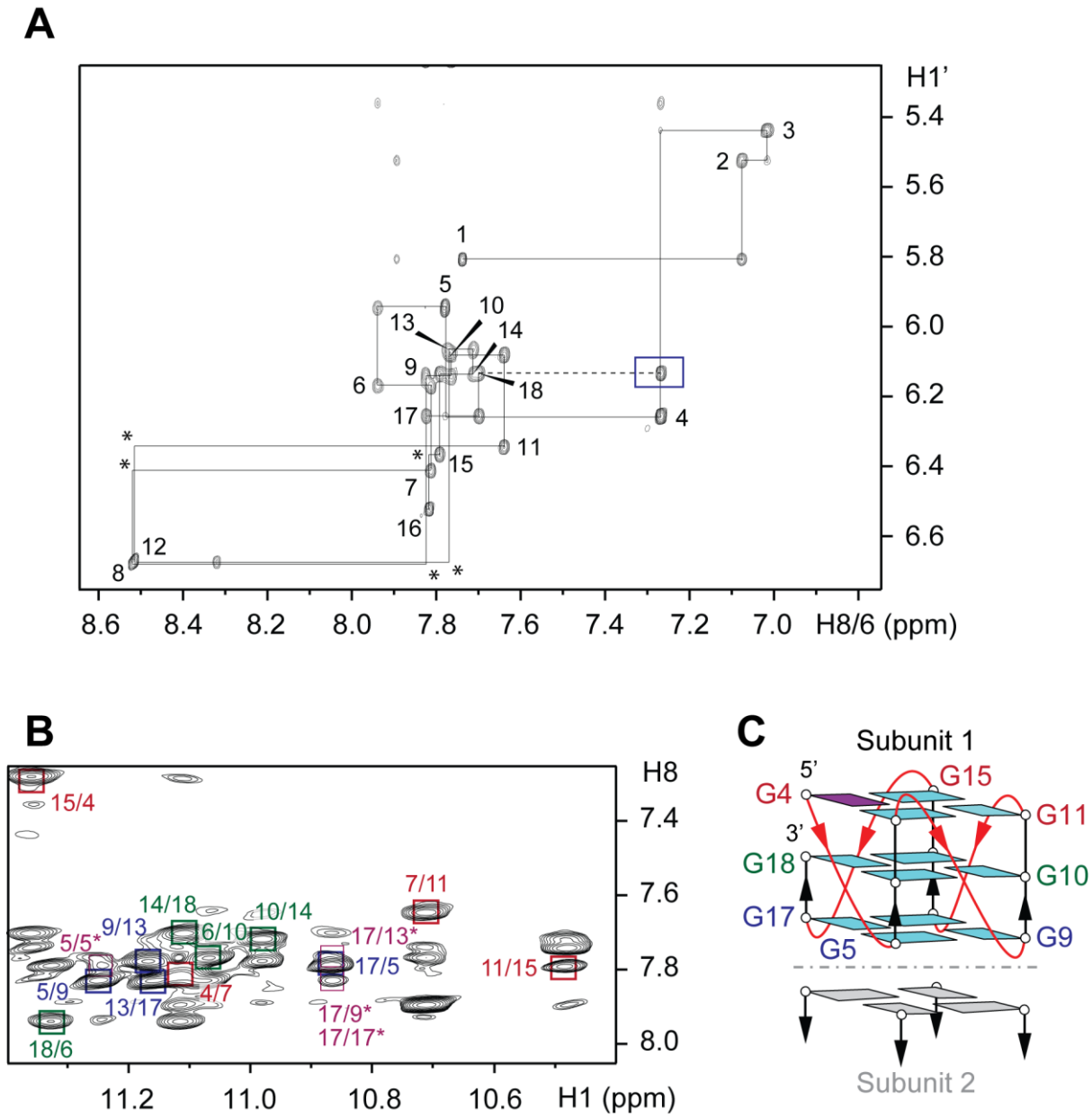

Figure S9. Folding topology of *G(2,3)T*. (A) The H8/H6-H1' sequential connectivity on NOESY spectrum (mixing time, 300 ms). Intra-residue H8/6-H1' cross-peaks are labeled with residue numbers. Missing cross peaks are marked with asterisks. A moderate-intensity cross peak corresponds to the NOE from G18(H1') and G4(H8) is shown by the blue square. (B) The imino-H8 cyclic connectivities on NOESY spectrum (mixing time, 300 ms). The tetrad arrangements were identified from framed cross-peaks with associated label of the residue number of imino proton in the first position and that of H8 proton in the second position. Inter-molecular cross peaks are colored in magenta. (C) Schematic representation of stacked dimeric *G(2,3)T* satisfying the connectivities shown in (A) and (B). *Anti* and *syn* tetrad guanines are colored cyan and magenta, respectively. The backbones of the core and loops are in black and red, correspondingly.

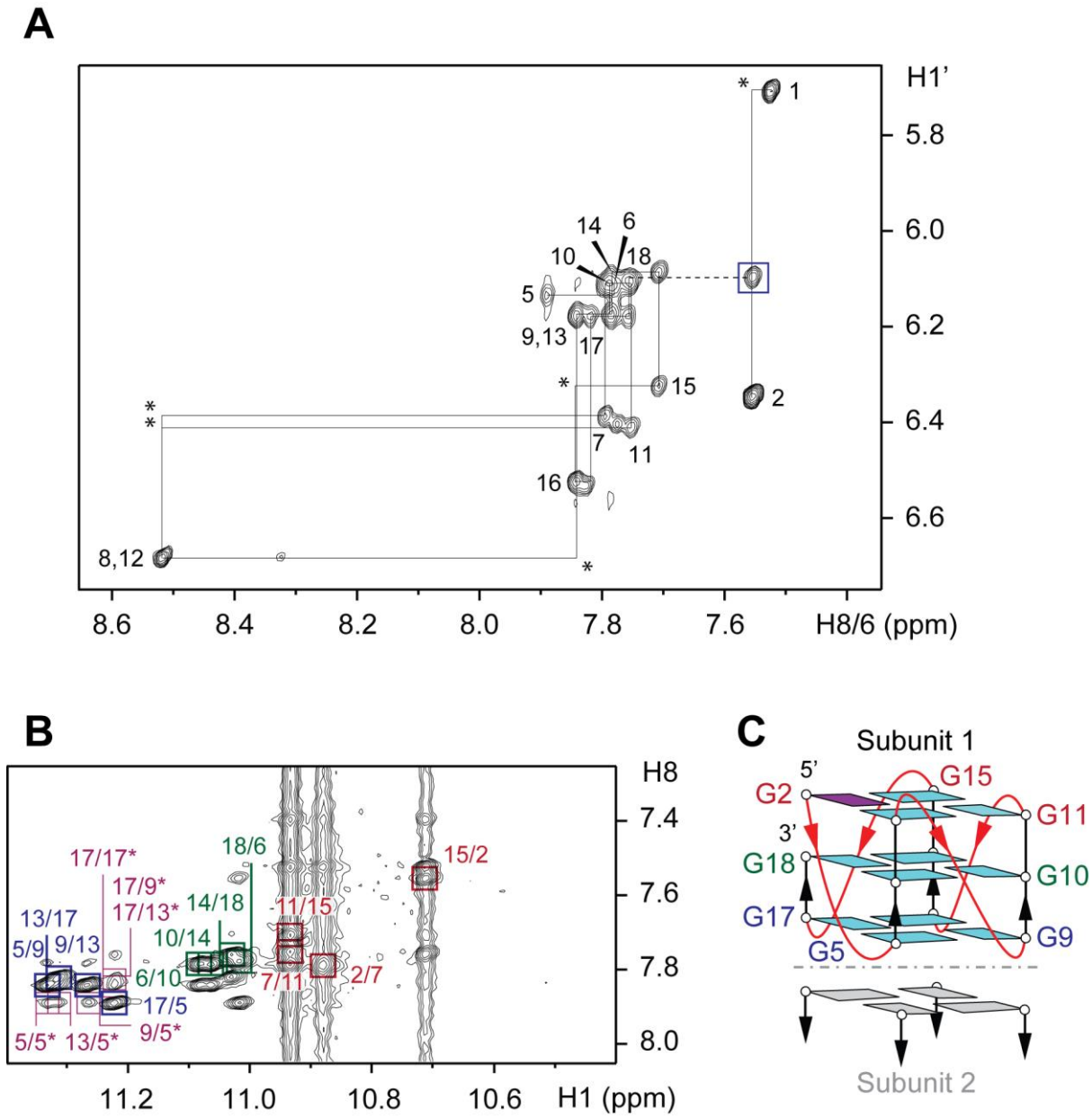

Figure S10. Folding topology of *G(3,4)T*. (A) The H8/H6-H1' sequential connectivity on NOESY spectrum (mixing time, 300 ms). Intra-residue H8/6-H1' cross-peaks are labeled with residue numbers. Missing cross peaks are marked with asterisks. A moderate-intensity cross peak corresponds to the NOE from G18(H1') and G2(H8) is shown by the blue square. (B) The imino-H8 cyclic connectivities on NOESY spectrum (mixing time, 300 ms). The tetrad arrangements were identified from framed cross-peaks with associated label of the residue number of imino proton in the first position and that of H8 proton in the second position. Inter-molecular cross peaks are colored in magenta. (C) Schematic representation of stacked dimeric *G(3,4)T* satisfying the connectivities shown in (A) and (B). *Anti* and *syn* tetrad guanines are colored cyan and magenta, respectively. The backbones of the core and loops are in black and red, correspondingly.

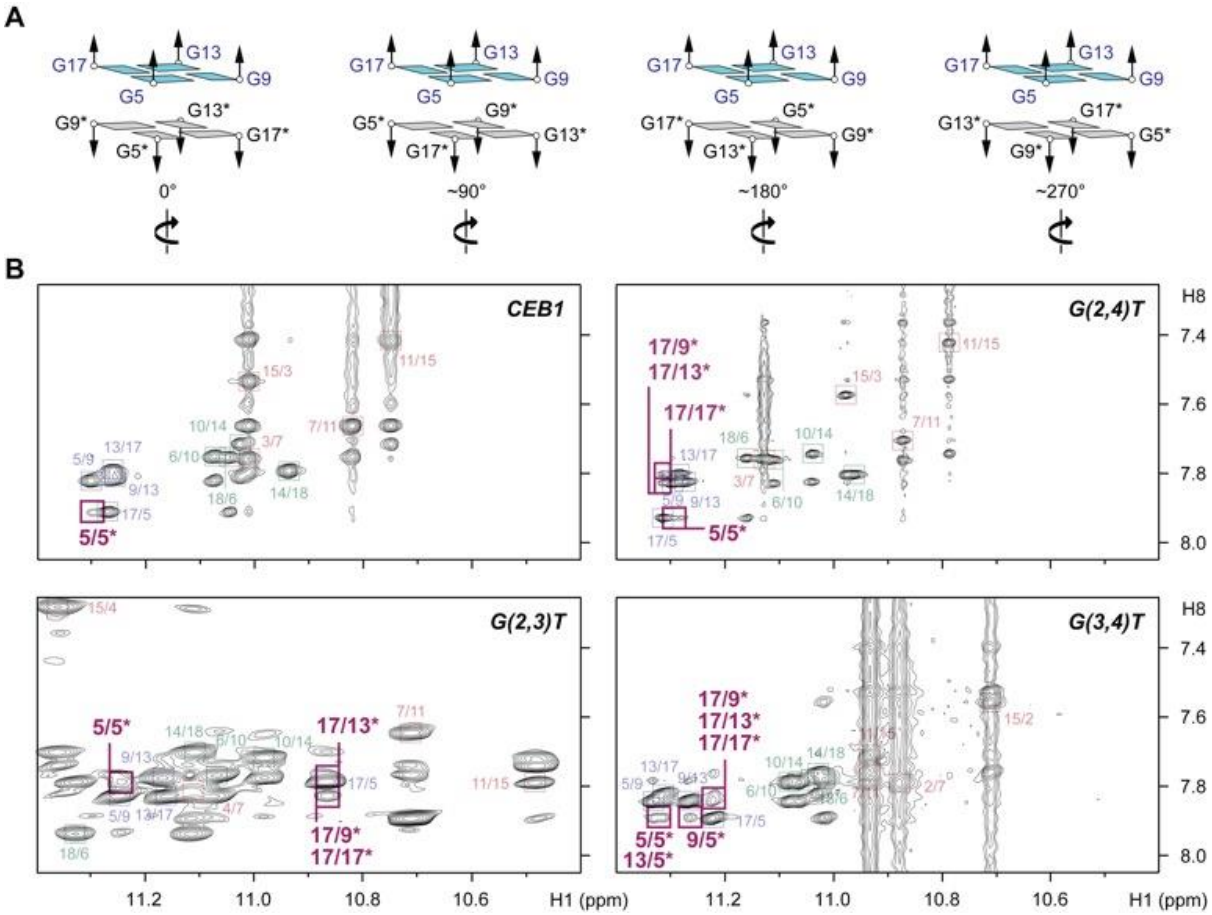

Figure S11. (A) Schematics of possible stacking orientations in *CEB1*. Rotation angle of the bottom subunit (colored gray) relative to the upper subunit (cyan) is shown. Other G-tetrads from each subunit are not drawn for clarity. (B) G-tetrad stacking modes in *CEB1*, *G(2,4)T*, *G(2,3)T*, and *G(3,4)T*. (A) The imino-H8 cyclic connectivities on NOESY spectra (mixing time, 300 ms) of *CEB1* and mutant samples showing folding similarity between *CEB1* and mutant sub-units. The arrangements of G-tetrad were identified from framed cross-peaks with associated label of the residue number of imino proton in the first position and that of H8 proton in the second position. Inter-molecular cross peaks associated to two different stacking modes are highlighted in magenta squares.

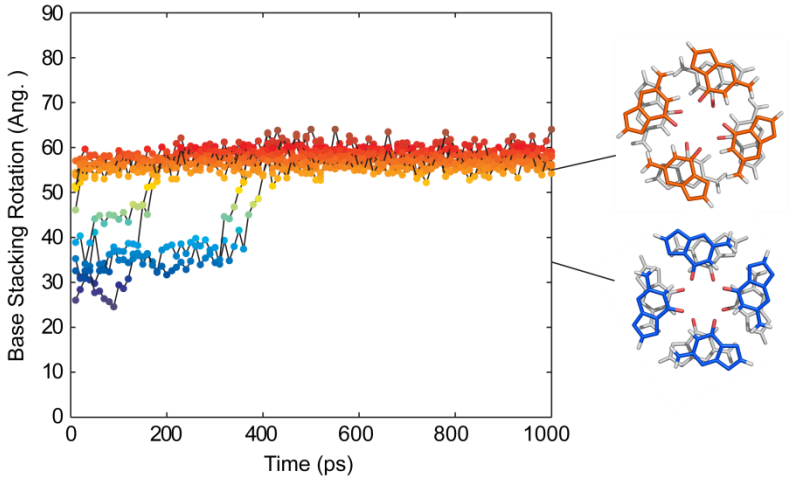

Figure S12. The convergent of stacking orientation within one tetrad quadrant at rotation angle of ~57° was observed during distance-restrained molecular dynamics refinement in explicit solvent using the program AMBER 10. At the beginning of refinement, four out of ten initial 'in vacuo' structures started from base stacking rotation angles of ~25° to ~40°. Representative stacking tetrads at rotation angle of ~35° (colored in blue) and ~57° (orange) show maximum overlap of the six-membered rings of guanine bases and partial overlap of the five- and six-membered rings of guanine bases, respectively.

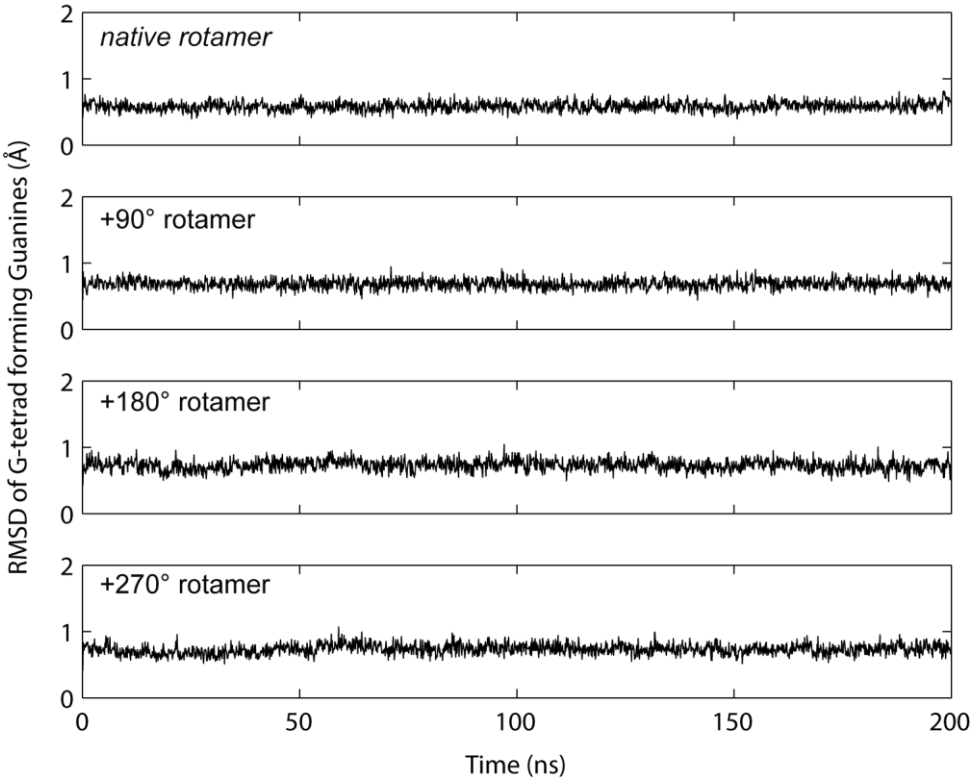

Figure S13. Heavy atom (C,N,O) RMSD of guanine bases within the G-tetrad core of the dimeric *CEB1* G-quadruplex. Stable RMSD values of <1 are observed in 200-ns simulations of all four rotamers of the *CEB1* dimer.

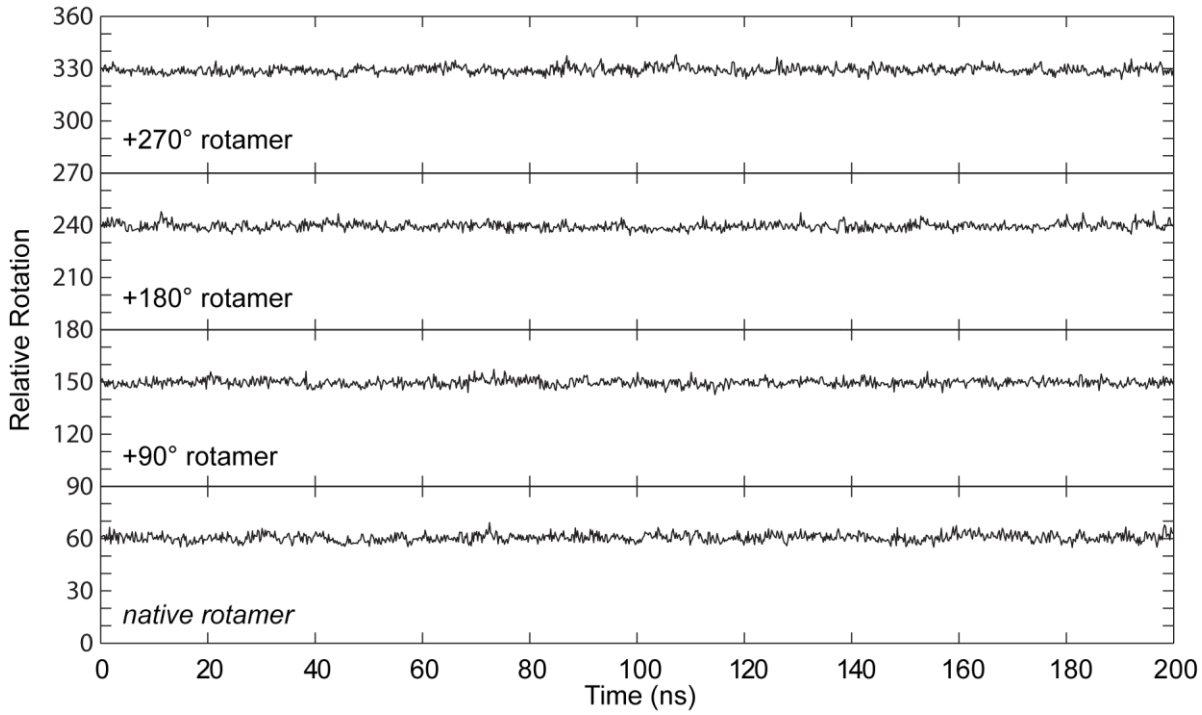

Figure S14. Characterization of interface rotation geometry in MD trajectories: A stable 5/6-ring arrangement of stacked guanine bases is observed over the entire 200-ns trajectory for each of the four stacking quadrants.

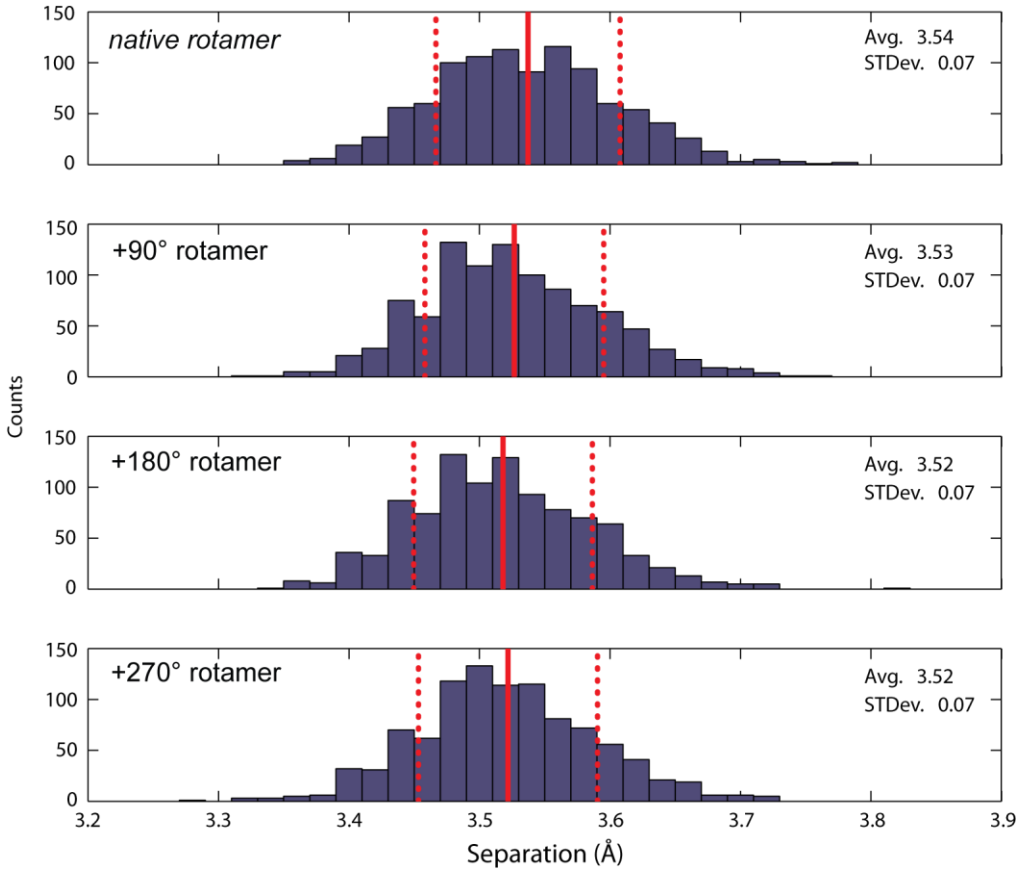

Figure S15. Separation values computed for stacked guanines at the interface of *CEB1* monomers in MD trajectories. Solid line indicates the average separation value while dotter lines indicate the 1st standard deviation from the average (N=1000).

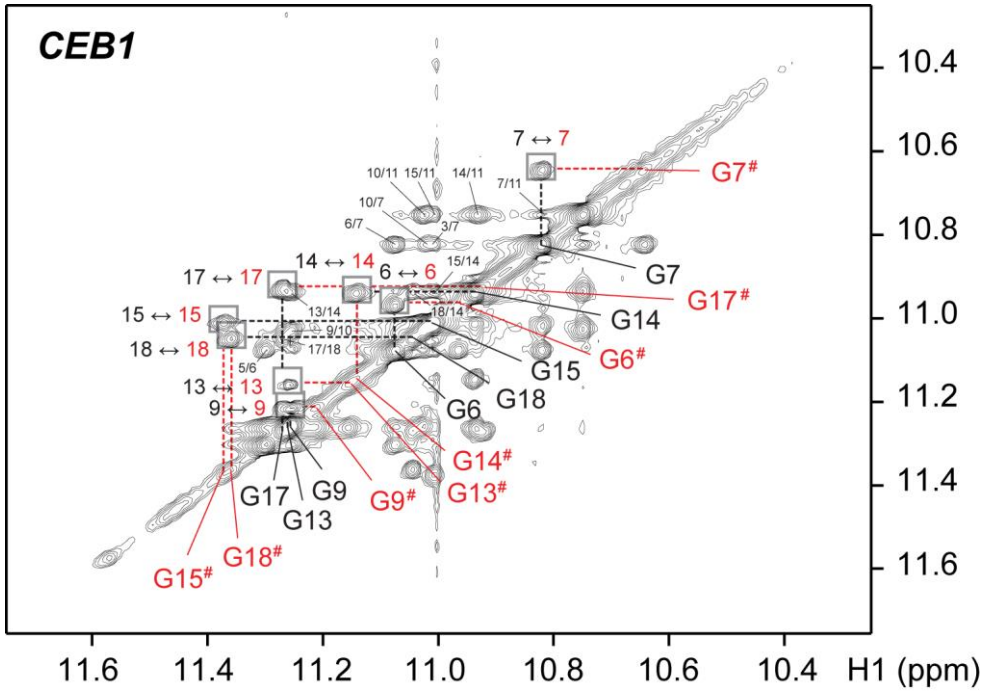

Figure S16. Expanded NOESY spectra (mixing time, 300 ms) of *CEB1* showing imino proton exchange cross peaks (framed) between major and minor forms. NOE cross peak assignments in major form are shown on the spectrum.

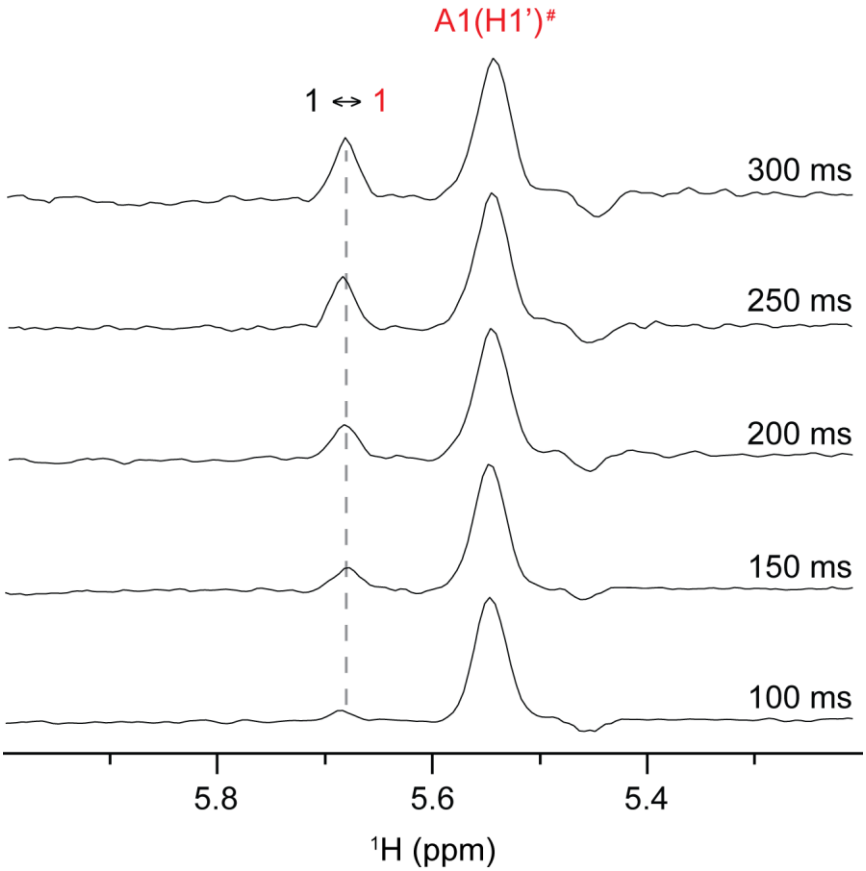

Figure S17. One-dimensional NOESY cross-section spectra of different indicated mixing times showing the profiles of exchange cross-peak and diagonal peak of the minor form (i.e., from A1(H1)$^{\#}$ proton).

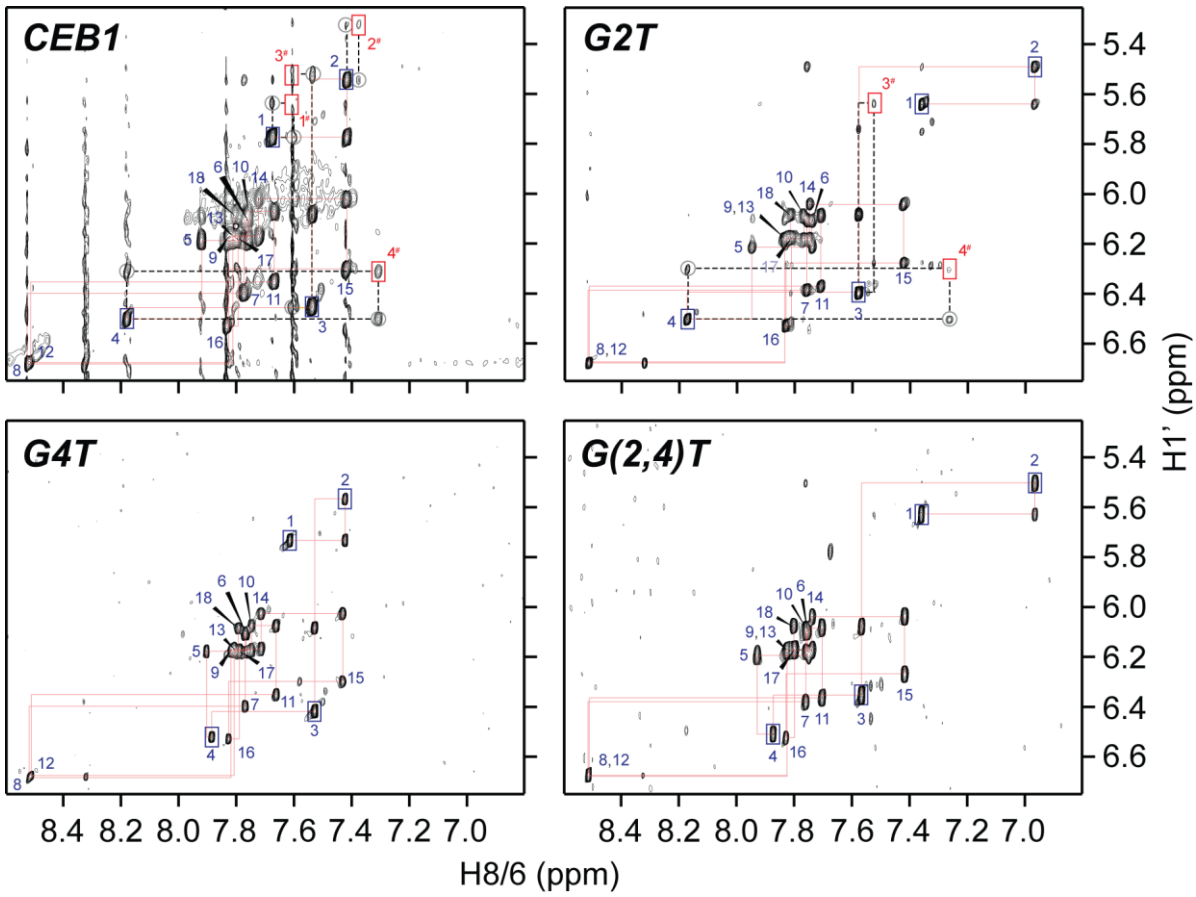

Figure S18. The H8/H6-H1' sequential connectivities on NOESY spectra (mixing time, 300 ms) of *CEB1*, *G2T*, *G4T* and *G(2,4)T*. Cross peaks generated by NOE from dynamic residues of the major form are marked by blue squares; by NOE from the minor form, red squares; by both NOE and chemical exchange, gray circles.