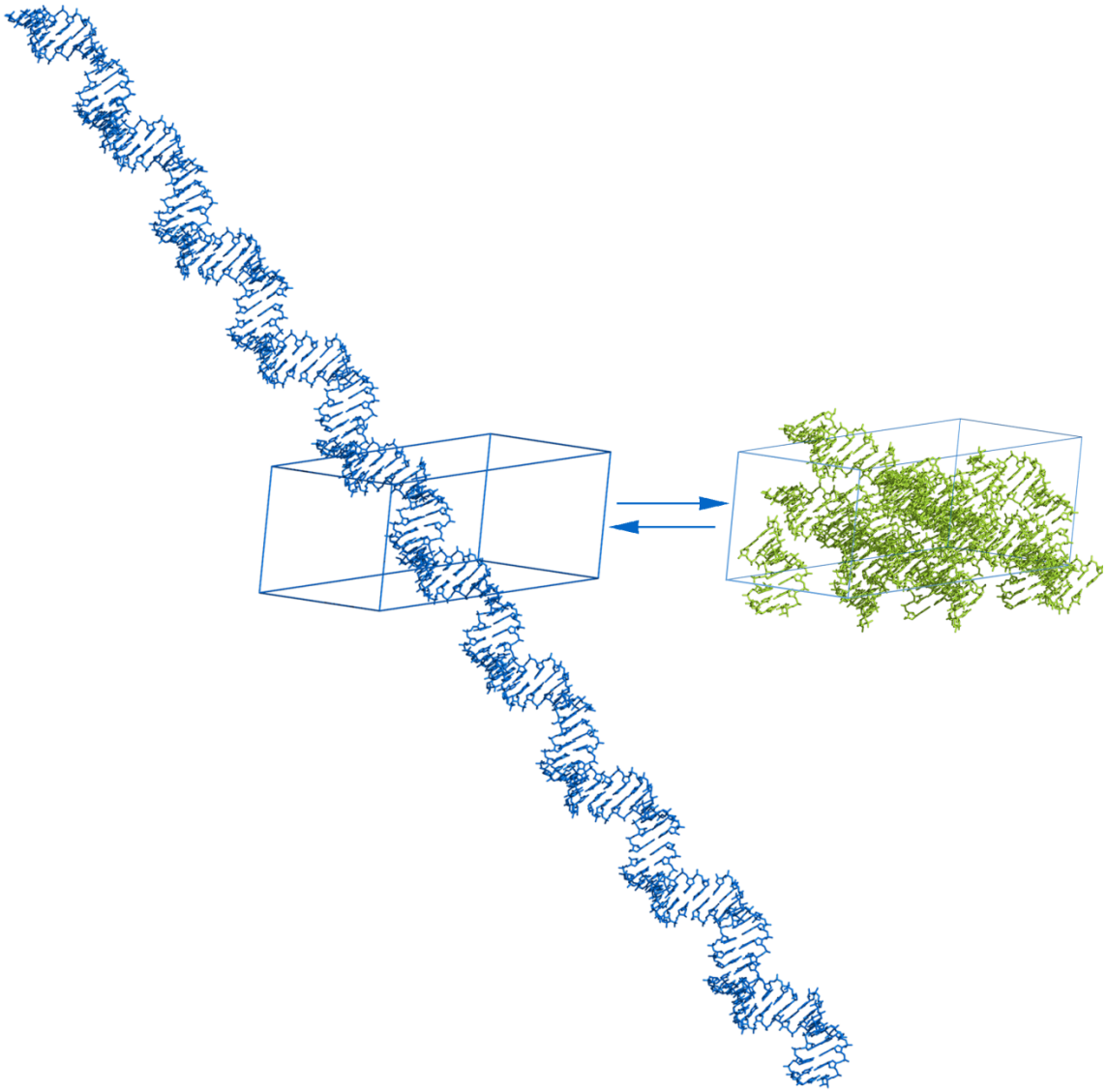
Supplementary Materials

to

Crystal structures of CGG RNA repeats with implications for fragile X-associated tremor ataxia syndrome

Agnieszka Kiliszek, Ryszard Kierzek, Wlodzimierz J. Krzyzosiak, Wojciech Rypniewski

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland



Supplementary Figure 1. The packing of $[GCGGCGGC]_2$ duplexes in the P1 crystal lattice. All the 18 independent molecules stack end-to-end and form semi-infinite parallel columns. The equivalent arrangement shown on the right corresponds more closely to the dimensions of the unit cell.

Supplementary Table 1. Displacement[¥] (Å) for the 18 duplexes in the crystal structure of $(GCGGCGGC)_2$ (a), $[GC(8\text{-}BrG)GCGGCGGC]_2$ (b) and $[GC(8\text{-}BrG)GCGGC]_2$ (c). Multiple conformations are detailed in (b) and (c).

a)

chain	AB	CD	EF	GH	IJ	KL	MN	OP	QR	SY	TU	VZ	WX	ab	cd	ef	gh	ij
G-C	6.4	7.0	7.2	6.5	6.6	6.7	7.4	6.8	7.3	7.1	7.1	6.9	7.0	6.7	6.5	6.6	6.5	6.7
C-G	6.7	6.9	7.0	6.3	6.8	6.7	7.2	6.7	7.0	6.9	6.8	6.7	7.2	7.1	6.7	7.0	6.7	7.0
G-G	6.3	6.8	6.3	6.2	6.2	6.2	6.4	6.3	6.3	6.4	6.4	6.1	6.6	6.6	6.3	6.8	6.5	6.5
G-C	6.5	6.5	6.7	6.8	6.5	6.6	6.6	6.7	6.6	6.6	6.6	6.7	6.9	6.7	6.8	7.1	6.9	6.4
C-G	6.0	7.1	6.9	6.5	6.8	6.8	6.4	6.5	6.4	6.7	6.9	6.8	7.1	6.7	7.0	6.7	6.4	6.6
G-G	6.1	6.8	6.3	6.1	6.4	6.3	6.3	6.4	6.2	6.5	6.4	6.5	6.9	6.4	6.4	6.6	6.2	6.2
G-C	6.9	6.9	6.7	6.8	6.7	6.4	6.8	6.8	6.7	7.0	6.9	6.7	7.2	6.9	6.8	7.1	6.8	6.6
C-G	6.6	6.5	6.4	6.7	6.2	6.1	6.2	6.8	6.2	6.5	7.0	6.7	6.7	7.0	6.6	6.9	6.8	6.4
average	6.4	6.8	6.7	6.5	6.5	6.5	6.7	6.6	6.6	6.7	6.8	6.6	7.0	6.8	6.6	6.8	6.6	6.5
s.d.	0.3	0.2	0.3	0.3	0.3	0.3	0.4	0.2	0.4	0.3	0.3	0.2	0.2	0.2	0.2	0.2	0.2	0.2

b)

chain	AB	CD	EE'
G-C	7.9 / 7.5	7.7 / 7.6 / 6.8 / 6.7	8.1 / 8.0 / 8.4 / 8.2
C-G	8.2 / 7.7	8.0 / 7.9 / 7.4 / 7.2	7.8 / 7.7 / 8.3 / 8.2
GBr-G	7.8 / 7.9 / 7.3 / 7.3	7.3 / 7.1 / 6.8 / 6.6	7.2 / 7.0 / 7.7 / 7.5
G-C	7.6 / 7.7 / 7.3 / 7.4	7.0 / 6.9 / 6.6 / 6.5	8.3 / 8.2 / 8.6 / 8.5
C-G	6.4 / 6.5 / 6.4 / 6.5	6.0 / 5.4 / 6.0 / 5.4	8.0 / 8.3 / 8.3 / 8.5
G-G	7.0 / 6.9	5.9 / 5.4 / 5.9 / 5.3	8.3 / 8.6
G-C	7.8 / 7.7 / 7.7 / 7.5	6.5 / 6.6 / 6.4 / 6.5	8.3 / 8.0 / 8.5 / 8.3
C-G	8.1 / 7.9 / 8.0 / 7.8	6.9 / 7.0 / 6.9 / 7.0	8.2 / 8.3 / 8.5 / 8.6
G-GBr	6.5 / 6.3 / 6.5 / 6.2	5.6 / 5.7 / 5.7 / 5.8	7.0 / 7.2 / 7.5 / 7.7
G-C	6.2 / 6.3 / 6.4 / 6.4	5.1 / 5.4	7.7 / 7.8 / 8.2 / 8.3
C-G	6.1 / 5.5 / 6.5 / 5.9	5.0 / 4.9 / 5.3 / 5.3	8.0 / 8.1 / 8.2 / 8.4
average	7.1	6.3	8.1
s.d.	0.7	0.9	0.4

c)

chain	AB
G-C	5.5
C-G	5.8
GBr-G	5.7
G-C	6.2
C-G	6.6
G-GBr	6.0 / 5.7
G-C	6.2
C-G	6.7
average	6.0
s.d.	0.4

Supplementary Table 2. Angle[¥] (°) for the 18 duplexes in the crystal structure of $(GCGGCGGC)_2$ (a), $[GC(8\text{-}BrG)GCGGCGGC]_2$ (b) and $[GC(8\text{-}BrG)GCGGC]_2$ (c). Multiple conformations are detailed in (b) and (c).

a)

chain	AB	CD	EF	GH	IJ	KL	MN	OP	QR	SY	TU	VZ	WX	ab	cd	ef	gh	ij
G-C	6.5	8.8	7.3	6.2	9.6	10.6	5.0	3.6	5.5	4.2	5.7	4.2	6.8	7.5	6.0	5.2	4.2	4.4
C-G	3.8	7.0	5.9	4.3	10.1	10.6	4.7	3.1	5.5	3.8	4.4	4.0	6.2	4.9	5.8	3.8	3.5	3.1
G-G	5.3	7.1	6.8	5.9	8.6	9.0	7.8	4.8	8.2	8.3	6.1	6.2	6.0	5.0	5.9	3.4	4.4	5.7
G-C	4.6	4.3	7.2	5.7	4.3	5.4	8.5	3.7	8.9	5.6	5.3	6.4	3.2	5.9	2.6	1.8	2.7	6.2
C-G	5.8	4.6	6.6	6.5	3.6	4.3	6.9	3.8	7.7	7.3	5.1	5.3	3.0	3.9	1.8	2.6	5.5	4.6
G-G	6.2	7.7	8.0	5.8	6.0	7.2	7.0	5.1	8.0	7.4	5.7	6.8	5.2	6.5	5.6	4.6	5.1	5.2
G-C	2.3	7.9	7.6	3.6	7.2	9.5	4.6	2.7	5.2	5.4	3.1	4.4	5.2	3.3	6.3	5.1	3.4	5.9
C-G	5.2	7.3	10.2	5.3	9.3	10.6	7.9	3.5	8.4	8.8	4.4	7.3	8.1	3.2	7.3	5.4	3.7	6.9
Average	5.0	6.8	7.4	5.4	7.3	8.4	6.5	3.8	7.2	6.4	5.0	5.6	5.5	5.0	5.2	4.0	4.0	5.2
s.d.	1.4	1.6	1.3	1.0	2.5	2.5	1.5	0.8	1.5	1.9	1.0	1.3	1.7	1.6	1.9	1.3	0.9	1.2

b)

chain	AB	CD	EE'
G-C	2.2 / 2.6 / 5.0 / 5.5	0.9 / 1.0 / 3.1 / 3.2	9.1 / 8.9 / 5.6 / 5.5
C-G	2.3 / 2.3 / 7.4 / 7.6	0.9 / 0.9 / 3.8 / 3.8	14.1 / 13.9 / 8.3 / 8.1
GBr-G	6.2 / 6.0 / 10.2 / 10.1	4.0 / 4.0 / 6.5 / 6.4	18.9 / 18.6 / 12.9 / 12.7
G-C	7.4 / 6.9 / 6.7 / 6.2	4.6 / 4.5 / 4.5 / 4.4	10.6 / 10.4 / 8.9 / 8.7
C-G	10.0 / 9.5 / 9.5 / 8.8	6.2 / 7.8 / 5.8 / 7.5	6.9 / 6.1 / 6.1 / 5.2
G-G	7.2 / 6.7 / 7.3 / 6.7	6.3 / 6.8 / 6.4 / 7.0	7.0 / 6.0
G-C	1.3 / 0.9 / 2.1 / 1.6	2.5 / 2.4 / 3.1 / 3.1	6.1 / 6.9 / 5.2 / 6.1
C-G	2.3 / 2.2 / 3.6 / 3.5	2.0 / 2.0 / 2.9 / 3.0	10.4 / 10.6 / 8.7 / 8.9
G-GBr	8.9 / 9.0 / 10.3 / 10.4	6.7 / 7.7	18.6 / 18.9 / 12.7 / 12.9
G-C	10.9 / 10.4 / 12.1 / 11.7	9.5 / 9.6 / 10.2 / 10.3	13.9 / 14.1 / 8.1 / 8.3
C-G	12.5 / 11.5 / 13.2 / 12.1	10.3 / 10.4 / 10.6 / 10.7	8.9 / 9.1 / 5.5 / 5.6
Average	7.1	5.5	9.7
s.d.	3.6	3.0	4.0

c)

chain	AB
G-C	13.2 / 13.3
C-G	12.9 / 12.8
GBr-G	13.5 / 13.2
G-C	9.0 / 8.7
C-G	6.0 / 5.8
G-GBr	10.8 / 10.4
G-C	10.0
C-G	9.5 / 9.7
average	10.5
s.d.	2.5

Supplementary Table 3. Helical twist[¥] (°) for the 18 duplexes in the crystal structure of $(GCGGCGGC)_2$ (a), $[GC(8\text{-}BrG)GCGGCGGC]_2$ (b) and $[GC(8\text{-}BrG)GCGGC]_2$ (c). Multiple conformations are detailed in (b) and (c).

a)

chain	AB	CD	EF	GH	IJ	KL	MN	OP	QR	SY	TU	VZ	WX	ab	cd	ef	gh	ij
G-C	31.0	34.8	29.6	32.3	29.7	28.8	30.5	32.1	30.7	31.8	31.3	33.6	31.8	29.3	30.3	30.9	32.0	31.3
C-G	29.4	31.1	27.6	33.5	32.4	33.6	27.2	25.3	26.6	27.8	25.6	27.2	33.8	29.3	33.2	33.5	26.4	27.4
G-G	33.5	26.2	32.7	26.8	27.3	28.2	32.6	33.8	33.5	32.8	33.1	32.9	25.8	30.9	28.4	27.0	35.4	32.7
G-C	30.7	34.1	31.1	33.4	31.8	32.7	29.5	31.9	30.8	30.6	32.1	31.7	27.2	32.0	29.5	27.7	32.1	31.9
C-G	32.7	25.6	29.5	32.7	28.1	29.6	34.8	32.4	34.6	32.3	34.5	32.6	27.9	32.8	27.8	25.3	30.1	29.6
G-G	27.4	32.2	32.8	26.9	34.0	33.7	28.2	26.1	28.8	30.3	24.8	26.7	33.3	27.5	34.1	33.5	28.3	32.6
G-C	30.1	31.1	28.0	31.6	29.1	28.2	32.6	31.3	32.4	29.9	30.7	31.1	30.6	30.8	28.3	31.7	31.2	32.4
C-G	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
average	30.7	30.7	30.2	31.0	30.3	30.7	30.8	30.4	31.1	30.8	30.3	30.8	30.1	30.4	30.2	29.9	30.8	31.1
s.d.	2.1	3.6	2.1	2.9	2.4	2.5	2.7	3.3	2.8	1.7	3.7	2.8	3.1	1.8	2.5	3.3	2.9	2.0

b)

chain	AB	CD	EE'
G-C	35.1	27.8 / 29.9	32.0 / 34.7
C-G	32.4 / 30.8	31.8 / 33.2	36.8 / 33.7
GBr-G	25.3 / 27.8	29.1 / 29.3	20.4 / 16.7
G-C	32.1 / 35.8	26.9 / 33.5 / 31.0 / 37.6	35.7 / 32.2 / 33.4 / 30.0
C-G	27.1	28.03 / 28.6	17.7 / 31.9
G-G	27.0	34.5 / 27.2	31.9 / 17.7
G-C	33.9	33.6	32.2 / 35.7 / 30.0 / 33.4
C-G	21.8	28.5	16.7 / 20.4
G-GBr	36.1 / 33.1	35.1	33.7 / 36.8
G-C	30.5 / 39.2	28.6	34.7 / 32.0
C-G	---	---	---
average	30.2	30.8	29
s.d.	4.8	2.9	7.3

c)

chain	AB
G-C	32.6
C-G	36.7
GBr-G	27.6
G-C	34.0 / 33.9
C-G	28.0 / 25.5
G-GBr	33.1 / 35.6
G-C	33.8
C-G	---
average	32.4
s.d.	3.8

Supplementary Table 4. Rise[¥] (Å) for the 18 duplexes in the crystal structure of $(GCGGCGGC)_2$ (a), $[GC(8\text{-}BrG)GCGGCGGC]_2$ (b) and $[GC(8\text{-}BrG)GCGGC]_2$ (c). Multiple conformations are detailed in (b) and (c).

a)

chain	AB	CD	EF	GH	IJ	KL	MN	OP	QR	SY	TU	VZ	WX	ab	cd	ef	gh	ij
G-C	3.1	3.1	3.0	3.2	3.4	3.3	3.1	3.1	3.0	3.1	3.2	3.2	3.4	3.0	3.3	3.4	3.0	3.1
C-G	3.0	3.0	2.8	2.7	3.1	3.0	2.8	3.0	2.7	2.9	2.9	2.7	3.1	3.0	3.1	3.2	3.1	2.9
G-G	3.1	2.8	2.9	3.1	2.7	2.7	3.1	3.2	3.1	2.8	3.0	2.9	2.9	3.0	2.9	3.0	3.1	3.1
G-C	3.4	3.0	3.4	3.5	3.1	3.0	3.5	3.5	3.5	3.3	3.5	3.6	3.2	3.3	3.2	3.2	3.6	3.2
C-G	2.9	2.8	2.8	3.0	2.9	2.8	3.0	3.1	2.9	3.1	3.1	3.1	2.9	3.0	2.9	3.0	2.9	3.0
G-G	3.0	3.1	2.8	2.9	3.0	3.0	3.0	3.0	2.9	2.9	3.0	2.8	3.1	3.0	3.1	3.2	2.9	3.0
G-C	3.2	3.1	3.2	3.1	3.2	3.2	2.9	3.2	2.9	2.9	3.0	3.0	3.3	3.1	3.3	3.4	3.2	3.1
C-G	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
average	3.1	3.0	3.0	3.1	3.0	3.0	3.0	3.1	3.0	3.0	3.1	3.1	3.1	3.1	3.1	3.2	3.1	3.1
s.d.	0.2	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.2	0.1	0.2	0.2	0.2	0.1

b)

chain	AB	CD	EE'
G-C	3.1 / 3.0 / 3.4 / 3.4	3.3 / 3.3 / 3.4 / 3.4	3.0 / 3.1 / 2.8 / 2.8
C-G	3.2	3.3	2.5 / 2.6
GBr-G	3.0 / 2.6	3.1 / 3.1 / 2.8 / 2.8	2.3 / 2.3 / 2.7 / 2.8
G-C	3.0	3.0 / 2.8 / 3.0 / 2.8	2.8 / 2.9 / 3.0 / 3.1
C-G	2.8 / 2.8 / 2.8 / 2.9	2.7	3.1 / 2.7 / 3.2 / 2.8
G-G	2.8	2.9 / 3.1 / 2.9 / 3.1	2.7 / 3.1 / 2.8 / 3.2
G-C	3.3 / 3.2	3.5	2.9 / 2.8 / 3.1 / 3.0
C-G	3.0 / 2.9	3.1	2.3 / 2.3 / 2.8 / 2.7
G-GBr	2.6 / 2.8 / 2.6 / 2.7	2.9 / 2.8 / 2.8 / 2.8	2.5 / 2.5 / 2.6 / 2.6
G-C	3.0 / 2.9	2.7	3.1 / 3.0 / 2.8 / 2.8
C-G	--- --- --- ---	--- --- --- ---	--- --- --- ---
average	3.0	3.0	2.8
s.d.	0.2	0.3	0.3

c)

chain	AB
G-C	2.9
C-G	3.0
GBr-G	2.6 /2.7
G-C	2.9
C-G	2.9
G-GBr	2.5
G-C	2.9
C-G	---
average	2.8
s.d.	0.2

[¥] Definitions of displacement, angle, twist and rise according to Dickerson *et al.* J. Mol. Biol. (1989) 205, 787-791.

Supplementary Table 5. The distribution of syn-anti conformation along the duplexes in the $(GCGGCGGC)_2$ (a), $[GC(8\text{-}BrG)GCGGCGGC]_2$ (b) and $[GC(8\text{-}BrG)GCGGC]_2$ (c).

a)

chain	AB	CD	EF	GH	IJ	KL	MN	OP	QR	SY	TU	VZ	WX	ab	cd	ef	gh	ij
G-C C-G																		
G-G	anti-syn	syn-anti	anti-syn	disorder	syn-anti	syn-anti	anti-syn	anti-syn	anti-syn	anti-syn	anti-syn	anti-syn	syn-anti	anti-syn	syn-anti	syn-anti	anti-syn	anti-syn
G-C C-G																		
G-G	syn-anti	anti-syn	anti-syn	syn-anti	anti-syn	anti-syn	syn-anti	syn-anti	syn-anti	syn-anti	syn-anti	syn-anti	anti-syn	syn-anti	anti-syn	anti-syn	disorder	anti-syn
G-C C-G																		

b)

chain	AB	CD	EE'
G-C C-G			
GBr-G	syn-anti	syn-anti	syn-anti
G-C C-G			
G-G	anti-syn	disorder	disorder
G-C C-G			
G-GBr	anti-syn	anti-syn	anti-syn
G-C C-G			

c)

chain	AB
G-C C-G	
GBr-G	syn-anti
G-C C-G	
G-GBr	anti-syn
G-C C-G	

Supplementary Table 6. Comparison of characteristic features of double helical CGG, CAG (1) and CUG (2) repeats.

Feature	CGG	CAG	CUG
Helix form	A	A	A
Helical twist* (°)	30-32	28.5 ± 5.7	33.6 ± 4.1
Major groove width† (Å)	17.9 ± 0.9,17.8 ± 2.5, 14.3 (three models)	23.8 ± 0.2	12.7 ± 2.3
Minor groove width† (Å)	15.8 ± 0.5, 15.4 ± 0.5, 16.1	15.3 ± 0.1	15.7 ± 0.4
Average C1'-C1' distance for N-N§ (and for the other pairs) (Å)	11.3 (10.7)	11.0 (10.7)	10.4 (10.5)
Local effect of N-N on helicity	local unwinding, compensated elsewhere along the duplex	unwinding	not observed
N-N pairing interaction	N1H···O6 and *exo*-N2H···N7 and *intra*-molecular: *exo*-N2H···O2	C2-H2···N1 hydrogen bond	N3-H3···O4 hydrogen bond
Manner of accommodating N-N (according to direction of the glycosidic bond, λ)	One G is in *syn* conformation (inclined towards the major groove)	One A turned towards major groove ("thumbs-up")	One U inclined towards minor groove
Effect of N-N conformation on neighbouring N-N	Symmetric arrangements are clearly favoured	Co-operativity: A-A pairs in consecutive repeats have alternative conformations	No effect: each U-U takes one of two possible conformation independently
Electrostatic profile	Alternating stripes of positive and negative potential due to C-G pairs	Alternating stripes of positive and negative potential due to C-G pairs	Alternating stripes of positive and negative potential due to C-G pairs
Observed ligand affinity of N-N	Sulphate or Ca^{2+} binding in major groove	Sulphate binding in major groove	Sulphate or glycerol through-water binding in major groove
Exposed functional groups of N-N:			
- major groove	G(*anti*) O6 carbonyl G(*syn*) O6 carbonyl, N1 amino, N2 *exo*-amino	First A N6 amino Second A N1 imino, N6 amino	First U O4 carbonyl
- minor groove	G(*anti*) N3 imino, N2 *exo*-amino	First A N3 imino Second A N3 imino	First U O2 carbonyl Second U O2 carbonyl, N3 amino

* Average for A-RNA is 33.1 (3).
† The values given are the "refined" widths, according to the program 3DNA (4).
§ G-G or A-A or U-U base pair.

References:

1. Kiliszek, A., Kierzek, R., Krzyzosiak, W.J. and Rypniewski, W. (2009) Structural insights into CUG repeats containing the 'stretched U-U wobble': implications for myotonic dystrophy. *Nucleic Acids Res.*, 37, 4149-4156.

2. Kiliszek, A., Kierzek, R., Krzyzosiak, W.J. and Rypniewski, W. (2010) Atomic resolution structure of CAG RNA repeats: structural insights and implications for the trinucleotide repeat expansion diseases. *Nucleic Acids Res.*, 38, 8370-8376.

3. Bloomfield,V.A., Crothers,D.M. and Tinoco,I. (2000) Nucleic Acids: Structures, Properties, and Functions. University Science Books, Sausalito.

4. Olson, W.K., Bansal, M., Burley, S.K., Dickerson, R.E., Gerstein, M., Harvey, S.C., Heinemann, U., Lu, X.J., Neidle, S., Shakked, Z. *et al.* (2001) A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.*, 313, 229-237.