



BIOLOGICAL
CRYSTALLOGRAPHY

Volume 71 (2015)

Supporting information for article:

The novel double-folded structure of d(GCATGCATGC): a possible model for triplet-repeat sequences

Arunachalam Thirugnanasambandam, Selvam Karthik, Pradeep Kumar Mandal and Namasivayam Gautham

Table S1 Torsion angles of the phosphate-sugar backbone and sugar puckering. Large deviations from the ‘standard’ A or B models are shown in **bold**

Backbone torsion (°)	G1	C2	A3	T4	G5	C6	A7	T8	G9	C10	Ideal B-type	Ideal A-type
α	--	-68.9 (<i>g</i> ⁻)	-57.5 (<i>g</i> ⁻)	-145.6 (t)	-72.2 (<i>g</i> ⁻)	-54.0 (<i>g</i> ⁻)	-56.8 (<i>g</i> ⁻)	68.1 (g⁺)	-77.8 (<i>g</i> ⁻)	65.7 (g⁺)	-28.9 (<i>g</i> ⁻)	-78.4 (<i>g</i> ⁻)
β	--	169.8 (<i>t</i>)	145.2 (<i>t</i>)	139.0 (<i>t</i>)	-128.4 (<i>t</i>)	161.7 (<i>t</i>)	145.7 (<i>t</i>)	-143.8 (<i>t</i>)	170.5 (<i>t</i>)	112.3 (g⁺)	146.8 (<i>t</i>)	-171.1 (<i>t</i>)
γ	47.0 (<i>g</i> ⁺)	52.7 (<i>g</i> ⁺)	49.5 (<i>g</i> ⁺)	-45.4 (g⁻)	63.1 (<i>g</i> ⁺)	42.6 (<i>g</i> ⁺)	50.6 (<i>g</i> ⁺)	59.8 (<i>g</i> ⁺)	40.0 (<i>g</i> ⁺)	164.5 (t)	23.6 (<i>g</i> ⁺)	63.9 (<i>g</i> ⁺)
δ	148.9 (<i>t</i>)	104.9 (<i>g</i> ⁺)	143.8 (<i>t</i>)	163.1 (<i>t</i>)	146.0 (<i>t</i>)	109.9 (<i>g</i> ⁺)	139.9 (<i>t</i>)	142.9 (<i>t</i>)	87.0 (<i>g</i> ⁺)	87.2 (<i>g</i> ⁺)	150.7 (<i>t</i>)	80.6 (<i>g</i> ⁺)
ε	-177.2 (<i>t</i>)	-125.2 (<i>t</i>)	-134.7 (<i>t</i>)	-121.9 (<i>t</i>)	-166.8 (<i>t</i>)	-122.4 (<i>t</i>)	-78.7 (g⁻)	-143.5 (<i>t</i>)	-144.5 (<i>t</i>)	--	-142.8 (<i>t</i>)	-145.4 (<i>t</i>)
ζ	-90.0 (<i>g</i> ⁻)	150.9 (<i>t</i>)	122.3 (<i>t</i>)	-48.0 (<i>g</i> ⁻)	-106.8 (<i>g</i> ⁻)	136.0 (<i>t</i>)	-131.5 (<i>t</i>)	-83.9 (<i>g</i> ⁻)	55.8 (g⁺)	--	-160.7 (<i>t</i>)	-78.7 (<i>g</i> ⁻)
χ (glycosidic torsion)	-97.72	-123.05	-154.88	-96.28	-99.62	-117.62	-160.41	-115.72	-87.73	-124.74	-92.0	-174.2
sugar torsion (°)												
v0	-2.41	-36.52	-22.25	0.68	-15.15	-39.93	-21.11	-18.5	-9.84	-12.52	-28.2	11
v1	21.11	25.67	35.81	23.99	30.94	32.28	33.04	32.31	-12.42	-9.09	44.4	-33
v2	-30.67	-6.38	-35.07	-38.02	-34.19	-13.33	-31.84	-33.23	28.31	25.68	-43	41.2
v3	30.16	-14.28	23.21	39.53	26.32	-9.06	20.53	23.44	-34.83	-33.65	28.2	-35.1
v4	-17.68	31.97	-0.84	-25.5	-7.25	30.68	0.13	-3.33	28.39	29.34	-0.2	15.7
P	194.45	100.28	162.51	199.62	173.22	110.12	161	166.77	34.84	40.04	161.51	3.1
τm	31.67	35.79	36.77	40.37	34.43	38.74	33.67	34.14	34.49	33.54	45.3	41.23
Puckering	C3'-exo	O4'-endo	C2'-endo	C3'-exo	C2'-endo	C1'-exo	C2'-endo	C2'-endo	C3'-endo	C4'-exo	C2'-endo	C3'-endo

Table S2 Inter- and intra-molecular stacking interactions, calculated using the distance restraints server <http://rns.ucsc.edu/pdbrestraints/pdbtorestraints.html> (Laurberg *et al.*, 2008).

Bases	Distance (Å)
G1-C2	3.44
C2-A7*	3.78
A3-C6*	3.81
A3-G5	3.09
G5-C6	3.41
C6-A3*	3.81
A7-G1	3.10
A7-C2*	3.78
T8-G9	4.00
C10-A7	3.59

*symmetry mate [-x+1,-y+1,z]

Table S3 Comparison of ‘bi-loop’ region of the present decamer with previously reported structures.

Groups	Sequence	PDB ID	Metal ion	Bases superposed		Backbone atoms	Minor groove tetrad	Linear/ Cyclic	References
				Previously reported structures	Present structure				
X-ray									
1	d(GCATGCT)	1QZL	Co ²⁺	1-6	5-10	0.36	G:C:G:C	Linear	-
	d(GCATGCT)	1R2O	Ni ²⁺	1-6	5-10	0.39	G:C:G:C	Linear	-
2	d(GCATGCT)	1QYL	V ³⁺	1-7	1-7	0.80	G:C:G:C	Linear	-
	d(GCATGCT)	1QYK	Ba ²⁺	1-7	1-7	0.84	G:C:G:C	Linear	-
	d(GCATGCT)	3T86	Ca ²⁺	1-7	1-7	0.72	G:C:G:C	Linear	-
	d(GCATGCT)	1MF5	[Co(NH ₃) ₆] ³⁺	1-7	1-7	0.78	G:C:G:C	Linear	(Thorpe <i>et al.</i> , 2003)
	d(GCATGCT)	184D	Mg ²⁺	1-7	1-7	0.75	G:C:G:C	Linear	(Leonard <i>et al.</i> , 1995)
3	d<ATTCCATTC>	284D	Na ²⁺ , Ba ²⁺	1-7	1-7	0.60	A:T:A:T	Cyclic	(Salisbury <i>et al.</i> , 1997)
NMR									
4	d(TCGTTGCT)	2K8Z	-	2-8	1-7	1.17	C:G:G:C	Linear	(Viladoms <i>et al.</i> , 2009)
	d(TGCTTCGT)	2K90	-	2-8	1-7	1.36	G:C:C:G	Linear	(Viladoms <i>et al.</i> , 2009)
5	d<TGCTCGCT>	1EU2	-	2-8	1-7	0.59	G:C:G:C	Cyclic	(Escaya <i>et al.</i> , 2000)
	d<CATTCCATT>	1EU6	-	2-8	1-7	0.84	A:T:A:T	Cyclic	(Escaya <i>et al.</i> , 2000)

d<CGCTCATT>	1N96	-	2-8	1-7	1.67	G:C:A:T	Cyclic	(Escaja <i>et al.</i> , 2003)
d<CCGTCCGT>	2HK4	-	2-8	1-7	1.07	C:G:C:G	Cyclic	(Escaja <i>et al.</i> , 2007)
d<CGCTCCGT>	2K97	-	2-8	1-7	0.97	G:C:C:G	Cyclic	(Viladoms <i>et al.</i> , 2009)

Table S4 DNA-water interactions.

DNA	Water molecule	Distance (Å)
Hydrogen bonds with unpaired bases.		
A3(N1)	O24	2.81
T8(O4)	O12	2.91
G9(N1)	O12	2.76
C10(N4)	O18	3.30
C10(N3)	O18	3.16
C10(N3)	O13B	3.19
C10(O2)	O13B	2.79
C10(O2)	O13A	2.91
Hydrogen bonds with other DNA atoms		
C2(N4)	O7	3.02
C10(O3')	O5	3.26
C10(O3')	O8	3.37
G1(N7)	O1	3.09
G1(O6)	O1	2.42
G9(N7)	O1	2.99
G1(N7)	O2	3.02
T8(OP2)	O2	2.67
G9(N7)	O2	3.07
G9(O6)	O2	2.73
G1(N7)	O3	3.19
A7(O3')	O3	2.73
G9(N7)	O3	3.19
G1(O6)	O4	3.63
C10(OP2)	O4	2.64
G9(OP2)	O5	2.46
C10(O3')	O5	3.26
G1(N7)	O7	3.25
G9(N7)	O7	3.34
G9(O6)	O7	3.33
C10(O3')	O8	3.33
C10(OP1)	O8	2.78
A7(N3)	O10	2.59
G9(O5')	O10	3.30
G9(OP1)	O10	2.80
G5(N7)	O11	3.14
G5(OP2)	O11	2.41

A7(N6)	O13B	3.40
A3(O3')	O14	2.83
G5(N7)	O14	3.25
A3(N3)	O15	2.87
G5(N7)	O16	2.98
G5(N7)	O17	3.42
G5(O6)	O17	2.28
G9(OP1)	O19	2.74

Table S5 Previously reported triplet repeat sequence structures.

S.No	PDB ID	Repeat	Sequence	Conformation	Reference
RNA					
1	4YN6	CAG	UUGGGCCAGCAGCAGGUCC	X-ray A-duplex	-
2	2MS5	CAG	CCGCAGCGG	NMR A-duplex	-
3	3NJ6	CAG	GGCAGCAGCC	X-ray A-duplex	(Kiliszek <i>et al.</i> , 2010)
4	3NJ7	CAG	GGCAGCAGCC	X-ray A-duplex	(Kiliszek <i>et al.</i> , 2010)
5	4J50	CAG	UUGGGCCAGCAGCAGGUCC	X-ray A-duplex	(Yildirim <i>et al.</i> , 2013)
6	1ZEV	CUG	CUGCUGCUGCUGCUG	X-ray A-duplex	(Mooers <i>et al.</i> , 2005)
7	3GM7	CUG	CUGCUGCUGCUGCUG	X-ray A-duplex	(Kiliszek <i>et al.</i> , 2009)
8	3GLP	CUG	GCUGCUGC	X-ray A-duplex	(Kiliszek <i>et al.</i> , 2009)
9	4FNJ	CUG	CUGCUGCUAAGGCAUGAAAGUGCUAUGCUCUGCUG	X-ray A-duplex	(Coonrod <i>et al.</i> , 2012)
10	4E48	CUG	GCUGCUGCUGCUGCUGCUG	X-ray A-duplex	(Tamjar <i>et al.</i> , 2012)
11	3SYW	CUG	UUGGGCCUGCUGCUGGUCC	X-ray A-duplex	(Kumar <i>et al.</i> , 2011)
12	3SZX	CUG	UUGGGCCUGCUGCUGGUCC	X-ray A-duplex	(Kumar <i>et al.</i> , 2011)
13	4J39	CAG/CUG	CAGCAGCAGCCUGCUGCUG + protein	X-ray A-duplex	-
14	4E5C	CGG/CUG	GGCGGGGGCGGCUGCUGGCC	X-ray A-duplex	(Tamjar <i>et al.</i> , 2012)
15	4E6B	CGG/CUG	CGGCGGGGCCUGCUGCUG	X-ray A-duplex	(Tamjar <i>et al.</i> , 2012)
16	4J5V	CAG/CCG	CAGCAGCAGCCCCGCCGCG + protein	X-ray A-duplex	-
17	2RSK	GGA	GGAGGAGGAGGA	NMR Quadruplex	(Mashima <i>et al.</i> , 2013)
DNA					
18	1MYQ	GGA	GGAGGAGGAGGA	NMR Quadruplex	(Matsugami <i>et al.</i> , 2001)
19	1OZ8	GGA	GGAGGAGGAGGAGGAGGAGGAGGA	NMR Quadruplex	(Matsugami <i>et al.</i> , 2003)
20	1B3P	GGA	GGAGGAT	NMR Arrowhead (folded)	(Kettani <i>et al.</i> , 1999)
21	4PZQ	CCG	TCCGCCGCCGA	X-ray i-motif	(Chen <i>et al.</i> , 2014)

22	1NOQ	CCG	CCGCCG	NMR	e-motif (duplex C-flipped)	(Zheng <i>et al.</i> , 1996)
23	1A6H	CGG	GCGGTTGCGG	NMR	Quadruplex	(Kettani <i>et al.</i> , 1995)
24	1NP5	GAC	GACGACGAC	NMR	Parallel duplex	(Zheng <i>et al.</i> , 1996)
25	2MJJ	GGG/GCG	GGGAGCGAGGGAGCG	NMR	Tetraplex	(Kocman & Plavec, 2014)

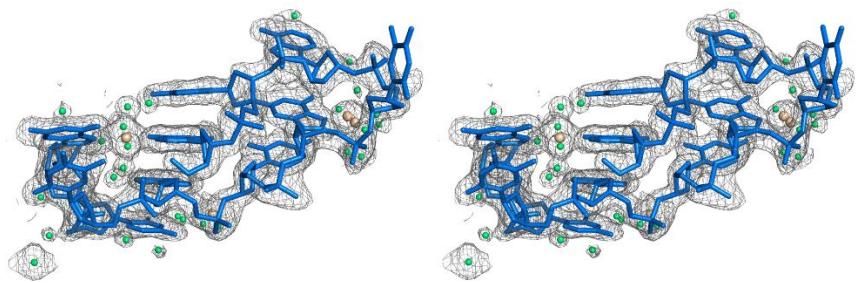


Figure S1 Fully refined structure, shown in electron density, $(2\text{Fo}-\text{Fc})$ map contoured at 1σ .

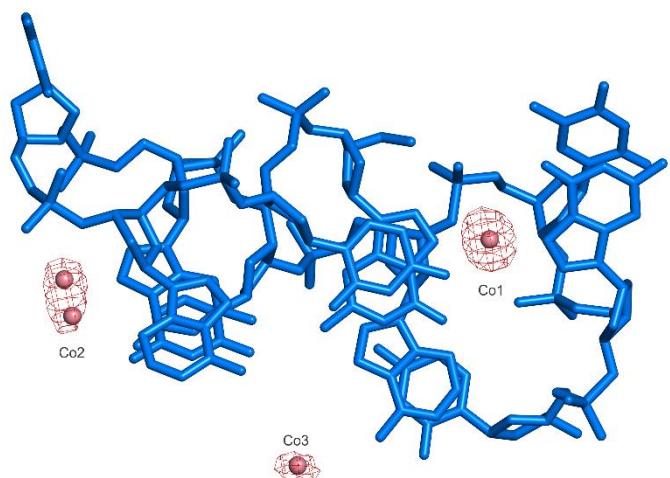


Figure S2 Anomalous difference Fourier map around the Co^{2+} ions, contoured at 5σ .

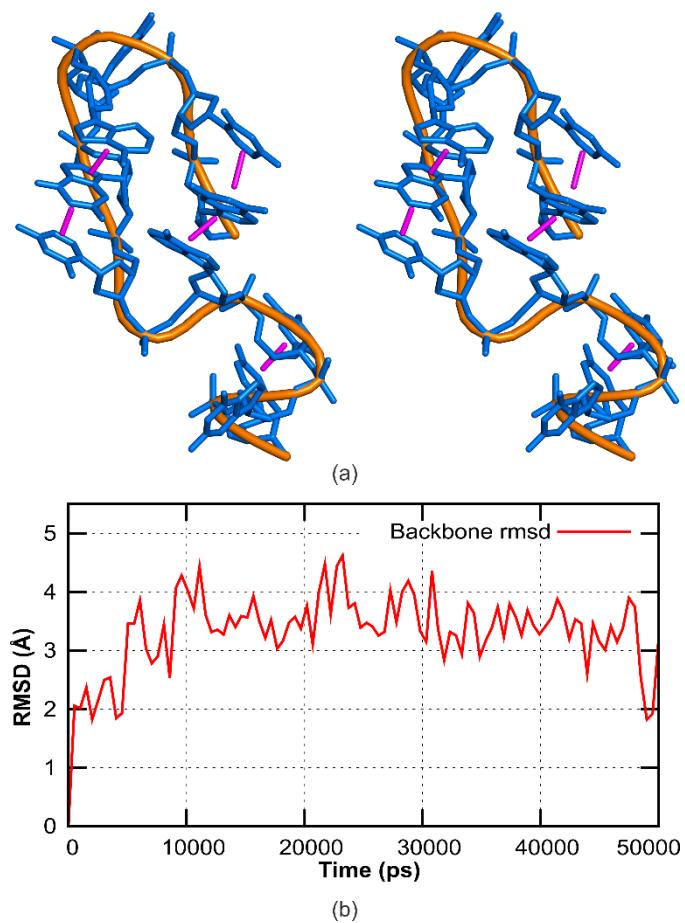


Figure S3 The structure of the single strand double-folded decamer, without the Co^{2+} ions, Co at the end of 50 ns MD simulation. **a)** Stereo-view showing most of the intra-strand interactions intact. **b)** Trajectory of the backbone RMSD.

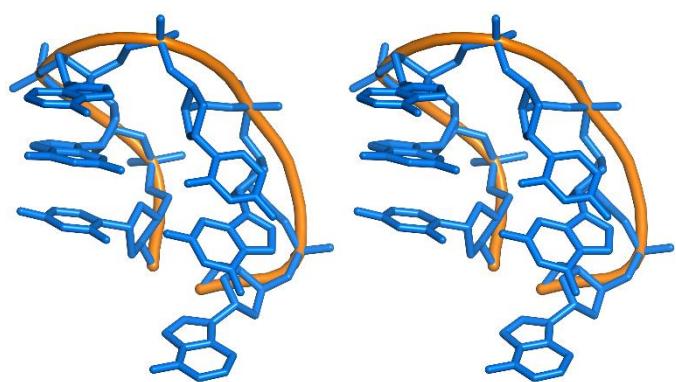


Figure S4 Stereo-view after energy minimization of the single strand model for d(AGCAGC) built from the double-folded decamer.

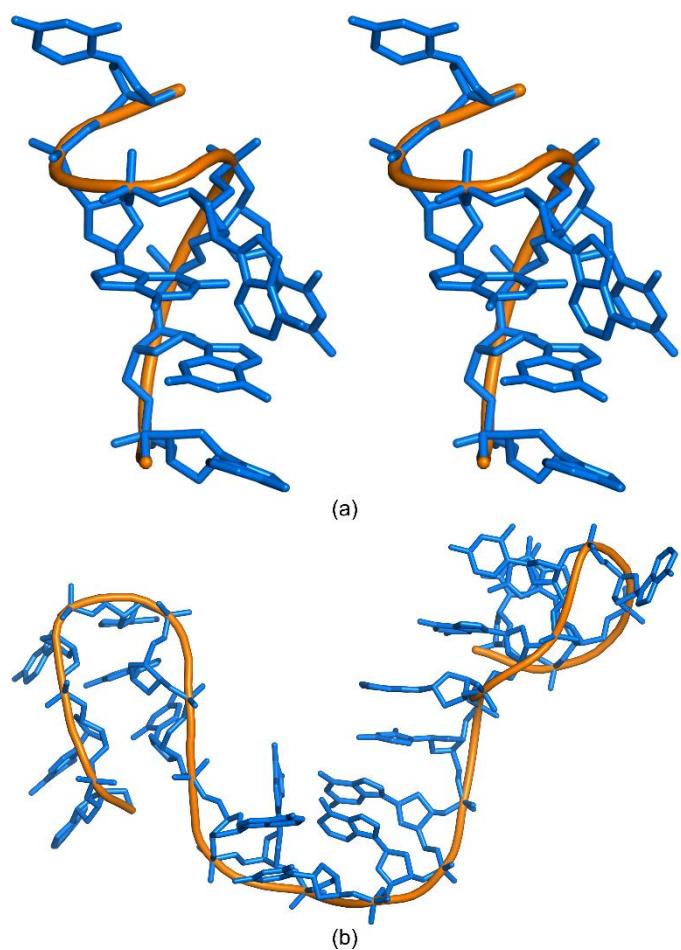


Figure S5 The model structures at the end of 50 ns MD simulations. **a)** Single strand 6mer d(AGC)₂ and **b)** Single strand 18mer d(AGC)₆.

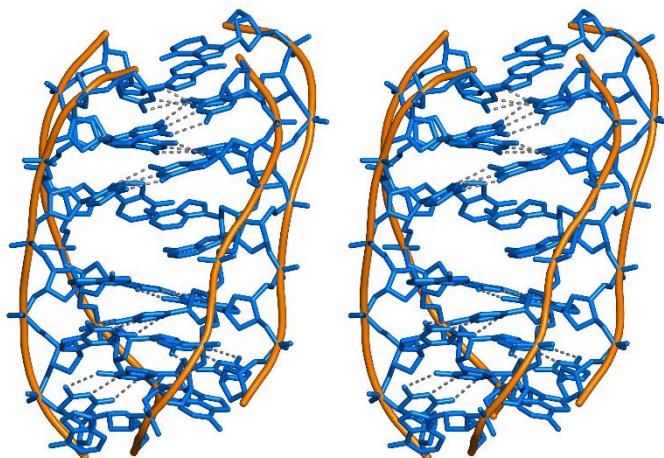


Figure S6 The structure of the tetraplex model at the end of 50 ns MD simulations. Most of the inter- and intra-strand interactions are retained.

References

- Chen, Y.-W., Jhan, C.-R., Neidle, S. & Hou, M.-H. (2014). *Angewandte Chemie International Edition* **53**, 10682-10686.
- Coonrod, L. A., Lohman, J. R. & Berglund, J. A. (2012). *Biochemistry* **51**, 8330-8337.
- Escaya, N., Gelpi, J. L., Orozco, M., Rico, M., Pedroso, E. & Gonzalez, C. (2003). *Journal of the American Chemical Society* **125**, 5654-5662.
- Escaya, N., Gomez-Pinto, I., Pedroso, E. & Gonzalez, C. (2007). *Journal of the American Chemical Society* **129**, 2004-2014.
- Escaya, N., Pedroso, E., Rico, M. & González, C. (2000). *Journal of the American Chemical Society* **122**, 12732-12742.
- Kettani, A., Bouaziz, S., Skripkin, E., Majumdar, A., Wang, W., Jones, R. A. & Patel, D. J. (1999). *Structure* **7**, 803-815.
- Kettani, A., Kumar, A. R. & Patel, D. J. (1995). *Journal of Molecular Biology* **254**, 638-656.
- Kiliszek, A., Kierzek, R., Krzyzosiak, W. J. & Rypniewski, W. (2009). *Nucleic acids research* **37**, 4149-4156.
- Kiliszek, A., Kierzek, R., Krzyzosiak, W. J. & Rypniewski, W. (2010). *Nucleic acids research* **38**, 8370-8376.
- Kocman, V. & Plavec, J. (2014). *Nature communications* **5**, 5831.
- Kumar, A., Park, H., Fang, P., Parkesh, R., Guo, M., Nettles, K. W. & Disney, M. D. (2011). *Biochemistry* **50**, 9928-9935.
- Laurberg, M., Asahara, H., Korostelev, A., Zhu, J., Trakhanov, S. & Noller, H. F. (2008). *Nature* **454**, 852-857.
- Leonard, G. A., Zhang, S., Peterson, M. R., Harrop, S. J., Helliwell, J. R., Cruse, W. B. T., Langlois d'Estaintot, B., Kennard, O., Brown, T. & Hunter, W. N. (1995). *Structure* **3**, 335-340.
- Mashima, T., Nishikawa, F., Kamatari, Y. O., Fujiwara, H., Saimura, M., Nagata, T., Kodaki, T., Nishikawa, S., Kuwata, K. & Katahira, M. (2013). *Nucleic acids research* **41**, 1355-1362.
- Matsugami, A., Okuzumi, T., Uesugi, S. & Katahira, M. (2003). *Journal of Biological Chemistry* **278**, 28147-28153.
- Matsugami, A., Ouhashi, K., Kanagawa, M., Liu, H., Kanagawa, S., Uesugi, S. & Katahira, M. (2001). *Journal of Molecular Biology* **313**, 255-269.
- Mooers, B. H. M., Logue, J. S. & Berglund, J. A. (2005). *Proceedings of the National Academy of Sciences of the United States of America* **102**, 16626-16631.

- Salisbury, S. A., Wilson, S. E., Powell, H. R., Kennard, O., Lubini, P., Sheldrick, G. M., Escaja, N., Alazzouzi, E., Grandas, A. & Pedroso, E. (1997). *Proceedings of the National Academy of Sciences of the United States of America* **94**, 5515-5518.
- Tamjar, J., Katorcha, E., Popov, A. & Malinina, L. (2012). *Journal of Biomolecular Structure and Dynamics* **30**, 505-523.
- Thorpe, J. H., Teixeira, S. C. M., Gale, B. C. & Cardin, C. J. (2003). *Nucleic acids research* **31**, 844-849.
- Viladoms, J., Escaja, N., Frieden, M., Gomez-Pinto, I., Pedroso, E. & Gonzalez, C. (2009). *Nucleic acids research* **37**, 3264-3275.
- Yildirim, I., Park, H., Disney, M. D. & Schatz, G. C. (2013). *Journal of the American Chemical Society* **135**, 3528-3538.
- Zheng, M., Huang, X., Smith, G. K., Yang, X. & Gao, X. (1996). *Journal of Molecular Biology* **264**, 323-336.