Rotational positioning of nucleosomes facilitates selective binding of p53 to response elements associated with cell cycle arrest

Feng Cui and Victor B. Zhurkin

Supplementary Material

Figure S1. Calculation of the W/S score for nucleosomal DNA sequences.
(**A**) Locations of the minor- and major-groove bending sites. The crystal structure of the 1kx5 nucleosome with 147-bp long DNA (49) is shown schematically: the DNA fragment is divided into two halves, separated by the dyad (black ball and diamond). The base-pair centers in the 'anterior' half are represented by large balls, and the sugar-phosphate backbone is shown as a yellow ribbon. For the 'posterior' half of the nucleosome, the base-pair centers are connected by sticks. The fragments whose minor grooves face the histone octamer (grey cylinder) are colored in blue, while the fragments whose minor grooves face away from the histone are colored in red. Seven minor-groove bending sites and six major-groove bending sites are shown in the 'anterior' half of the nucleosome; each site is 4 bp in length.
(**B**) Positions of the minor- and major-groove bending sites in the 'anterior' half of nucleosome. These sites are characterized by negative and positive Roll values, respectively (45,50). In the 'posterior' half, the 4-bp sites are positioned symmetrically with respect to the dyad (position 0). Overall, there are 14 minor-groove bending sites and 12 major-groove bending sites for a given 147-bp nucleosomal DNA.

Figure S2. Comparison of the CCA- and Apo-sites.
Shown are the distributions of distance to TSS (**A**), GC content (**B**), PHMM score (**C**) and spacer length (**D**) for 20 CCA-sites (blue, Supplementary Table S1) and 38 Apo-sites (red, Supplementary Table S2). The GC content is calculated for the 200-bp fragments with the REs at the center. The sequences flanking the REs are extracted from the human genome (Build 36). The other parameters are taken directly from Tables S1 and S2 in Riley et al. (4).

Figure S3. Profiles of NRP scores for nucleosomal DNA fragments mapped *in vitro.* Shown are the W/S and YR scores for the '601' sequence (**A**, **C**) and for 20 nucleosomal DNA sequences determined *in vitro* (50) (**B**, **D**). The scores for the 20 sequences are averaged and 'symmetrized' with respect to the dyad. The black diamonds denote the dyad positions.

Figure S4. Representative score profiles for the CCA-sites in the ambiguous or 'closed' states. (**A, C**) A CCA-site in the ambiguous state. The profiles of W/S and YR scores are shown for the sequence containing the *PLK1-2* RE (4). The score at position 0 (yellow circle) is neither a maximal score nor a minimal score, and it is not in phase or out of phase with neighboring peaks (black diamonds). This indicates that this RE is neither fully exposed nor completely buried. Its accessibility in the core particle is questionable (denoted as "±"). (**B, D**) A CCA-site in the closed state. The W/S and YR score profiles are shown for the sequence containing the *ARID3A* RE. The yellow circle is in phase with the neighboring peaks (black diamonds), indicating that the *ARID3A* RE is inaccessible to p53.

Figure S5. Representative score profiles for the Apo-sites in the ambiguous or open states. Shown are the W/S and YR score profiles for the sequences containing the *BCL2L14* RE (**A, C**) and *DKK1* RE (**B, D**). The notations are as in Supplementary Figure S4.

Figure S6. NRP score profiles for the p53 sites with spacer S = 0. Shown are the W/S and YR score profiles for the CCA-sites (**A, C**) and Apo-sites (**B, D**). The sites are listed in Supplementary Tables S1 and S2. The notations are as in Figures 3 and 4.

Figure S7. Nucleosome occupancy for genomic DNA flanking the 20 CCA-sites (blue) and 38 Apo-sites (red). Nucleosome occupancy at each position is averaged and symmetrized across the center of the p53 site. The analysis was based on the nucleosome occupancy data for human granulocytes (27) (584 million reads).

Figure S8. DNA distortion in the tetrameric complex with p53DBD (A) is similar to DNA distortions in nucleosome (B). In addition to bending into the minor groove (from the viewer), the DNA axis undergoes a lateral displacement, Slide ≈ 2.5 Å (inside the black circles).
(A) In crystallized complex with p53DBD (58) the lateral displacement Slide helps relieving the clashes between the two p53DBD dimers (see the red arrow on the left).
(B) In nucleosome (49), positive Slide widens the minor groove and accommodates the histone arginines penetrating the minor groove facing histone octamer (45,47).

Table S1. The CCA-REs used in this study and their predicted exposure states in nucleosomes

#	Gene Name(s)	Spacer length, bp	1st Half-site	Spacer	2nd Half-site	Exposure
1	*ARID3A*	0	GGACACGCTG	**TG|**	GGACATGCCT	–
2	*CCNG1*	0	GCACAAGCCC	**(CA:TG)**	AGGCTAGTCC	–
3	*CDKN1A, p21* 5'	0	GAACATGTCC	**|CA**	CAACATGTTG	+
4	*CDKN1A, p21* 3'	0	AGACTGGGCA	**CA|TG**	TGTCTGGGCA	+
5	*GML*	0	ATGCTTGCCC	**(CA:TG)**	AGGCATGTCC	–
6	*IBRDC2*	0	AGACAGGTCC	**|TG**	TGACAAGCAG	+
7	*MAD1L1*	0	ATTCAAGCTG	**TG|**	ATACTGAGTA	+
8	*PLK2_3*	0	AAACATGCCT	**(TG:CA)**	GGACTTGCCC	+
9	*S100A2*	0	GGGCATGTGT	**(TG:CA)**	GGGCACGTTC	+
10	*SFN, 14-3-3σ*	0	GCATTAGCCC	**(CA:TG)**	AGACATGTCC	+
11	*VCAN*	1	AGACTTGCCA	c	AGACAAGTCC	–
12	*CCNK*	2	AAACTAGCTT	gc	AGACATGCTG	+
13	*c13orf15*	3	AGGCGAGTTT	aag	CAGCTTGTCC	+
14	*PLK2_1*	3	AGACATGGTG	tgt	AAACTAGCTT	±
15	*PLK2_2*	3	GGTCATGATT	cct	TAACTTGCCT	±
16	*RB1*	4	GGGCGTGCCC	cgac	GTGCGCGCGC	–
17	*ODC1_1*	5	GGACCAGTTC	caggc	GGGCGAGACC	±
18	*PLK3*	6	TAACATGCCC	gggcaa	AAGCGAGCGC	+
19	*ODC1_2*	10	GGGCTCGCCT	tggtacagac	GAGCGGGCCC	+
20	*CDC25C*	11	GGGCAAGTCT	taccatttcca	GAGCAAGCAC	±

(+):(–):(±) = 11:5:4

The gene names and the RE sequences are taken from the Riley et al. dataset (4). The sites are sorted according to the spacer length. For the sites with spacer $S = 0$, the central dinucleotide is shown in red upper case letters if it is pyrimidine-purine (CA:TG or CG). The CA:TG dinucleotides are shown in blue letters when they are shifted by 1 bp from the center (denoted by a bar). For the sites with spacer $S > 0$, the spacer sequence is shown in black low case letters. Three exposure states of the REs wrapped in nucleosomes are assigned based on Set 1 score profile (Table SIII in (34)), with "+" standing for 'exposed' state, "–" for 'buried' state and "±" for the state that cannot be determined unambiguously. The occurrences of the three states are summed up in the last line. Typical profiles for these states are shown in Figure 3 and Supplementary Figure S4.

Table S2. The Apo-REs used in this study and their predicted exposure states in nucleosomes

#	Gene Name(s)	Spacer length, bp	1st Half-site	Spacer	2nd Half-site	Exposure
1	*BAX*	0	GGGCAGGCCC	(CG)	GGGCTTGTCG	±
2	*BBC3, PUMA*	0	CTGCAAGTCC		TGACTTGTCC	–
3	*BCL2L14**	0	AGCCAAGGCT	(TG:CA)	GGTCTTGAAC	±
4	*BID**	0	GGGCATGATG		GTGCATGCCT	–
5	*CTSD*	0	AACCTTGGTT		TGCAAGAGGC	–
6	*CTSD*	0	AAGCTGGGCC	(CG)	GGGCTGACCC	±
7	*FAS, Apo-1*	0	GGACAAGCCC		TGACAAGCCA	–
8	*GDF15*	0	AGCCATGCCC	(CG)	GGGCAAGAAC	–
9	*GDF15**	0	CATCTTGCCC	(CA:TG)	AGACTTGTCT	–
10	*IER3*	0	CCACATGCCT		CGACATGTGC	–
11	*NLRC4*	0	AGACATGTTC		CTGGTAGTTT	–
12	*p53AIP1*	0	TCTCTTGCCC	(CG)	GGGCTTGTCG	–
13	*PERP*	0	AGGCAAGCTC		CAGCTTGTTC	+
14	*SCN3B*	0	TGACTTGCTC		TGCCTTGCCT	–
15	*SCN3B*	0	TGGCAAGGCT	(TG:CA)	GAGCTAGTTC	±
16	*TNFRSF10A*	0	GGGCATGTCC	(CG)	GGGCAGGAGG	–
17	*TNFRSF10B, KILLER*	0	GGGCATGTCC	(CG)	GGGCAAGACG	–
18	*TNFRSF10C*	0	GGGCATGTCC	(CG)	GGGCAGGACG	–
19	*TNFRSF10D*	0	GGGCATGTCT	(TG)	GGGCAGGACG	–
20	*TP53INP1*	0	GAACTTGGGG		GAACATGTTT	–
21	*TRIAP1*	0	CTTCATGTCC	(CG)	GTGCATGCCT	–
22	*UBD*	0	AGGCATGCTC	(CA:TG)	AGTGGCGTGG	+
23	*PCBP4*	1	GGTCTTGGCC	c	AGACTTAGCA	–
24	*PYCARD*	2	GTGCAAGCCC	ag	AGACAAGCAG	±
25	*BIRC5,* survivin	3	GGGCGTGCGC	tcc	CGACATGCCC	–
26	*EphA2**	3	CACCATGTTG	gcc	AGGCATGTCT	–
27	*CASP6**	4	AGGCAAGGAG	tttg	AGACAAGTCT	+
28	*LGALS3*	4	GGGCTTGCAA	gctg	GAGCCTTGTT	–
29	*BNIP3L**	5	AAGCTAGTCT	cagtg	GCGCATGCCT	–
30	*CASP10*	5	AAACTTGCTG	gttta	AATCTTGGCT	–
31	*DKK1*	6	AGCCAAGCTT	ttaatg	AACCAAGTTC	+
32	*AIFM2, AMID**	8	GGTCTCGCTA	tgttgccc	AGGCTGGTCT	–
33	*LRDD*	8	AGGCCTGCCT	gcgtgctg	GGACATGTCT	±
34	*AIFM2, AMID**	10	AGGCATGAGC	caccgtgcct	GGCCATGCCC	+
35	*PCBP4*	10	GAACTTAAGA	ccgaggctct	GGACAAGTTG	+
36	*NDRG1*	12	CCACATGCAC	acgcacgagcgc	GCACATGAAC	±
37	*APAF1*	13	AGACATGTCT	ggagaccctagga	CGACAAGCCC	–
38	*PTEN*	14	GAGCAAGCCC	caggcagctacact	GGGCATGCTC	+

(+):(–):(±) = 7:24:7

All notations are as in Table S1. Typical profiles for the three exposure states are shown in Figures 3 and Supplementary Figure S5. In addition, the REs occurring in Alu repeats (16) are marked by asterisks.

Table S3. Conformations of the central YR steps in the tetrameric p53DBD-DNA cocrystals

PDB ID [Ref]	Central Hexamer	Central Step	Twist (°)	Tilt (°)	Roll (°)	Shift (Å)	Slide (Å)	Rise (Å)
3KZ8 [S1]	CCC|GGG	CG:CG	34.2	0.0	2.9	0.00	1.33	2.97
3KMD [S2]	CCT|AGG	TA:TA	50.5	0.7	-5.6	0.06	2.87	3.35
3TS8 [S3]	CCC|AGA	CA:TG	43.2	-0.5	-3.2	-0.43	2.62	3.22
3EXJ [S4]	CTC|GAG	CG:CG	41.8	0.0	-10.2	0.00	4.21	3.54

Base-pair step parameters were calculated using CompDNA/3DNA algorithm [S5,S6].

References

[S1] Kitayner,M., Rozenberg,H., Rohs,R., Suad,O., Rabinovich,D., Honig,B., Shakked,Z. (2010) Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs. *Nat. Struct. Mol. Biol.* **17**, 423-429.

[S2] Chen,Y., Dey,R. and Chen,L. (2010) Crystal structure of the p53 core domain bound to a full consensus site as a self-assembled tetramer. *Structure* **18**, 246-256.

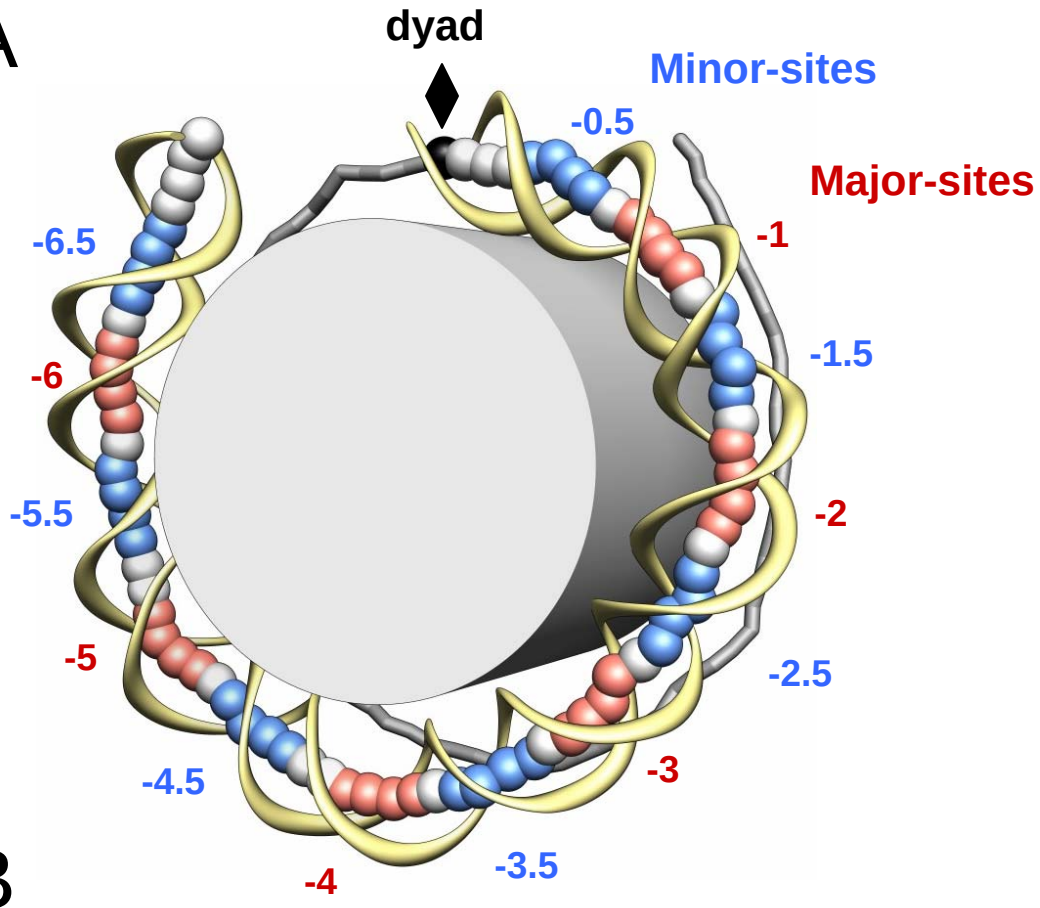
[S3] Malecka,K.A., Ho,W.C. and Marmorstein,R. (2009) Crystal structure of a p53 core tetramer bound to DNA. *Oncogene* **28**, 325–333.

[S4] Emamzadah,S., Tropia,L. and Halazonetis,T.D. (2011) Crystal structure of a multidomain human p53 tetramer bound to the natural CDKN1A (p21) p53 response element. *Mol. Cancer Res.* **9**, 1493-1499.

[S5] Gorin,A.A., Zhurkin,V.B. and Olson,W.K. (1995) B-DNA twisting correlates with base-pair morphology. *J. Mol. Biol.* **247**, 34-48.

[S6] Lu, X.J. and Olson,W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31**, 5108-21.

A



B

SHL	-6.5	-6	-5.5	-5	-4.5	-4	-3.5	-3	-2.5	-2	-1.5	-1	-0.5	dyad
4-bp sites	-69/-66	-64/-61	-59/-56	-53/-50	-48/-45	-42/-39	-37/-34	-32/-29	-27/-24	-22/-19	-17/-14	-12/-9	-7/-4	♦

Fig. S1

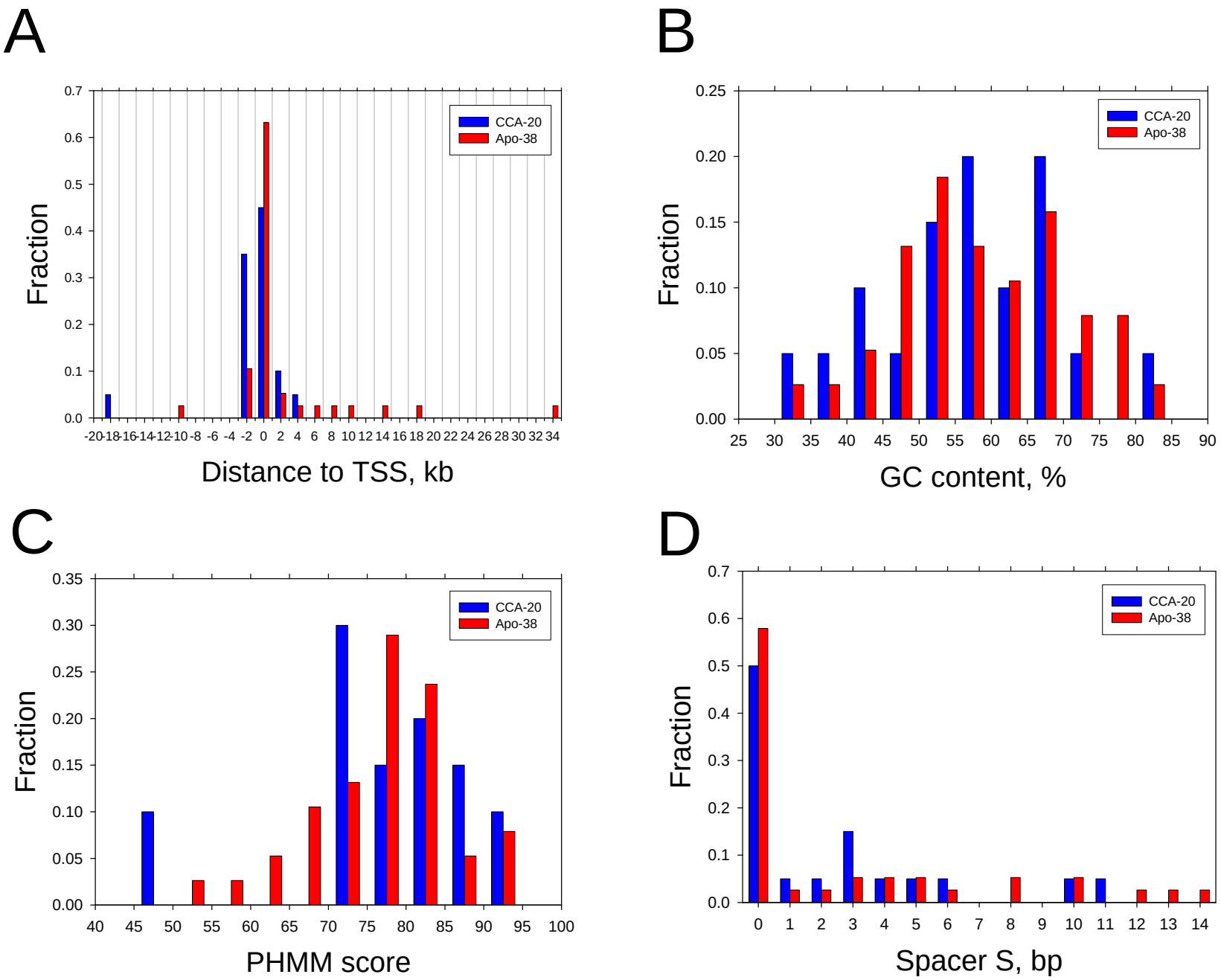

A
Fraction
Distance to TSS, kb
CCA-20
Apo-38
B
Fraction
GC content, %
C
Fraction
PHMM score
D
Fraction
Spacer S, bp

Fig. S2

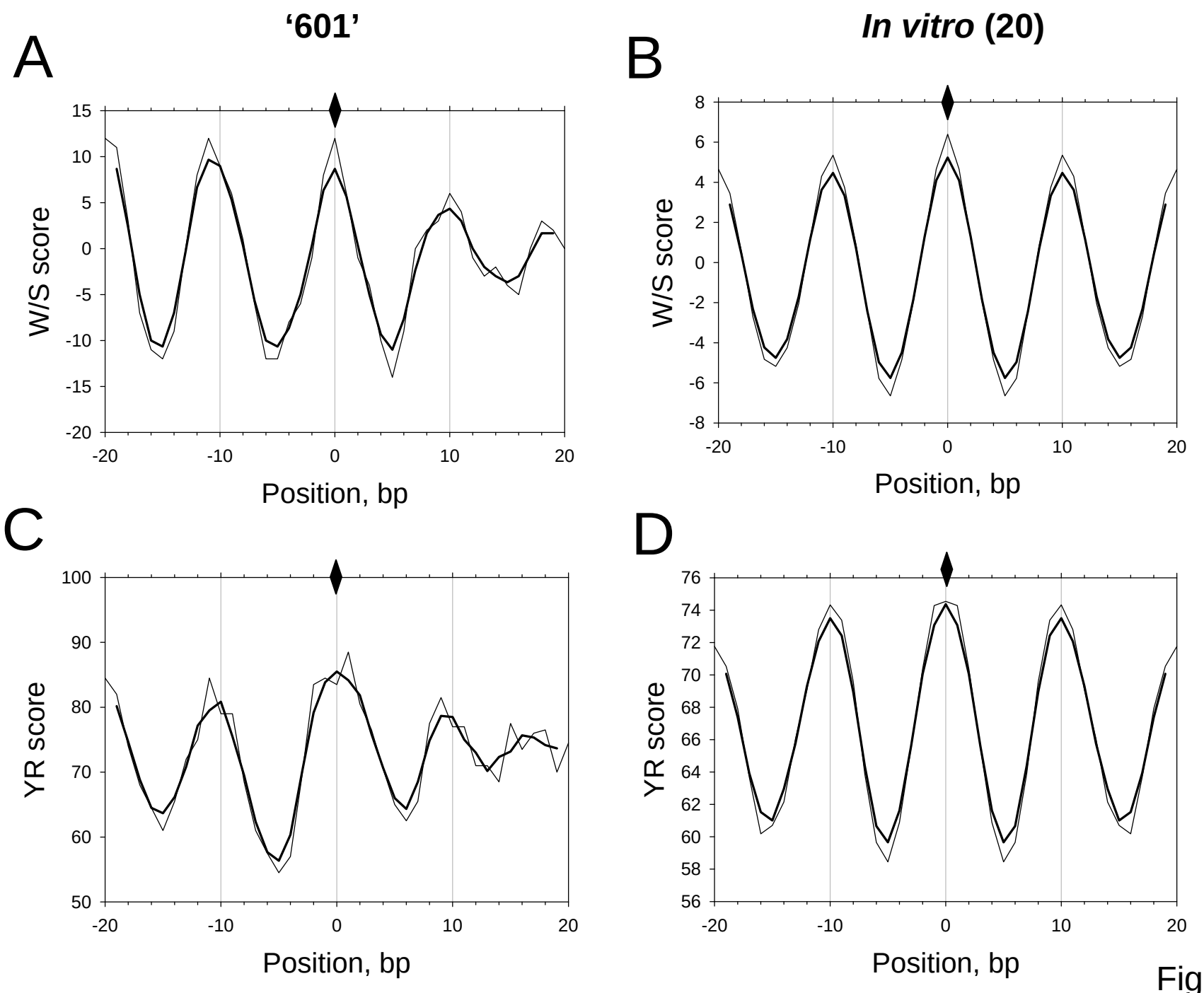

A
'601'
W/S score
Position, bp
B
In vitro (20)
W/S score
Position, bp
C
YR score
Position, bp
D
YR score
Position, bp

Fig. S3

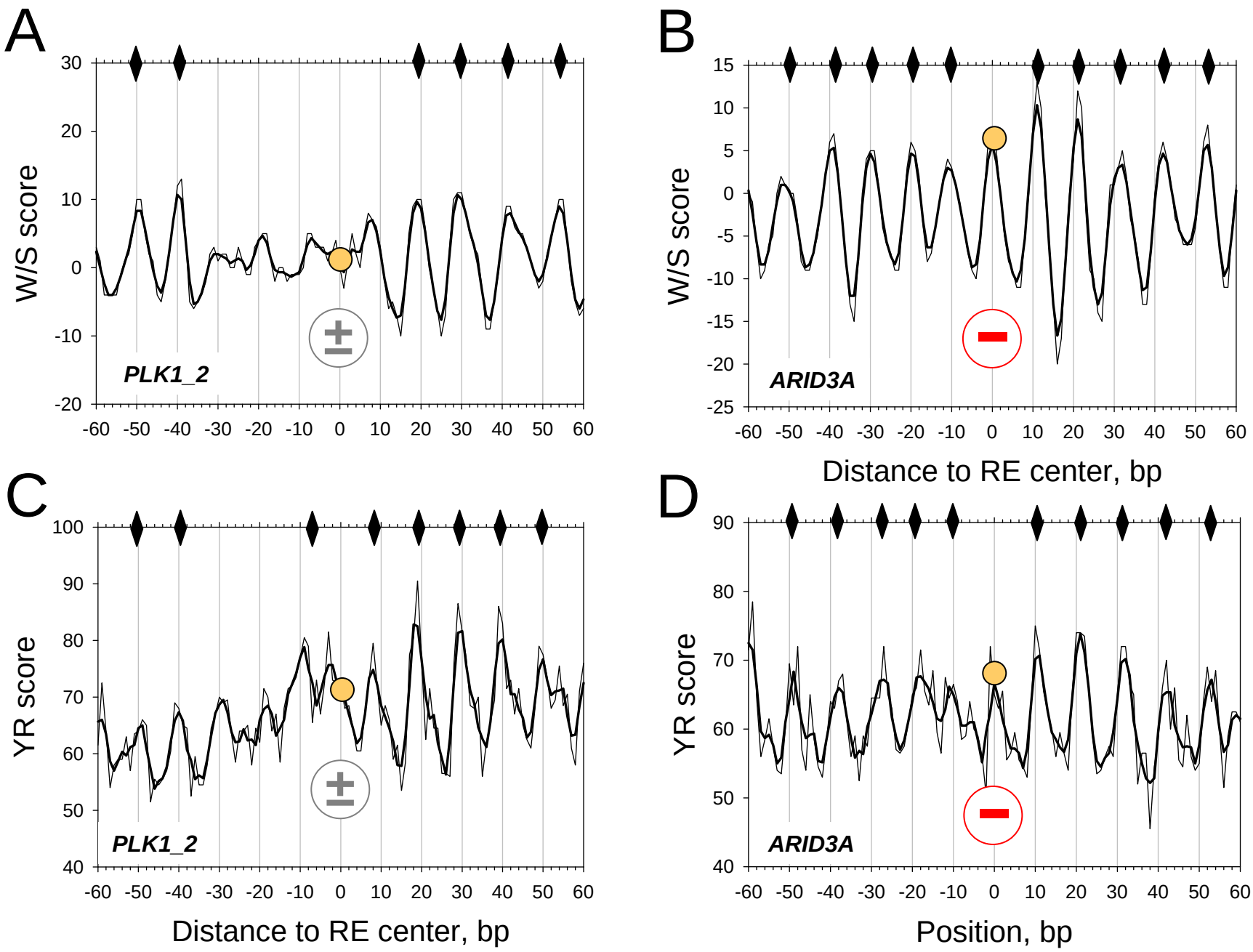

A
B
C
D
W/S score
YR score
PLK1_2
ARID3A
Distance to RE center, bp
Position, bp

Fig. S4

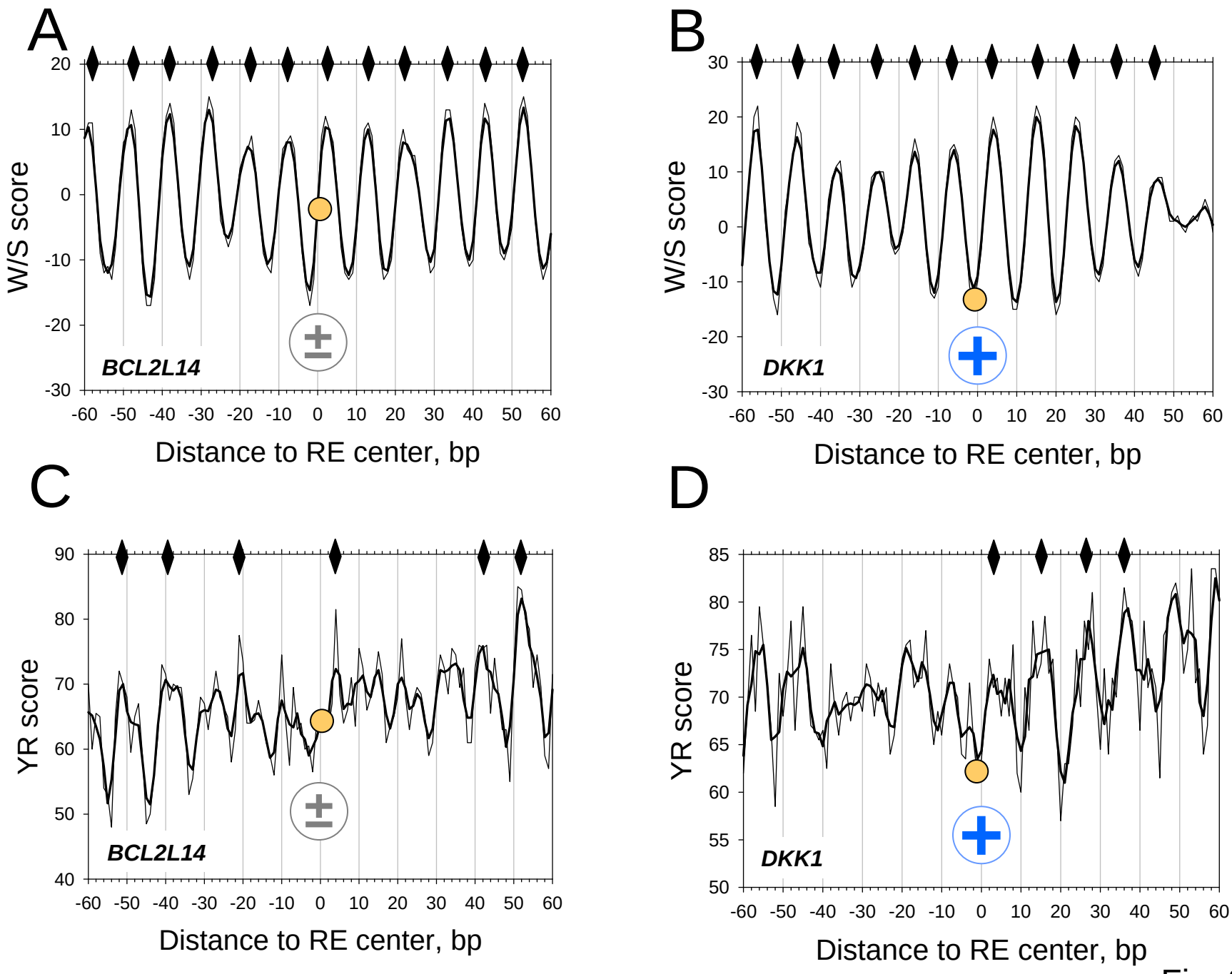
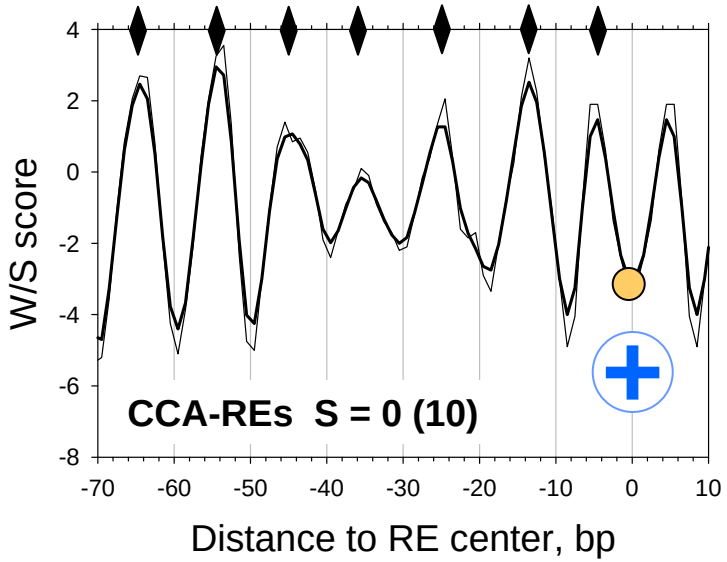
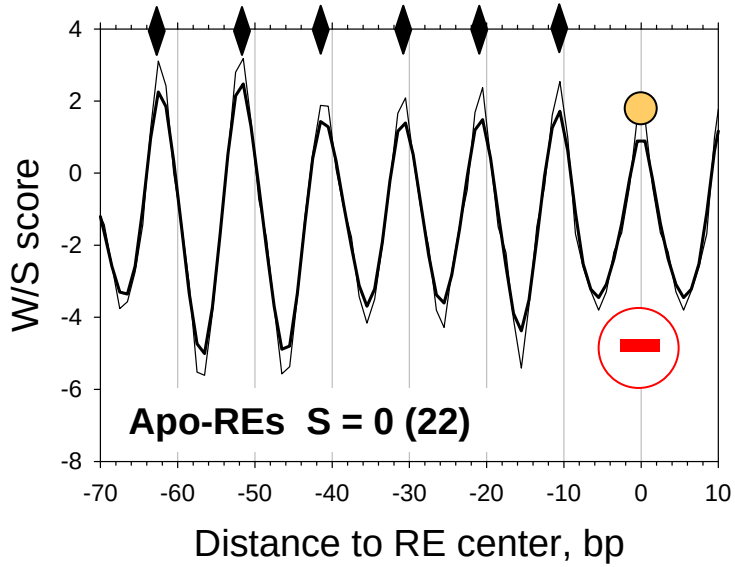

A
W/S score
BCL2L14
±
Distance to RE center, bp
B
W/S score
DKK1
+
Distance to RE center, bp
C
YR score
BCL2L14
±
Distance to RE center, bp
D
YR score
DKK1
+
Distance to RE center, bp

Fig. S5

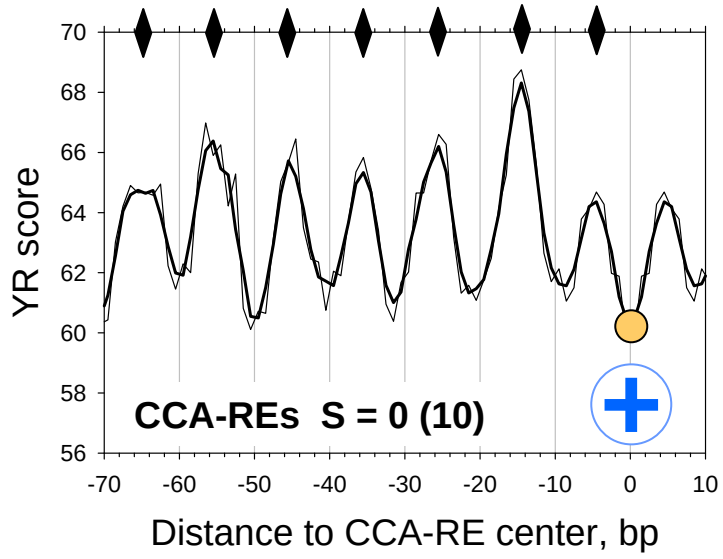
A


W/S score
CCA-REs S = 0 (10)
Distance to RE center, bp

B


W/S score
Apo-REs S = 0 (22)
Distance to RE center, bp

C


YR score
CCA-REs S = 0 (10)
Distance to CCA-RE center, bp

D

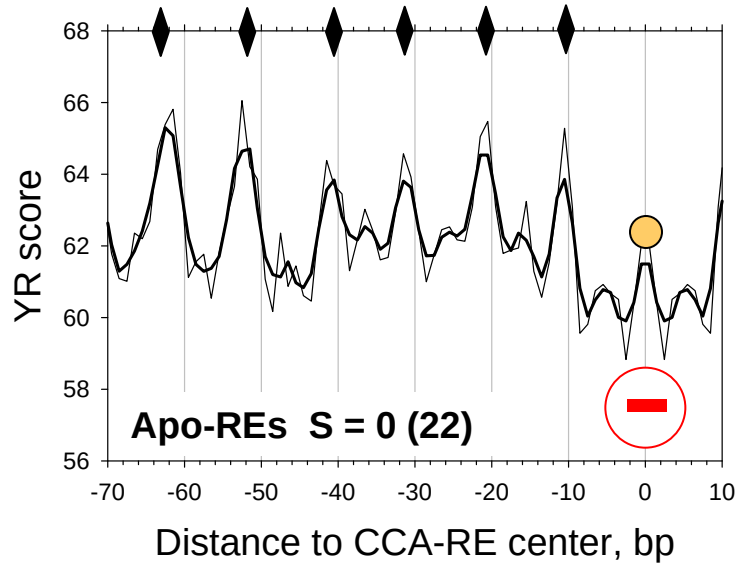
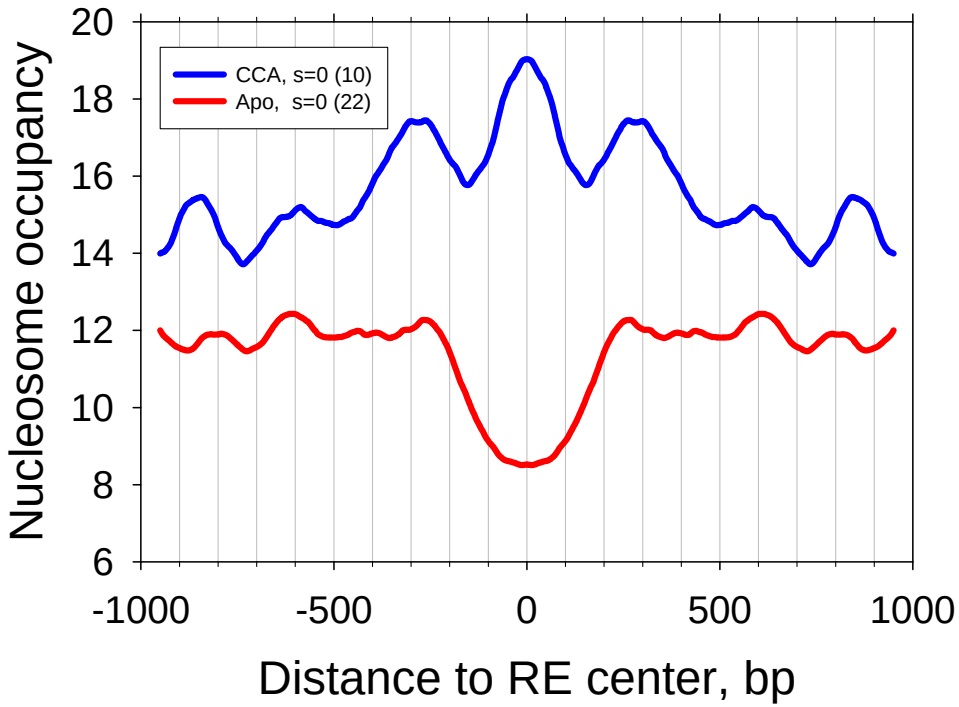

YR score
Apo-REs S = 0 (22)
Distance to CCA-RE center, bp

Fig. S6

A


Nucleosome occupancy
20
18
16
14
12
10
8
6
CCA, s=0 (10)
Apo, s=0 (22)
-1000
-500
0
500
1000
Distance to RE center, bp

B

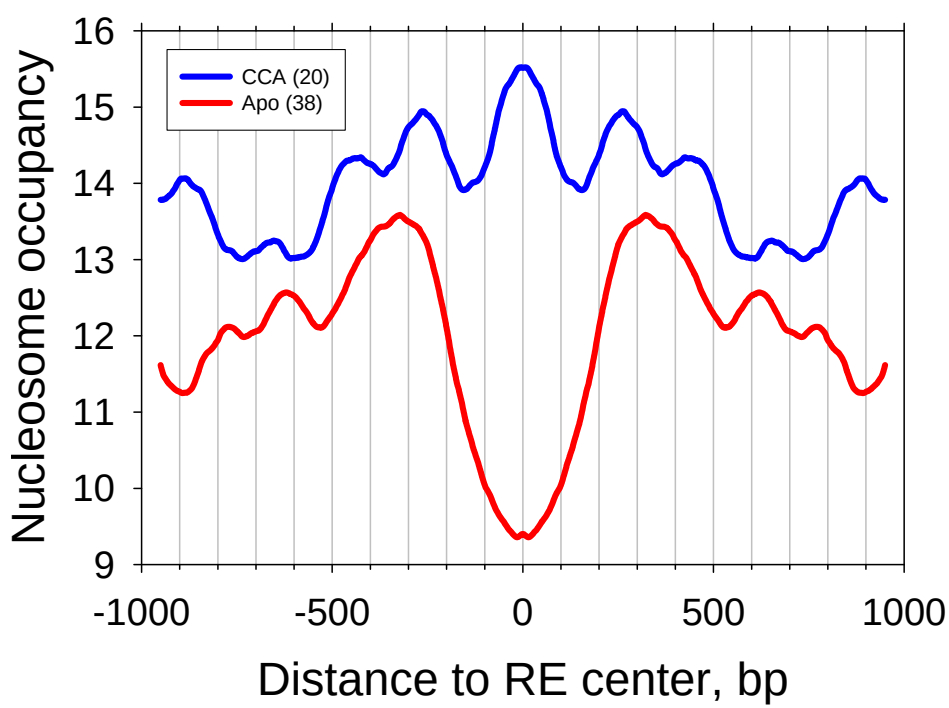
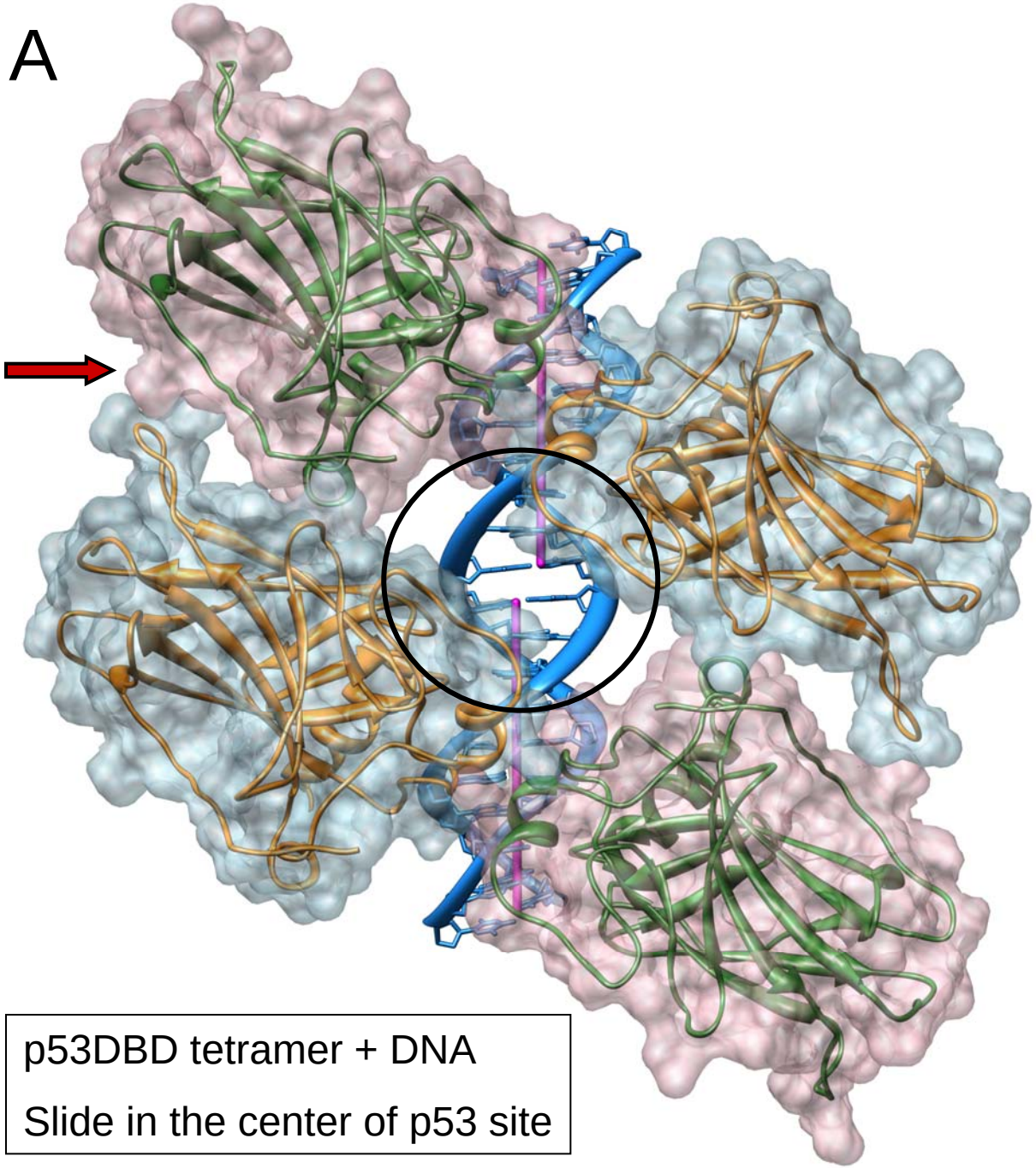

Nucleosome occupancy
16
15
14
13
12
11
10
9
CCA (20)
Apo (38)
-1000
-500
0
500
1000
Distance to RE center, bp

Fig. S7

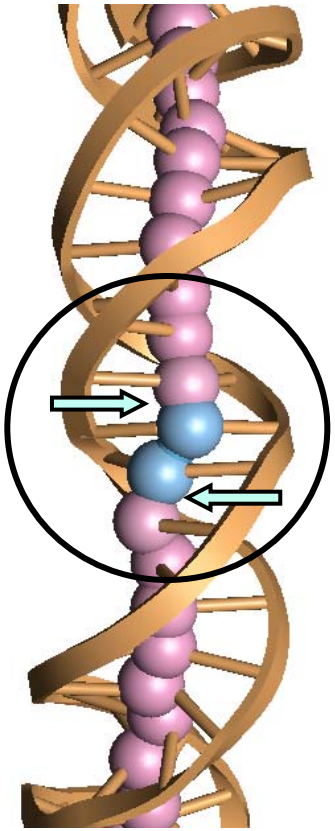
A



p53DBD tetramer + DNA

Slide in the center of p53 site

B



Kink-and-Slide
in nucleosome
at CA:TG & TA

Fig. S8