

Supporting information for:

**Physical Binding of the Tobacco Smoke
Carcinogen NNK Diazonium Ion to the Human
Tumor Suppressor Gene TP53 Exon 5**

Christos Deligkaris^{*,†} and Evan Millam[‡]

*†Department of Geology and Physics, University of Southern Indiana, Evansville IN 47712,
USA*

‡Department of Chemistry, University of Southern Indiana, Evansville IN 47712, USA

E-mail: Dr.Deligkaris@usi.edu

TP53 Exon 5

Table S1: Base pair step parameters used to construct the DNA model. The parameters are average values from experimentally determined structures.^{S1} The parameters were obtained using the 3-Dimensional Nucleic Acids program.^{S2} Distances are in Å and angles are in degrees. Note that complementary base pair steps (such as AA and TT) have opposite signs in Shift and Tilt.

Base Pair Step	Shift	Slide	Rise	Tilt	Roll	Twist
AA	-0.06	-0.16	3.28	+0.1	+1.4	35
TT	+0.06	-0.16	3.28	-0.1	+1.4	35
AC	+0.06	-0.43	3.23	-0.3	+1.4	32
GT	-0.06	-0.43	3.23	+0.3	+1.4	32
AG	+0.06	+0.34	3.27	+0.2	+5.5	28
CT	-0.06	+0.34	3.27	-0.2	+5.5	28
AT	+0.12	-0.57	3.30	+0.3	-1.2	31
CA	+0.02	+1.88	3.32	0.0	-1.2	43
TG	-0.02	+1.88	3.32	0.0	-1.2	43
CC	+0.05	+0.28	3.40	-0.1	+3.9	35
GG	-0.05	+0.28	3.40	+0.1	+3.9	35
CG	+0.06	+0.68	3.25	0.0	+6.2	31
GA	0.00	-0.01	3.43	0.0	+0.4	41
TC	0.00	-0.01	3.43	0.0	+0.4	41
GC	-0.30	+0.31	3.57	0.0	-6.8	40
TA	-0.17	+0.38	3.37	-1.4	-0.6	43

Table S2: Base pair parameters used to construct the DNA model. The parameters are average values from experimentally determined structures.^{S3} The parameters were obtained using the 3-Dimensional Nucleic Acids program.^{S2} Distances are in Å and angles are in degrees. Note that M-N and N-M base pairs (such as A-T and T-A) have opposite signs in Shear and Buckle.

Base Pair	Shear	Stretch	Stagger	Buckle	Propeller-Twist	Opening
A-T	-0.03	-0.18	-0.03	+0.83	-14.80	+0.12
T-A	+0.03	-0.18	-0.03	-0.83	-14.80	+0.12
G-C	+0.08	-0.15	+0.10	-1.52	-11.80	-0.88
C-G	-0.08	-0.15	+0.10	+1.52	-11.80	-0.88

Table S3: Methylated cytosines in exon 5 of TP53. All CpG cytosines in exon 5, both strands.^{S4}

Chain	Cytosine PDB residue number	Codon
A	180	152
A	184	153
A	191	156
A	193	156
A	197	158
A	199	158
A	234	170
A	248	175
B	588	152
B	584	154
B	577	156
B	575	157
B	571	158
B	569	159
B	534	170
B	520	175

Table S4: Base sequence used. Intron: 12215-12314. Exon: 12315-12498. Intron: 12499-12598.

```

Intron | ggtgct tacgcatggt tgtttctttg ctgccctctt ccagttgctt tatctgttca cttgtgccct gaacttcaac tetgtctcct tcctcttctt acag
Exon   | tactcc cctgccctca acaagatggt ttgccaactg gccaaagacct gccctgtgca gctgtggggt gattccacac ccccgcccgg caccgcgctc cgcgcatgg
Exon   | ccatctacaa gcagtcacag cacatgacgg aggttgtgag gcgctgcccc caccatgagc gctgctcaga tagcgatg
Intron | gt gagcagctgg ggctggagag acgacagggc tggttgccca ggtgccccag gcctctgatt cctcactgat tgctcttagg tctggcccct cctcagca

```

Table S5: Residue numbers, codons, and sequence information

RESIDUES	CODON	SEQUENCE	RESIDUES	CODON	SEQUENCE
101-103 (666-668)	126	TAC	197-199 (570-572)	158	CGC
104-106 (663-665)	127	TCC	200-202 (567-569)	159	GCC
107-109 (660-662)	128	CCT	203-205 (564-566)	160	ATG
110-112 (657-659)	129	GCC	206-208 (561-563)	161	GCC
113-115 (654-656)	130	CTC	209-211 (558-560)	162	ATC
116-118 (651-653)	131	AAC	212-214 (555-557)	163	TAC
119-121 (648-650)	132	AAG	215-217 (552-554)	164	AAG
122-124 (645-647)	133	ATG	218-220 (549-551)	165	CAG
125-127 (642-644)	134	TTT	221-223 (546-548)	166	TCA
128-130 (639-641)	135	TGC	224-226 (543-545)	167	CAG
131-133 (636-638)	136	CAA	227-229 (540-542)	168	CAC
134-136 (633-635)	137	CTG	230-232 (537-539)	169	ATG
137-139 (630-632)	138	GCC	233-235 (534-536)	170	ACG
140-142 (627-629)	139	AAG	236-238 (531-533)	171	GAG
143-145 (624-626)	140	ACC	239-241 (528-530)	172	GTT
146-148 (621-623)	141	TGC	242-244 (525-527)	173	GTG
149-151 (618-620)	142	CCT	245-247 (522-524)	174	AGG
152-154 (615-617)	143	GTG	248-250 (519-521)	175	CGC
155-157 (612-614)	144	CAG	251-253 (516-518)	176	TGC
158-160 (609-611)	145	CTG	254-256 (513-515)	177	CCC
161-163 (606-608)	146	TGG	257-259 (510-512)	178	CAC
164-166 (603-605)	147	GTT	260-262 (507-509)	179	CAT
167-169 (600-602)	148	GAT	263-265 (504-506)	180	GAG
170-172 (597-599)	149	TCC	266-268 (501-503)	181	CGC
173-175 (594-596)	150	ACA	269-271 (498-500)	182	TGC
176-178 (591-593)	151	CCC	272-274 (495-497)	183	TCA
179-181 (588-590)	152	CCG	275-277 (492-494)	184	GAT
182-184 (585-587)	153	CCC	278-280 (489-491)	185	AGC
185-187 (582-584)	154	GGC	281-283 (486-488)	186	GAT
188-190 (579-581)	155	ACC			
191-193 (576-578)	156	CGC			
194-196 (573-575)	157	GTC			

References

- (S1) Pérez, A.; Noy, A.; Lankas, F.; Luque, F. J.; Orozco, M. The relative flexibility of B-DNA and A-RNA duplexes: database analysis. *Nucl. Ac. Res.* **2004**, *32*, 6144.
- (S2) Lu, X.-J.; Olson, W. K. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protocols* **2008**, *3*, 1213–1227.
- (S3) Sadowska-Aleksiejew, A.; Rak, J.; Voityuk, A. A. Effects of intra base-pairs flexibility on hole transfer coupling in DNA. *Chem. Phys. Lett.* **2006**, *429*, 546 – 550.
- (S4) Tornaletti, S.; Pfeifer, G. P. Complete and tissue-independent methylation of CpG sites in the p53 gene: implications for mutations in human cancers. *Oncogene* **1995**, *10*, 1493–1499.