Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Stacking Interaction involving non Watson-Crick basepair: Dispersion Corrected Density Functional Theory Studies

Satyabrata Maiti, Dhananjay Bhattacharyya*

Computational Science Division, Saha Institute of Nuclear Physics, 1/AF Bidhannagar, Kolkata 700064, India

Homi Bhaba National Institute, Anushaktinagar, Mumbai - 400 094, India

*Address all correspondence to: Dhananjay Bhattacharyya

E-mail: dhananjay.bhattacharyya@saha.ac.in

FAX: +91-33-2337 4637

Supplementary Information

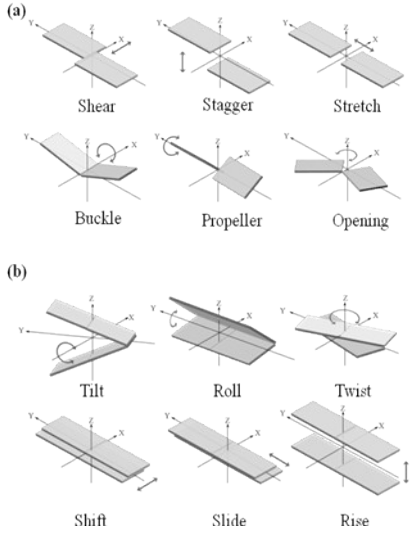


Figure SI-1: Standard nomenclature of (a) basepair parameter and (b) basepair step parameter, as suggested by IUPAC-IUB Cambridge Convension[53].

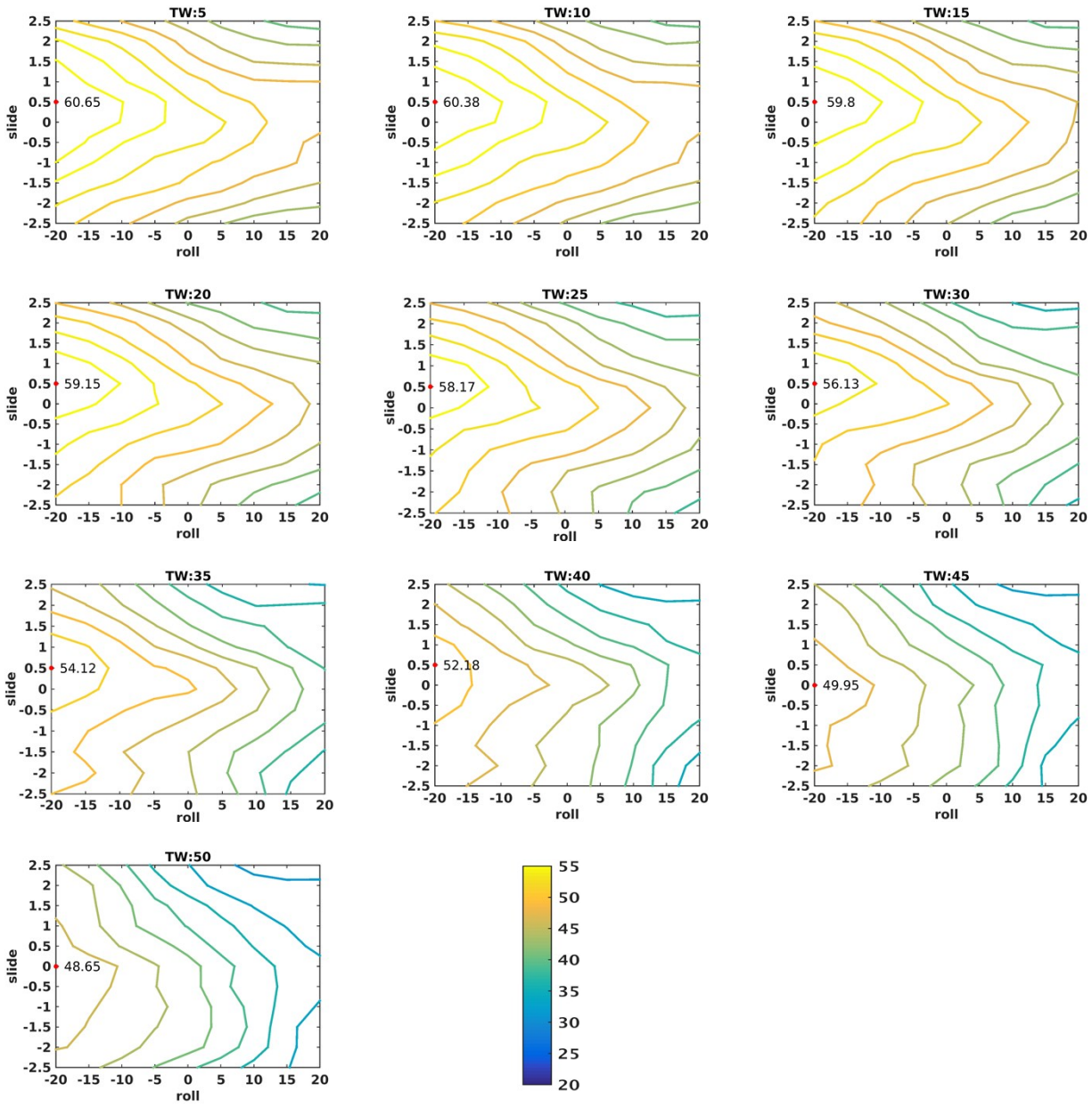
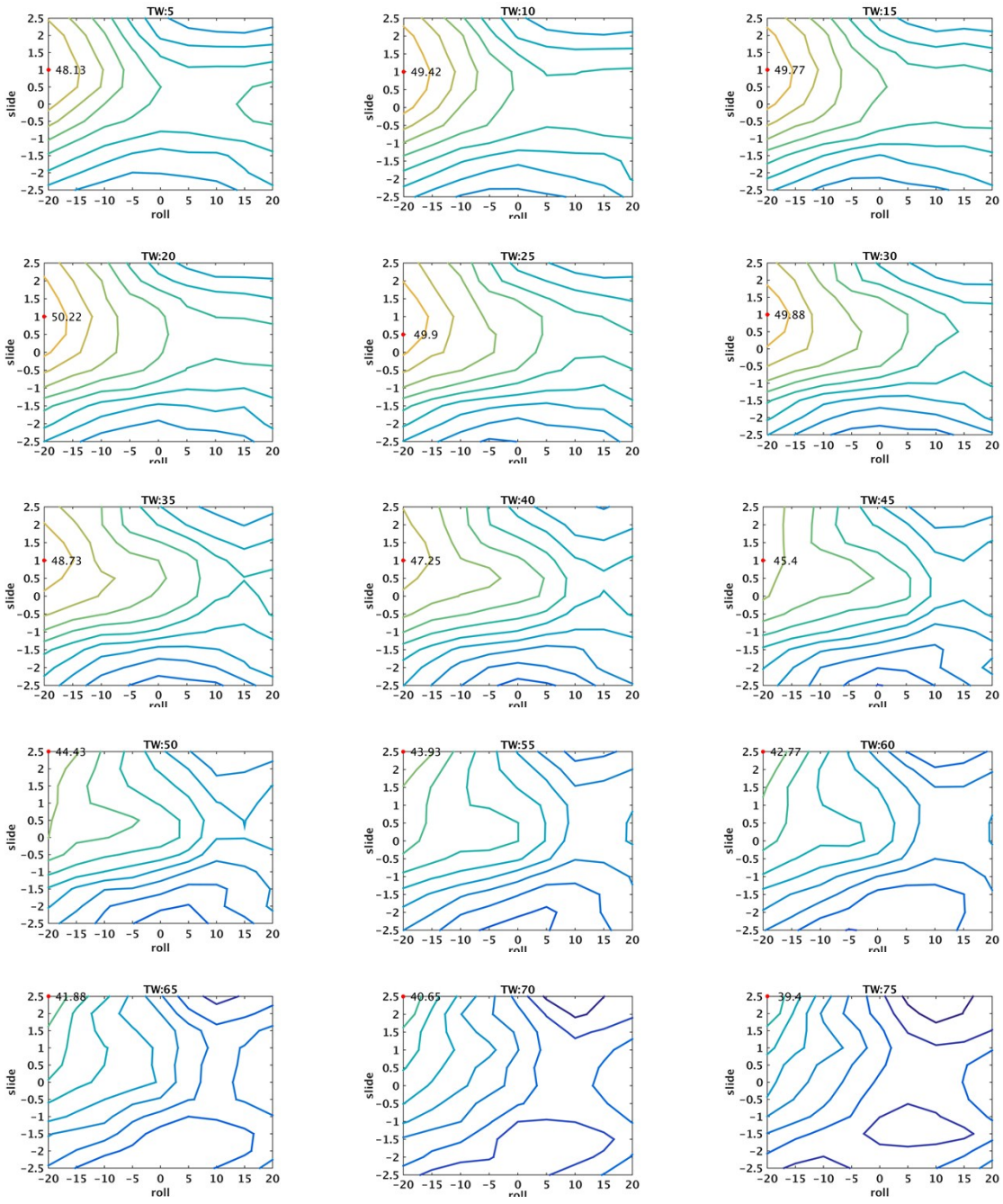


Figure SI-2: Stacking overlap area contours at different twist values for A:U W:WC::C:G:W:WC are shown here. Color bar is given in $Å^2$. Contour lines are 2.5 $Å^2$ apart.


TW:5
48.13
TW:10
49.42
TW:15
49.77
TW:20
50.22
TW:25
49.9
TW:30
49.88
TW:35
48.73
TW:40
47.25
TW:45
45.4
TW:50
44.43
TW:55
43.93
TW:60
42.77
TW:65
41.88
TW:70
40.65
TW:75
39.4
slide
roll

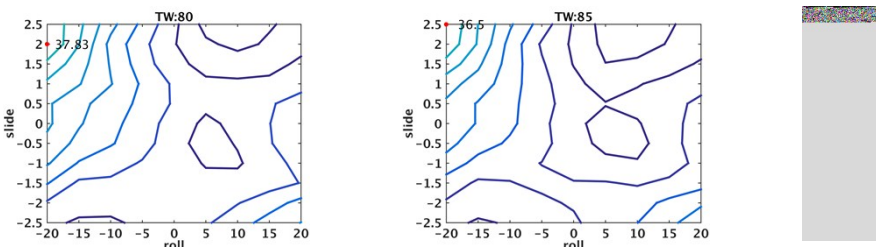
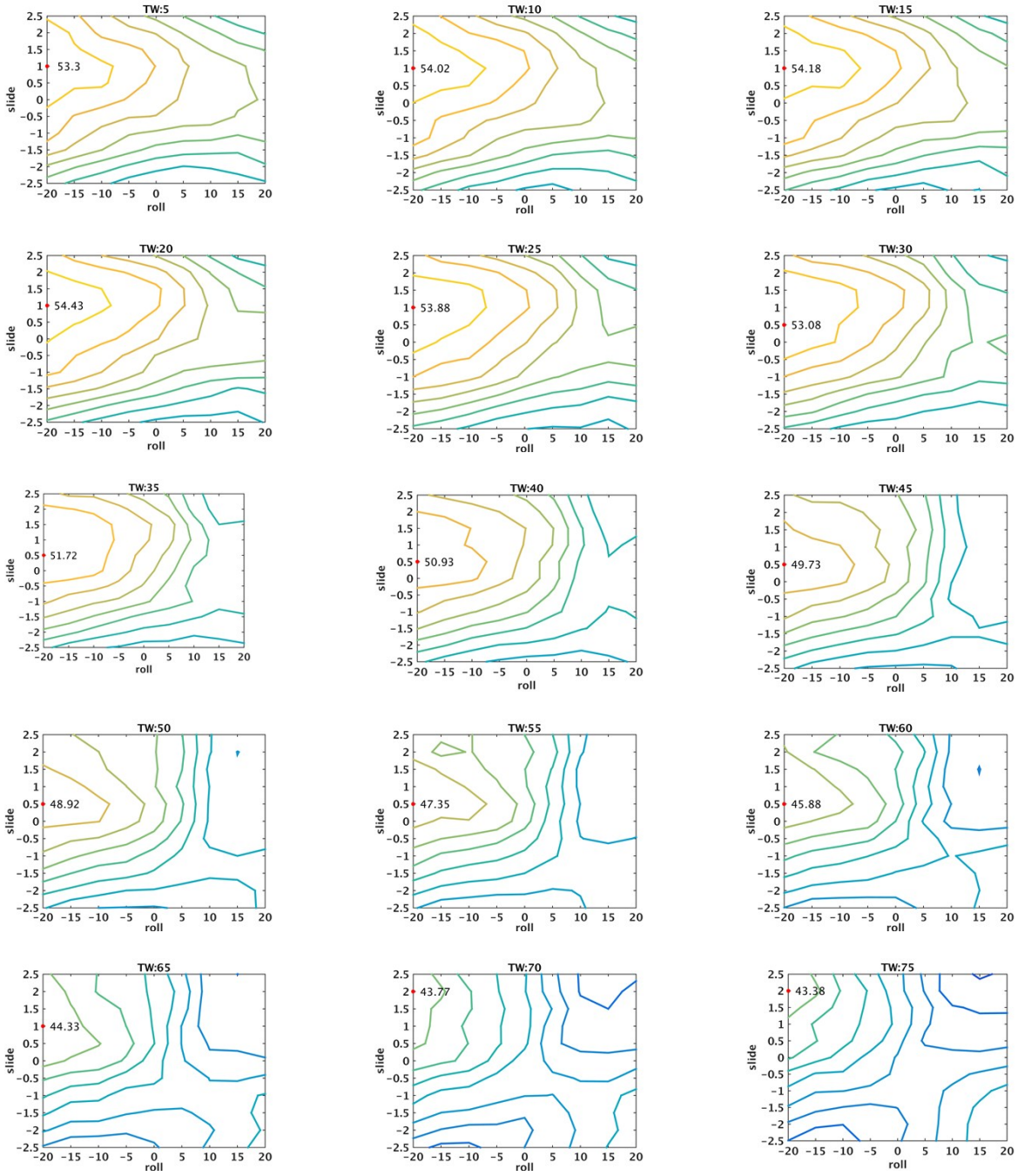


Figure SI-3: Stacking overlap area contours at different twist value for G:C W:WC::G:A S:HT are shown here. Color bar is given in Å^2. Contour lines are 2.5 Å^2 apart.


TW:5
53.3
TW:10
54.02
TW:15
54.18
TW:20
54.43
TW:25
53.88
TW:30
53.08
TW:35
51.72
TW:40
50.93
TW:45
49.73
TW:50
48.92
TW:55
47.35
TW:60
45.88
TW:65
44.33
TW:70
43.77
TW:75
43.38
slide
roll

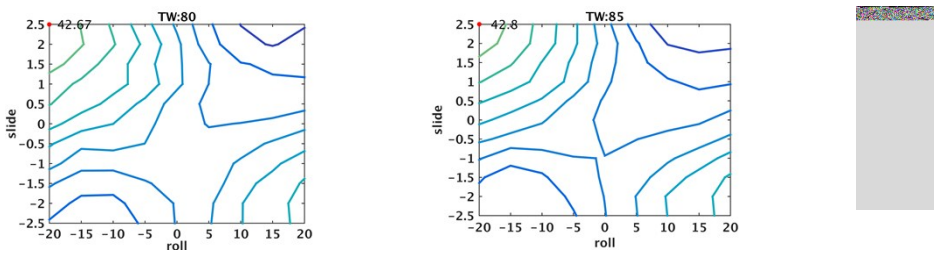


Figure SI-4: Stacking overlap area contours at different twist value for C:G W:WC::G:A S:HT are shown here. Color bar is given in $Å^2$. Contour lines are 2.5 $Å^2$ apart.

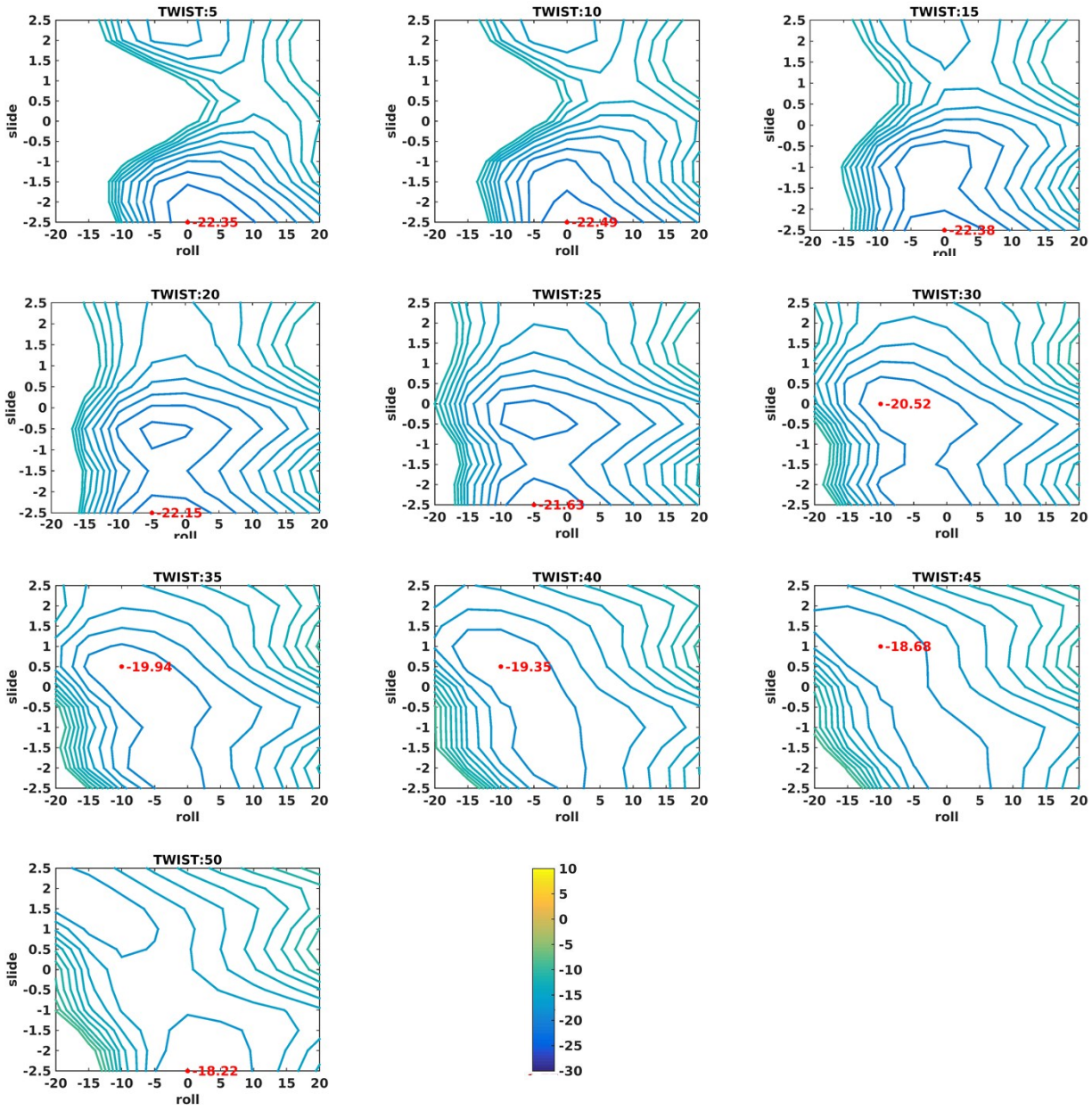
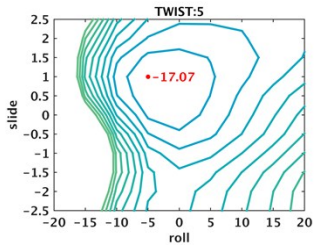
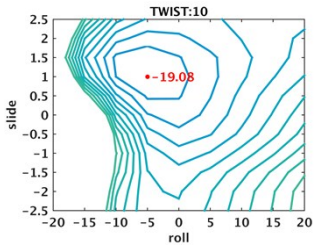
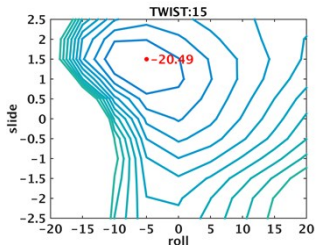
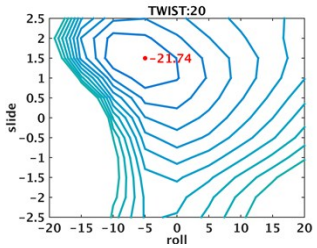


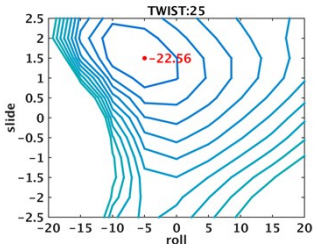
Figure SI-5: Intrinsic stacking energy contours considering DFT-D energy at different twist values for A:U W:WC::C:G W:WC dinucleotide step are shown here. Color bar is given in kcal/mol. The contour lines are 1kcal/mol apart.

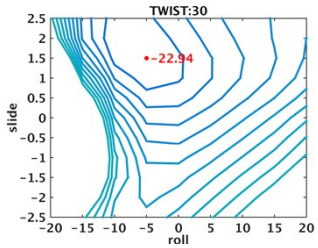

TWIST:5
-17.07
slide
roll

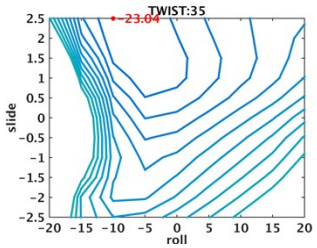

TWIST:10
-19.08
slide
roll

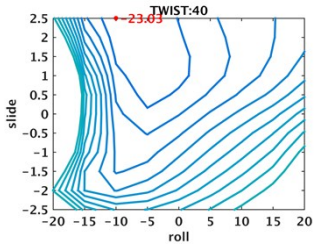

TWIST:15
-20.49
slide
roll

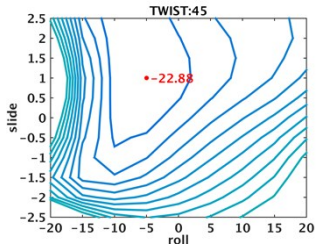

TWIST:20
-21.74
slide
roll

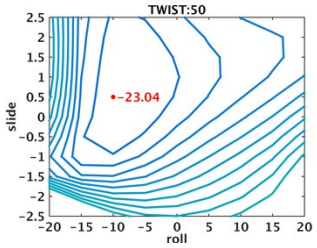

TWIST:25
-22.56
slide
roll

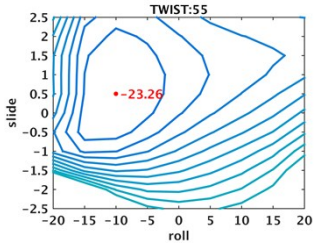

TWIST:30
-22.94
slide
roll

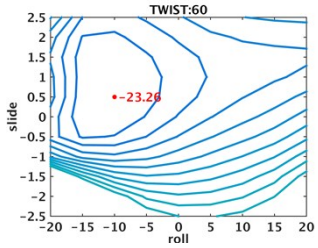

TWIST:35
-23.04
slide
roll

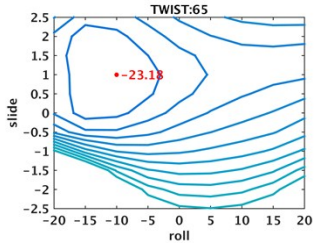

TWIST:40
-23.03
slide
roll

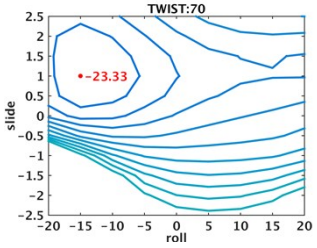

TWIST:45
-22.83
slide
roll

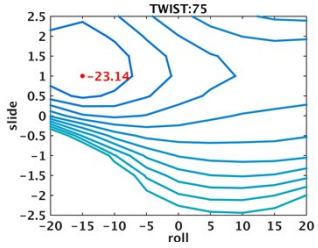

TWIST:50
-23.04
slide
roll


TWIST:55
-23.26
slide
roll


TWIST:60
-23.26
slide
roll


TWIST:65
-23.18
slide
roll


TWIST:70
-23.33
slide
roll


TWIST:75
-23.14
slide
roll

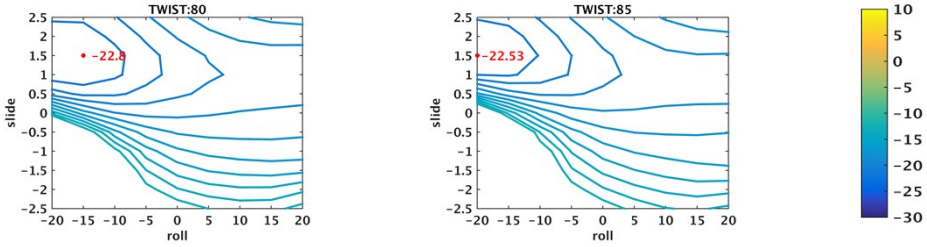
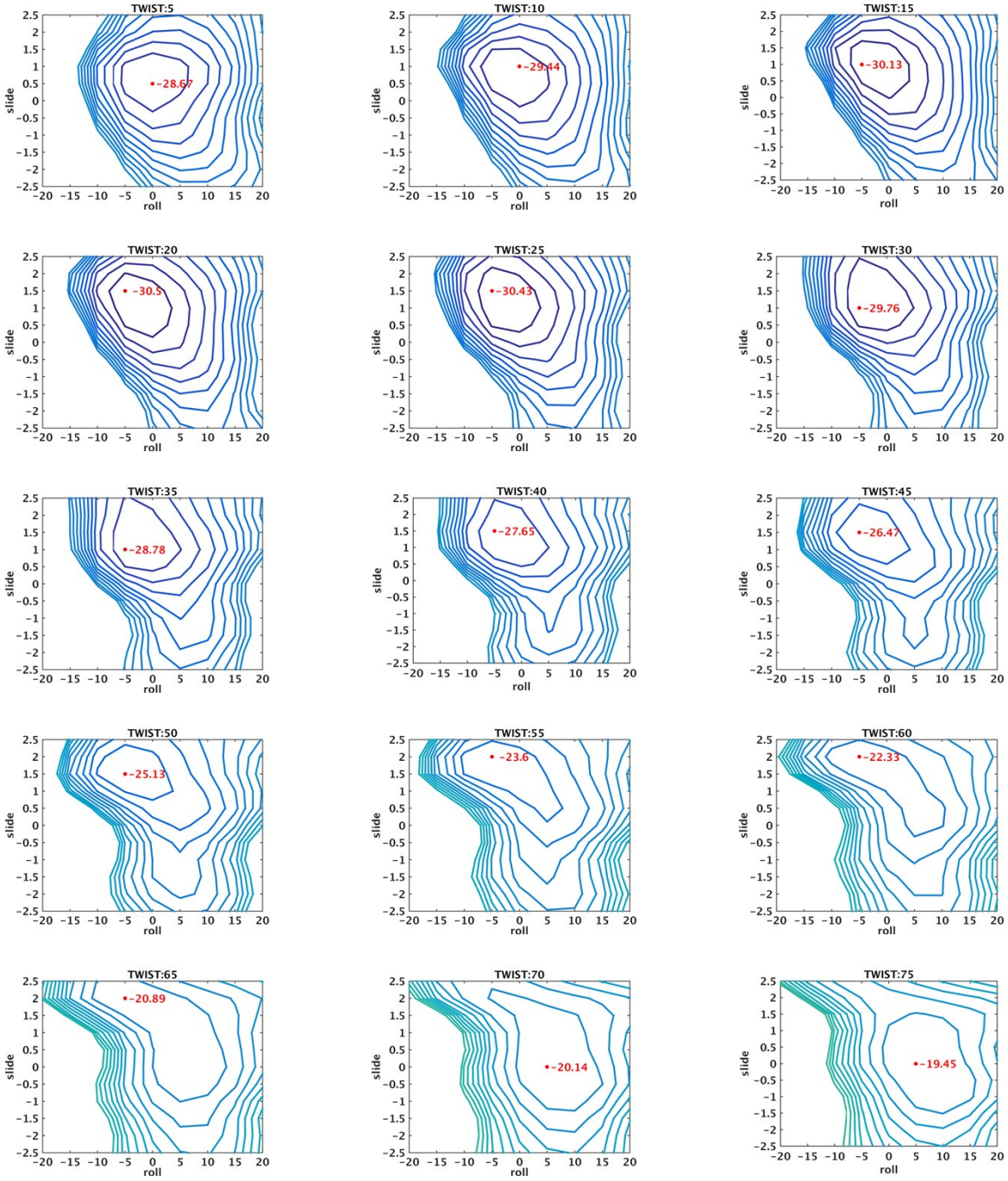


Figure SI-6: Intrinsic stacking energy contours considering DFT-D energy at different twist values for G:C W:WC::G:A S:HT dinucleotide step are shown here. Color bar is given in kcal/mol. The contour lines are 1kcal/mol apart.


TWIST:5
-28.67
TWIST:10
-29.44
TWIST:15
-30.13
TWIST:20
-30.5
TWIST:25
-30.43
TWIST:30
-29.76
TWIST:35
-28.78
TWIST:40
-27.65
TWIST:45
-26.47
TWIST:50
-25.13
TWIST:55
-23.6
TWIST:60
-22.33
TWIST:65
-20.89
TWIST:70
-20.14
TWIST:75
-19.45
slide
roll

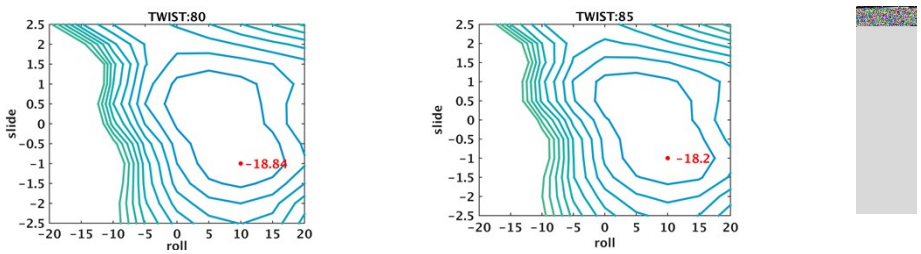
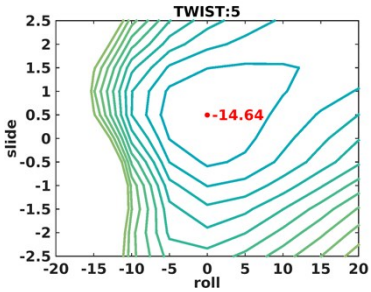
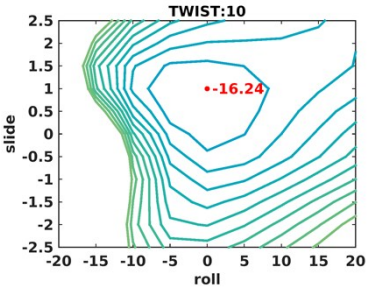
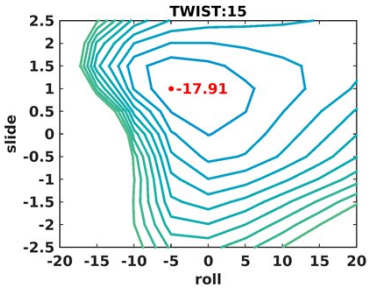
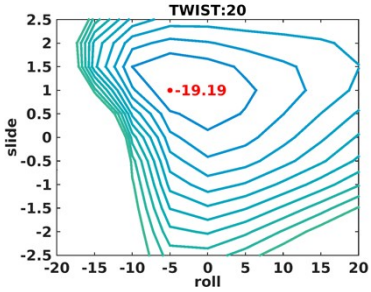


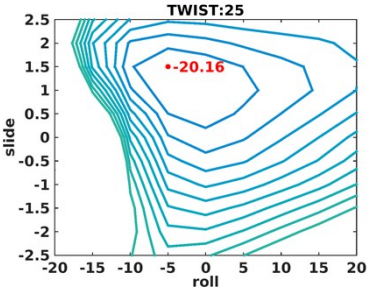
Figure SI-7: Intrinsic stacking energy contours considering DFT-D energy at different twist values for C:G W:WC::G:A S:HT dinucleotide step are shown here. Color bar is given in kcal/mol. The contour lines are 1kcal/mol apart.

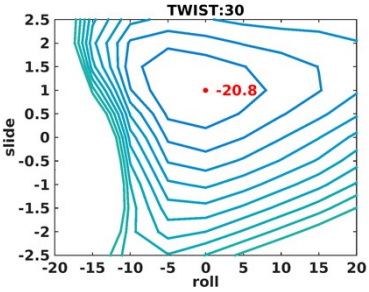

TWIST:5
•-14.64
slide
roll

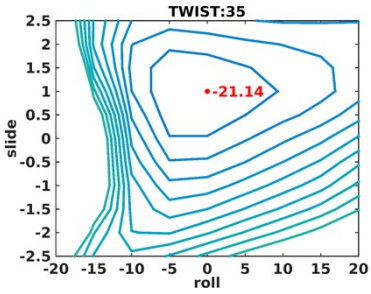

TWIST:10
•-16.24
slide
roll

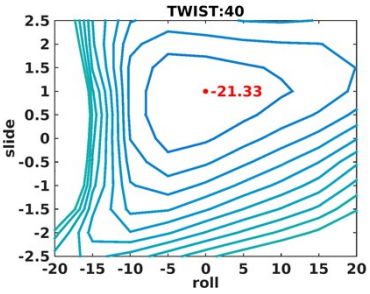

TWIST:15
•-17.91
slide
roll

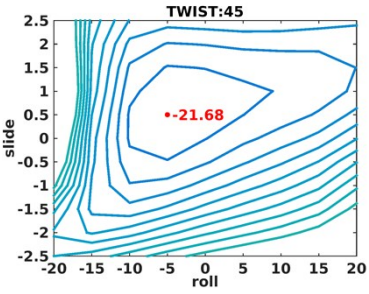

TWIST:20
•-19.19
slide
roll

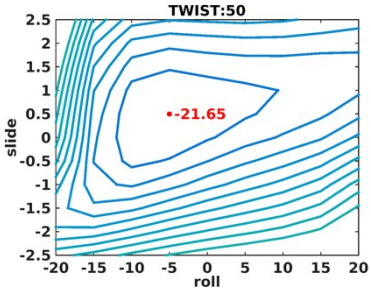

TWIST:25
•-20.16
slide
roll

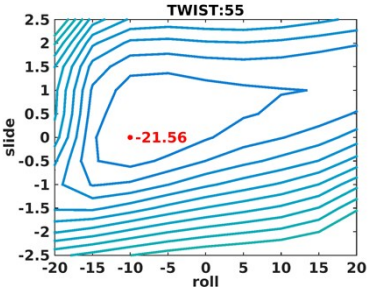

TWIST:30
• -20.8
slide
roll

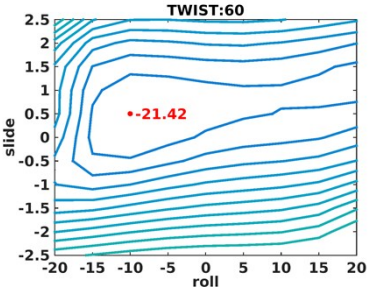

TWIST:35
•-21.14
slide
roll


TWIST:40
•-21.33
slide
roll


TWIST:45
•-21.68
slide
roll


TWIST:50
•-21.65
slide
roll


TWIST:55
•-21.56
slide
roll


TWIST:60
•-21.42
slide
roll

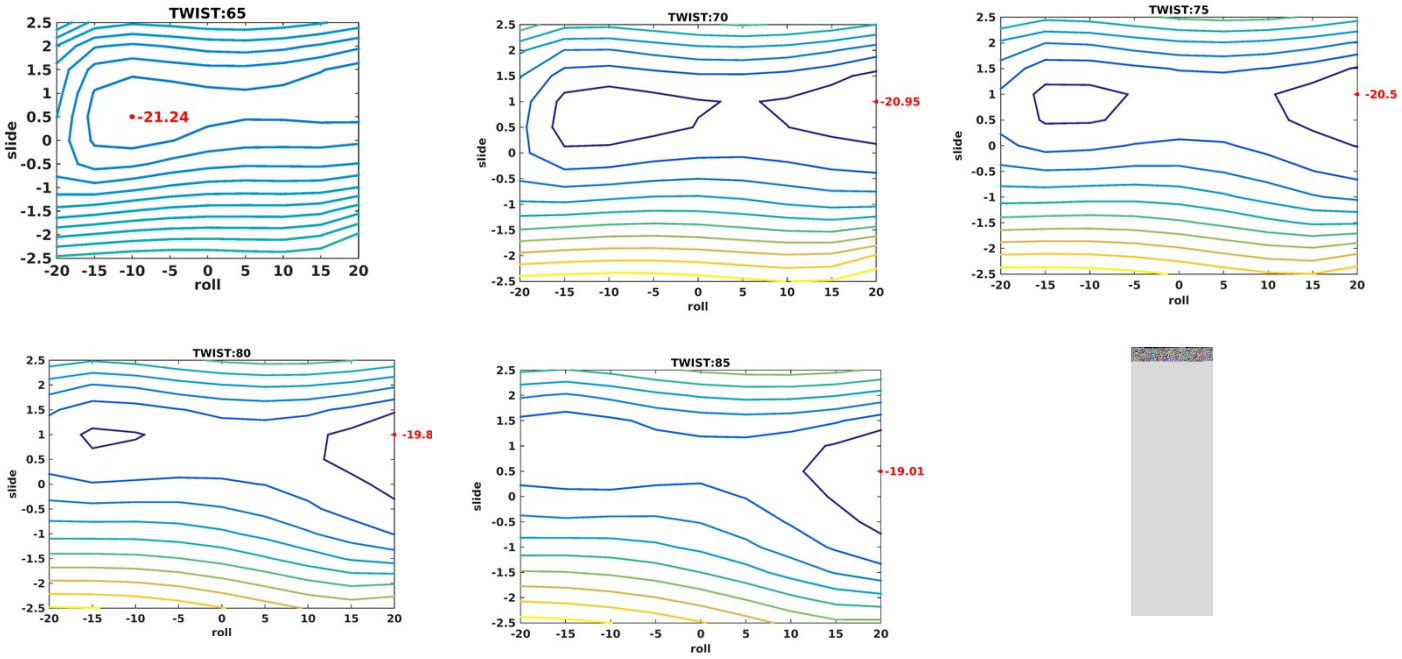
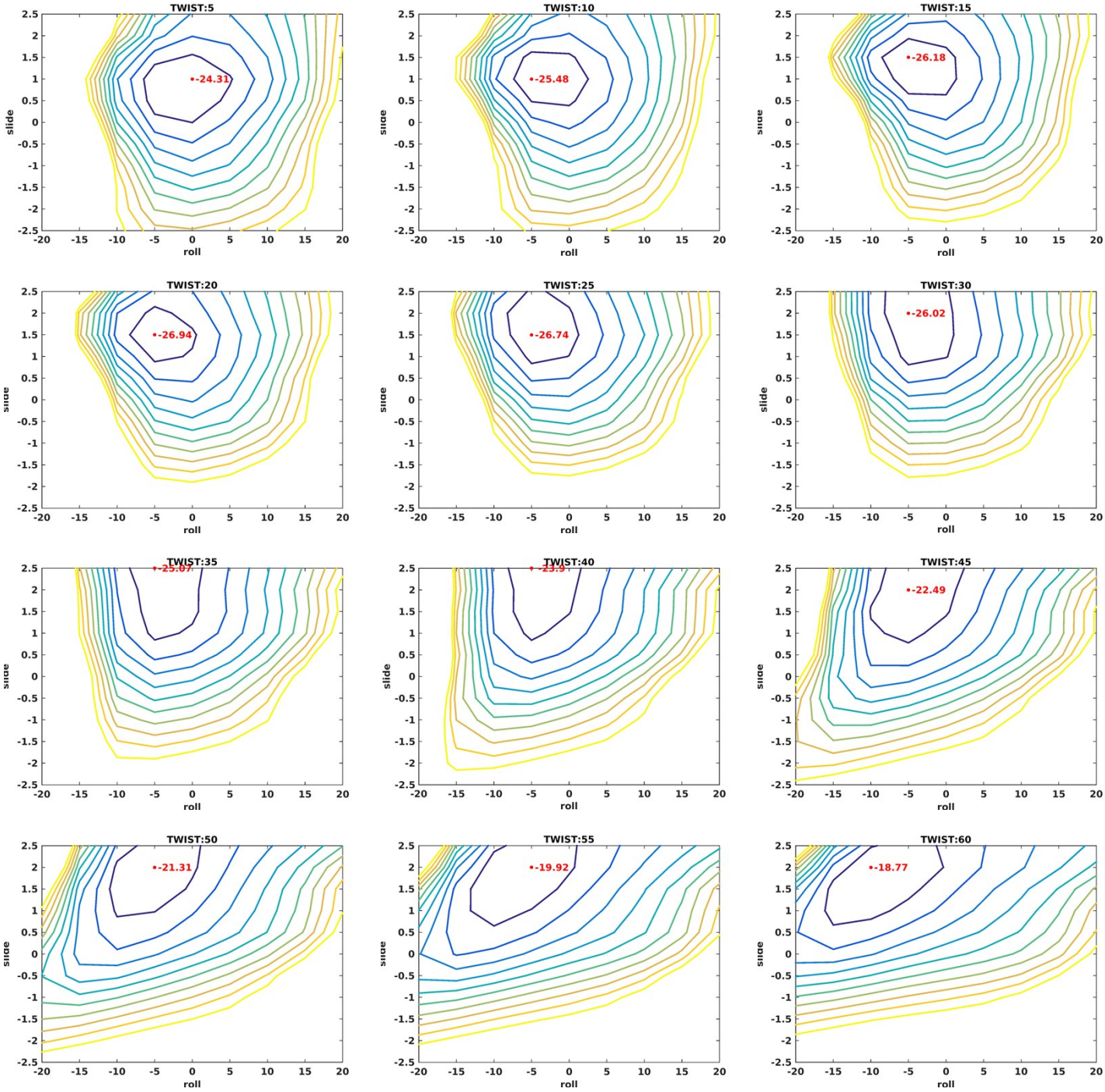


Figure SI-8: Intrinsic intra-strand stacking energy contours, i.e interaction energy between the two Gua residues and those between the Cyt and Ade residues (see Figure 1 for detail), considering DFT-D energy at different twist values for of G:C W:WC::G:A S:HT dinucleotide step are shown here. Color bar is given in kcal/mol. The contour lines are 1kcal/mol apart.


TWIST:5
•-24.31
TWIST:10
•-25.48
TWIST:15
•-26.18
TWIST:20
•-26.94
TWIST:25
•-26.74
TWIST:30
•-26.02
TWIST:35
•-25.07
TWIST:40
•-23.9
TWIST:45
•-22.49
TWIST:50
•-21.31
TWIST:55
•-19.92
TWIST:60
•-18.77
slide
roll

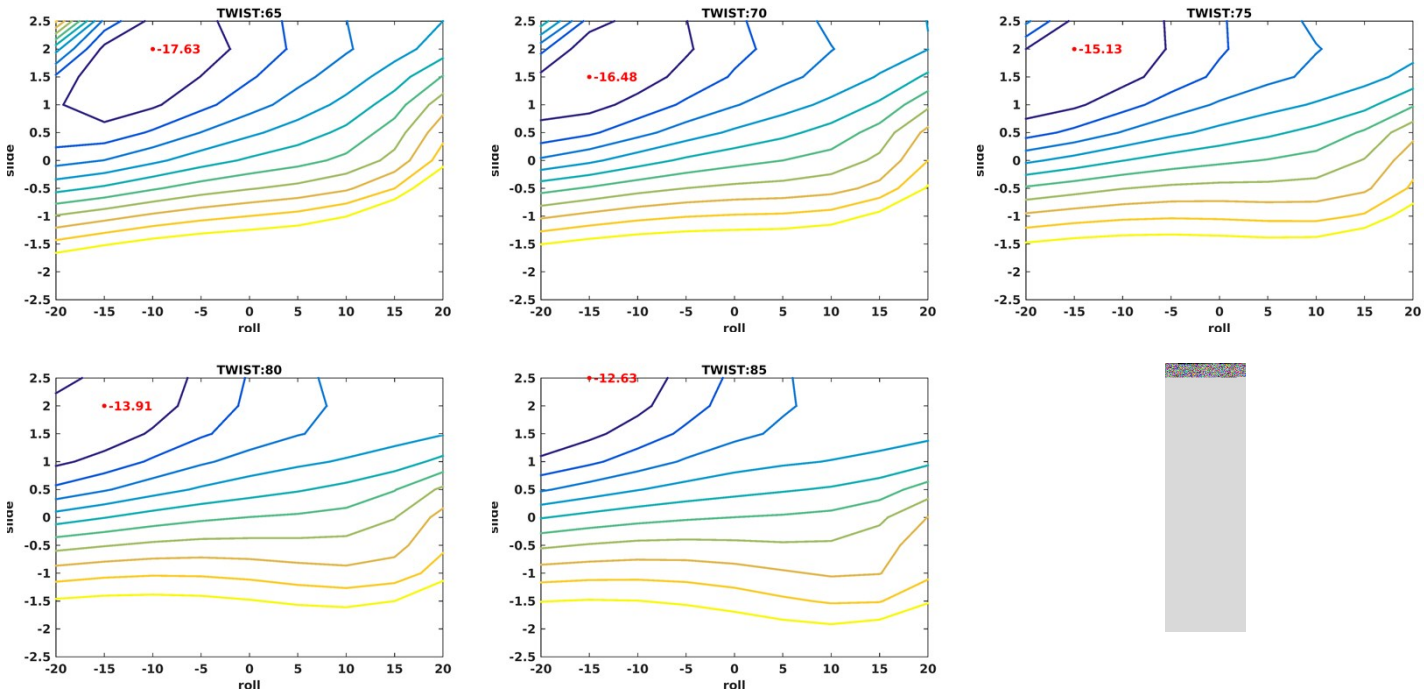
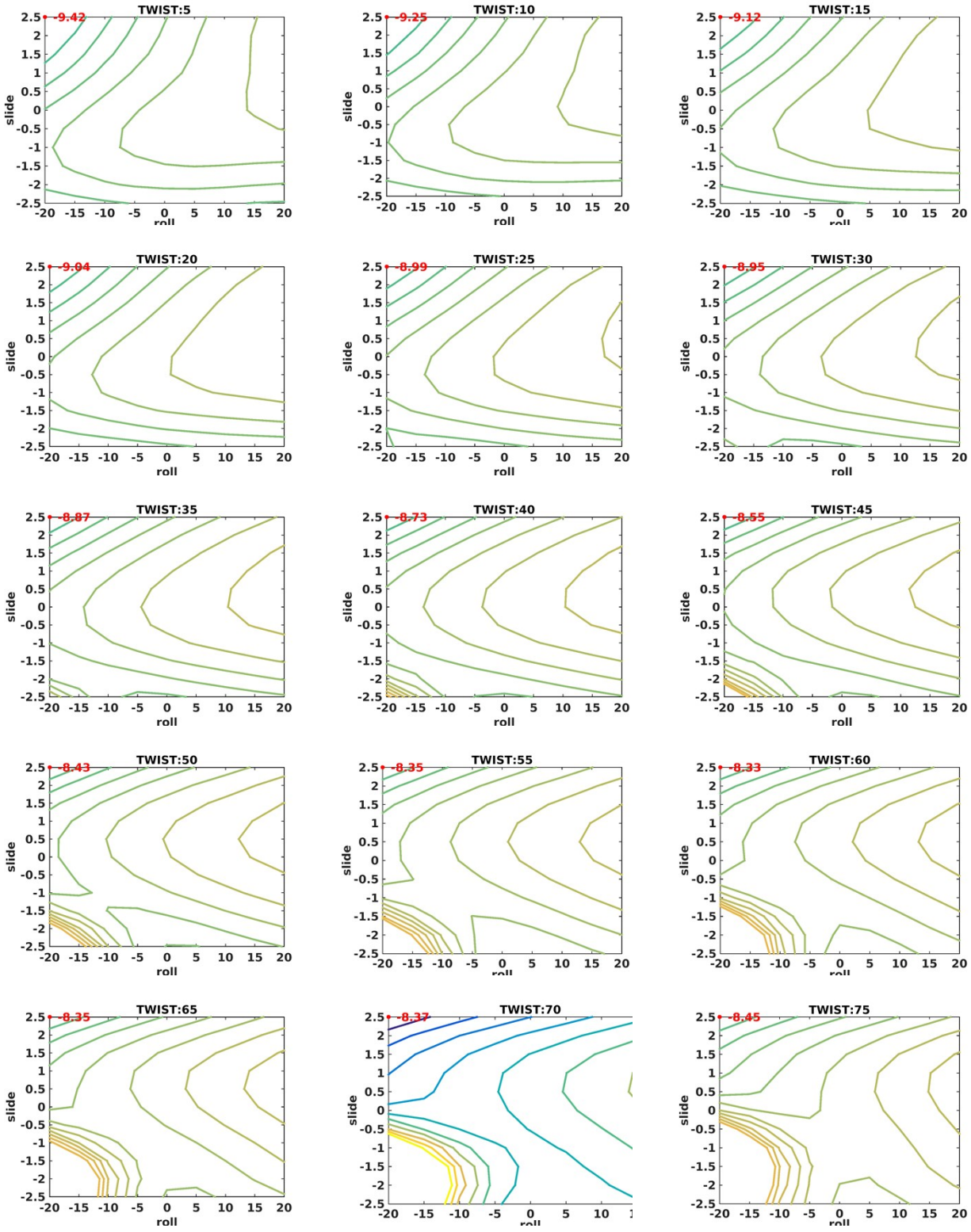


Figure SI-9: Intrinsic intra-strand stacking energy contours considering DFT-D energy at different twist values for C:G W:WC::G:A S:HT dinucleotide step are shown here. Color bar is given in kcal/mol. The contour lines are 1kcal/mol apart.


TWIST:5
-9.42
TWIST:10
-9.25
TWIST:15
-9.12
TWIST:20
-9.04
TWIST:25
-8.99
TWIST:30
-8.95
TWIST:35
-8.87
TWIST:40
-8.73
TWIST:45
-8.55
TWIST:50
-8.43
TWIST:55
-8.35
TWIST:60
-8.33
TWIST:65
-8.35
TWIST:70
-8.37
TWIST:75
-8.45
slide
roll

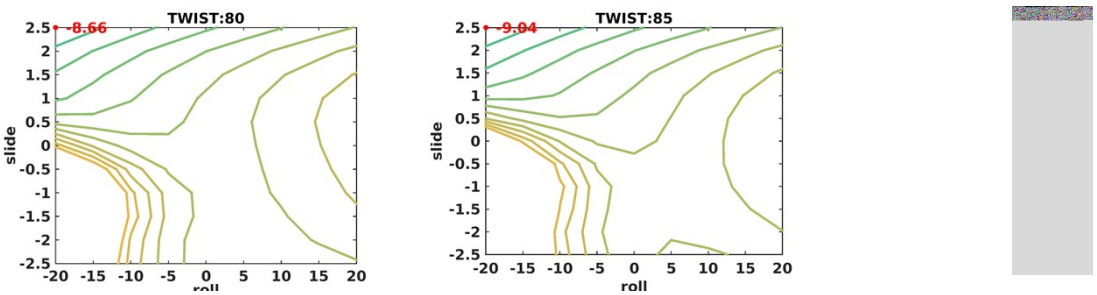
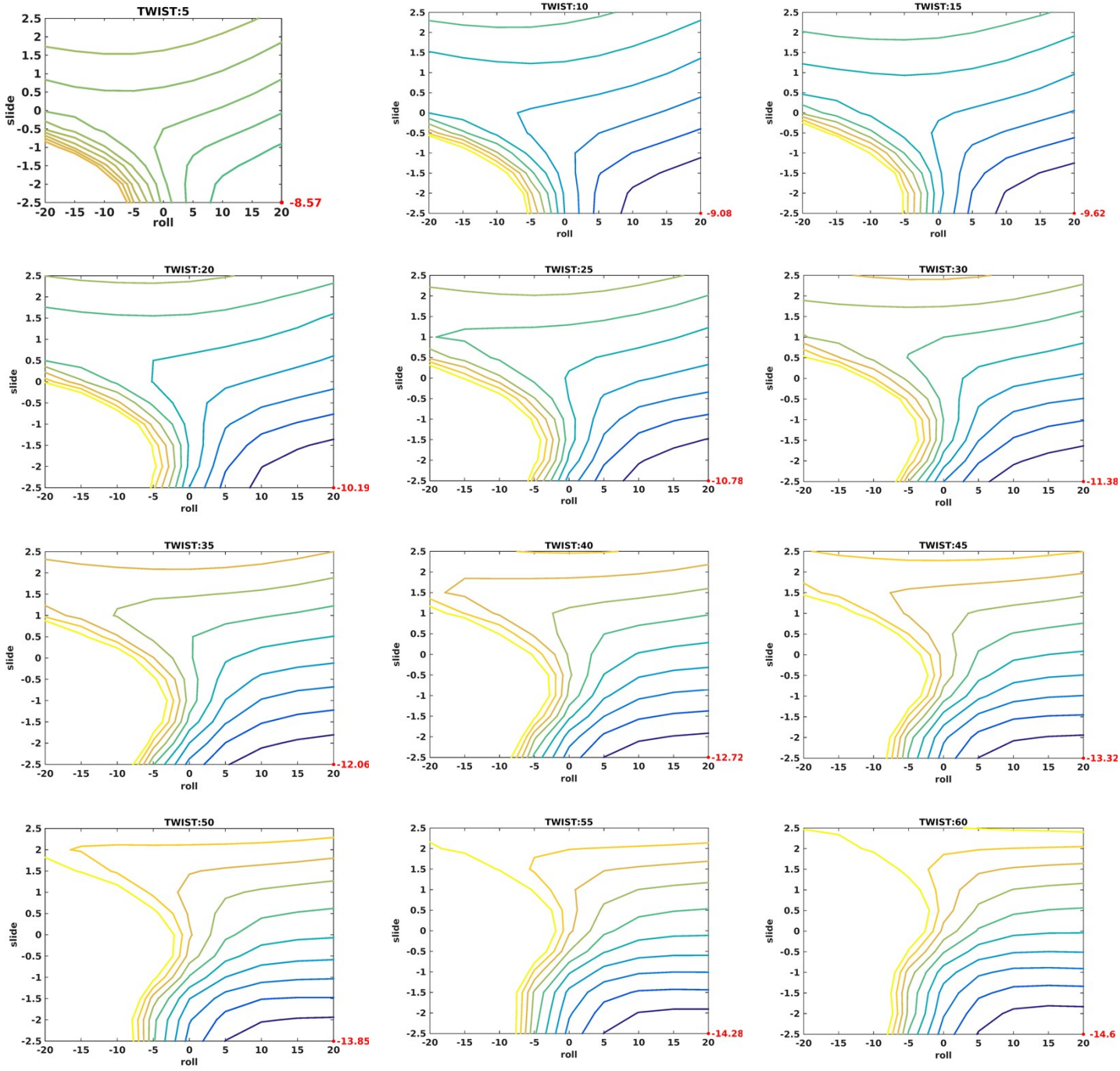


Figure SI-10: Intrinsic inter-strand stacking energy contours considering DFT-D energy at different twist values for of G:C W:WC::G:A S:HT dinucleotide step are shown here. Color bar is given in kcal/mol. The contour lines are 1kcal/mol apart.


TWIST:5
TWIST:10
TWIST:15
TWIST:20
TWIST:25
TWIST:30
TWIST:35
TWIST:40
TWIST:45
TWIST:50
TWIST:55
TWIST:60
slide
roll
-8.57
-9.08
-9.62
-10.19
-10.78
-11.38
-12.06
-12.72
-13.32
-13.85
-14.28
-14.6

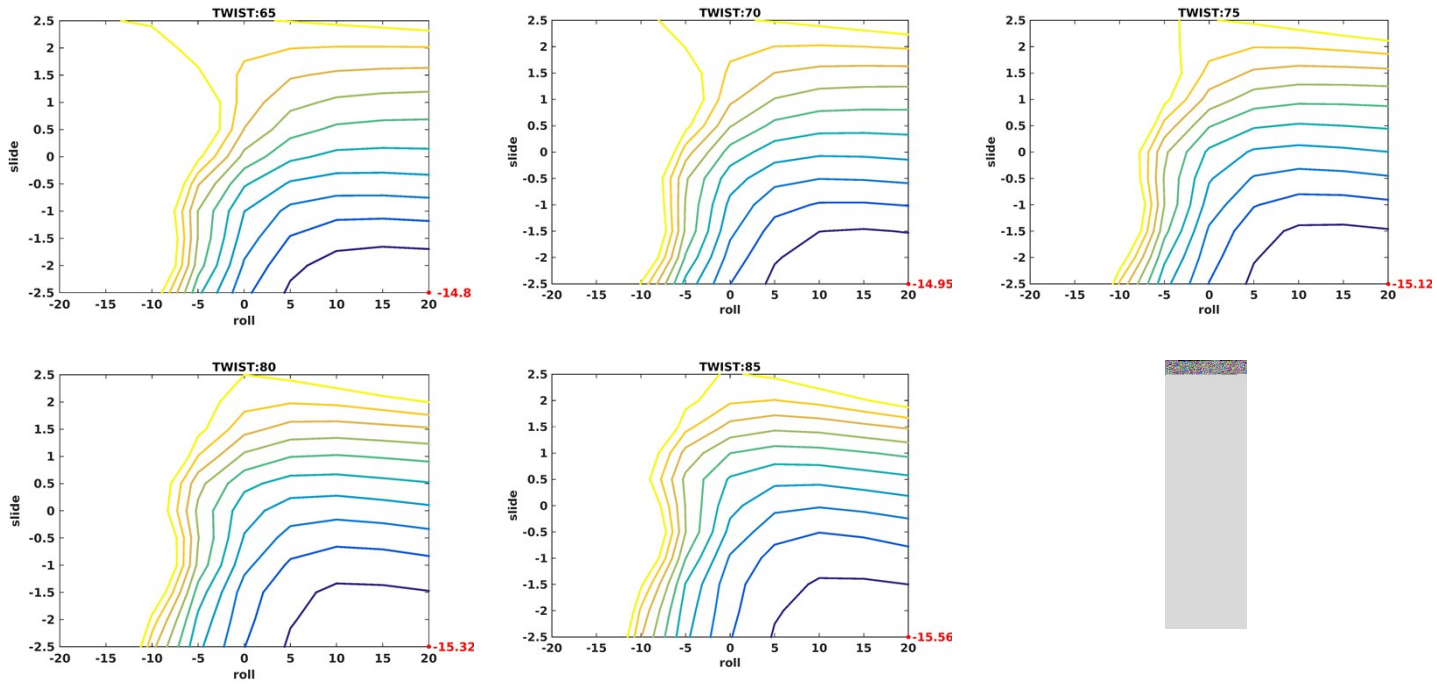


Figure SI-11: Intrinsic inter-strand stacking energy contours considering DFT-D energy at different twist values for of C:G W:WC::G:A S:HT dinucleotide step are shown here. Color bar is given in kcal/mol. The contour lines are 1kcal/mol apart.

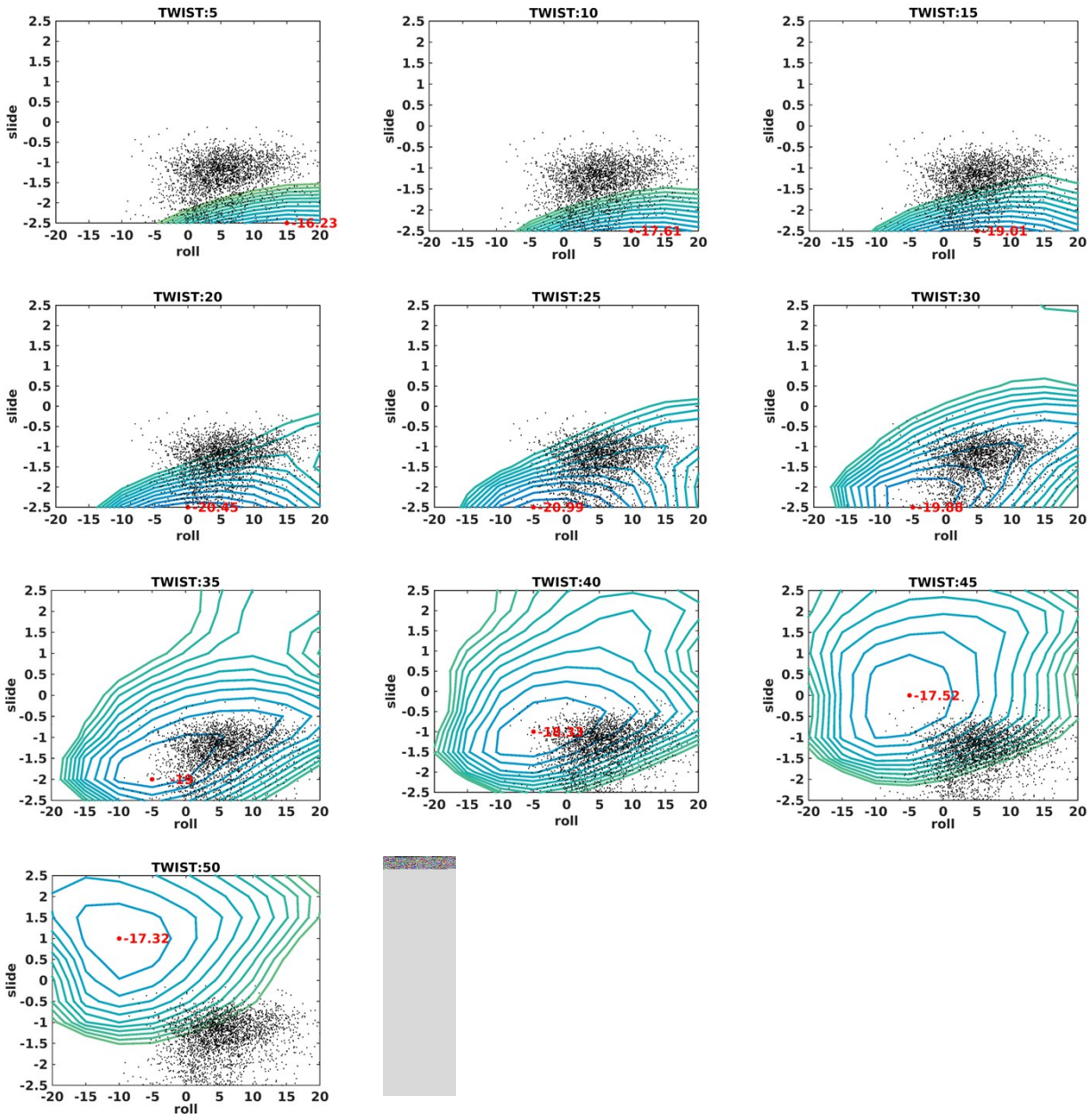
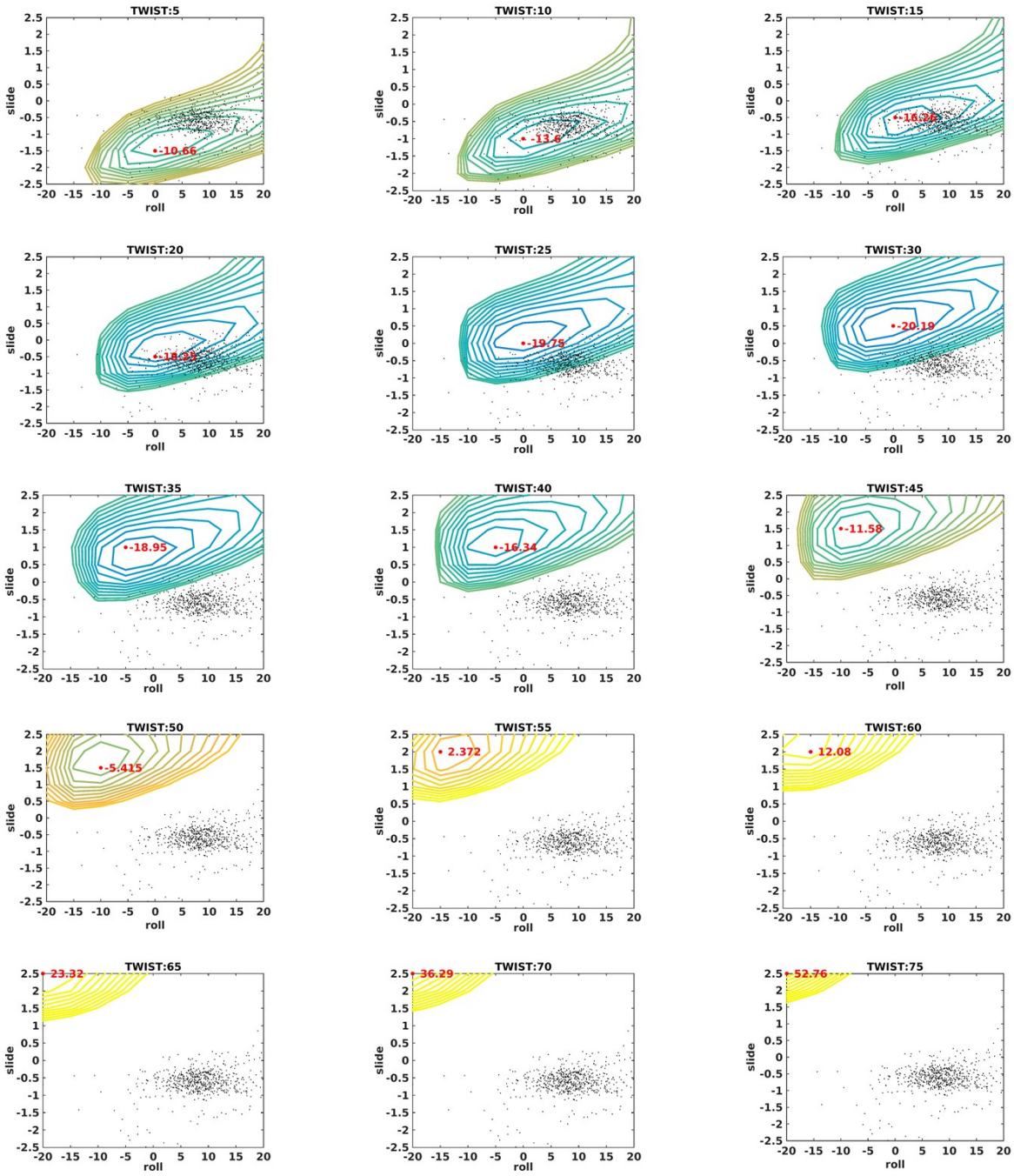


Figure SI-12: Intrinsic total stacking energy contours for A:U W:WC::C:G W:WC dinucleotide steps considering DFT-D energies and coarse grain penalty energy values at various twist values. Color bar is presented in kcal/mol. Difference between two adjacent contour lines is 1 kcal/mol. Roll and slide values from crystal structure database are marked by black point. The minimum energy conformations for every twist values are represented by red dots.


TWIST:5
-10.66
TWIST:10
-13.6
TWIST:15
-16.26
TWIST:20
-18.23
TWIST:25
-19.75
TWIST:30
-20.19
TWIST:35
-18.95
TWIST:40
-16.34
TWIST:45
-11.58
TWIST:50
-5.415
TWIST:55
2.372
TWIST:60
12.08
TWIST:65
23.32
TWIST:70
36.29
TWIST:75
52.76
slide
roll

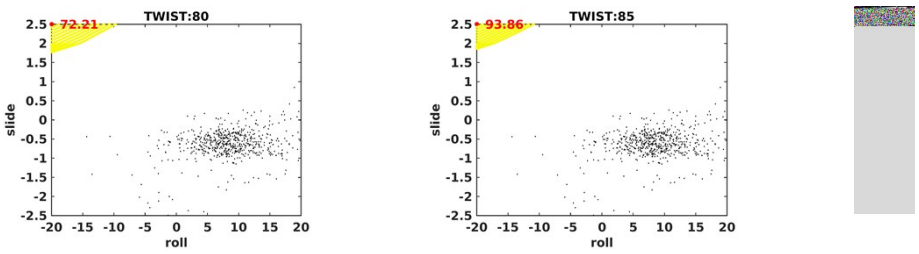
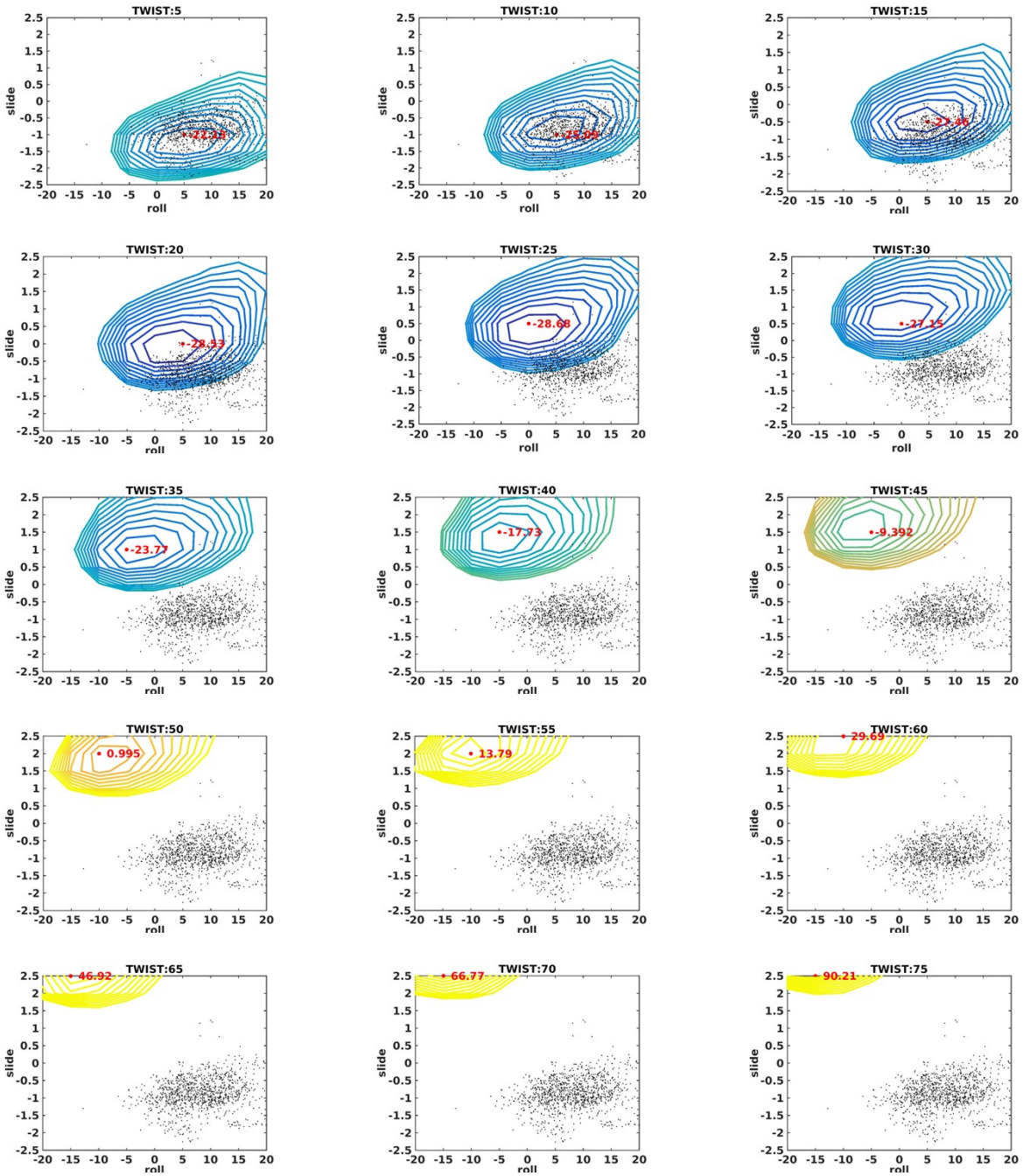


Figure SI-13: Total stacking energy contours considering DFT-D and coarse grain penalty energy at different twist values for G:C W:WC::G:A S:HT dinucleotide step are shown here. Color bar is given in kcal/mol. The contour lines are 1kcal/mol apart.


TWIST:5
TWIST:10
TWIST:15
TWIST:20
TWIST:25
TWIST:30
TWIST:35
TWIST:40
TWIST:45
TWIST:50
TWIST:55
TWIST:60
TWIST:65
TWIST:70
TWIST:75
slide
roll

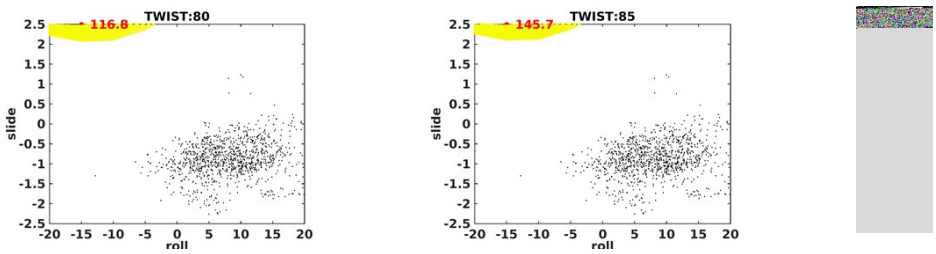


Figure SI-14: Total **s**tacking energy contours considering DFT-D energy at different twist values for C:G W:WC::G:A S:HT dinucleotide step are shown here. Color bar is given in kcal/mol. The contour lines are 1kcal/mol apart.

Table SI-1: Mean C1'…C1' distances, calculated from X-ray crystallographic database, along with their standard deviations (in parenthesis), for the two strands of the two sequences and the calculated force-constants values for stretching of the pseudo-bonds. See Figure 1 for definitions of strand directions.

Sequence	First Strand(C1'…C1')		Second Strand(C1'…C1')	
	Mean Distance (Å)	Force Constant (kcal/mol/Å^2)	Mean Distance (Å)	Force Constant (kcal/mol/Å^2)
C:G W:WC::G:A S:HT	C-G: 5.25(0.23)	15.72	A-G: 5.03(0.23)	16.00
G:C W:W:C::G:A S:HT	G-G: 5.36(0.20)	19.79	A-C: 5.13(0.25)	13.09
A:U W:WC::C:G W:WC	A-C: 5.38(0.30)	9.20	G-U: 5.63(0.32)	8.33

Table SI-2: Best stacking energy (DFT-D) for all twist values for the dinucleotide step sequences are shown below. Mean values of C1'…C1' distances (Å) from X-ray crystal structure database are also shown in parenthesis in the last two columns.

A:U W:WC::C:G W:WC:

Twist (°)	Roll (°)	Slide (Å)	Best Stacking Energy (DFT-D component) (kcal/mol)	C1'…C1' Distance(Å)	
				1st Strand (5.38)	2nd Strand (5.63)
5	0	-2.5	-22.35	4.45	4.30
10	0	-2.5	-22.49	4.66	4.5
15	0	-2.5	-22.38	4.91	4.77
20	-5	-2.5	-22.15	5.05	4.91
25	-5	-2.5	-21.63	5.37	5.24
30	-10	0	-20.52	4.23	4.05
35	-10	0.5	-19.94	4.45	4.26
40	-10	0.5	-19.35	4.82	4.64
45	-10	1	-18.68	5.08	4.90
50	0	-2.5	-18.22	7.33	7.20

Table SI-3: Geometrical and energy parameters for the structures of the dinucleotide step sequence having minimum energy (sum of DFT-D and coarse-grain energies) corresponding to each twist value are listed. Mean values of C1'…C1' distances (Å) from X-ray crystal structure database are also shown in parenthesis in the last two columns.

A:U W:WC::C:G W:WC:

Twist (°)	Roll (°)	Slide (Å)	Best Stacking Energy (DFT-D and coarse grain penalty component) (kcal/mol)	Penalty Value(Kcal/mol)	C1'…C1' Distance(Å)	
					1st Strand (5.38)	2nd Strand (5.63)
5	15	-2.5	-16.23	3.69	4.96	4.80
10	10	-2.5	-17.62	3.36	4.99	4.83
15	5	-2.5	-19.01	2.59	5.07	4.92
20	0	-2.5	-20.45	1.52	5.20	5.06
25	-5	-2.5	-21.00	0.64	5.37	5.24
30	-5	-2.5	-19.88	0.53	5.72	5.58
35	-5	-2	-19.00	0.63	5.75	5.61
40	-5	-1	-18.33	0.36	5.56	5.40
45	-5	0	-17.53	0.43	5.51	5.34
50	-10	1	-17.32	0.52	5.46	5.29

Table SI-4: Best stacking energy (DFT-D) and dispersion energy components in configuration having maximum stacking overlap area for all twist values for the dinucleotide step sequences are shown below.

a) **G:C W:WC::G:A S:HT**

Twist (°)	Roll (°)	Slide (Å)	Maximum Overlap Area (Å^2)	Best Stacking Energy (kcal/mol)	Dispersion Energy (kcal/mol)
5	-20	1.0	48.13	2.23	-21.66
10	-20	1.0	49.42	-0.58	-21.95
15	-20	1.0	49.77	1.84	-22.03
20	-20	1.0	50.22	4.26	-21.74
25	-20	0.5	49.90	16.96	-20.91
30	-20	1.0	49.88	4.11	-20.81
35	-20	1.0	48.73	4.13	-20.41
40	-20	1.0	47.25	1.53	-20.06
45	-20	1.0	45.40	-6.46	-19.56
50	-20	2.5	44.43	-14.37	-18.97
55	-20	2.5	43.93	-16.10	-18.47
60	-20	2.5	42.77	-18.17	-17.95
65	-20	2.5	41.88	-20.13	-17.38
70	-20	2.5	40.65	-21.39	-16.78
75	-20	2.5	39.40	-21.81	-16.18
80	-20	2.0	37.83	-22.19	-15.55
85	-20	2.5	36.50	-21.45	-15.09

b) **C:G W:WC :: G:A S:HT**

Twist (°)	Roll (°)	Slide (Å)	Maximum Overlap Area ($Å^2$)	Best Stacking Energy (kcal/mol)	Dispersion Energy (kcal/mol)
5	-20	1	53.30	3.46	-24.07
10	-20	1	54.02	-0.21	-24.29
15	-20	1	54.18	1.32	-24.29
20	-20	1	54.43	3.52	-23.94
25	-20	1	53.88	4.02	-23.42
30	-20	0.5	53.08	10.25	-22.35
35	-20	0.5	51.72	11.98	-21.82
40	-20	0.5	50.93	10.46	-21.26
45	-20	0.5	49.73	8.25	-20.63
50	-20	0.5	48.92	13.32	-20.06
55	-20	0.5	47.35	25.50	-19.50
60	-20	0.5	45.88	39.83	-18.76
65	-20	1	44.33	21.08	-18.49
70	-20	2	43.77	-8.60	-18.44
75	-20	2	43.38	-3.28	-17.87
80	-20	2.5	42.67	-8.38	-17.60
85	-20	2.5	42.80	-4.73	-17.20