Supporting Information: Partial base flipping is sufficient for strand slippage near DNA duplex termini

Nilesh K. Banavali[1,2]

1: Laboratory of Computational and Structural Biology

Division of Genetics, Wadsworth Center, NYS Department of Health

2: Department of Biomedical Sciences, School of Public Health

State University of New York at Albany

email: banavali@wadsworth.org

CMS 2008, 150 New Scotland Ave, Albany, NY 12208

Tel: 518-474-0569

Fax: 518-402-4623

Website: http://www.wadsworth.org/resnres/bios/banavali.htm

Keywords: base-pair separation, template slippage, base flipping, force fields, base stacking, torsions, molecular dynamics, umbrella sampling, potential of mean force

The Journal of the Americal Chemical Society

Spontaneous cytosine flipping, restacking-pairing, and slippage in the metastable state

The metastable state involving a triplet interaction between the partially flipped cytosine and a neighboring G:C base pair was identified through a potential of mean force (PMF) profile of cytosine base flipping constructed using restrained molecular dynamics (MD) simulations. Although a similar state is known to occur experimentally in the large ribosomal subunit RNA in *Haloarcula marismortui*,[1] an influence of the restraint force applied in these simulations on the perceived stability of this structural state cannot be ruled out completely. To understand the behavior of this structural state in the absence of the restraint and without a force field bias, multiple unrestrained MD simulations were performed using different DNA force fields starting from two umbrella sampling windows near this metastable state in the original PMF profile. The initial conformation of the solute and solvent in all the simulations from one window was the same, and at least 5 simulations were performed using the same force field, but with different random initial velocities assigned to all the atoms.

Flipping. The resulting dynamical behavior of the metastable state for unrestrained simulations of two windows from the original base flipping potential of mean force (PMF) simulations, monitored in terms of the pseudo-dihedral coordinate describing the flipping of the third cytosine, is shown in Supporting Figure 1 and Supporting Figure 2. The percentage breakdown of the sampling of the pseudo-dihedral in terms of possible restacked conformations, pre-stacked conformations that are intermediate between the metastable state and stacked conformations, conformations remaining near the starting metastable state, or conformations that demonstrate greater flipping of the cytosine, are shown in Supporting Table 1 and Supporting Table 2. Different MD simulations from the same starting conformation and force field did not necessarily show similar pseudo-dihedral coordinate dynamics. The simulations using the CHARMM27 force field showed maintenance of the metastable partially flipped conformation, further flipping of the cytosine, or possible motion towards restacking through the minor groove. For the AMBER94 force field simulations, sampling was mostly divided between maintenance of the metastable state or slight motion towards restacking through the minor groove. For the AMBER99 force field simulations, which were performed in 10 replicas, maintenance of the metastable state, further flipping of the cytosine, or possible motion towards restacking were

all observed. The AMBERBSC0 force field simulations had sampling divided between maintenance of the metastable state or motion towards restacking, with very little further flipping of the cytosine. The BMS force field simulations also showed maintenance of the metastable state or motion towards restacking with very little further flipping of the cytosine. The CHARMM36 force field simulations showed maintenance of the metastable state, motion towards restacking, and also further flipping of the cytosine, but seemed to show more tendency for restacking than the CHARMM27 force field. The percentages reported in Supporting Table 1 and Supporting Table 2 show that the CHARMM27 force field allowed further flipping the most, while the AMBERBSC0 and BMS fields favored it least. Correspondingly, the AMBERBSC0 and BMS force fields seemed to favor motion towards restacking, and a generally lower range of motion in the pseudo-dihedral coordinate.

Restacking-pairing. A pseudo-dihedral coordinate value near 10 degrees does not necessarily indicate a restacked and Watson-Crick paired conformation, as the coordinate measures relative orientation of four sets of atoms, and can attain these values even in the absence of proper restacking with Watson-Crick base pairing, especially in contexts other than standard B-form DNA duplexes. To assess restacking combined with proper base pairing (restacking-pairing), a coordinate composed of the following 5 distances was used: (1) between the O2 atom in template strand C3 and the N2 atom in primer strand G7, (2) between the N3 atom in template strand C3 and the N1 atom in primer strand G7, (3) between the N4 atom in template strand C3 and the O6 atom in primer strand G7, (4) between the centers-of-masses of template strand bases C2 and C3, and (5) between the centers-of-masses of template strand bases C3 and G4. The cutoff for the presence of base-pairing hydrogen bonds in the first three distances was 3.5 Å and the cutoff for presence of stacking interactions through the last two distances was 5 Å. Satisfaction of each of these cutoffs amounted to a value of 0.2 in the restacking coordinate, which meant complete restacking could be represented by coordinate values approaching 1. The restacking-pairing coordinate time series, in unrestrained simulations of the two windows from the original PMF simulations, is shown in Supporting Figure 3 and Supporting Figure 4. Proper restacking-pairing does not occur in all simulations where the pseudo-dihedral coordinate values approaching the stacked state values near 10 degrees. A total of 42 simulations (panels A2, B2, B4, C1, C2, C6, C8, C9, C10, D1, D2, D5, E1, E3, E5, F2, F3 in Supporting Figure 1, and panels A2, A3, A4,

A5, B1, B2, B3, B4, B5, C1, C2, C3, C4, C5, C6, C7, C8, D1, D3, D5, E2, F1, F2, F3, F5 in Supporting Figure 2) seem to show at least some tendency to restack in the pseudo-dihedral coordinate. As judged through the restacking-pairing coordinate, however, proper restacking-pairing only occurs for four of these simulations (panel E5 in Supporting Figure 3, panels C6, D3, F2 in Supporting Figure 3). The rest of the simulations might be restacking incorrectly, i.e. without reformation of Watson-Crick base pairing, but this possibility was not assessed in the present analysis.

Template slippage. To evaluate spontaneous template slippage in these unrestrained simulations, a coarse-grained coordinate composed of six distances was used. These distances were between: (1) the N3 atom in template strand C1 and the N1 atom in primer strand G9, (2) the N3 atom in template strand C2 and the N1 atom in primer strand G8, (3) the N3 atom in template strand C1 and the N1 atom in primer strand G8, (4) the N3 atom in template strand C2 and the N1 atom in primer strand G7, (5) the center-of-masses of template strand C1 and G4 bases, and (6) the center-of-masses of template strand C2 and G4 bases. The following cutoffs for each of these distances were imposed to classify them as template slipped (in the same order as above): (1) greater than 3.4 Å, (2) greater than 3.4 Å, (3) less than 3.4 Å, (4) less than 3.4 Å, (5) less than 9.0 Å, and (6) less than 6.0 Å. If any one of these distance cutoffs was satisfied, then the coordinate was incremented by 0.17, which showed the stages of template slippage as a function of time. The behavior of the slippage coordinate for the same MD simulations is monitored in Supporting Figure 5 and Supporting Figure 6 and shows that template slippage occurs unambiguously in three simulations started from the 60 degree umbrella sampling window (panels B2, C8, and C10 in Supporting Figure 5), and in three simulations started from the 90 degree umbrella sampling window (A2, A5, and B2 in Supporting Figure 6). Thus, spontaneous template slippage was observed in response to partial base flipping of a cytosine two bases removed from the 5' terminus in simulations using three different force fields: CHARMM27, AMBER94, and AMBER99. In addition, almost complete or transient template slippage, defined as values at or greater than 0.83 in the slippage coordinate, was also observed in two additional simulations from the 60 degree window (panels A5 and B1 in Supporting Figure 5) and four simulations from the 90 degree window (panels C4, C10, D2, and E2 in Supporting Figure 6). These data suggest that template slippage due to partial base flipping near the DNA duplex termini is likely not an artifact of added restraints or force field bias.

Spontaneous guanine flipping, restacking-pairing, and primer slippage from a minor groove flipped state

The possibility of single base addition mutations during replication of the same hot-spot sequence[2] suggests the possibility of primer slippage in addition to template slippage. A possible mechanism for such primer slippage could be a partially flipping of a base in the primer strand followed by slippage of the neighboring primer strand bases to fill the gap. The flipping PMF for the guanine opposite the third template strand cytosine does not show a minor groove metastable state in the same vicinity as the flipping PMF for the cytosine (Figure 1 in the main article). This presents an opportunity to probe the effect of the lack of a metastable state for the same extent of partial flipping on the propensity for slippage. Just as for cytosine flipping, a guanine minor groove flipping intermediate window corresponding to the pseudo-dihedral coordinate value of 335 degrees was chosen, and multiple unrestrained MD simulations using different DNA force fields were started from this conformation. This particular window was chosen because it was roughly the same minor groove deviation in pseudo-dihedral coordinate values from the stacked state as the windows used for the cytosine flipped unrestrained simulations (around 50 degrees). The group of atoms representing the first center-of-mass in the pseudo-dihedral restraint (i.e. the C4:G6 base pair) was kept the same for both cytosine and guanine flipping. Since this is the 5' base pair for the guanine and the 3'-base pair for the cytosine, the DNA groove direction of flipping represented by the pseudo-dihedral coordinate was opposite for the two bases. Flipping through the minor groove for the guanine occurred in the range below 15 degrees and above 330 degrees in the periodic pseudo-dihedral. As before, the initial conformation of the solute and solvent in all the simulations was identical, and at least 5 simulations were performed for each force field, with different random initial velocities assigned to all the atoms in each simulation.

Flipping. The corresponding pseudo-dihedral coordinate time series describing guanine flipping is shown in Supporting Figure 7. The percentage breakdown of the sampling of the pseudo-dihedral in terms of possible restacked conformations, pre-stacked conformations intermediate between the metastable state and stacked conformations, conformations remaining near the starting state, or conformations that demonstrate greater guanine flipping, are shown in Supporting Table 3. In contrast to cytosine flipping, simulations

using the CHARMM27 force field showed substantial possible restacking in 3 simulations, substantial residence near the starting state in 2 simulations, and substantial further flipping in only 1 simulation. For the AMBER94 force field simulations, substantial restacking occurred in 1 simulation, substantial flipping in another, and the guanine mostly remained close to the starting state in the other 3 simulations. In the AMBER99 simulations, 5 simulations showed rapid and stable possible restacking, 3 showed substantial residence near the starting state, and 2 showed substantial further flipping. In the AMBERBSC0 simulations, 2 simulations showed rapid and stable possible restacking, 2 showed substantial residence near the starting state, and 2 showed further flipping. In the BMS force field simulations, 2 showed rapid and stable possible restacking, 2 showed almost complete maintenance of sampling near the starting state, and only 1 showed substantial further flipping. In the CHARMM36 simulations, 4 simulations showed substantial possible restacking, while only one showed substantial further flipping. The results suggest that guanine restacking from a minor groove flipped intermediate state that is not metastable seems to be much more favored as compared to its partner cytosine.

Restacking-pairing. To assess restacking in conjunction with Watson-Crick base pairing, a coordinate composed of the following 5 distances was used: (1) between the O2 atom in template strand C3 and the N2 atom in primer strand G7, (2) between the N3 atom in template strand C3 and the N1 atom in primer strand G7, (3) between the N4 atom in template strand C3 and the O6 atom in primer strand G7, (4) between the centers-of-mass of primer strand bases C6 and G7, and (5) between the centers-of-mass of primer strand bases G7 and G8. As before, the cutoffs used to determine presence of base-pairing hydrogen bonds and stacking interactions were 3.5 Å and 5 Å, respectively. The restacking behavior in unrestrained simulations is shown in Supporting Figure 8. More than half the simulations (19 out of 35: panels A2-A4, B2, B5, C6-C10, D1, D4, E1-E3, F1-F3, F5) in Supporting Figure 7 seem to show a tendency to restack in the pseudo-dihedral coordinate. Of these, 16 are confirmed to undergo proper restacking through the restacking-pairing coordinate, with only panels E3, F1, F2 in Supporting Figure 8 being the exceptions. In addition, panel C5 in Supporting Figure 8 shows proper restacking-pairing towards the end of the simulation, suggesting that the sampling time for this simulation might increase if the simulation was extended further. This close correspondence between the restacking suggested by the pseudo-dihedral coordinate and the

restacking-pairing coordinate suggests that proper guanine restacking through the minor groove in this sequence context is much more favorable.

Primer slippage. Spontaneous primer slippage was evaluated in these unrestrained simulations through a similar coarse-grained coordinate as template slippage composed of six distances between: (1) the N1 atom in primer strand G9 and the N3 atom in template strand C1, (2) the N1 atom in primer strand G8 and the N3 atom in template strand C2, (3) the N1 atom in primer strand G9 and the N3 atom in template strand C2, (4) the N1 atom in primer strand G8 and the N3 atom in template strand C3, (5) the center-of-masses of primer strand G9 and C6 bases, and (6) the center-of-masses of primer strand G8 and C6 bases. The cutoffs for each of these distances to classify them as primer slipped were also similar (in the same order as above): (1) greater than 3.4 Å, (2) greater than 3.4 Å, (3) less than 3.4 Å, (4) less than 3.4 Å, (5) less than 9.0 Å, (6) less than 6.0 Å. Satisfaction of any one of these distance cutoffs resulted in an increment of 0.17 in the primer slippage coordinate, as shown in Supporting Figure 9. Primer slippage occurs in three simulations (panels A5, C5, and E5 in Supporting Figure 9) using three different force fields: CHARMM27, AMBER99, and BMS. This suggests that spontaneous primer slippage can also occur due to partial base flipping of a guanine two bases removed from the 3'-terminus. In addition, almost complete primer slippage, defined as values at or greater than 0.83 in the slippage coordinate, was also observed in one simulation (panel B5 in Supporting Figure 9) for the AMBER94 force field. An important distinction from the behavior observed during template slippage is that the stability of the primer slipped state seems to be much lower than the template slipped state. This is indicated by the low residence times in the primer slipped state for most of the simulations that attain it. This instability is possibly a consequence of the the lack of non-canonical minor groove base pairing with the neighboring GC base pair for the partially flipped guanine in this sequence context.

Additional pairing/stacking distortions.

Specific structural changes such as restacking-pairing or further flipping of the partialy flipped base and primer or template slippage that are analyzed in the present study may not account for all the structural

heterogeneity that can result from a single base partial flipping near the DNA terminus. The structures observed at the end of the individual 5 ns unrestrained simulations visually illustrate some additional possibilities of pairing and stacking structural distortions. These are shown for the 60 degree cytosine, 90 degree cytosine, and the 335 degree guanine flipped intermediate unrestrained simulations in Supporting Figure 10, Supporting Figure 11, and Supporting Figure 12, respectively. In addition to local distortions in paired bases, and helical distortions such as twisting and bending, two possibilities for coarser pairing or stacking distortions revealed by these structures are: (a) base pair separation and flipping in terminal base pairs beyond the single flipped base that is not associated with slippage (panels A2, A5, F5 in Supporting Figure 10; panels A4, D3 in Supporting Figure 11; and panels A3, A4, in Supporting Figure 12) (b) base pair separation and stacking of two separated partner bases on each other (panels A1, B5, F2 in Supporting Figure 10; panels A3, F3 in Supporting Figure 11; and panels A1, A2, F1 in Supporting Figure 12). Although these distortions are also observed in multiple force fields (CHARMM27, AMBER94, and AMBERBSC0), they appear more frequently in the CHARMM27 force field simulations. Further analysis of experimental data, quantum chemical studies, and MM studies can be used to rule out the possibility that such distortions are due to specific force field inaccuracies, or to identify the source of the inaccuracies if they exist.

Additional material

Supporting Figure 13 illustrates the coarse-grained pseudo-dihedral parameter, in the stacked and single base flipped structural states, that is used in the present study to represent base flipping in a one-dimensional coordinate.

Supporting Table 4 shows the steric overlap in overlaying partially flipped or strand slipped structures into two different polymerase active sites. Without additional optimization, it is clear that the B-family polymerase active site is much more restrictive in allowing the partial flipping or strand slippage distortions. Upon optimization, however, these structures could be accommodated without significant change in protein structure.

Supporting videos 1 to 4 show trajectories illustrating cytosine restacking-pairing, template strand slippage, guanine restacking-pairing, and primer strand slippage, respectively.

References

1. Ban, N.; Nissen, P.; Hansen, J.; Moore, P.; Steitz, T. *Science* **2000,** *289,* 905–920.

2. Potapova, O.; Grindley, N.; Joyce, C. *Journal of Biological Chemistry* **2002,** *277,* 28157–28166.

3. Wilson, R.; Pata, J. *Molecular Cell* **2008,** *29,* 767–779.

4. Xia, S.; Konigsberg, W.; Wang, J. *Journal of the American Chemical Society* **2011,** *133,* 10003–10005.

5. Foloppe, N.; MacKerell Jr, A. *Journal of Computational Chemistry* **2000,** *21,* 86–104.

6. MacKerell, A.; Banavali, N. *Journal of Computational Chemistry* **2000,** *21,* 105–120.

7. Cornell, W.; Cieplak, P.; Bayly, C.; Gould, I.; Merz, K.; Ferguson, D.; Spellmeyer, D.; Fox, T.; Caldwell, J.; Kollman, P. *Journal of the American Chemical Society* **1995,** *117,* 5179–5197.

8. Cheatham III, T.; Cieplak, P.; Kollman, P. *Journal of Biomolecular Structure and Dynamics* **1999,** *16,* 845–862.

9. Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham III, T.; Laughton, C.; Orozco, M. *Biophysical Journal* **2007,** *92,* 3817–3829.

10. Langley, D. *Journal of Biomolecular Structure and Dynamics* **1998,** *16,* 487–509.

Supporting video captions:

Supporting video 1: Visual depiction of a trajectory of a partially flipped template strand cytosine (C3) restacking and Watson-Crick pairing with its primer strand partner guanine base (G7). The terminal C1:G9 base pair is in red, the C2:G8 base pair is in blue, the C3:G7 base pair is in green, and other base pairs are shown in orange. This trajectory is obtained from simulation 5 using the BMS force field and starting from the 60 degree cytosine umbrella sampling window, and has been smoothed at an interval of 1 ps for clarity.

Supporting video 2: Visual depiction of a trajectory starting from a partially flipped template strand cytosine (C3) and resulting in template strand slippage by one base pair. The terminal C1:G9 base pair is in red, the C2:G8 base pair is in blue, the C3:G7 base pair is in green, and other base pairs are shown in orange. This trajectory is obtained from simulation 10 using the AMBER99 force field and starting from the 60 degree cytosine umbrella sampling window, and has been smoothed at an interval of 1 ps for clarity.

Supporting video 3: Visual depiction of a trajectory of a partially flipped primer strand guanine (G7) restacking and Watson-Crick pairing with its partner template strand cytosine base (C3). The terminal C1:G9 base pair is in red, the C2:G8 base pair is in blue, the C3:G7 base pair is in green, and other base pairs are shown in orange. This trajectory is obtained from simulation 5 using the AMBER94 force field and starting from the 335 degree guanine umbrella sampling window, and has been smoothed at an interval of 1 ps for clarity.

Supporting video 4: Visual depiction of a trajectory starting from a partially flipped primer strand guanine (G7) and resulting in primer strand slippage by one base pair. The terminal C1:G9 base pair is in red, the C2:G8 base pair is in blue, the C3:G7 base pair is in green, and other base pairs are shown in orange. This trajectory is obtained from simulation 5 using the AMBER99 force field and starting from the 335 degree guanine umbrella sampling window, and has been smoothed at an interval of 1 ps for clarity.

Table 1: Percentage distribution of the pseudo-dihedral coordinate describing the flipping of the third cytosine in unrestrained MD simulations starting from the 60 degree umbrella sampling window for the original PMF using different force fields

Force field	Simulation number	Possible restacking	Pre-stacking	Starting state	Further flipping
CHARMM27	1	0.0	0.4	99.4	0.1
	2	9.9	6.7	40.9	42.4
	3	0.0	0.1	21.2	78.6
	4	0.0	0.8	99.1	0.0
	5	0.0	0.2	99.8	0.1
AMBER94	1	0.0	1.0	98.9	0.1
	2	16.5	16.8	62.3	4.4
	3	1.3	39.6	59.1	0.0
	4	2.9	10.6	86.5	0.0
	5	0.2	1.2	64.6	34.1
AMBER99	1	49.1	8.1	8.5	34.2
	2	10.9	46.6	42.6	0.0
	3	0.1	2.4	95.4	2.1
	4	0.8	12.2	86.8	0.2
	5	0.9	3.6	68.0	27.4
	6	49.1	8.1	8.5	34.2
	7	0.0	0.1	95.7	4.2
	8	27.1	35.0	37.9	0.0
	9	5.4	45.6	49.0	0.0
	10	40.4	16.3	38.2	5.0
AMBERBSC0	1	25.5	30.0	44.5	0.0
	2	23.9	26.6	49.4	0.1
	3	0.0	0.0	99.9	0.1
	4	0.0	0.3	99.7	0.0
	5	3.1	41.3	55.2	0.3
BMS	1	2.2	5.3	92.3	0.2
	2	0.0	0.0	99.7	0.3
	3	38.6	25.2	36.2	0.0
	4	0.0	0.0	99.7	0.3
	5	17.0	17.8	55.0	10.2
CHARMM36	1	40.7	38.0	20.8	0.5
	2	1.7	11.2	86.2	0.9
	3	1.0	5.2	54.7	39.0
	4	31.0	50.4	18.5	0.1
	5	28.3	33.3	38.3	0.0

The cutoffs in degrees for the different pseudo-dihedral ranges are as follows: Possible restacking: 345-360 and 0-25, Pre-stacking: 25-40, starting metastable state: 40-80, further flipping: 80-345. All reported values are percentages obtained from MD simulations that were each performed for 5 ns, yielding a total of sampling time of 150 ns.

Table 2: Percentage distribution of the pseudo-dihedral coordinate describing the flipping of the third cytosine in unrestrained MD simulations starting from the 90 degree umbrella sampling window for the original PMF using different force fields

Force field	Simulation number	Possible restacking	Pre-stacking	Starting state	Further flipping
CHARMM27	1	0.0	0.0	45.3	54.7
	2	5.9	3.8	41.4	49.0
	3	0.9	6.7	43.1	49.3
	4	75.3	7.7	15.9	1.1
	5	15.4	6.0	35.4	43.2
AMBER94	1	2.7	53.9	43.4	0.0
	2	9.8	18.1	57.1	15.0
	3	0.8	46.7	51.5	1.0
	4	10.8	33.0	55.1	1.1
	5	6.7	63.0	30.3	0.0
AMBER99	1	5.8	12.1	81.9	0.1
	2	1.4	9.9	78.1	10.6
	3	0.1	2.8	88.2	8.9
	4	6.8	18.2	73.6	1.5
	5	4.3	85.3	10.3	0.0
	6	78.8	7.8	13.3	0.1
	7	54.1	41.4	4.5	0.0
	8	0.0	23.0	75.6	1.4
	9	0.0	0.9	98.1	1.0
	10	0.0	3.6	96.2	0.1
AMBERBSC0	1	32.1	33.1	34.8	0.0
	2	0.1	2.2	97.1	0.6
	3	77.1	9.2	13.3	0.5
	4	0.0	1.6	95.2	3.2
	5	1.6	5.0	64.6	28.8
BMS	1	0.0	0.0	99.6	0.4
	2	1.9	58.6	39.4	0.1
	3	0.0	1.2	98.1	0.7
	4	0.0	0.0	99.6	0.4
	5	0.0	0.0	99.5	0.5
CHARMM36	1	10.8	27.9	35.1	26.2
	2	90.5	0.6	7.7	1.2
	3	52.2	40.3	7.5	0.0
	4	7.9	10.8	38.6	42.6
	5	4.5	0.2	43.4	51.9

The cutoffs in degrees for the different pseudo-dihedral ranges are as follows: Possible restacking: 345-360 and 0-25, pre-stacking: 25-40, starting metastable state: 40-80, further flipping: 80-345. All reported values are percentages obtained from MD simulations that were each performed for 5 ns, yielding a total of sampling time of 150 ns.

Table 3: Percentage distribution of the pseudo-dihedral coordinate describing the flipping of guanine7 in unrestrained MD simulations starting from the 335 degree umbrella sampling window for the original PMF using different force fields

Force field	Simulation number	Possible restacking	Pre-stacking	Starting state	Further flipping
CHARMM27	1	0.0	0.0	47.5	52.5
	2	97.6	0.7	1.7	0.0
	3	97.9	0.9	1.2	0.0
	4	43.9	13.2	18.4	24.5
	5	0.0	0.0	89.7	10.3
AMBER94	1	0.0	0.2	37.8	62.0
	2	24.7	9.4	60.1	5.8
	3	0.0	0.3	91.9	7.8
	4	1.6	0.2	81.2	16.9
	5	81.3	1.4	17.3	0.1
AMBER99	1	0.2	3.2	20.8	75.7
	2	0.0	0.4	98.9	0.8
	3	0.0	0.1	24.7	75.2
	4	0.3	1.6	90.3	7.8
	5	9.1	36.7	54.2	0.0
	6	98.2	1.8	0.0	0.0
	7	97.7	2.3	0.0	0.0
	8	98.6	1.4	0.0	0.0
	9	98.4	1.6	0.0	0.0
	10	98.6	1.4	0.0	0.0
AMBERBSC0	1	97.9	0.2	1.9	0.0
	2	0.0	0.0	48.3	51.7
	3	0.0	0.1	72.9	27.1
	4	97.8	0.2	2.0	0.0
	5	0.0	0.1	9.8	90.0
BMS	1	95.9	0.1	4.0	0.0
	2	94.7	2.7	2.6	0.0
	3	23.6	3.6	21.3	51.5
	4	0.0	0.0	97.6	2.4
	5	0.0	0.8	98.2	1.1
CHARMM36	1	86.7	5.1	6.3	1.9
	2	66.9	11.5	21.6	0.1
	3	87.8	3.9	8.3	0.0
	4	0.8	2.5	16.4	80.3
	5	82.0	2.7	15.0	0.3

The cutoffs in degrees for the different pseudo-dihedral ranges are as follows: Possible restacking: 0-60, pre-stacking: 350-360, minor groove intermediate: 320-350, further flipping: 60-320. All reported values are percentages obtained from MD simulations that were each performed for 5 ns, yielding a total of sampling time of 150 ns.

Table 4: Steric overlap of initially overlaid partially flipped and strand slipped structures in two polymerase active sites prior to optimization.

Polymerase	State	protein*	DNA*
Dbh	Cyt-F	0	0
Dbh	TS	0	0
Dbh	Gua-F	0	0
Dbh	PS	0	0
RB69	Cyt-F	1	1
RB69	TS	3	3
RB69	Gua-F	7	10
RB69	PS	5	7

*Values are for number of non-hydrogen atoms within 1.5 Å of any non-hydrogen atom of the bound partner. The abbreviations used for the states are: (a) Cyt-F: state with partially flipped cytosine, (b) TS: template strand slipped state, (c) Gua-F: state with partially flipped guanine, (d) PS: primer strand slipped state. Dbh is a Y-family lesion bypass polymerase[3] and RB69 is a B-family replicative polymerase.[4]

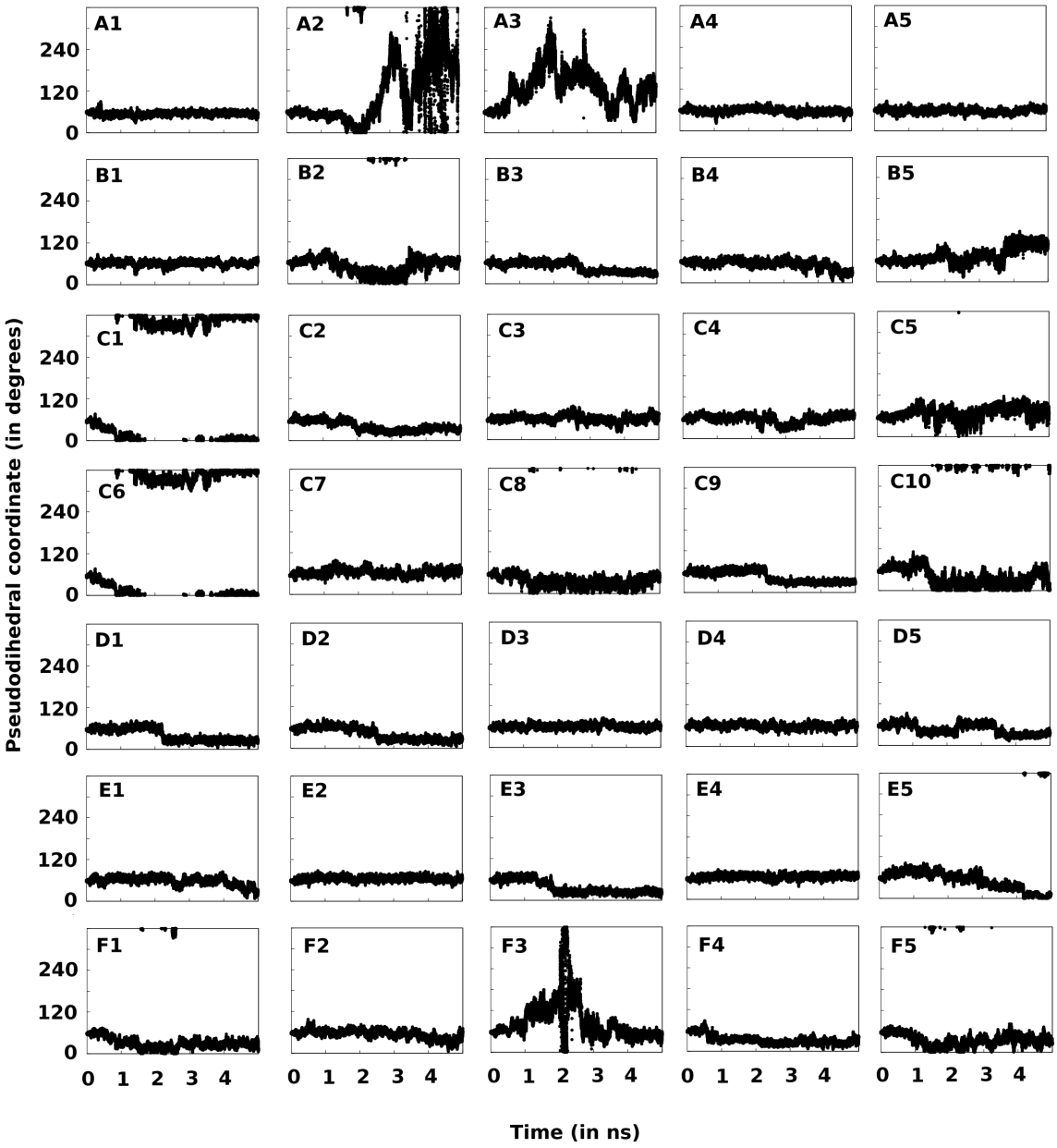

Figure 1: Time series of the pseudo-dihedral coordinate in unrestrained 5 ns MD simulations started from the 60 degree umbrella sampling window for cytosine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5: AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10]

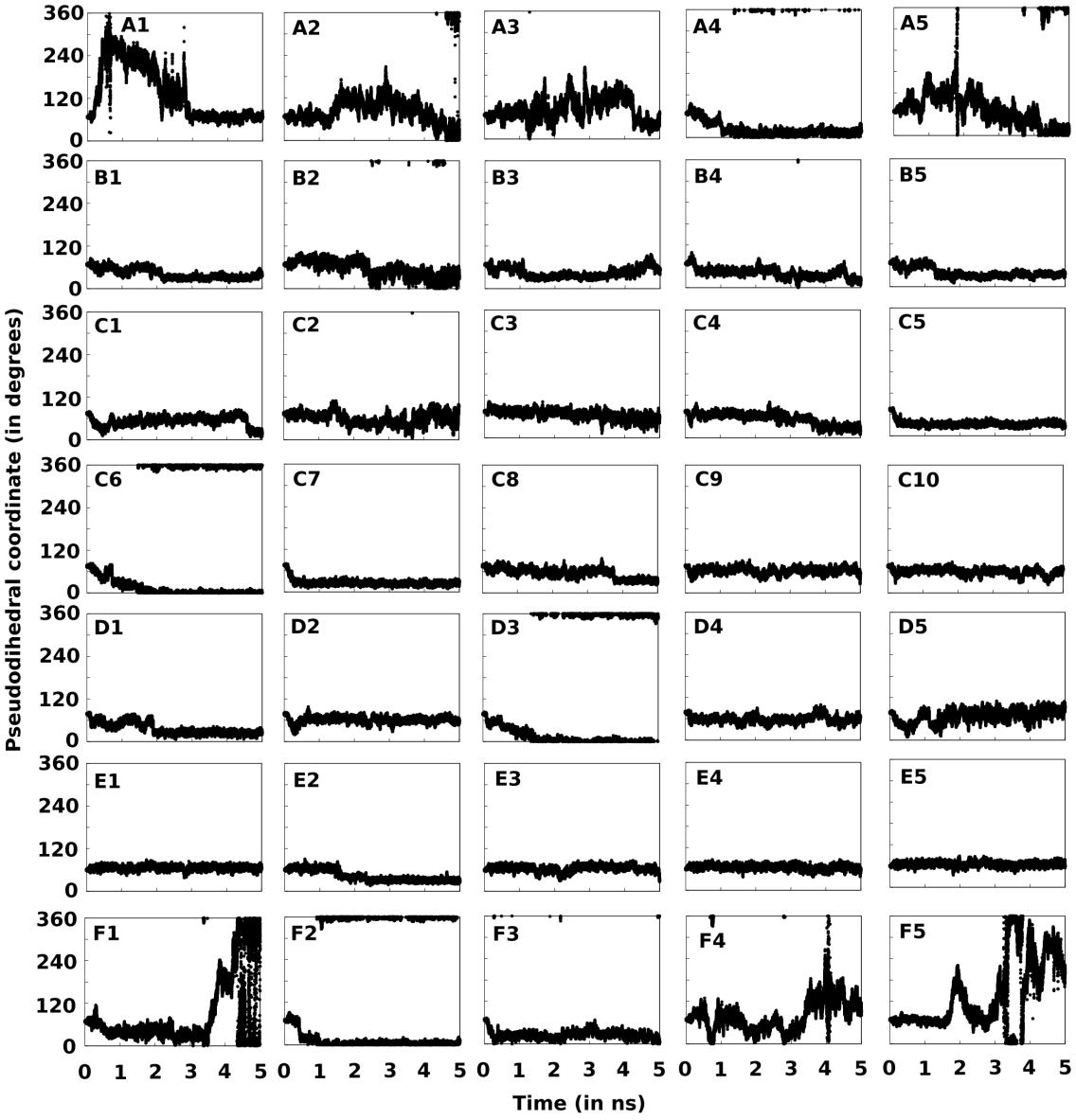

Figure 2: Time series of the pseudo-dihedral coordinate in unrestrained 5 ns MD simulations started from the 90 degree umbrella sampling window for cytosine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5 AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10]

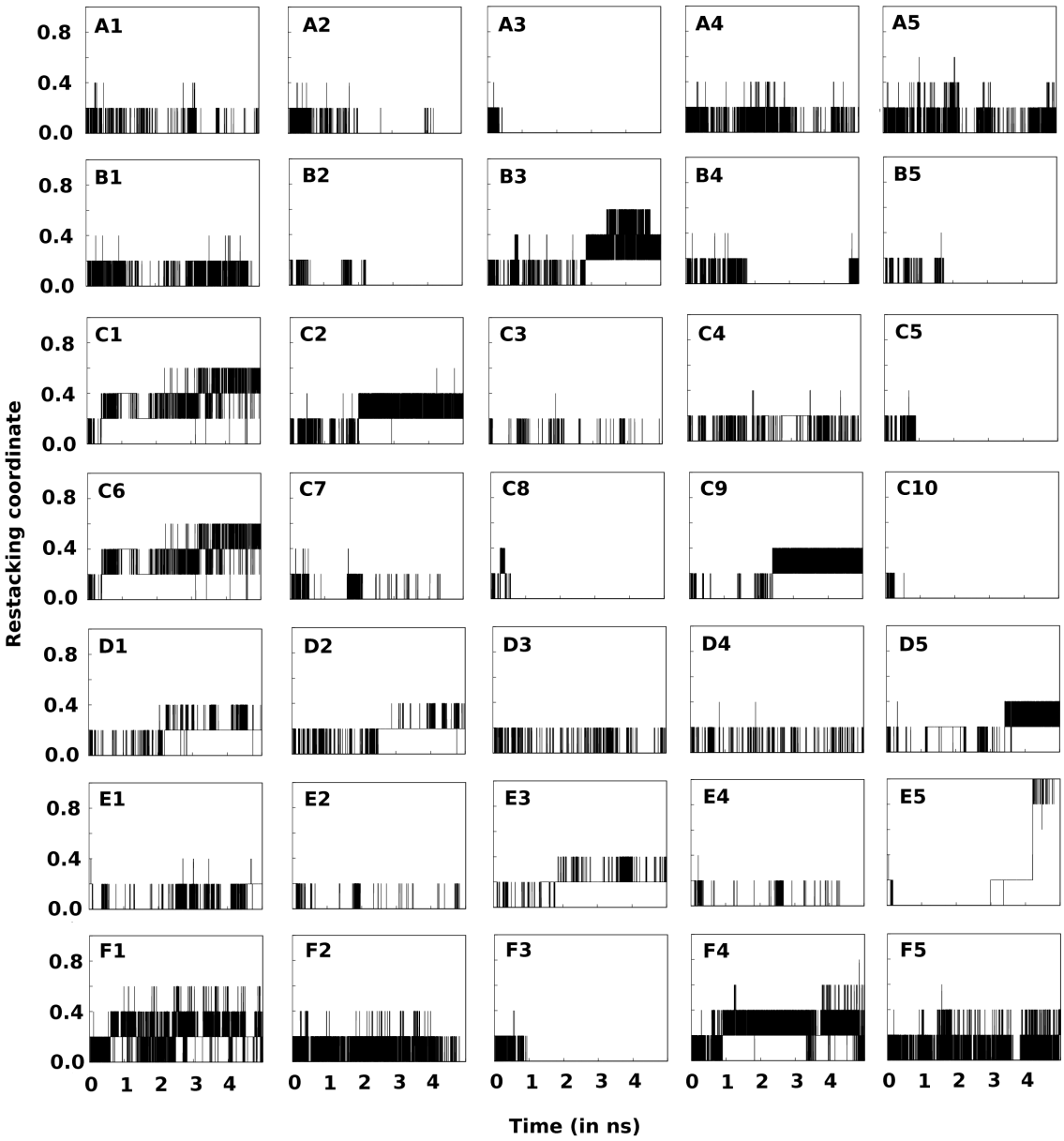

Figure 3: Time series of the restacking-pairing coordinate in unrestrained 5 ns MD simulations started from the 60 degree umbrella sampling window for cytosine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5 AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10]

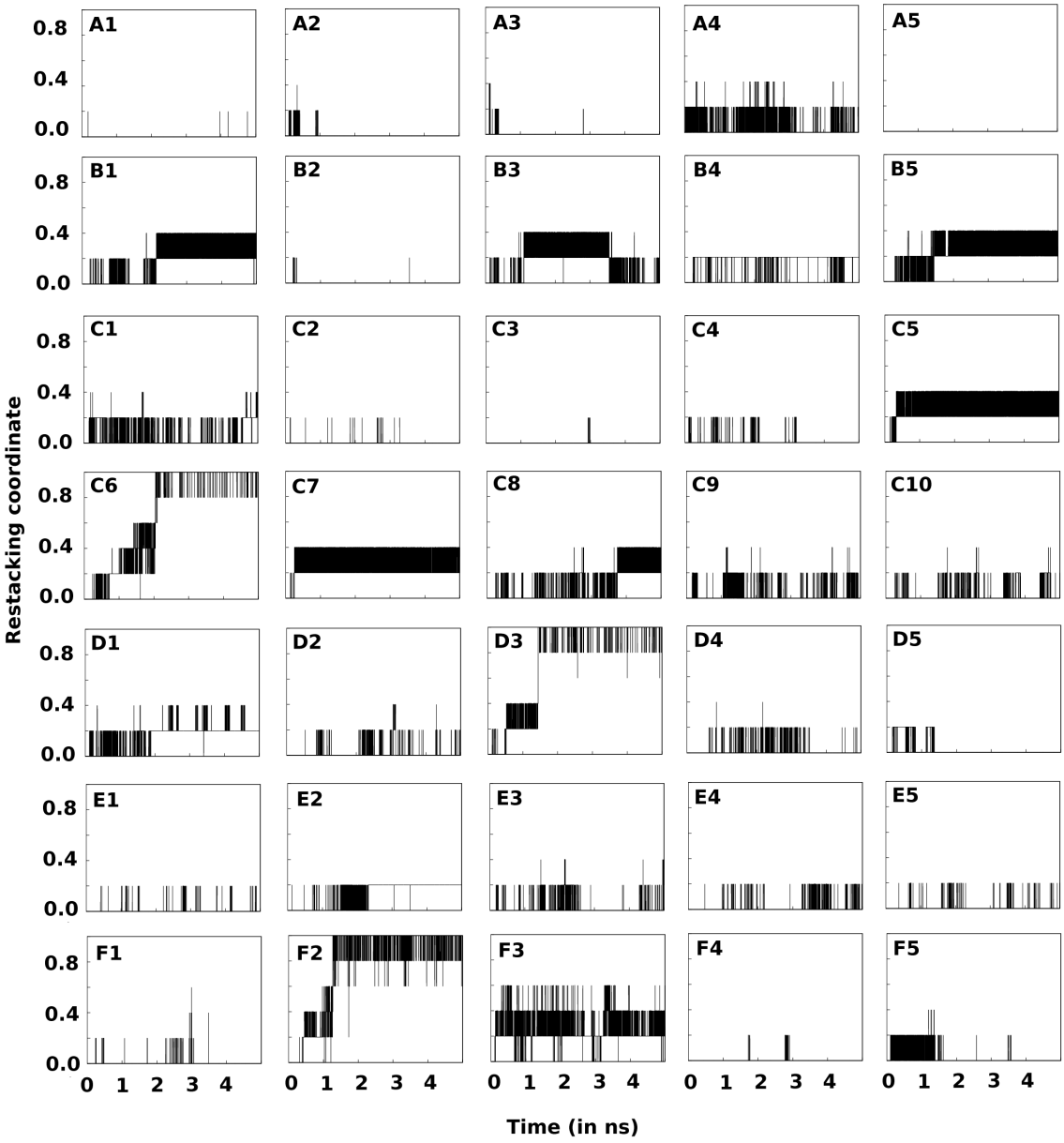

Figure 4: Time series of the restacking-pairing coordinate in unrestrained 5 ns MD simulations started from the 90 degree umbrella sampling window for cytosine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5 AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10]

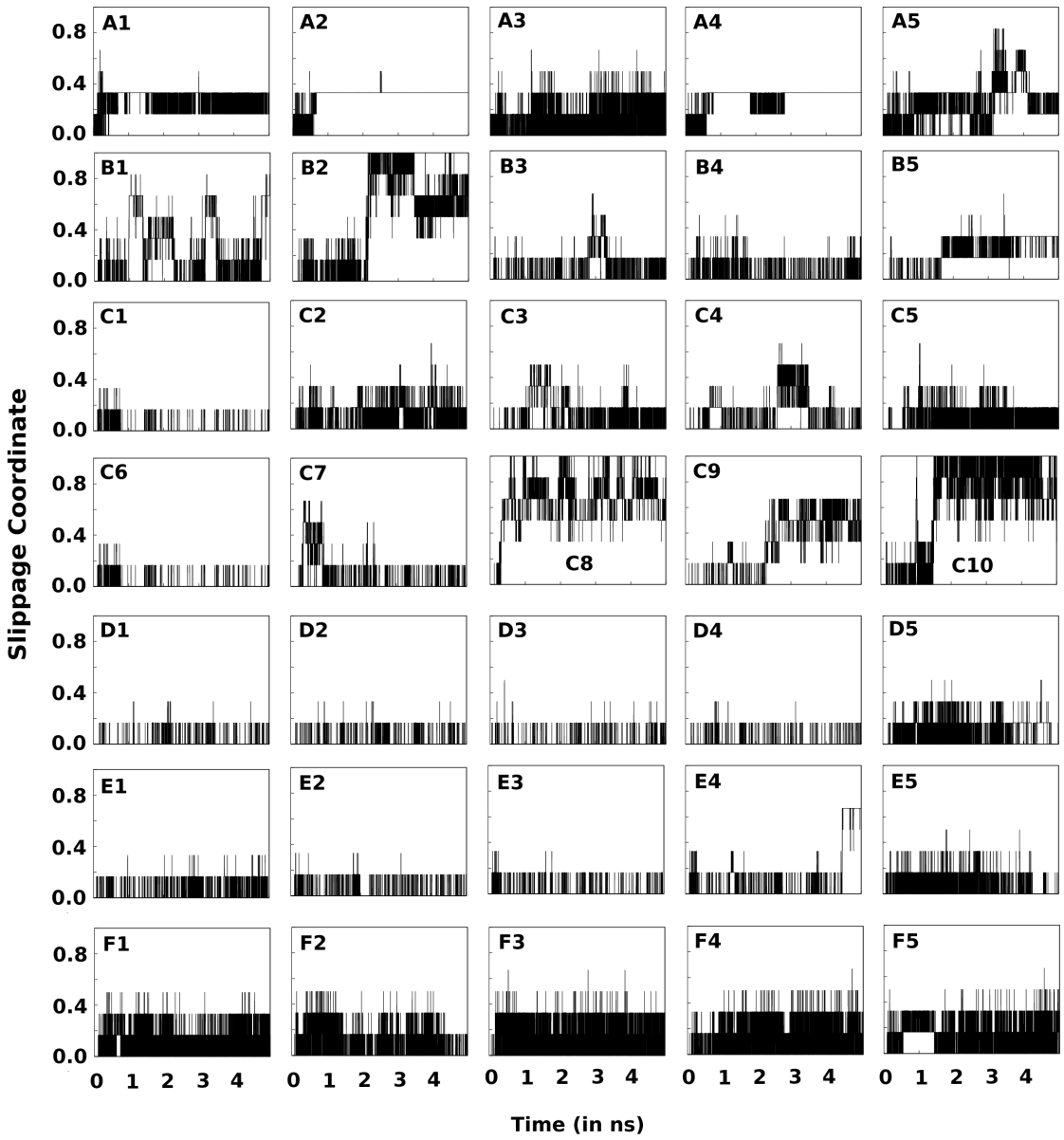

Figure 5: Time series of the slippage coordinate in unrestrained 5 ns MD simulations started from the 60 degree umbrella sampling window for cytosine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5 AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10]

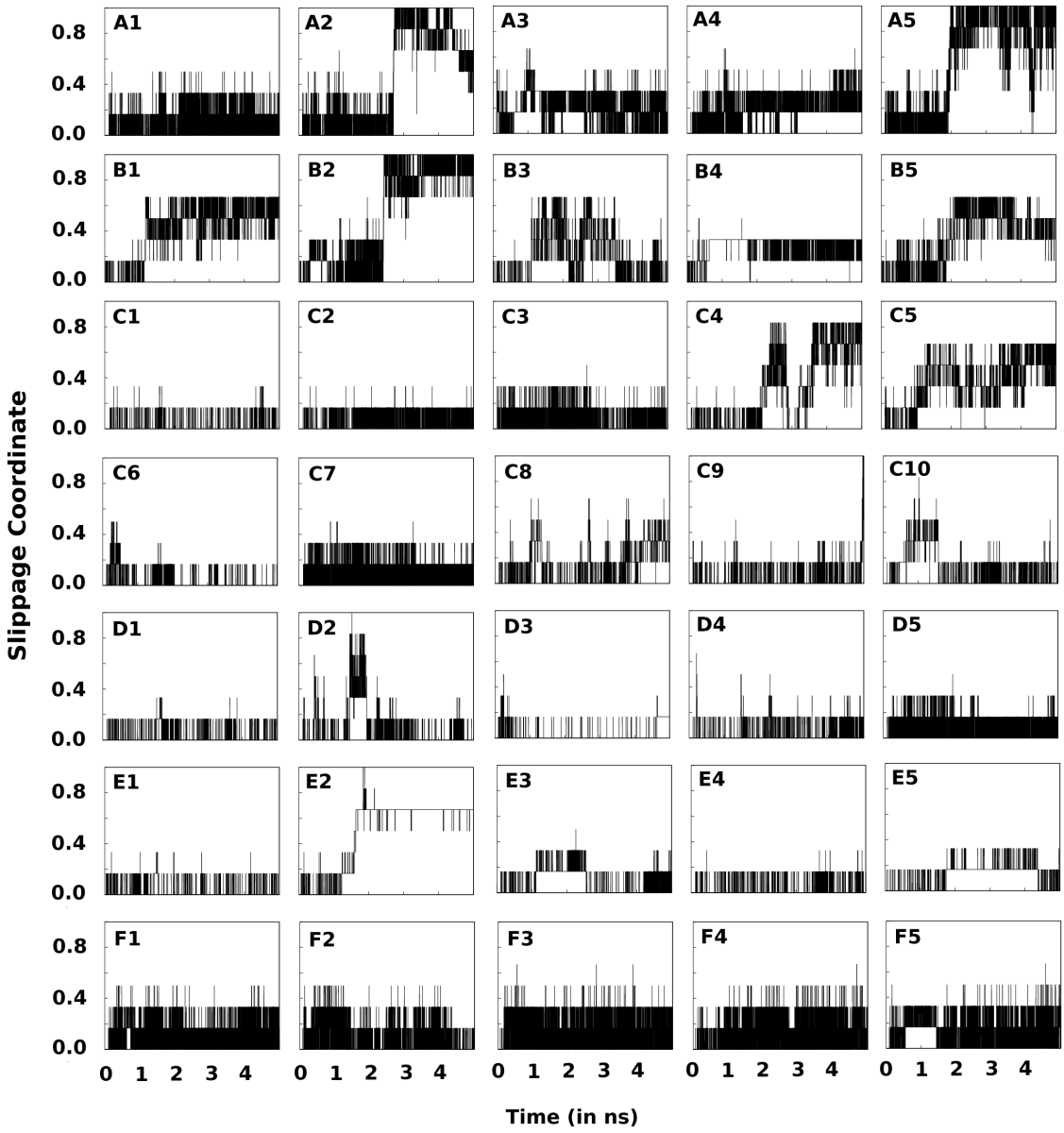

Figure 6: Time series of the slippage coordinate in unrestrained 5 ns MD simulations started from the 90 degree umbrella sampling window for cytosine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5 AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10]

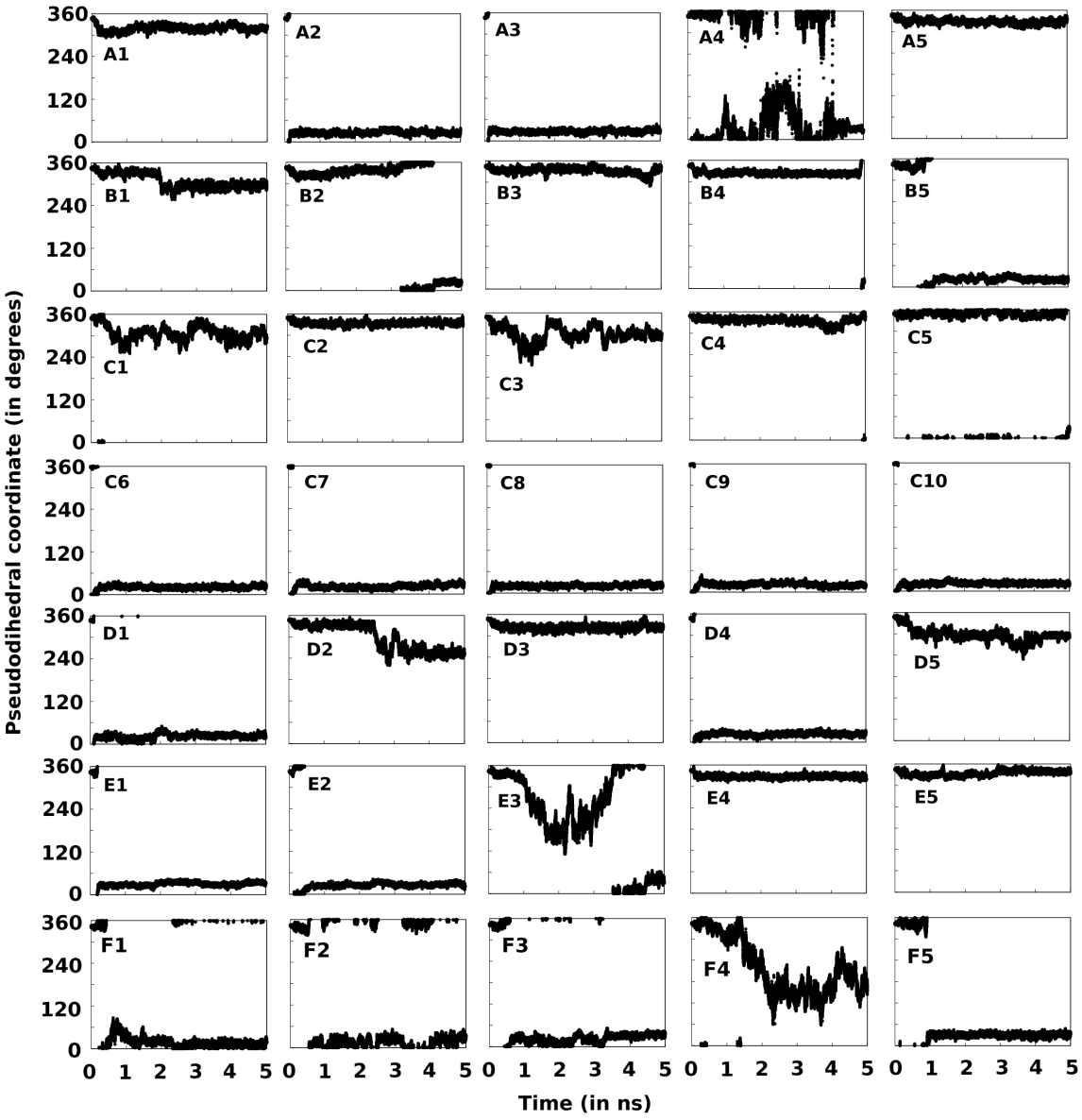

Figure 7: Time series of the pseudo-dihedral coordinate in unrestrained 5 ns MD simulations started from the 335 degree umbrella sampling window for guanine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5: AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10]

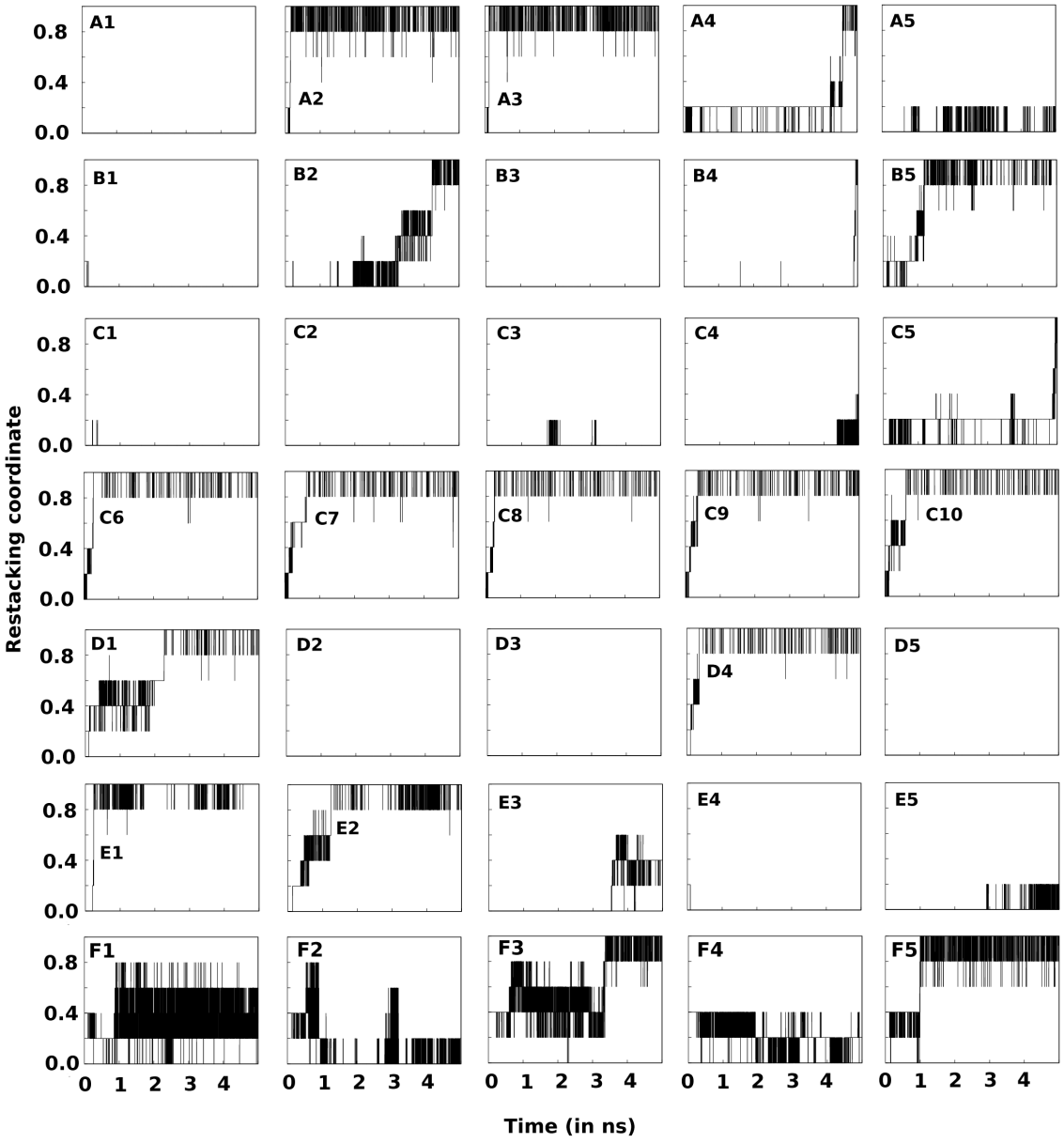

Figure 8: Time series of the restacking-pairing coordinate in unrestrained 5 ns MD simulations started from the 335 degree umbrella sampling window for guanine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5 AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10]

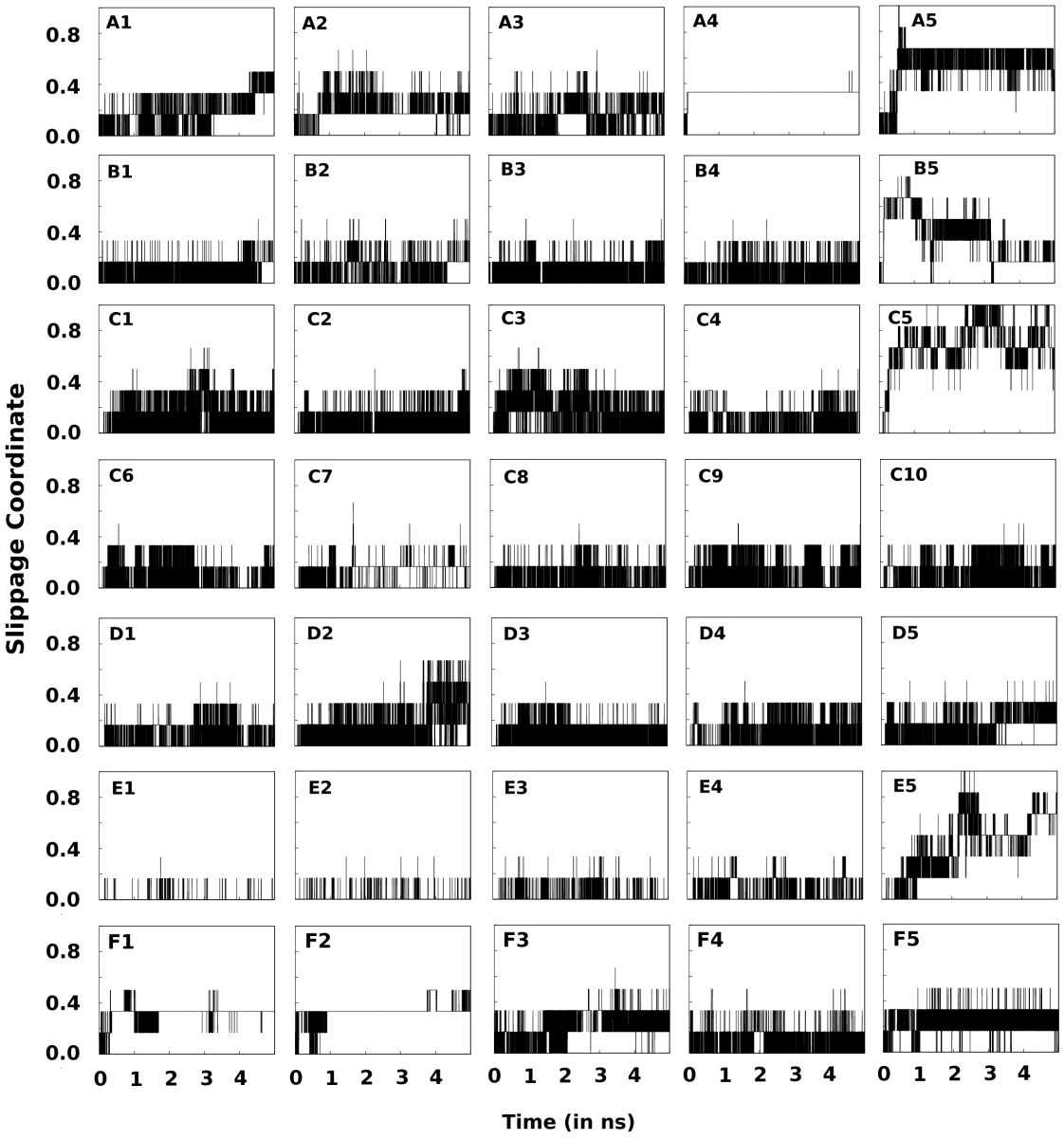

Figure 9: Time series of the slippage coordinate in unrestrained 5 ns MD simulations started from the 335 degree umbrella sampling window for guanine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5 AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10]

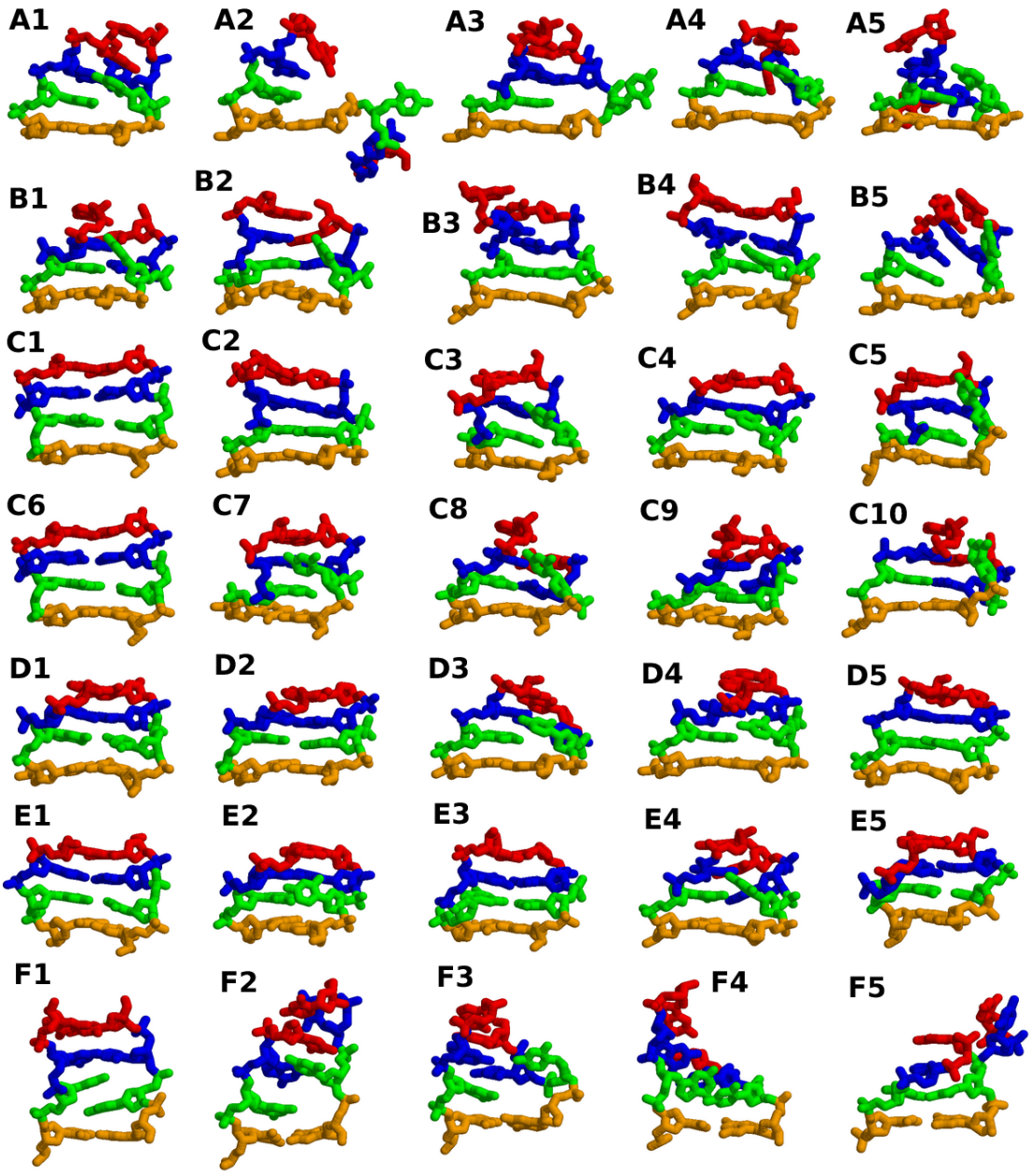

Figure 10: The final structures for the terminal four base pairs in unrestrained 5 ns MD simulations started from the 60 degree umbrella sampling window for cytosine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5 AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10] The first base pair is in red, the second in blue, the third in green, and the fourth in orange, so slippage can be detected by identifying multi-colored base pairs.

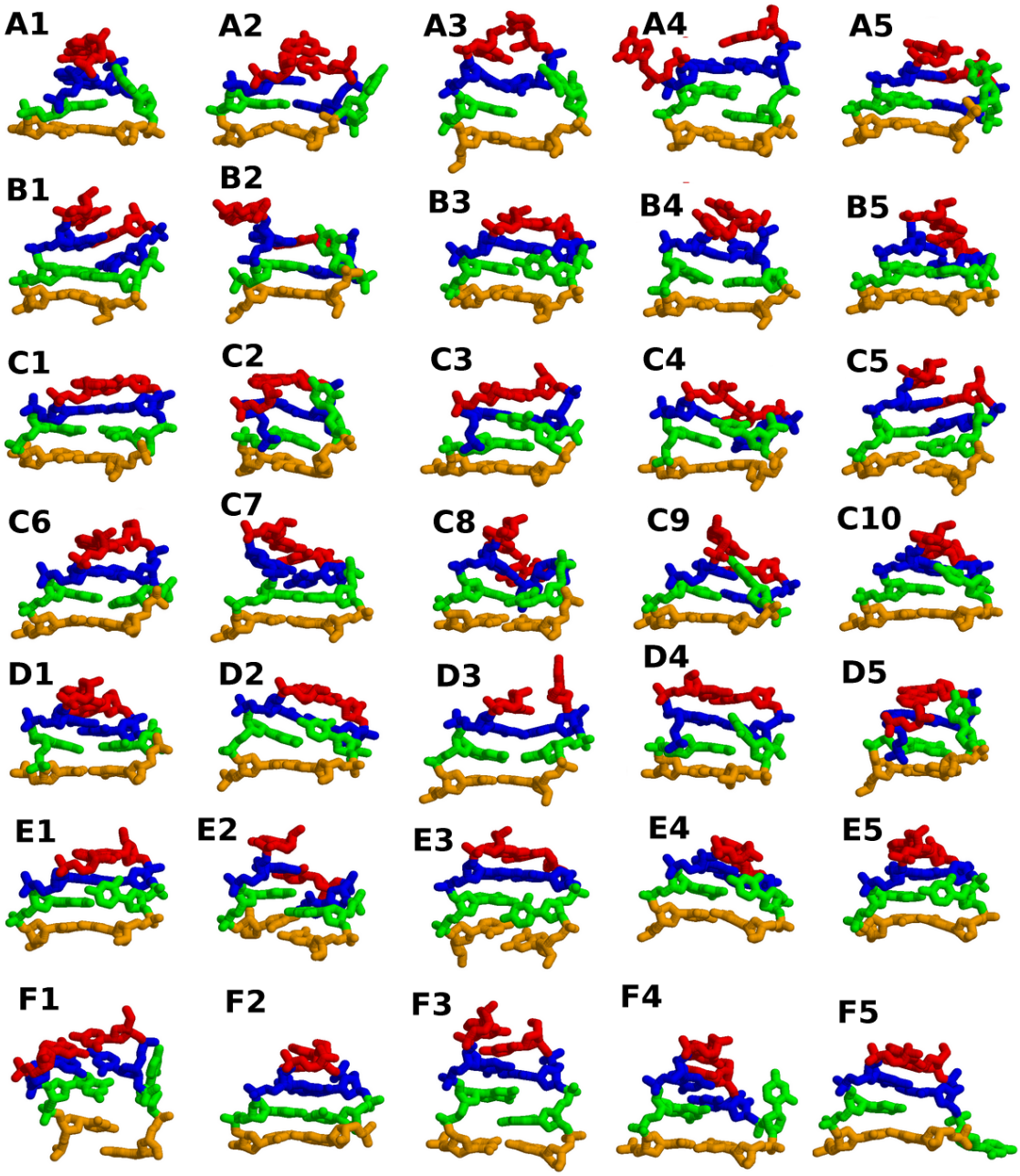

Figure 11: The final structure for the terminal four base pairs in unrestrained 5 ns MD simulations started from the 90 degree umbrella sampling window for cytosine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5 AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10] The first base pair is in red, the second in blue, the third in green, and the fourth in orange, so slippage can be detected by identifying multi-colored base pairs.

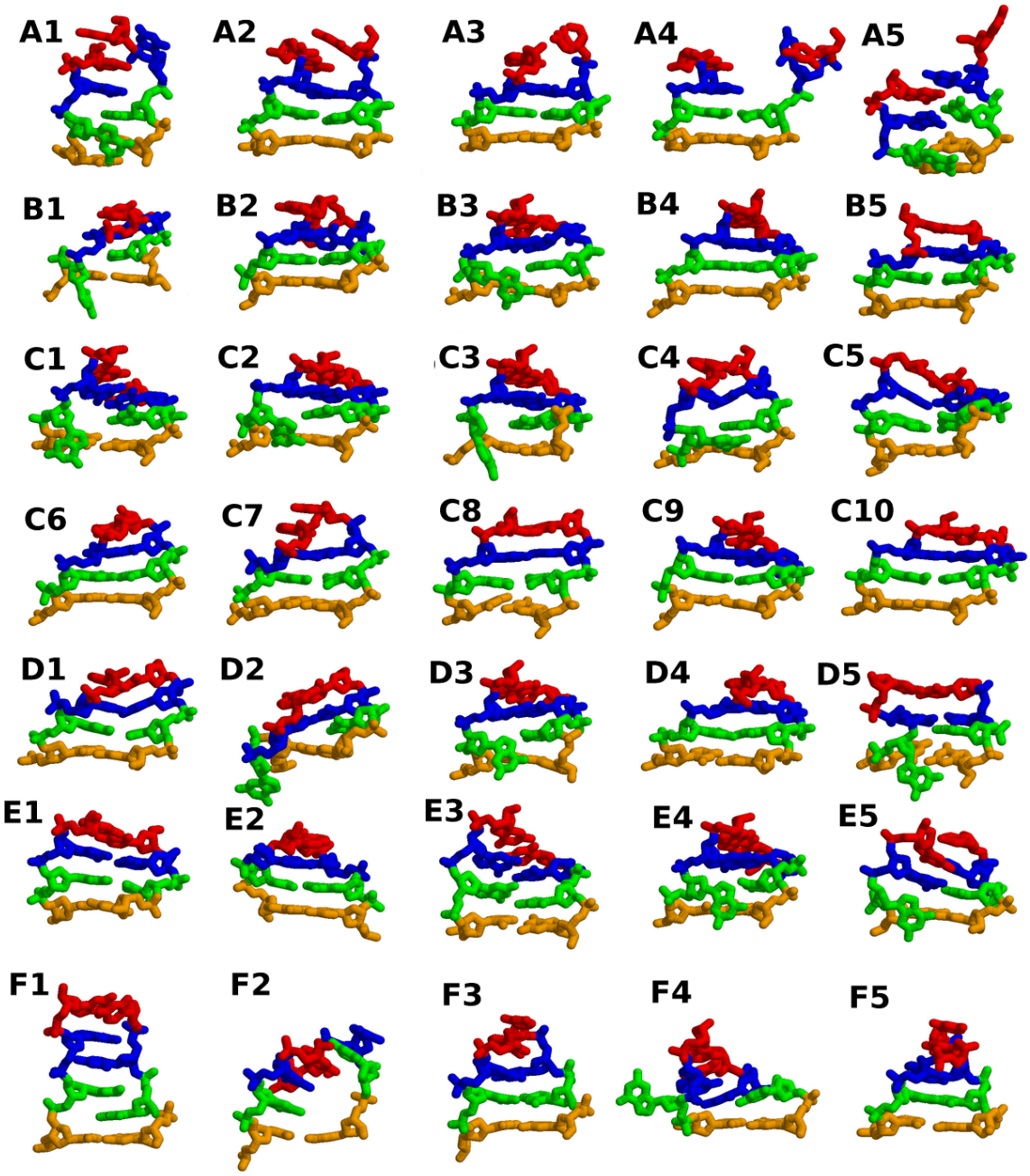

Figure 12: The final structure for the terminal four base pairs in unrestrained 5 ns MD simulations started from the 335 degree umbrella sampling window for guanine flipping using different force fields. The force fields used for the simulations are as follows: A1-A5: CHARMM27,[5,6] B1-B5 AMBER94,[7] C1-C10: AMBER99,[8] D1-D5: AMBERBSC0,[9] E1-E5: BMS.[10] The first base pair is in red, the second in blue, the third in green, and the fourth in orange, so slippage can be detected by identifying multi-colored base pairs.

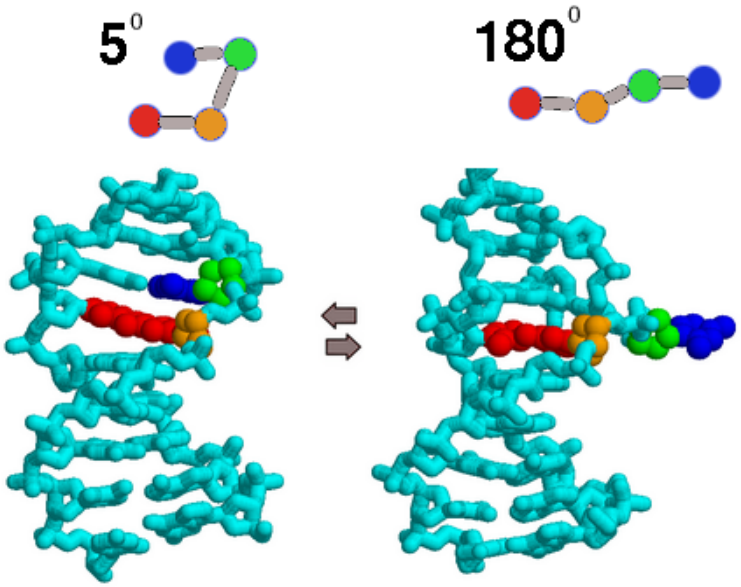

Figure 13: Illustration of the coarse-grained pseudo-dihedral coordinate used to enforce or monitor single base flipping. The coordinate consists of the dihedral angle between four center-of-masses for four sets of atoms: (1) non-hydrogen base atoms of a neighboring base pair (shown in red), (2) non-hydrogen atoms of a neighboring sugar moiety in the same strand (shown in orange), (3) non-hydrogen atoms of the sugar moiety of the single base that is flipping (shown in green), and (4) non-hydrogen atoms of the flipping base itself (shown in blue). The stacked and fully flipped conformations are shown along with their corresponding coarse-grained center-of-mass depictions, and their pseudo-dihedral values.