The intimate knowledge of DNA crystals revealed by molecular dynamics simulations

Antonija Kuzmanic[1,2], Pablo D. Dans[1,2], and Modesto Orozco[1,2,3,*]

[1]Institute for Research in Biomedicine (IRB Barcelona). The Barcelona Institute of Science and Technology. Baldiri Reixac 10-12, 08028 Barcelona, Spain.
[2]Joint BSC-IRB Research Program in Computational Biology. Baldiri Reixac 10-12, 08028 Barcelona, Spain.
[3]Department of Biochemistry and Molecular Biology. University of Barcelona, 08028 Barcelona, Spain.

[*] To whom correspondence should be addressed: Prof. Modesto Orozco, Tel: +34 934037155, Email: modesto.orozco@irbbarcelona.org.

INTRODUCTION

Ever since Linus Pauling laid the foundations of structural biology, X-ray crystallography has been the cornerstone method for revealing biomolecular 3D structures at atomic resolution to the world. Hundreds of thousands of structures deposited in the Protein Data Bank (PDB) are the best witnesses of the power and scope of this technique, which has also become the golden standard for validating other structural determination methods (e.g., NMR, electron microscopy, and other diffraction techniques). However, just as any other method, X-ray crystallography has its own shortcomings and obtaining diffractable crystals is not always an easy task. Typically, hundreds of different conditions are tested in unsuccessful attempts to crystallize a specific biomolecule and the impact that buffer conditions have in determining the symmetry of the unit cell geometry, the overall structure, and the physical properties of the biomolecule are ignored.

Nucleic acids have benefited greatly from X-ray crystallography and thousands of DNA structures deposited in the PDB have been crucial in understanding the fine details of isolated and protein-bound DNAs. Unfortunately, artifacts in crystal DNA structures are known, with the most important ones related to lattice effects and crystallization conditions that, for example, bias duplexes to the A-form which is otherwise undetected in solution. Even in well-known cases, such as Drew-Dickerson dodecamer (DDD), where crystal structures resemble the solution ones reasonably well, it is not clear how crystallization conditions modulate the DNA conformation, and more intriguingly, the symmetry space in which the molecules crystallize. Knowing which molecular interactions stabilize the crystals would help

understand the real properties of DNA in such systems and the mechanisms through which highly charged molecules are packed in severely crowded conditions, e.g. the cellular nucleus.

Here we take advantage of the new generation of supercomputers and a recently developed force field, which is able to generate duplex structures indistinguishable from the experimental ones, to analyze the nature of DDD in three known crystal lattices and explore the effects of crystallization additives, particularly spermine, in the stabilization of the crystal. Through extensive unbiased MD simulations on the millisecond timescale, we demonstrate how the stability of DNA crystals depends on subtle interactions between the packed DNA molecules and the components of the crystallization buffer. For the first time, we obtained stable simulations of DNA crystals in various symmetry groups (all in which DDD has been crystallized) and in different solvent environments allowing us to understand with unprecedented level of detail the nature of the intermolecular interactions that stabilize the crystals. Our results are a proof of concept that expands the actual limits of the field, opening the door to better initial crystallization conditions. Detailed analysis of more than 1.4 milliseconds of simulations allowed us to characterize for the first time at atomic resolution the physical properties of DNA in a highly crowded environment that resembles the highly packed conformation of DNA in the nucleus.

METHODS

Systems setup. X-ray structures of the Drew-Dickerson dodecamer used as the starting conformations in the unbiased MD simulations are those deposited under PDB codes 1EHV, 1FQ2, and 463D. Both 1EHV and 1FQ2 entries contain the full DNA duplex, while the 5' cytosines are missing in 463D and were modelled in with PyMOL v1.8.2[1] by using 1EHV as a template. The dodecahedral simulation boxes for solution systems (labelled sol) were created by solvating a single DNA duplex with TIP3P water molecules,[2,3] leaving a 10-Å solvent layer around the duplex (~20,000 atoms). K^+, Na^+, and Cl^- ions (Smith and Dang parameters)[4] were added to reach neutrality (with K^+ ions) and the final concentration of 400 mM (with 4:3 potassium to sodium ratio). In the case of 1FQ2 1:6 sol system, 6 spermine molecules were randomly added prior to solvation with TIP3P water and neutrality was reached with Cl^- ions. For the crystal systems, the simulation boxes were built following the dimensions of the unit cells reported in the PDB using PyMOL v1.8.The Na/KCl and $MgCl_2$systems were completed by first adding TIP3P water and then the mentioned ions to reach neutrality and the final concentrations listed in Table 1 (using Smith and Dang[4] parameters for Na/KCl systems and Allnér-Nilsson-Villa[5] parameters for $MgCl_2$). The $SpmCl_4$and $MgCl_2{\cdot}6H_2O$ crystal systems were completed by first randomly adding spermine or hexacoordinated

Mg ions (using Yoo and Aksimentiev[6,7] parameters, quantities in Table 1), then TIP3P waters followed by counter ions. System sizes ranged from 44,000-60,000 atoms for the smaller crystals, while the 1EHV systems with 81 duplexes had ~170,000 atoms.

Relaxation and equilibration of the systems. Energy of the simulated systems (listed in Table 1) was initially minimized through steepest-descent energy minimization. The initial velocities for the atoms were taken from Maxwell distribution at 100 K and the systems were subsequently heated to the final temperature (Table 1) in five steps of ~50 K simulated for either 100 or 1000 ps at constant volume and temperature using velocity rescale thermostat.[8] In parallel, atomic position restraints for the DNA molecules were uniformly relaxed in ~1 ns pre-equilibration. To achieve internal pressure close to 1 bar in crystal simulations, additional TIP3P water molecules were randomly placed within each system, which were then re-optimized, thermalized, pre-equilibrated and subjected to 50 ns equilibration runs, in a trial and error procedure until a suitable pressure was obtained.

Production trajectories. Once simulation conditions matched those expected in the crystal or in solution, production runs (1-4 μs long) were generated using GROMACS 5.0.4 biomolecular simulation package[9] with a 2-fs integration step and coordinate output of 10 ps. Constant volume conditions were employed for the crystal simulations, while the constant pressure (1 bar) was used in solution simulations. All simulations were done at constant temperature (284-295 K depending on crystallization conditions). Parrinello-Rahman barostat[10] and velocity rescale thermostat were used.[8] Periodic boundary conditions and particle mesh Ewald[11,12] were used in conjunction with state-of-the-art simulation conditions. DNA was described using the recently developed parmbsc1 force-field.[13] See Methods for additional details of the simulations.

Analysis. Center of mass (COM) displacements were calculated after the optimal roto-translational fitting of the whole system to the ideal crystal. RMSD values were calculated for each duplex after the optimal roto-translational fitting to the original X-ray structures (see Suppl. Methods for details). Diffusion coefficients were calculated from a linear fit of mean square displacements over time using Einstein's equation. The effective temperature of waters in the crystals were estimated from the corresponding diffusiveness in the simulations,[14] while effective DNA temperature was inferred at the base pair step level from the analysis of helical covariance matrices, which provided the stiffness matrices associated to harmonic deformations in the helical space.[15] Essential dynamics were determined by diagonalization of the Cartesian covariance matrices. Similarity in essential deformation modes were computed using Hess[16] and Pérez et al metrics.[17] Additional details on the analysis and the associated packages are shown in Suppl. Methods.

Data availability. All trajectories reported here are deposited in the MuG-BigNASim database and can be freely downloaded.

RESULTS AND DISCUSSION

Are the simulations of DNA crystals stable in the μs regime with "standard" protocols? We decided to focus our study on the paradigmatic Drew-Dickerson dodecamer, the first DNA molecule ever crystallized, which has been determined in various space groups and under diverse conditions at high resolution. We choose to perform our simulations on structures deposited in the PDB with accession codes 1EHV, 1FQ2, and 463D, which were respectively crystallized in 3 different space groups: $P3_212$, $P2_12_12_1$, and H3. Following the multistep protocol developed by the Case group, which includes a careful stabilization of the internal pressure by calibrating the number of water molecules in the system (ref, and see Supporting Methods), we ran multi-microsecond long simulations of each crystal cell, which contained respectively 27, 24 and 36 copies of DDD (Table 1). To our surprise, two out of three systems showed a systematic degradation of the cell lattices, as shown by the center-of-mass (COM) displacements of each duplex with respect to the ideal positions in the crystal (Figure 1). The distortion of the cell was particularly visible along the z-axis and got worse with simulation time for 1EHV and 1FQ2, while the 463D system was completely stable thorough the trajectory. Not only was the lattice intact in 463D, but the MD conformations sampled in the helical space at the base pair step (bps) level were in complete agreement with the average X-ray values (Figure 2), giving us confidence about the capacity of our force field to represent local and global features of crystal structures.

We thus centered our efforts on stabilizing the other two crystal cells starting first with the smallest system - 1EHV (3 DDD copies in the asymmetric unit cell). We simultaneously ran simulations of three times larger 1EHV crystals (with 81 DDD copies) with 5 different Na/KCl concentrations (Table 1), ranging from 15 mM to 800 mM, which is more than 5 times the assumed physiological condition of the cellular nucleus, without being able to stabilize the crystal lattice (Supp. Figure S1). Considering that the increase in crystal's size had little effect on its stability (Supp. Figure S2), we decided to proceed with the smaller system (with 27 DDD copies). We then changed the monovalent cations by Mg^{2+} at 400 mM in an attempt to reduce the inter-phosphate repulsion. Although we observed a perceptible improvement in the lattice integrity, it still slowly and continuously degraded along the y/z-axes during the 4-μs-long simulations, even when we used a recent parameterization of Mg^{2+} developed to accurately reproduce the osmotic pressure of crowded DNA systems (Supp. Figure S3).

We checked if the global degradation of the crystals was due to internal distortions of the double helices by analyzing the 2D distributions of the bps helical

parameters from the MD simulations and comparing them to X-ray values for 1EHV. The excellent agreement between the structure of the duplexes from crystal MD and X-ray (Supp. Figure S4), discards the notion that the lattice distortions were caused by the deterioration of the duplexes. Even when we analyzed the flexible ends (residue pairs 1-24 and 12-13), which are subject to fast but moderate fraying in solution, the average RMSD of the capping base pairs with respect to the X-ray position was in most cases under 4 Å, oscillating between 3D conformations compatible with the experimental electron density (Supp. Figure S5). Similar results were obtained when trying to stabilize the $P2_12_12_1$ 1FQ2 crystal without success (data not shown), which allowed us to discard the possibility of facing a force-field artifact that is space-group specific. This turned our attention towards the buffer components used in the crystallization of 1EHV and 1FQ2 crystals.

How important are the buffer components in the stabilization of DNA crystals? When analyzing the chemical additives used to obtain the DDD crystals in the studied space groups, we noticed that 463D was crystallized without using spermine (SPM). Spermine is a polyamine (+4 charge at pH 7) involved in cellular metabolism and normally found in all eukaryotic cells, which was used as an "inert molecular glue" for obtaining thousands of DNA crystals in the major forms (A, B, Z, etc.), although the molecular basis of its mechanism of action was mostly ignored. Accordingly, we found that spermine was used in the crystallization process for 1EHV and 1FQ2 crystals. We ran simulations of both crystals at three different spermine concentrations ranging from 1:3 (3 spermine molecules per duplex) and up to 1:12 (Table 1). As shown in Figure 3, according to COM displacements, the lattice integrity for both space groups was now preserved in a consistent spermine-concentration dependent manner. At the medium spermine concentration tested, 1:9 for 1EHV and 1:6 for 1FQ2, the observed stability is already comparable to the one obtained for 463D with no perceptible drift of the cell lattice in the y/z-axes (Supp. Figures S6 and S7).

As in the case of 463D, once properly stabilized, an impressive agreement is obtained between the crystals' (1EHV/1FQ2) average values and the MD results when comparing the rotational helical space (Figure 4), or even when looking at the groove dimensions, a property which has been always difficult to reproduce accurately by modern force fields (Supp. Figure S8). The dynamics of the ends also seems to be stable according to the RMSD values computed for the terminal residues, which are in the majority of cases below 2 Å with respect to the crystal position (Supp. Figures S9 and S10). A global 3D structural comparison of the average MD crystal simulations to the experimental crystal configuration for the three systems stabilized in the μs regime is shown in Supp. Figure S11.

We used the Curvilinear Helicoidal Coordinate (CHC) method as implemented in Canion/Curves+ to follow each ion and water molecule along the trajectories of

each duplex, allowing for the calculation of populations and concentrations in the inner and outer areas of major and minor grooves. We found that sequence-dependent binding sites observed for K^+ and the amino groups of spermine in solution were the same and were maintained in the three studied crystals, although the dependence was less pronounced in crystals, especially inside the minor groove where the preference for d(CpG) steps observed in solution was almost not present (Supp. Figure S12). A clear pattern of intercalated sites was observed inside the major groove with cation preferences for d(GpC) and d(GpA)/d(TpC) steps, although most of the K^+ and SPM amino groups were located outside the major grooves interacting with the phosphate groups in the central part of the duplexes without a specific sequence preference (Supp. Figures S13-15). SPM can concentrate up to 12 M all around the phosphate groups in the central portion of the duplexes, where no DNA intermolecular contacts are present, which in contrast are particularly pronounced in terminal bases (Supp. Figure S16).

While the temperature for both experiments and MD simulations was 284, 295, and 293 K for 1EHV, 1FQ2, and 463D, respectively, the intermolecular DNA contacts at the termini froze the duplex structure. DNA's internal effective temperature can decrease by 200 K (with respect to the thermostat temperature) at the end of the duplexes, while decreasing by only ~50 K in the central portion (Supp. Figure S17). Water molecules and SPM are also affected by the crowded and highly negatively charged environment produced by the phosphate groups of the backbone chains. The diffusion coefficient of water molecules, as computed from Einstein's equation (Supp. Figure S18), was reduced by one order of magnitude - from $4.34 \cdot 10^{-5}$ cm^2 s^{-1} in solution DNA simulations to $4.27 \cdot 10^{-6}$ cm^2 s^{-1} in crystal simulations (note that the diffusion coefficient of pure water at 298.15 K and 1 atm is $2.30 \cdot 10^{-5}$ cm^2 s^{-1}). This reduction in water mobility was also noticeable when computing the effective temperature of water molecules in crystal simulations, which gave very similar values ranging from 255 to 258 K depending on the crystal type (30 K lower that the thermostat temperature). The effect was even more pronounced for SPM when comparing the 1FQ2 1:6 solution simulations with the 1FQ2 crystal at the same SPM concentration: the diffusion coefficient was reduced by three orders of magnitude in the crystal - from $1.90 \cdot 10^{-6}$ cm^2 s^{-1} to $2.62 \cdot 10^{-9}$ cm^2 s^{-1}. Although the mobility of SPM is greatly reduced in the crystal, the lack of specificity in its binding to DNA (extensive and non-specific binding outside the major groove) has made its detection in the experimental electron densities almost impossible. As expected, the diffusion of water and SPM in the crystal was correlated and exhibited an SPM-concentration dependent effect as shown in Supp. Figure S19 for the case of 1EHV.

How do the structural and dynamical properties of DNA crystals compare with solution? X-ray structures, usually a single average configuration obtained from many copies present in the crystal cell, have always been the cornerstone of

structural biology. Moreover, the 3D conformations obtained have also been the gold standard for comparison and validation of other techniques like solution NMR, cryoEM, or atomistic MD simulations. In all cases, within certain quality indicators, the obtained structure is taken at face value and the expected distortions that could arise from packing effects (among others) are assumed to be low or even negligible. We computed the global RMSD of MD simulations (crystal and solution) with respect to the experimental crystal configuration (Table 2), confirming earlier results showing that the average MD structure in solution deviates much more from the crystal configuration (with larger deviation in the backbone) than does the average structure from the crystal simulation. Furthermore, we compared the helical space and groove dimensions of the crystals with solution MD simulations, first confirming that different unbiased solution trajectories, starting from different initial structures, all converged to exactly the same structure (Supp. Figure S20 and S21). As shown in Figure 5, the solution simulations were generally able to sample the X-ray values within the first standard deviation, but a close observation clearly shows small discrepancies between the average MD solution simulations and the X-ray values at several bps. As expected from the global RMSD discussed above, when we compared the X-ray values with the crystal simulations, most of these small discrepancies at the bps level disappeared, in particular for shift, rise (Supp. Figure S22), roll and twist (Supp. Figure S23). The differences seen between the average X-ray values and solution MD could then be due to crystal packing effects and the crowed environment that modifies the solvent-solute interactions.

We also compared the solution simulations to crystal MD by analyzing the flexibility of each duplex and their dynamical behavior. The force constants associated with the deformation of a single helical degree of freedom per base pair step were compared between solution and crystal simulations. As expected, the DNA in the crystals was much more rigid than in solution, with a large increase in rigidity at the ends of the duplexes (Supp. Figure S24), in agreement with the effective temperatures discussed in the previous section. Interestingly, the central portion of the DNA always followed the same sequence-dependent pattern of flexibility as the one observed for solution simulations, for all degrees of freedoms and for all studied crystals. Furthermore, we obtained the essential dynamics of DNA through principal component analysis of MD trajectories and we built similarity indexes based on the overlap of the first 10 eigenvectors. According to those indexes, the simulations shared a high degree of similarity: from 93 to 98% for solution *vs* solution simulations, from 76 to 95% for crystal *vs* crystal MD, and from 65 to 84% when comparing crystal to solution (Supp. Figure S25), suggesting that the dynamics of the principal movements of the duplexes in the crystals slightly differ from the ones in solution. These results support the idea that the small structural discrepancies in the average helical parameters discussed in the

previous paragraph arose from the dynamic, rather than static, properties of the DNA.

We finally focused on two paradigmatic structural polymorphisms that seem to be causally connected, namely the backbone BI/BII polymorphism (related to ε and ζ torsions), and the bimodal behavior of the twist helical parameter observed for d(CpG) steps. BII percentages have been computed in the past from NMR in solution and compared with those extracted from crystal structures of free B-DNA, showing that the solution and crystal states often differ for steps with a significant propensity for populating BII without a clear sequence dependent pattern. Likewise, discrepancies are observed when comparing BII percentages coming from solution or crystal MD simulations (Supp. Figure 26). In agreement with experiments, less BII % was observed for d(CpG) and d(GpA) in the crystal than in the solution simulations, while the inverse situation was observed for d(GpC), and quite similar results were observed for d(ApA) and d(ApT). This would suggest that the differences in BII propensities between X-ray crystals and NMR solutions are mainly due to crystal packing effects. Polymorphisms at the base level were much more difficult to detect experimentally, something that was attributed to the resolution of the crystals and the "substate-washing-effect" produced by the averaging of structures present in the crystal cell. We analyzed the twist polymorphism at the d(CpG) step for the three simulated crystals and compared the results to the solution trajectories. We found that the bimodal behavior was mitigated in the crystals due to both averaging effects and a lower propensity for the bimodal twist (Supp. Figures 27-29). The two and three substates observed in the 2D density plots of roll vs twist and twist vs tilt, respectively, were all found among the duplexes present in the crystal units, but not simultaneously in a unique duplex as in the solution simulations. This correlated perfectly with the reduced amount of cations found inside the minor groove of d(CpG) steps in the crystals simulations (see previous section), since K^+ and Na^+ were proposed as the main driving forces of the polymorphic twist behavior. These results suggest that the solvent environment created around the duplexes in the crowded crystal affects DNA's dynamics and limits its conformational space compared to the free DNA in solution.

REFERENCES

1. Schrodinger, L.L.C. The PyMOL Molecular Graphics System, Version 1.8. (2015).
2. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.***79**, 926-935 (1983).

3. Mahoney, M.W. & Jorgensen, W.L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **112**, 8910-8922 (2000).
4. Smith, D.E. & Dang, L.X. Computer simulations of NaCl association in polarizable water. *The Journal of Chemical Physics* **100**, 3757-3766 (1994).
5. Allnér, O., Nilsson, L. & Villa, A. Magnesium ion-water coordination and exchange in biomolecular simulations. *J. Chem. Theory Comput.* **8**, 1493-1502 (2012).
6. Yoo, J. & Aksimentiev, A. Improved parametrization of Li^+, Na^+, K^+, and Mg^{2+} ions for all-atom molecular dynamics simulations of nucleic acid systems. *The Journal of Physical Chemistry Letters* **3**, 45-50 (2012).
7. Yoo, J. & Aksimentiev, A. Improved parameterization of amine-carboxylate and amine-phosphate interactions for molecular dynamics simulations using the CHARMM and AMBER force fields. *J. Chem. Theory Comput.* **12**, 430-443 (2016).
8. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
9. Abraham, M.J., et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19-25 (2015).
10. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182-7190 (1981).
11. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald - an N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092 (1993).
12. Essmann, U., et al. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577-8593 (1995).
13. Ivani, I., et al. Parmbsc1: A refined force-field for DNA simulations. *Nature methods* **13**, 55-58 (2016).
14. Hospital, A., Candotti, M., Gelpi, J.L. & Orozco, M. The multiple roles of waters in protein solvation. *Journal of Physical Chemistry B* **121**, 3636-3643 (2017).
15. Drsata, T. & Lankas, F. Multiscale modelling of DNA mechanics. *Journal of physics. Condensed matter : an Institute of Physics journal* **27**, 323102 (2015).
16. Hess, B. Similarities between principal components of protein dynamics and random diffusion. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics* **62**, 8438-8448 (2000).
17. Lankaš, F., Šponer, J., Hobza, P. & Langowski, J. Sequence-dependent elastic properties of DNA. *J. Mol. Biol.* **299**, 695-709 (2000).

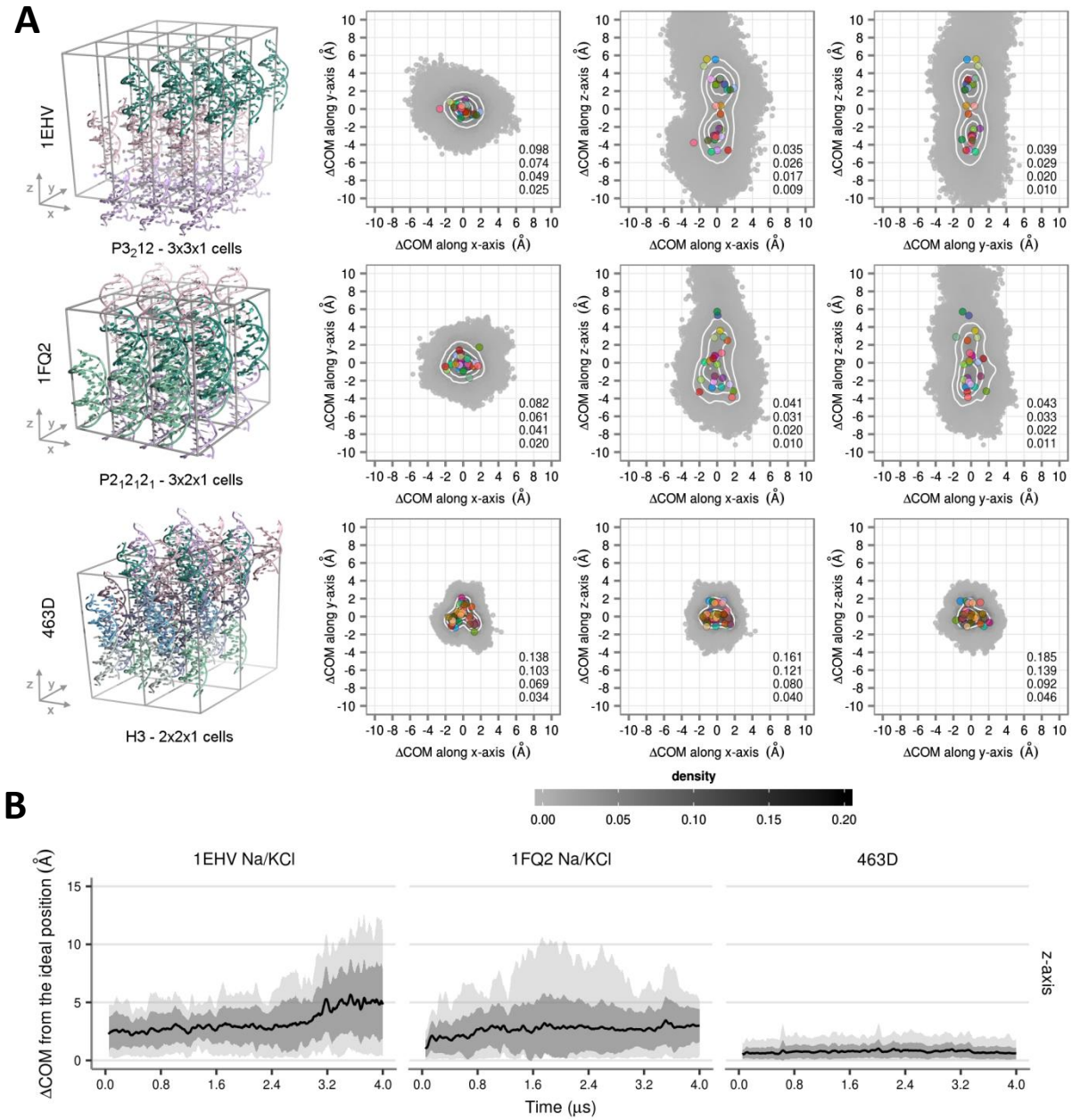

Figure 1. A) COM displacements from the ideal crystallographic positions shown for each duplex in the simulated supercell systems (1EHV, 1FQ2, 463D) along principal axes. Points are colored in greyscale according to the smoothed 2D densities, estimated by fitting the observed distributions to a bivariate normal kernel (evaluated on a square grid of 90 × 90 bins). Four isodensity curves are shown in white and are quantified on the bottom right side of each plot. The average positions of the duplexes are given as colored circles. B) COM displacements in time along the z-axis shown for all three systems. Mean values are shown as black lines, standard deviations are shown as dark grey ribbons, while the extreme values are shown with light grey ribbons (lower panels).

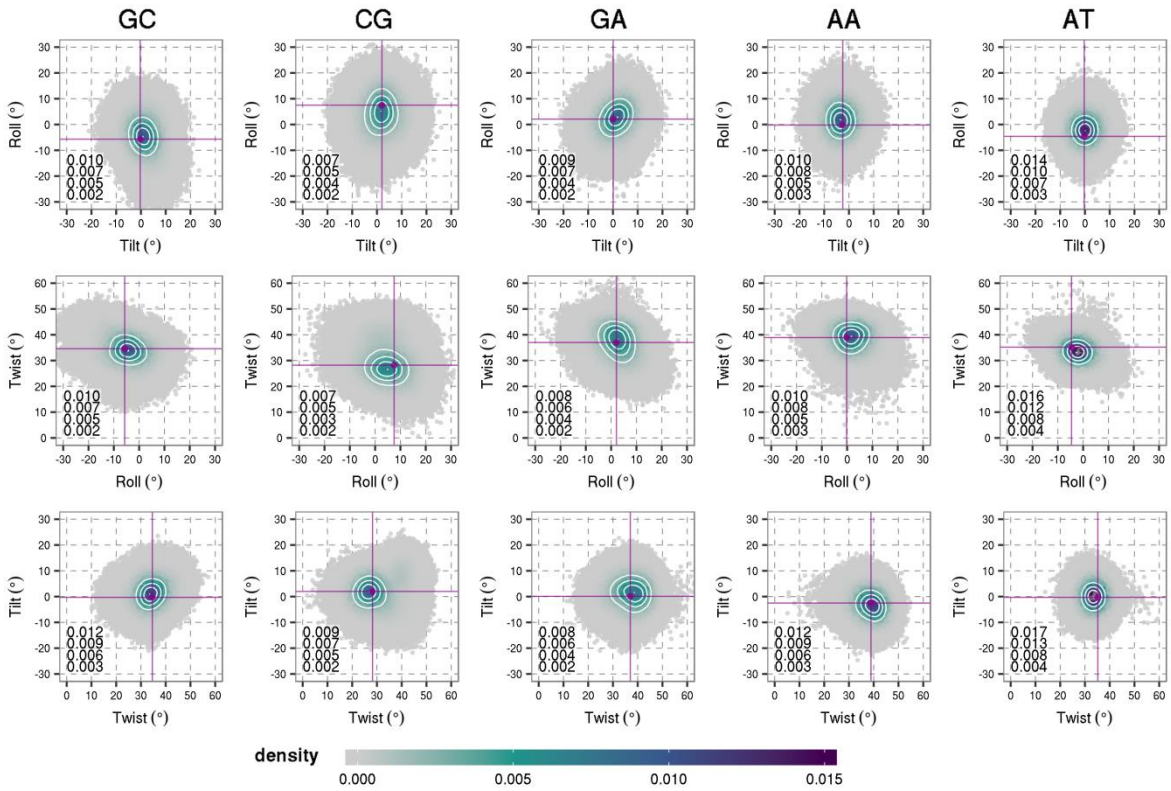

Figure 2. Rotational helical parameters (roll, twist, tilt) of all the duplexes in the 463D system (given for distinct base pair steps in DDD after removing the ends). The values calculated from the X-ray structure are shown with purple lines. Points are colored according to the smoothed 2D densities, estimated by fitting the observed distributions to a bivariate normal kernel (evaluated on a square grid of 90 × 90 bins). Four isodensity curves are shown in white and are quantified on the bottom right side of each plot.

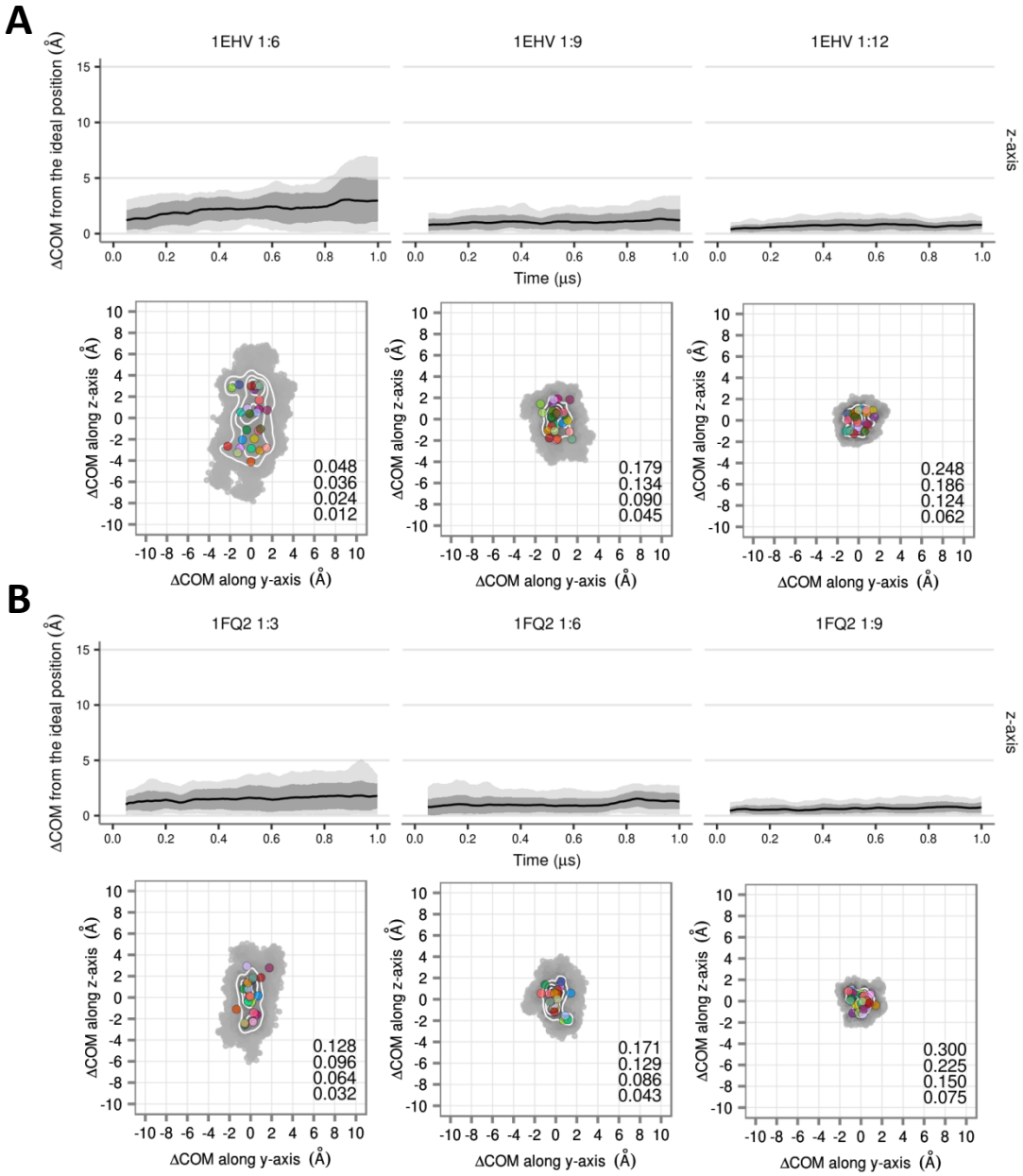

Figure 3. COM displacements from the ideal crystallographic positions for different spermine concentrations in A) 1EHV and B) 1FQ2 systems. The upper panels show the COM displacements in time along the z-axis where mean values are shown as black lines, standard deviations are shown as dark grey ribbons, while the extreme values are shown with light grey ribbons. The lower panels show COM displacements shown for each duplex along y and z axes with the average positions of the duplexes given as colored circles. Points are colored in greyscale according to the smoothed 2D densities, estimated by fitting the observed distributions to a bivariate normal kernel (evaluated on a square grid of 90 × 90 bins). Four isodensity curves are shown in white and are quantified on the bottom right side of each plot.

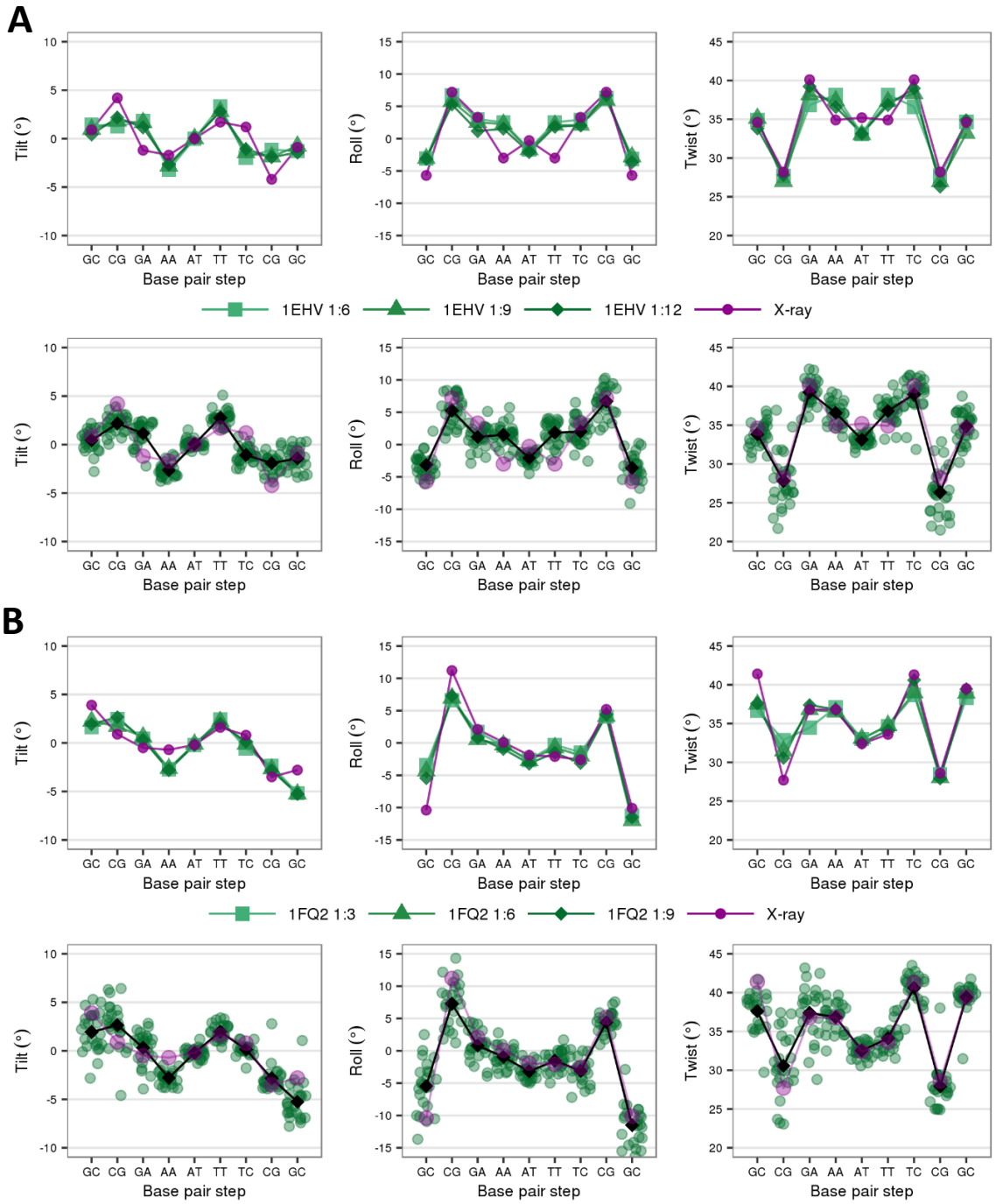

Figure 4. Rotational helical parameters (roll, twist, tilt) (given for all base pair steps in DDD after removing the ends) for different spermine concentrations in A) 1EHV and B) 1FQ2 systems. The upper panels show average values across all duplexes for varying spermine concentrations (colored green), with the purple line giving values calculated for the X-ray structure. The lower panels show the average values for each duplex in the systems with the highest spermine concentration (shown as green circles), while the average over all duplexes is shown as a black line and the X-ray values in purple.

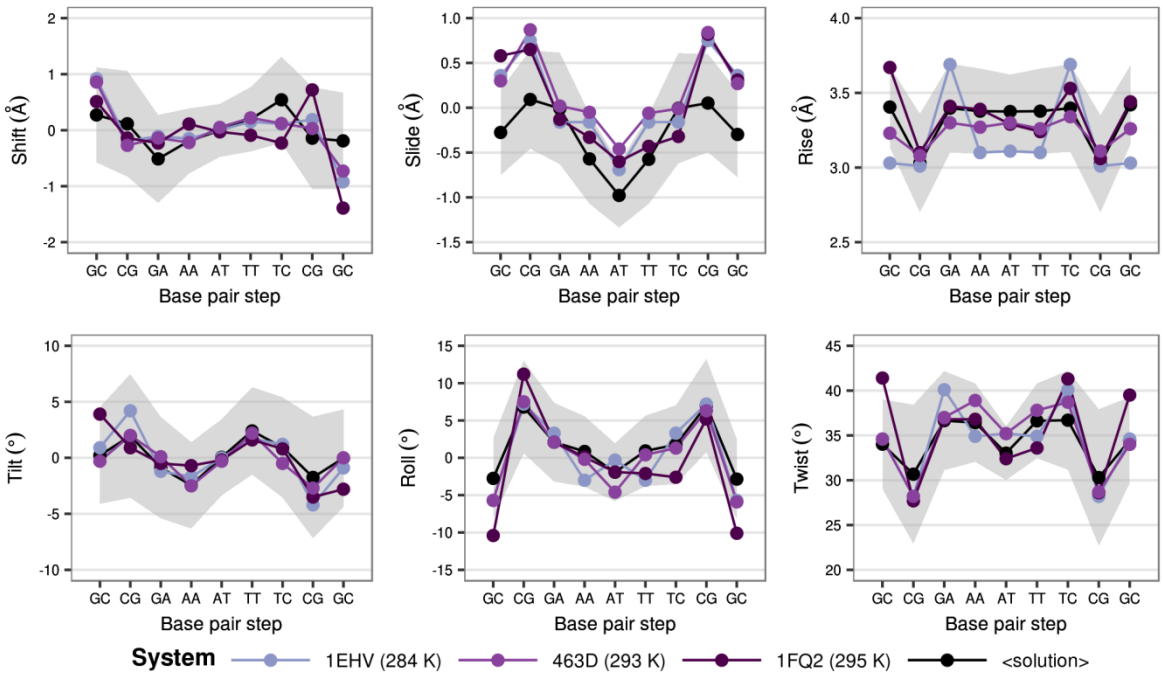

Figure 5. Translational (shift, slide, rise) and rotational (roll, twist, tilt) helical parameters (given for all base pair steps in DDD after removing the ends) for the solution simulations of single duplexes starting from different X-ray structures. The average values across the simulations are shown in black, while the average standard deviations are shown as grey ribbons. Values calculated from X-ray models are shown as colored lines.

TABLES

Table 1. Studied systems and simulated conditions.

System name	PDB	Environment	Space group	Nº of unit cells	Nº of DNA duplexes	Salt type	[Salt] (mM)	dsDNA to Spm ratio	Temp. (K)	Equilibration time (ps)	Simulation time (μs)
1EHV sol	1EHV	solution	---	---	1	Na/KCl	400	---	284	600	1
15 mM	1EHV	crystal	$P3_212$	3x3x3	81	Na/KCl	15	---	284	600	1
200 mM	1EHV	crystal	$P3_212$	3x3x3	81	Na/KCl	200	---	284	600	1
400 mM	1EHV	crystal	$P3_212$	3x3x3	81	Na/KCl	400	---	284	600	4
600 mM	1EHV	crystal	$P3_212$	3x3x3	81	Na/KCl	600	---	284	600	1
800 mM	1EHV	crystal	$P3_212$	3x3x3	81	Na/KCl	800	---	284	600	1
1EHV Na/KCl	1EHV	crystal	$P3_212$	3x3x1	27	Na/KCl	400	---	284	600	4
1EHV Mg	1EHV	crystal	$P3_212$	3x3x1	27	$MgCl_2$	400	---	284	600	4
1EHV Mg hex	1EHV	crystal	$P3_212$	3x3x1	27	$MgCl_2{\cdot}6H_2O$	400	---	284	600	4
1EHV 1:6	1EHV	crystal	$P3_212$	3x3x1	27	$SpmCl_4$	---	1:6	284	6000	1
1EHV 1:9	1EHV	crystal	$P3_212$	3x3x1	27	$SpmCl_4$	---	1:9	284	6000	1
1EHV 1:12	1EHV	crystal	$P3_212$	3x3x1	27	$SpmCl_4$	---	1:12	284	6000	1
1FQ2 sol	1FQ2	solution	---	---	1	Na/KCl	400	---	295	600	1
1FQ2 1:6 sol	1FQ2	solution	---	---	1	$SpmCl_4$	---	1:6	295	600	1
1FQ2 Na/KCl	1FQ2	crystal	$P2_12_12_1$	3x2x1	24	Na/KCl	400	---	295	600	4
1FQ2 Na/KClequ	1FQ2	crystal	$P2_12_12_1$	3x2x1	24	Na/KCl	400	---	295	6000	1
1FQ2 1:3	1FQ2	crystal	$P2_12_12_1$	3x2x1	24	$SpmCl_4$	---	1:3	295	6000	1
1FQ2 1:6	1FQ2	crystal	$P2_12_12_1$	3x2x1	24	$SpmCl_4$	---	1:6	295	6000	1
1FQ2 1:9	1FQ2	crystal	$P2_12_12_1$	3x2x1	24	$SpmCl_4$	---	1:9	295	6000	1
1FQ2 1:9 CUFIX	1FQ2	crystal	$P2_12_12_1$	3x2x1	24	$SpmCl_4$ CUFIX	---	1:9	295	6000	1
1FQ2 1:9 Mg	1FQ2	crystal	$P2_12_12_1$	3x2x1	24	$SpmCl_4$ + $MgCl_2{\cdot}6H_2O$	---	1:9	295	6000	1
463D sol	463D	solution	---	---	1	Na/KCl	400	---	293	600	1
463D	463D	crystal	H3	2x2x1	36	Na/KCl	400	---	293	6000	4

Table 2. Root mean square deviations (RMSD, Å) from the deposited crystal structures for 463D, 1EHV, and 1FQ2. The statistics in each box are heavy atoms, backbone atoms, and base atoms RMSD of average MD structure with respect to the crystal structure. The values in parentheses exclude the terminal residues.

	All heavy atoms	Backbone	Base	All heavy atoms (sol)	All heavy atoms (sol + SPM)
463D	1.72 (0.86)	1.39 (0.99)	2.03 (0.65)	3.20 (1.20)	-
1EHV 1:12	0.96 (0.83)	1.00 (0.95)	0.90 (0.64)	3.20 (1.15)	-
1FQ2 1:9	0.57 (0.52)	0.64 (0.60)	0.47 (0.41)	1.45 (1.19)	1.49 (1.23)