

# Understanding nucleic acid structural changes by comparing wide-angle x-ray scattering (WAXS) experiments to molecular dynamics simulations

Suzette A. Pabit,<sup>1</sup> Andrea M. Katz,<sup>1</sup> Igor S. Tolokh,<sup>2</sup> Aleksander Drozdetski,<sup>3</sup>  
Nathan Baker,<sup>4</sup> Alexey V. Onufriev,<sup>2,3</sup> and Lois Pollack<sup>1</sup>

<sup>1</sup>*School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853, USA*

<sup>2</sup>*Department of Computer Science, Virginia Tech, Blacksburg, Virginia 24061, USA*

<sup>3</sup>*Department of Physics, Virginia Tech, Blacksburg, Virginia 24061, USA*

<sup>4</sup>*Pacific Northwest National Laboratory, Richland, Washington 99352, USA*

(Received 2 December 2015; accepted 5 May 2016; published online 24 May 2016)

Wide-angle x-ray scattering (WAXS) is emerging as a powerful tool for increasing the resolution of solution structure measurements of biomolecules. Compared to its better known complement, small angle x-ray scattering (SAXS), WAXS targets higher scattering angles and can enhance structural studies of molecules by accessing finer details of solution structures. Although the extension from SAXS to WAXS is easy to implement experimentally, the computational tools required to fully harness the power of WAXS are still under development. Currently, WAXS is employed to study structural changes and ligand binding in proteins; however, the methods are not as fully developed for nucleic acids. Here, we show how WAXS can qualitatively characterize nucleic acid structures as well as the small but significant structural changes driven by the addition of multivalent ions. We show the potential of WAXS to test all-atom molecular dynamics (MD) simulations and to provide insight into understanding how the trivalent ion cobalt(III) hexammine (CoHex) affects the structure of RNA and DNA helices. We find that MD simulations capture the RNA structural change that occurs due to addition of CoHex. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4950814>]

## I. INTRODUCTION

The accessibility of synchrotron radiation sources for Small Angle X-ray Scattering (SAXS) experiments has enabled new methods for measuring low resolution structures of biological macromolecules.<sup>1,2</sup> The spatial resolution  $d$  of a SAXS measurement is largely determined by the highest  $q$  value of the measurement, given by  $d = 2\pi/q$ , where  $q$  is the momentum transfer,  $q = (4\pi/\lambda) \cdot \sin(2\theta/2)$ ,  $\lambda$  is the x-ray wavelength and  $2\theta$  is the scattering angle.<sup>3</sup> Extension from small-angle to wide-angle x-ray scattering (WAXS) is appealing because scattering at wider angles yields higher resolution structural features.<sup>3,4</sup> For example, even a modest maximum  $q$  at around  $1.0 \text{ \AA}^{-1}$  corresponds to features that are in the  $5\text{--}10 \text{ \AA}$  spatial range. Implementation of WAXS is simple and often requires a trivial modification to a SAXS beamline: the detector is moved closer to the sample to capture x-rays scattered to higher angles. It is even possible to simultaneously acquire SAXS and WAXS data by placing a detection window near the sample.<sup>5</sup> WAXS provides information about length scales of critical structural importance and is particularly relevant for studies of nucleic acid structures. For example, in DNA, peaks in the WAXS regime of  $0.4 < q < 1.0 \text{ \AA}^{-1}$  arise from interstrand pair distance correlations from major and minor groove spacing and helix radius.<sup>6</sup>

Despite the relative ease of acquiring WAXS data, the challenge in its application arises from the interpretation of the measured scattering profiles. On these shorter length scales, for example, the contribution of the solvent, including the

hard-to-model hydration shell around biomolecules, becomes significant.<sup>7</sup> A number of experimental and computational tools for the analysis of WAXS structures of biomolecules are available but all rely on the availability of high resolution atomic coordinates for comparison.<sup>3,8,9</sup> Since atomic coordinates are not readily available for many biomolecules, this limits the applicability of WAXS analysis software. Moreover, most of these analysis tools are geared towards studies of protein structural fluctuations and protein ligand binding,<sup>3,10,11</sup> where the hydration layer is reasonably well described. In contrast, because of the highly charged nature of nucleic acids, accurate models of the dense solvent layer adjacent to the molecule are still under development. A method to calculate nucleic acid SAXS and WAXS profiles that include not only the hydration shell but also charge compensating counterions has recently been introduced<sup>12</sup> but still requires input of matching atomic coordinates of nucleic acid structures.

The goal of this paper is to extend the use of WAXS to studies of nucleic acids. We discuss two cases where interpretation of WAXS experimental data is possible. First, when key structural features are present in the nucleic acid scattering profiles, we can compare the location of the peaks and valleys to those calculated from existing structural models of nucleic acids based on canonical helices and Molecular Dynamics (MD) simulation results. In this way, WAXS data can provide experimental benchmarks that can lead to development and refinement of MD simulations. When data and models agree, MD results can help to visualize WAXS-resolved structural features. The second way to interpret

nucleic acid WAXS data is to look at changes in the WAXS profiles induced by the addition of ligands or multivalent ions. Intensity difference curves have been used in the study of WAXS structures in proteins, particularly for comparing different time points in time-resolved experiments.<sup>13,14</sup> Here, we show that a comparison of intensity difference curves (MD to experiment) is a useful way to interpret changes in the nucleic acid structure on the nanometer length scale.

## II. EXPERIMENTAL AND COMPUTATIONAL DETAILS

### A. Background on wide-angle X-ray scattering (WAXS)

Small Angle X-ray Scattering (SAXS) is a useful experimental technique that typically provides low resolution structural information: a measure of the radius of gyration, pair-distance distribution function, and computation of shape envelopes of proteins and nucleic acids in solution.<sup>1,2</sup> A standard SAXS setup is shown schematically in the top panel of Figure 1. X-rays are incident on a biological sample; the scattered x-rays pass through a 1–2 m vacuum flight path and are collected by a photon detector. The collected scattered intensity  $I$  is displayed as a function momentum transfer  $q$ , given by  $q = (4\pi/\lambda) \cdot \sin(2\theta/2)$ , where  $\lambda$  is the x-ray wavelength and  $2\theta$  is the scattering angle.<sup>3</sup>

Wide-angle x-ray scattering (WAXS) experiments, shown schematically in the bottom panel of Figure 1, typically use the same setup as SAXS but with a shorter scattering flight path which allow access to higher resolution structural information.<sup>3,4</sup> The spatial resolution  $d$  of a scattering experiment is largely determined by the highest  $q$  value of the measurement, given by  $d = 2\pi/q$ . For example, shortening the sample-to-detector distance to 0.4–0.5 m allows access to a higher  $q$ -range;  $q$  at around  $1.0 \text{ \AA}^{-1}$  corresponds to features that are in the 5–10  $\text{\AA}$  spatial range. WAXS data acquisition is becoming a popular addition to SAXS experiments by simple installation of a WAXS detection window closer to the sample.<sup>5</sup> Therefore, there is a need for the development of more analytical tools to understand WAXS data.

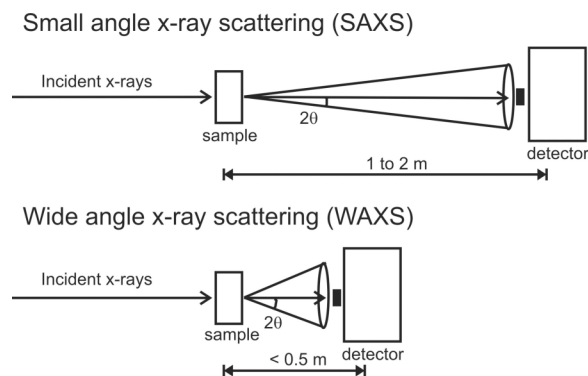


FIG. 1. Schematic of solution-based x-ray scattering experiments in the small-angle (top) and wide-angle (bottom) regimes. The main difference between SAXS and WAXS setups is the sample-to-detector distance. The wider angles in WAXS experiments allow access to high resolution information in the 5–10  $\text{\AA}$  spatial range.

### B. Sample preparation and experimental conditions

Double-stranded 25 base-pair (bp) DNA and RNA constructs were annealed from single-stranded nucleic acids purchased from IDT (Coralville, IA). We used the same mixed sequence described in our previous publications, GCA TCT GGG CTA TAA AAG GGC GTC G (U replacing T for the RNA strands).<sup>15,16</sup> Cobalt(III) hexammine chloride,  $\text{Co}(\text{NH}_3)_6\text{Cl}_3$  (CoHex), sodium chloride, NaCl, and the buffers used were purchased from Sigma-Aldrich (St. Louis, MO). The duplexes were extensively dialyzed using Amicon Ultra Concentrators (Millipore, Billerica, MA) in pH 7.0 1 mM Na-MOPS buffer containing either 100 mM NaCl only or 100 mM NaCl and 0.8 mM  $\text{Co}(\text{NH}_3)_6\text{Cl}_3$ . The monovalent salt concentrations were chosen to prevent trivalent CoHex induced precipitation and focus the study on CoHex-nucleic acid interactions in the pre-condensed solution phase. The WAXS experiments were carried out at the G1 station at the Cornell High Energy Synchrotron Source (CHESS). Using a scattering flight path with a 0.455 m sample to detector distance, we reached a  $q_{\text{max}} = 0.95 \text{ \AA}^{-1}$ , which is sensitive to correlation lengths of  $d_{\text{min}} = 2\pi/q_{\text{max}} = 6.6 \text{ \AA}$ . This allows us to target the WAXS region of  $0.4 < q < 0.95 \text{ \AA}^{-1}$  which corresponds to the length scales of the helix radius and the minor and major groove spacing.<sup>6</sup> We purposely avoided  $q > 1.0 \text{ \AA}^{-1}$  where contributions from solvent scattering become more intense.<sup>3</sup> The x-ray energy used was 10.60 keV and scattered x-rays were imaged using a low-noise photon counting area detector (Pilatus 100 K, Dectris, Baden, Switzerland). Duplex concentrations were about 450  $\mu\text{M}$ . The samples, with volumes of about 30  $\mu\text{l}$ , were kept in a 2-mm quartz capillary (Hampton Research, Aliso Viejo, CA) and radiation damage was prevented by oscillating the sample during the data collection. Signals from the buffer background were subtracted from the data and absolute calibration was performed using water as a calibrant as described previously.<sup>12,17</sup> Data analysis was done using code written in MATLAB (Mathworks, Natick, MA). A more detailed description of the experimental measurements is provided in Ref. 12.

### C. Molecular Dynamics (MD) simulations and comparison with WAXS data

MD simulations of RNA and DNA constructs in explicit water and monovalent ions (NaCl) were carried out in the presence and absence of CoHex ions using AMBER 12<sup>20</sup> and ff99bsc0 force fields<sup>21,22</sup> which includes monovalent ion parameters.<sup>23,24</sup> All simulated systems included about 16 880 TIP3P water molecules.<sup>25</sup> The systems with CoHex included 16 CoHex ions using parameters from Cheatham and Kollman<sup>26</sup> which are part of the AMBER12 distribution. The systems without CoHex were neutralized by 48  $\text{Na}^+$  ions. Monovalent salt background in all the systems was accounted for by additional 24 NaCl ion pairs. The starting structures for the MD simulations were generated from canonical helices (A-form for RNA and B-form for DNA) built using Nucleic Acid Builder (NAB)<sup>18</sup> and shown in Figure 2. All the systems were initially equilibrated for 0.5 ns in canonical ensemble (NVT) and 0.5 ns in isothermal-isobaric ensemble (NPT) using 1 fs time step and achieving 1 atm pressure and 300 K temperature.



FIG. 2. DNA and RNA double helices have the same charge,  $-2e$  per base pair, but different helical structures. Shown here are 25-base-pair long canonical double helices of B-form DNA (left) and A-form RNA (right) generated by Nucleic Acid Builder (NAB)<sup>18</sup> using the sequences used in the experiment. These helices are the launching point of the molecular dynamics simulations used and the basis for comparison of the WAXS experimental spectra to the canonical spectra generated using the program CRY SOL.<sup>19</sup> The canonical structures shown are drawn using Accelrys Discovery Studio program suite.

The following simulations were carried out in NVT ensemble using 2 fs time step, periodic boundary conditions, the particle mesh Ewald method at 300 K maintained using Langevin dynamics with the collision frequency of  $1 \text{ ps}^{-1}$ . In the first 140 ns of simulation, a restraining harmonic potential with a force constant of  $50 \text{ kcal/mol/\AA}^2$  was applied to the canonical

forms of the RNA and DNA molecules. The restraining force was removed and the simulations were allowed to proceed for another 100-200 ns.

To compare the MD results with the WAXS data, we extracted 100-500 snapshots<sup>27</sup> from each trajectory in PDB format and generated WAXS profiles using the program CRY SOL.<sup>19</sup> CRY SOL is a commonly used program for calculating the solution scattering profiles of macromolecules with known atomic coordinates. Though other WAXS solvers are available and a simple test of the program WAXSiS<sup>9</sup> yielded similar profiles, we chose CRY SOL for speed of the calculations and ease of handling hundreds of PDB snapshots. The CRY SOL input settings were configured so that the output scattering profiles are fitted to the experimental data. All other parameters were set to the CRY SOL default values (e.g., solvent density set to pure water). We report all CRY SOL results in absolute intensity units ( $e^2$ ). To focus on the nucleic acid structural features, the ions were generally not included in the CRY SOL calculations. The inclusion of the ions is discussed in the latter part of the paper (see Section IV A). Profiles generated from hundreds of PDB snapshots were pooled together into two groups, MD profiles with the restraining force in place (“restrained” profiles) and MD profiles after the restraining force was released (“free” profiles). The pools were averaged and these averages were compared to the data.

### III. WAXS RESULTS

#### A. Nucleic acids in NaCl

Figure 3 compares background subtracted experimental WAXS profiles of nucleic acids in NaCl to canonical helices and MD simulation profiles. We show intensity  $I$  vs.

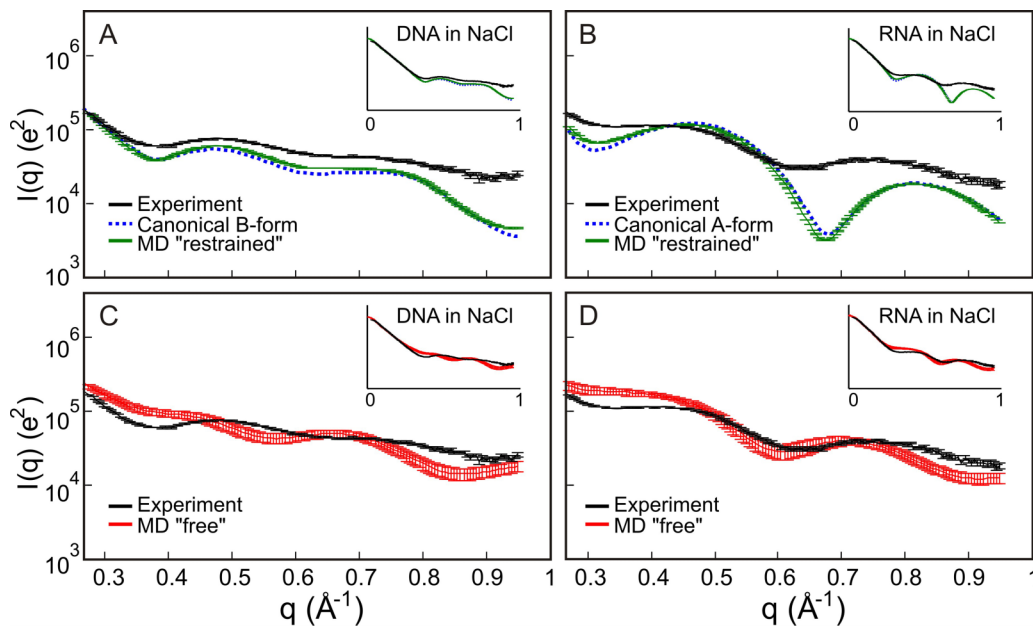


FIG. 3. Comparison of wide-angle scattering data from (a) DNA and (b) RNA in 100 mM NaCl to canonical helices built using Nucleic Acid Builder (NAB) and structural ensembles generated by “restrained” MD simulations. The small differences between NAB and “restrained” MD reflect the small deviation from the restraining harmonic potential. A similar comparison is made between (c) DNA and (d) RNA data and “free” MD structures after the restraining force is removed. We magnified the features in the WAXS regime,  $q > 0.4 \text{ \AA}^{-1}$ , and show the full scattering profile as insets.

$q$  curves in logarithmic scale to enhance structural features in the WAXS regime, the higher- $q$  region of the dataset ( $q > 0.4 \text{ \AA}^{-1}$ ). Displaying the data in this form allows us to easily assess whether the general features and peaks of the experimental profiles agree or disagree with the simulations and the canonical helices built using NAB. In Figure 3(a), we compare our experimental profiles of DNA in NaCl to a canonical B-form helix and MD simulations in NaCl and explicit water with the restraining force in place (“restrained” profiles). Since the “restrained” MD models are derived from NAB, the calculated profiles are very similar with very small deviations due to the restraining harmonic potential. Note the relative agreement between the experimental and the calculated “restrained” MD and NAB profiles. They exhibit the characteristic features that were seen in previous measurements<sup>6</sup> on B-form DNA with a peak measured at  $q = 0.48 \text{ \AA}^{-1}$  and a shallow peak detected at  $q = 0.72 \text{ \AA}^{-1}$ . Although agreement is qualitative, the peak positions are in the right place. Thus, scattering profiles for DNA in NaCl compare favorably with predictions of “restrained” MD simulations. In contrast, the scattering profile of RNA in NaCl shown in Figure 3(b) looks substantially different from the “restrained” MD profiles. As expected, the MD simulation results using “restrained” RNA recapitulate the features exhibited in the NAB canonical A-form structure because of the restraining potential. However, both curves disagree with the experimental data. The first peak in the experimental WAXS profiles for RNA is too shallow (more like a shoulder than a peak) and shifted to the left compared to the “restrained” MD profile; the second peak is also shifted to the left.

In Figures 3(c) and 3(d), we compare the experimental data with the MD simulation results after the restraining force has been released (“free” profiles). Interestingly, the WAXS profiles computed for “free” DNA, Figure 3(c), no longer capture the agreement with experiment as seen in the “restrained” models. A close examination of the WAXS profiles around  $q = 0.5 \text{ \AA}^{-1}$  and  $q = 0.7 \text{ \AA}^{-1}$  of unrestrained DNA curves in Figure 3(c) shows that the peak positions are completely different and the simulation curves appear “out-of-phase” with the data. On the other hand, the features of the RNA WAXS data in NaCl are better captured by the MD simulations after the restraining force has been released. In Figure 3(d), we note that the shoulder in the “free” MD near  $q = 0.4 \text{ \AA}^{-1}$  also appears in the experimental curves. However, the second peak position is not exactly in the same place. Nonetheless, at  $q = 0.71 \text{ \AA}^{-1}$ , it is closer to the experimental peak at  $q = 0.74 \text{ \AA}^{-1}$  than the peak value of  $q = 0.82 \text{ \AA}^{-1}$  for the “restrained” MD. The peak positions in Figures 3(c) and 3(d) show that there is definite improvement in the simulated RNA spectra when unrestrained MD is performed but not in DNA. This result suggests that in an NaCl solution, duplex RNA does not have a rigid canonical A-form structure while DNA maintains the B-form canonical structure in NaCl (Figure 3(a)).

## B. Nucleic acids in CoHex

In Figure 4, we examine structural changes in DNA and RNA induced by the addition of multivalent ions. Cobalt(III) hexammine,  $[\text{Co}(\text{NH}_3)_6]^{3+}$  or CoHex, is a trivalent ion that

is widely used to facilitate and study DNA condensation.<sup>28,29</sup> This highly charged spherical ion has a dramatically different effect on short double helices made of DNA and RNA. The addition of a small amount of CoHex ions precipitates short DNA helices from solution, while RNA molecules of similar sequence remain soluble.<sup>15,16</sup> The efficiency of condensation appears to be primarily related to the mode of CoHex binding, which is mostly determined by the helix geometry, although other factors like nucleotide sequence do contribute.

In a related study,<sup>16</sup> we used UV spectroscopy and circular dichroism (CD) to monitor CoHex-induced precipitation of DNA and RNA helices and compared MD simulations of the different duplexes. We showed that simulations of B-form DNA suggest that CoHex ions decorate the outer surface of the structure, while MD simulations of A-like RNA suggest that the ions are drawn deep into the negatively charged major groove. There, we discussed how ion placement influences precipitation. However, a limitation of that previous study<sup>16</sup> lies in our inability to directly compare MD simulations to experimental data. Although CD reports spectral changes in RNA, there is no straightforward way to correlate the changes in nucleic acid CD spectra with MD predictions. In fact, our attempts to use the CD modeling software Dichrocalc,<sup>30</sup> which was written for proteins, yielded results that substantially differ from the CD data. Here, we explore the use of WAXS as an experimental way to check if changes in the structures of DNA and RNA duplexes due to CoHex are predicted by simulation.

Although the signal from the CoHex ions is too small to be detected experimentally in the WAXS profiles, CoHex-driven changes in local structure, suggested by the CD data, should be noticeable. Figure 4 shows WAXS profiles of the nucleic acids in a CoHex-NaCl solution and compares the profiles with simulations. Although we observe some lineshape broadening in the WAXS peaks of the DNA samples, the peak positions do not move after addition of CoHex. The same B-form DNA structure is still captured in Figure 4(a) with “restrained” MD. This behavior is expected from DNA molecules since Circular Dichroism (CD) measurements showed that CD changes in DNA with and without CoHex are minimal while a large spectral shift in the CD spectra for RNA was observed upon the addition of CoHex.<sup>16</sup> In agreement with CD, distinct changes are observed in the RNA WAXS profiles upon addition of CoHex. In fact, in the presence of CoHex, the WAXS structures of RNA appear more like canonical A-form helix and have peaks in similar positions to the “restrained” MD (Figure 4(b)). Similar to DNA in NaCl, the “free” MD profiles for DNA in CoHex show disagreement in peak positions (Figure 4(c)). For RNA, peak positions in the “free” MD profiles qualitatively agree with the data (Figure 4(d)).

## C. Intensity difference profiles before and after CoHex

To better compare changes in structural features in WAXS profiles due to the addition of CoHex, we look at differences in the scattering intensity following the method used in Refs. 3, 13, 14, and 31. In Figure 5, we show experimental intensity difference profiles,  $\log(I_{\text{NaCl}}) - \log(I_{\text{CoHex}})$ , and compare them to the difference profiles computed from the simulated curves. Since all the experimental data shown have absolute

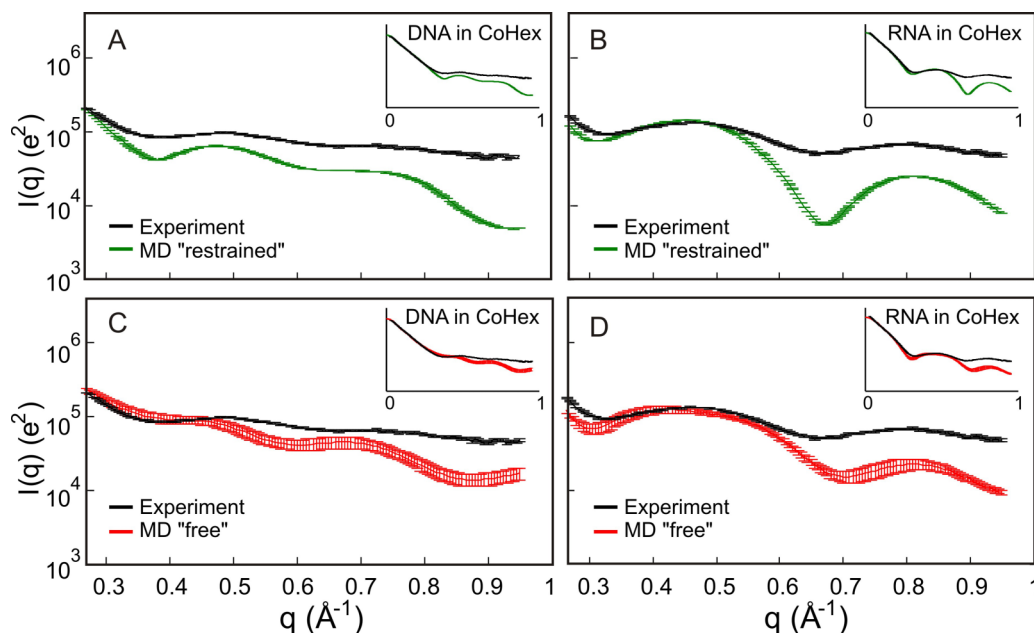


FIG. 4. Comparison of wide-angle scattering data from nucleic acids in 0.8 mM CoHex and 100 mM NaCl solution to “restrained” and “free” MD simulations. (a) and (c) are for DNA and (b) and (d) are for RNA. The number of CoHex and NaCl in the simulation box complements the experiment. We put emphasis in the features in the WAXS regime,  $q > 0.4 \text{\AA}^{-1}$ , by magnification and show the full scattering profile as insets.

calibration and the MD profiles calculated using CRYSOLE are made to fit the data, no scaling was applied prior to the subtraction. Looking at all panels in Figure 5, we immediately see that the key features in the intensity difference curves are best captured by MD results shown in Figure 5(d) where unrestrained or “free” MD difference profiles are compared to the experimental difference profiles for RNA. The difference peaks at  $q = 0.3 \text{\AA}^{-1}$  and at  $q = 0.7 \text{\AA}^{-1}$  in the RNA experimental curves also appear in the MD simulations albeit with different amplitudes. Makowski and co-workers<sup>3,31</sup> have shown that when intensity difference profiles in WAXS are

calculated using CRYSOLE,<sup>19</sup> the difference curves correspond more closely to experimental results than the absolute intensities. This result is quite interesting: it suggests that CoHex-induced changes in the RNA structure are well-represented by the “free” MD simulations after the restraining force on RNA has been removed.

MD models of DNA are harder to describe because features in the difference profiles are not as distinct as they are for RNA. However, even with less pronounced features, we find that the DNA data are more consistent with MD results from “restrained” DNA, not “free” DNA. The

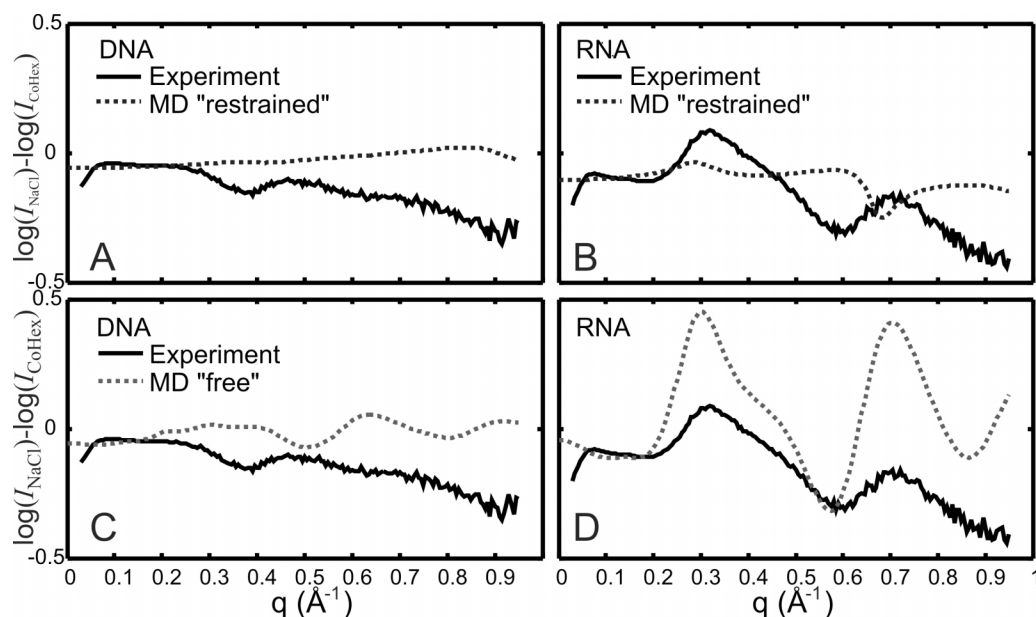


FIG. 5. Intensity difference profiles,  $\log(I_{\text{NaCl}}) - \log(I_{\text{CoHex}})$ , comparing nucleic acid experimental data to MD results when the molecules are “restrained” by a harmonic potential ((a) and (b)) and when molecules are “free” ((c) and (d)). The figure legends shown in panel (c) for DNA also apply to panel (d) for RNA. The best qualitative agreement with experiment happens when RNA is unrestrained in the MD simulations.

“restrained” MD-derived difference curve is largely featureless (Figure 5(a)) while the “free” MD-derived difference curve (Figure 5(c)) has peaks and valleys whose locations are inconsistent with the data. Figure 5(b) for “restrained” RNA also shows disagreement in the position of peaks and valleys. Thus, we find that comparison of intensity difference curves to models is not only a useful way to denote changes in the nucleic acid structure, but is also a robust method to identify good models and discard models that do not reflect experimental data. The agreement in Figure 5(d) between MD and experiment is still qualitative but the disagreement in Figures 5(b) and 5(c) is quite notable.

## IV. DISCUSSION

### A. RNA structure changes upon addition of CoHex

These results can now be used to gain physical insight into the origin of the structural changes. The need to use unrestrained “free” MD simulations for RNA suggests that CoHex introduces structural changes to the RNA helix. In Figure 6, we show experimental WAXS profiles,  $I$  vs.  $q$ , in logarithmic scale, comparing RNA to “free” MD results and the canonical A-form helix from NAB. Without CoHex, Figure 6(a) shows RNA experimental profiles favoring the “free” MD profiles with the shoulder around  $q = 0.45 \text{ \AA}^{-1}$ . However, the shallow peak seen in the data at  $q = 0.74 \text{ \AA}^{-1}$  is shifted to the left in the MD simulations (peak at  $q = 0.71 \text{ \AA}^{-1}$ ). In the presence of CoHex (Figure 6(b)), both RNA experimental curves and the results from the “free” MD appear to have features more similar to the canonical A-form profile. When the ions are turned off in CRY SOL (our default setting), the second peak around  $q = 0.81 \text{ \AA}^{-1}$  becomes more pronounced in the “free” MD profile with CoHex, which implies that CoHex changes the actual structure of the RNA, not just the counterion cloud or the hydration layer. The smearing of the  $q = 0.81 \text{ \AA}^{-1}$  peak when ions are turned on suggests diffuse positioning of the CoHex ions in the vicinity of  $d = 2\pi/q = 7.8 \text{ \AA}$ , consistent with our picture of CoHex

ions deep in the A-form major groove.<sup>16</sup> Figure 6 insets show representative snapshots of the RNA structure from the unrestrained MD simulation with and without CoHex. From the pool of MD-derived PDB snapshots, we measured the average distance between the 3'-end phosphates of the base-paired strands. The end-to-end phosphate distance for the RNA in NaCl was  $75 \pm 3 \text{ \AA}$  vs.  $66 \pm 2 \text{ \AA}$  for the RNA in CoHex/NaCl, suggesting a shortening of the structure in the presence of CoHex.

### B. Current MD force-fields work better for RNA

Here, we investigate the source of the disagreement between WAXS experiments and the unrestrained MD simulations of DNA. We compare the structural parameters of canonical A-form and B-form helices to the MD trajectories of the DNA and RNA sequences used in the study. Table I shows parameters from the first 100 ns following the release of the positional restraints and relaxation of the free structure. This corresponds to simulation time of 160–260 ns, the time scale consistent with CRY SOL-generated WAXS profiles averaged in Figures 3 and 4. We also show the parameters of a DNA dodecamer (DD) before and after simulations are performed. These parameters are also displayed in Table I.

In Table II, we show the same parameters for DNA and RNA after a longer simulation set, up to an additional 280 ns. The parameters appear to be quite stable during the entire length of the trajectories, and there is no significant difference between the time frame used for CRY SOL snapshots (Table I) and the next 100 ns.

These tables show convergence of MD simulations. The helical parameters reach equilibrium for the snapshots used for CRY SOL. The tables also show that on time scales of hundreds of nanoseconds, the structural parameters of unrestrained simulated mixed DNA duplexes drift away from the canonical B-form helix. The canonical parameters were originally reported by Arnott and Hukins.<sup>32</sup> The simulated structural parameters of the mixed sequence 25-bp DNA (expected to be in B-form) and DNA dodecamer (DD) both

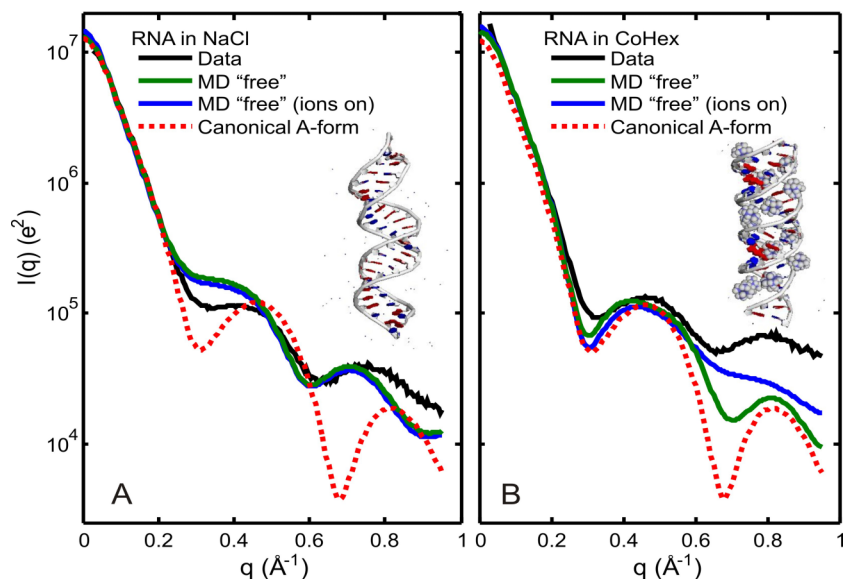


FIG. 6. WAXS profiles comparing RNA data to “free” MD and the canonical A-form helix from NAB in (a) NaCl only and (b) NaCl/CoHex solution. Insets show representative structures of RNA from MD snapshots. Double-stranded RNA with CoHex appears to have a shorter end-to-end distance and a WAXS profile that has features more similar to the canonical A-form helix.

TABLE I. Structural parameters for the canonical and simulated helices and the B-DNA dodecamer (DD, PDB ID 1BNA). The simulated DNA and RNA are the 25-bp fragments with the same sequence as those used in the experiment, simulated with and without CoHex counterions. The DD in NaCl was simulated unrestrained for 100 ns using the same MD simulation conditions described in the main text, and the entire trajectory analyzed. To extract structural parameters, cpptraj tool in the AmberTools package was used, with the average values reported in the table.

System	X-displacement (Å)	Rise (Å)	Twist (deg)
Canonical A-form helix	-4.73	2.56	32.70
RNA in NaCl	-4.94	2.72	31.11
RNA in NaCl/CoHex	-4.54	2.42	33.64
Canonical B-form helix	0.11	3.38	36.00
DNA in NaCl	-2.39	3.14	32.63
DNA in NaCl/CoHex	-2.32	3.08	33.11
DD (PDB:1BNA)	0.13	3.35	36.09
DD in NaCl for 100 ns	-1.37	3.22	34.17

TABLE II. Structural parameters for the simulated unrestrained DNA and RNA analyzed with a longer trajectory, up to an additional 280 ns. The parameters for A-form and B-form helices are also shown for comparison.

System	X-displacement (Å)	Rise (Å)	Twist (deg)
Canonical A-form helix	-4.73	2.56	32.70
RNA in NaCl	-4.90	2.74	31.08
RNA in NaCl/CoHex	-4.64	2.40	33.40
Canonical B-form helix	0.11	3.38	36.00
DNA in NaCl	-2.40	3.15	32.59
DNA in NaCl/CoHex	-2.35	3.08	33.02

deviate significantly from the canonical parameters of B-form DNA and appear to drift toward A-form parameters, with the deviation from canonical B-form increasing (significantly for X-displacement, and twist) for the longer fragment. This deviation for DD is consistent with simulation results reported previously<sup>33</sup> and implies that the force field used in this study is not ideal for reproducing subtle features of B-form DNA. Future work should look at the effects of other force fields on DNA and RNA helices and investigate the effects of dynamical motions such as base-pair fraying at the ends of the helices when comparing MD simulations to WAXS data.

At the same time, the addition of CoHex produces no significant effect (within the above set of parameters) on the simulated mixed sequence DNA structure. Thus, while the force-field may be biased such that canonical B-form of DNA is not the preferred structure for the mixed 25-bp fragments (which would explain the discrepancy between unrestrained simulations and experimental wide-angle scattering data), it still appears to reproduce ion-DNA interaction reasonably well, as evidenced by minimal change in DNA structure upon addition of CoHex, which is in agreement with CD<sup>16</sup> and WAXS measurements. Interestingly, the decrease in RNA rise with the addition of CoHex from 2.74 Å to 2.40 Å corresponds to the decrease in the end-to-end distance (75 Å to 66 Å) and the change in both CD<sup>16</sup> and WAXS spectra.

The limitations of MD force-fields for DNA have also been discussed in a paper by Tiede and co-workers where they benchmarked the WAXS profiles to MD solutions and found a need to develop experimentally validated, supramolecular force fields.<sup>6</sup> Our findings reiterate theirs, and further the field by reporting results with RNA. What is surprising here is

the better agreement between the unrestrained MD and the experimental data in RNA. The disagreement between our DNA WAXS data and the structures from the “free” MD for DNA arises from deviations of key structural features in the MD-generated ensemble from the values expected for canonical B-form DNA. Force field modification is beyond the scope of this work; however, we illustrate the use of WAXS as an experimental test. While the WAXS curves are not exactly predicted (on an absolute scale) by the force field, they present qualitative features that accurately reflect structural features of these duplexes, hence provide a useful experimental metric.

## V. CONCLUSIONS

In summary, we compared MD-derived structures to experimental WAXS data and showed the power of WAXS in investigating nucleic acid structural features and structural changes. Although our analysis of WAXS data is currently limited by the need to compare to accessible PDB structures, comparison of intensity difference profiles from experimental data and those generated from MD-derived models provides a test of MD simulation predictions. Changes in RNA WAXS data with and without CoHex ions were well represented by changes in profiles computed from structures generated in unrestrained MD simulations. Changes in experimental DNA WAXS profiles are smaller and give confidence to the previous assumption<sup>16</sup> to restrain DNA structures during MD simulations; restrained DNA is best used in situations when the study focuses on DNA surroundings, e.g., distribution of solvent components such as water or ions. Our study here showed that there are limitations in using unrestrained MD for

DNA in comparison to WAXS profiles. However, even with the limitations shown, we were able to gain insight into molecular conformations by comparing unrestrained MD of RNA to experimental WAXS data. We demonstrate that CoHex affects RNA helices by shortening the end-to-end distance and forcing the molecule to adapt a more A-like conformation. The future application of computational methods that allow MD simulations to be guided by SAXS and WAXS data, a process that is currently being developed for proteins,<sup>34</sup> holds great potential for studies of nucleic acids.

## ACKNOWLEDGMENTS

We acknowledge funding support from NIH Grant No. R01 GM-099450 and, in part, by NSF Grant No. CNS-0960081 and the HokieSpeed supercomputer at Virginia Tech. A.M.K. was supported by the NSF Graduate Research Fellowship Grant No. DGE-1144153. We thank Arthur Woll, Steve Meisburger, Huimin Chen, and Pollack Lab Members for experimental assistance. WAXS data were acquired at the Cornell High Energy Synchrotron Source (CHESS). CHESS is supported by the NSF and NIH/NIGMS via NSF Award No. DMR-1332208, and the MacCHESS resource is supported by NIGMS Award No. GM-103485.

- <sup>1</sup>J. Pérez and Y. Nishino, *Curr. Opin. Struct. Biol.* **22**, 670 (2012).
- <sup>2</sup>S. Skou, R. E. Gillilan, and N. Ando, *Nat. Protoc.* **9**, 1727 (2014).
- <sup>3</sup>L. Makowski, *J. Struct. Funct. Genomics* **11**, 9 (2010).
- <sup>4</sup>M. A. Graewert and D. I. Svergun, *Curr. Opin. Struct. Biol.* **23**, 748 (2013).
- <sup>5</sup>M. Allaire and L. Yang, *J. Synchrotron Radiat.* **18**, 41 (2011).
- <sup>6</sup>X. Zuo, G. Cui, K. M. Merz, L. Zhang, F. D. Lewis, and D. M. Tiede, *Proc. Natl. Acad. Sci. U. S. A.* **103**, 3534 (2006).
- <sup>7</sup>S. Park, J. P. Bardhan, B. Roux, and L. Makowski, *J. Chem. Phys.* **130**, 134114 (2009).
- <sup>8</sup>J. Bardhan, S. Park, and L. Makowski, *J. Appl. Crystallogr.* **42**, 932 (2009).
- <sup>9</sup>C. J. Knight and J. S. Hub, *Nucl. Acids Res.* **43**, W225 (2015).
- <sup>10</sup>J. J. Virtanen, L. Makowski, T. R. Sosnick, and K. F. Freed, *Biophys. J.* **101**, 2061 (2011).
- <sup>11</sup>D. D. L. Minh and L. Makowski, *Biophys. J.* **104**, 873 (2013).
- <sup>12</sup>H. T. Nguyen, S. A. Pabit, S. P. Meisburger, L. Pollack, and D. A. Case, *J. Chem. Phys.* **141**, 22D508 (2014).
- <sup>13</sup>M. Cammarata, M. Levantino, F. Schotte, P. A. Anfinrud, F. Ewald, J. Choi, A. Cupane, M. Wulff, and H. Ihee, *Nat. Methods* **5**, 881 (2008).
- <sup>14</sup>D. Arnlund, L. C. Johansson, C. Wickstrand, A. Barty, G. J. Williams, E. Malmerberg, J. Davidsson, D. Milathianaki, D. P. DePonte, R. L. Shoeman, D. Wang, D. James, G. Katona, S. Westenhoff, T. A. White, A. Aquila, S. Bari, P. Berntsen, M. Bogan, T. B. van Driel, R. B. Doak, K. S. Kjær, M. Frank, R. Fromme, I. Grotjohann, R. Henning, M. S. Hunter, R. A. Kirian, I. Kosheleva, C. Kupitz, M. Liang, A. V. Martin, M. M. Nielsen, M. Messerschmidt, M. M. Seibert, J. Sjöhamn, F. Stellato, U. Weierstall, N. A. Zatsepin, J. C. H. Spence, P. Fromme, I. Schlichting, S. Boutet, G. Groenhof, H. N. Chapman, and R. Neutze, *Nat. Methods* **11**, 923 (2014).
- <sup>15</sup>L. Li, S. A. Pabit, S. P. Meisburger, and L. Pollack, *Phys. Rev. Lett.* **106**, 108101 (2011).
- <sup>16</sup>T. S. Tolokh, S. A. Pabit, A. M. Katz, Y. Chen, A. Drozdetski, N. Baker, L. Pollack, and A. V. Onufriev, *Nucl. Acids Res.* **42**, 10823 (2014).
- <sup>17</sup>D. Orthaber, A. Bergmann, and O. Glatter, *J. Appl. Crystallogr.* **33**, 218 (2000).
- <sup>18</sup>T. Macke and D. A. Case, in *Molecular Modeling of Nucleic Acids*, edited by N. B. Leontes and J. SantaLucia, Jr. (American Chemical Society, Washington, DC, 1998), pp. 379–393.
- <sup>19</sup>D. Svergun, C. Barberato, and M. H. Koch, *J. Appl. Crystallogr.* **28**, 768 (1995).
- <sup>20</sup>D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, *J. Comput. Chem.* **26**, 1668 (2005).
- <sup>21</sup>T. E. Cheatham, P. Cieplak, and P. A. Kollman, *J. Biomol. Struct. Dyn.* **16**, 845 (1999).
- <sup>22</sup>A. Pérez, I. Marchán, D. Svozil, J. Sponek, T. E. Cheatham, C. A. Laughton, and M. Orozco, *Biophys. J.* **92**, 3817 (2007).
- <sup>23</sup>J. Aqvist, *J. Phys. Chem.* **94**, 8021 (1990).
- <sup>24</sup>D. E. Smith and L. X. Dang, *J. Chem. Phys.* **100**, 3757 (1994).
- <sup>25</sup>W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
- <sup>26</sup>T. E. Cheatham and P. Kollman, *Structure* **5**, 1297 (1997).
- <sup>27</sup>P.-C. Chen and J. S. Hub, *Biophys. J.* **107**, 435 (2014).
- <sup>28</sup>J. Widom and R. L. Baldwin, *J. Mol. Biol.* **144**, 431 (1980).
- <sup>29</sup>J. Pelta, F. Livolant, and J. L. Sikorav, *J. Biol. Chem.* **271**, 5656 (1996).
- <sup>30</sup>B. M. Bulheller and J. D. Hirst, *Bioinformatics* **25**, 539 (2009).
- <sup>31</sup>R. F. Fischetti, D. J. Rodi, D. B. Gore, and L. Makowski, *Chem. Biol.* **11**, 1431 (2004).
- <sup>32</sup>S. Arnott and D. W. L. Hukins, *Biochem. Biophys. Res. Commun.* **47**, 1504 (1972).
- <sup>33</sup>A. Pérez, F. Lankas, F. J. Luque, and M. Orozco, *Nucl. Acids Res.* **36**, 2379 (2008).
- <sup>34</sup>P.-C. Chen and J. S. Hub, *Biophys. J.* **108**, 2573 (2015).