5028–5037 *Nucleic Acids Research, 2021, Vol. 49, No. 9* *Published online 1 May 2021*
doi: 10.1093/nar/gkab290

Solution structure(s) of trinucleosomes from contrast variation SAXS

Alexander W. Mauney[1], Uma M. Muthurajan[2], Karolin Luger[2] and Lois Pollack[1,*]

[1]School of Applied and Engineering Physics, Cornell University, Ithaca, NY 14853, USA and [2]Department of Biochemistry, University of Colorado, Boulder, CO 80303, USA

Received July 28, 2020; Revised April 02, 2021; Editorial Decision April 05, 2021; Accepted April 11, 2021

ABSTRACT

Nucleosomes in all eukaryotic cells are organized into higher order structures that facilitate genome compaction. Visualizing these organized structures is an important step in understanding how genomic DNA is efficiently stored yet remains accessible to information-processing machinery. Arrays of linked nucleosomes serve as useful models for understanding how the properties of both DNA and protein partners affect their arrangement. A number of important questions are also associated with understanding how the spacings between nucleosomes are affected by the histone proteins, chromatin remodelers, or other chromatin-associated protein partners. Contrast variation small angle X-ray scattering (CVSAXS) reports the DNA conformation within protein-DNA complexes and here is applied to measure the conformation(s) of trinucleosomes in solution, with specific sensitivity to the distance between and relative orientation of linked nucleosomes. These data are interpreted in conjunction with DNA models that account for its sequence dependent mechanical properties, and Monte-Carlo techniques that generate realistic structures for comparison with measured scattering profiles. In solution, trinucleosomes segregate into two dominant populations, with the flanking nucleosomes stacked or nearly equilaterally separated, e.g. with roughly equal distance between all pairs of nucleosomes. These populations are consistent with previously observed magnesium-dependent structures of trinucleosomes with shorter linkers.

INTRODUCTION

In eukaryotic cells, DNA is hierarchically organized for compaction and organization, yet remains accessible for transcription, translation, repair and other processes. The fundamental unit of DNA organization is the nucleosome, a complex containing two sets of the four histone proteins H2A, H2B, H3, and H4. The octamer core formed by these proteins is wrapped by ∼145 bp of DNA (1). Additional proteins, such as the H1 linker histone, as well as divalent cations such as Mg^{2+}, help condense nucleosomes into higher order structures, eventually forming chromosomes (2).

Arrays of nucleosomes are used as model systems to understand the organizational hierarchy of these higher order structures. They commonly consist of a few, typically 3–12, nucleosomes connected by DNA linkers of variable lengths (3). Nucleosome arrays are generally used to probe the formation of the 30-nm fiber, the next level of organization, and to determine its structure, which is still not fully understood (2). They can also potentially report the effect of added (remodelers) or modified (histone variant) components (4). The higher order structures are more easily evaluated if the relative orientations and distances between the nucleosomes can be derived.

Multiple conformations and structures of nucleosomal arrays have already been identified by prior crystallographic and cryo-EM studies (5–7). Structural details and order of the arrays depend on the linker lengths, ionic conditions, histone composition, as well as the presence of additional chromatin architectural proteins. Both methods are limited in their ability to report on highly diverse or mobile structures, as they rely on the presence of many identical (crystallography) or a handful of classes of similar (cryo-EM) conformation. In nucleosomal arrays, the intrinsic disorder of the linker DNA introduces a high degree of variance in the structures, which limits the linker lengths accessible with these methods. Solution measurements such as fluorescence resonance energy transfer (FRET, both in bulk and at the single molecule level), as well as analytical ultracentrifugation (AUC) can detect these motions (8–10), but these techniques are limited in other ways. AUC can be label free, but only provides information on the global shape and size of the array, while FRET requires sample labeling, and is limited by the distance between the carefully placed labels.

Thus, the application of a solution technique which is sensitive to the many length scales present in nucleosome arrays would be advantageous, and also amenable to future

*To whom correspondence should be addressed. Tel: +1 607 255 8695; Fax: +1 607 255 7658; Email: lp26@cornell.edu

studies monitoring changes in the positioning of the nucleosomes where the feature of interest is the nucleosome spacing. Small-angle X-ray scattering (SAXS) has the potential to report on structural details of nucleosomal arrays in solution, even when there are large fluctuations or high variance in the sample population. SAXS allows for a full sampling of the conformational space of a biomolecule, without requiring any modifications (such as tags) to the structure of the molecule. Similarly, no constraints are imposed on the composition of the buffer. SAXS data reflect macromolecular length scales of 50–500 Å, a good match to the major features present in trinucleosome arrays. Despite these strengths, interpretation of standard SAXS curves for protein-nucleic acid complexes is difficult because signals from the nucleic acid and the proteins, as well as their combination (the so-called interference term) contribute to what is measured (11). Fortunately, contrast variation SAXS (CVSAXS) can be applied to 'mask' the contribution of the protein, thus measuring the scattering from only the nucleic acid portion of the complex, as has been previously demonstrated (12). Although removing all but one component simplifies the interpretation, CVSAXS experiments provide the most information when detailed models are available for comparison or fitting.

Here, we report a model-guided interpretation of CVSAXS data on minimal nucleosome arrays containing three nucleosomes connected by 60 base pairs of linker DNA, which is the average separation for nucleosomes in the cell. Comparing the experimental results with calculated scattering results from models allows extraction of key distances and orientational details of nucleosomes in arrays. These structural details are consistent with previously determined cryo-EM structures of trinucleosomes with shorter linker lengths, and in the presence of stabilizing divalent ions (6).

MATERIALS AND METHODS

Trinucleosome assembly

561 bp DNA used for assembling Non-linker ended (NLE) trinucleosomes.

ATCGGAGAATCCCGGTGCCGAGGCCGCTCAATT
GGTCGTAGACAGCTCTAGCACCGCTTAAACGCA
CGTACGC**G**CTGTCCCCCGCGTTTTAACCGCCAA
GGGGATTACTCCCTAGTCTCCAGGCACGTGTCA
GATATATACATCGATTGCATGTGGA
TCCGAATTCATATTAATCATATCT
AATACTAGGACCCTATACGCGGCC
GCATCGGAGAATCCCGGTGCCGAGGCCGCTCA
ATTGGTCGTAGACAGCTCTAGCACCGCTTAAAC
GCACGTACGC**G**CTGTCCCCCGCGTTTTAACCGC
CAAGGGGATTACTCCCTAGTCTCCAGGCACGTG
TCAGATATATACATCGATTGCATGTG
GATCCGAATTCATATTAATCATATCT
AATACTAGGACCCTATACGCGGCC
GCATCGGAGAATCCCGGTGCCGAGGCCGCTCA
ATTGGTCGTAGACAGCTCTAGCACCGCTTAAAC
GCACGTACGC**G**CTGTCCCCCGCGTTTTAACCG
CCAAGGGGATTACTCCCTAGTCTCCAGGCACGT
GTCAGATATATACATCGAT

Here, the positioning sequence is underlined, the linker regions are not, and the base at the dyad in each nucleosome is in bold type.

DNA production and purification. Detailed protocols for expressing and purifying the 601 positioning sequences in *E. coli* are described elsewhere (13,14).

Histone octamer refolding and purification. Human histones were purchased from 'The Histone Source' at Colorado State University and octamers were assembled and purified as described in Dyer *et al*., 2004 (13).

Trinucleosome assembly. Trinucleosomes were assembled using well-established protocols previously published (13,14) using the DNA sequences mentioned above as well as histone octamers assembled with human histones. Briefly, the DNA was incubated with varying amounts of human histone octamer in 2 M NaCl buffer (10 mM Tris 7.5, 1 mM EDTA, 1 mM DTT) and the salt concentration was slowly reduced via continuous dilution with a buffer containing 250 mM NaCl with the rest of the components same as the 2M NaCl buffer. Eventually, the samples were dialyzed into a no salt buffer (TCS 20 mM Tris 7.5, 1 mM EDTA, 1 mM DTT).

The quality of trinucleosomes was assessed as follows: 1 μg of trinucleosomes was subjected to EcoRI digestion and analyzed on a 1% TAE agarose gel. The absence of a 207 bp DNA after digestion indicates complete assembly / saturation (Supplementary Figure S1). A 10 μg aliquot of the trinucleosomes was also subjected to analytical ultracentrifugation to evaluate the level of -saturation. Briefly, trinucleosomes were diluted in TCS buffer (20 mM Tris 7.5, 1 mM EDTA, 1 mM DTT) to a final volume ∼400 μl and subjected to sedimentation velocity analytical ultracentrifugation at 37 000 rpm and 20°C and absorbance mode until the sample had sedimented completely. The data were processed using Ultrascan III software as described previously (15). A mid-point S value of ∼17–18S was deemed as a saturated trinucleosome array (Supplementary Figure S2). Several such preparations with similar saturation levels were then prepared and combined to achieve the higher concentrations needed for SAXS experiments. Trinucleosomes were concentrated using the Amicon ultra devices with a 100 kDa cut-off. The final concentration of trinucleosomes was 55 μM (∼37 mg/ml).

Contrast variation small angle X-ray scattering

Small angle X-ray scattering (SAXS) provides global structural information for macromolecules in solution. Scattering profiles report the scattering intensity (I) as a function of the reciprocal-space momentum transfer, $q = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle and λ is the X-ray wavelength. The q value can be related to the real space distance between scatterers, $r = 2\pi/q$.

For a system containing multiple components, such as a protein-nucleic acid complex, scattering arises not only from each macromolecular component, but also from the interference between them, through a so-called cross term

(16). SAXS profiles of complexes can be challenging to interpret as the scattering of the components combines nonlinearly. However, if there is a significant difference between the electron density of the components, the electron density of the solvent can be increased to match one of them. As the SAXS signal is proportional to difference in electron density between solute and solvent, this process removes the scattering from the matched components when the buffer scattering is subtracted. For nucleosomes, the protein density can be 'matched' by the addition of 50% (w/w) sucrose to the surrounding buffer (12). Sucrose has previously been shown to negligibly impact nucleosome stability (17,18). Under these conditions, CVSAXS reports only the DNA conformation within the trinucleosome array.

SAXS data were collected at the Cornell High Energy Synchrotron Source G1 station (now sector 7). Monochromatic X-rays at 11.166 keV were incident on the sample. Profiles were normalized by measuring the intensity transmitted through a semitransparent molybdenum beamstop and dividing to remove the effect of variation in X-ray beam strength. Scattering profiles were imaged using a PILATUS 300K (DECTRIS) detector located 2.0 m away from the sample, giving a q range of 0.006 to 0.260 Å^{-1}. This range was calibrated using a silver-behenate standard. Scattering profiles were processed in MATLAB (MathWorks), and background subtraction was done by subtracting the measured profile for the sucrose buffer from that of the sample, following previous protocols (19).

Monte-Carlo model generation

To model the trinucleosomes we extend earlier work focusing on the conformation(s) of DNA in single nucleosome core particles (12). This approach does not model the histone proteins and is a good match to contrast variation SAXS, where only the signal from the DNA is measured; it will be especially valuable when assessing changes in the relative positions/orientations of nucleosomes, following specific changes to the composition of the system, including the addition/variation of protein partners. In (12), we compared the measured scattering profiles of DNA in a single nucleosome to profiles computed from models constructed using the cgDNA package. The details of model construction are provided previously (12), but briefly, we first extract the DNA structure from a nucleosome crystal structure (1) and create a coarse-grained model, with individual base resolution. The sequence information is edited to match the sample of interest. cgDNA is then used to construct a stiffness matrix for the sequence. Portions of the sequence are released ('freed') from their positions, and the stiffness matrix is used to calculate the minimum energy position of the free bases, accounting for the sequence used. Possible thermodynamically accessible conformations are then randomly sampled. For each conformation, each grain is then replaced with the appropriate atomistic base and backbone model.

We used the following automated approach to generate the various conformations. To model the trinucleosome system the DNA was broken into three parts: the center nucleosome plus the full linkers comprised one part, and the two flanking nucleosomes comprised the other parts. A two base-pair section of the linkers was added to each flanking nucleosome to facilitate the subsequent merging of structures. Thermodynamic variants of each section were generated as described above. Up to 10 bp were allowed to unwrap from each end of flanking nucleosomes, while up to 50 bp were allowed to unwrap from each end of the center nucleosome. Once a set of models was generated, but prior to reinsertion of the bases, the models were merged to create a low-resolution structure of the full trinucleosome. The small linker piece on each flanking nucleosome was aligned to the same bases on the central nucleosome and then merged by averaging their positions

RESULTS AND DISCUSSION

Contrast variation SAXS

To measure the DNA conformation(s) within trinucleosome arrays, we performed CVSAXS, adding 50% (w/v) sucrose to buffers to contrast match the histone core as in past studies (18). Data were acquired at three trinucleosome concentrations (4.8, 2.4 and 1.2 μM; data shown in SI), and at buffer conditions of 20 mM Tris 7.5, 1 mM EDTA. All three experimental curves matched well and shared common features for q values in the 0.05–0.17 Å^{-1} range. Curves acquired at the highest nucleosome concentration showed signs of interparticle interference at low q, while the lowest concentration curves did not have sufficient signal to noise at high q as sucrose increases the background (Supplementary Figure S3). The sample containing nucleosomes at concentrations of 2.4 μM evaded both of these issues and was used for further analysis.

The measured CVSAXS signal from this trinucleosome sample is shown in Figure 1A. The well-defined features in the profile signal the presence of distinct length scales in the system. To focus on these length scales, we converted the data from reciprocal to real space (Figure 1B), using GNOM (20) to compute the pair distance distribution function, $P(r)$. The quality of this computation can be assessed by back calculating the $I(q)$ curve from $P(r)$. The result of this calculation is shown as the red curve in Figure 1A. This curve can be interpreted as a histogram of the distances between all pairs of scatterers in the sample. The presence of features such as sharp maxima or minima in the distance distribution correspond to length scales present in the sample. In this measurement two distinct peaks can be seen: a sharp maximum ∼80 Å and a broader one ∼260 Å.

To understand the structural basis of features present in the experimental $P(r)$ curves, we employed a bottom up strategy that starts by considering the computed P(r) for DNA from a single nucleosome. Each feature in this $P(r)$ curve is easily interpretable to reflect a meaningful macroscopic length scale. We then add a second (unconnected) set of nucleosomal DNA to the model. Features of the single nucleosome remain visible and new ones appear, reflecting distances between the two sets of DNA. This simple strategy ignores the effects of the linker DNA, however due to the relatively small size and lack of repeated distances in the linkers, they do not significantly change the features in the $P(r)$ curves. From the simple two nucleosome models we then move to models of three nucleosomes connected by linkers.

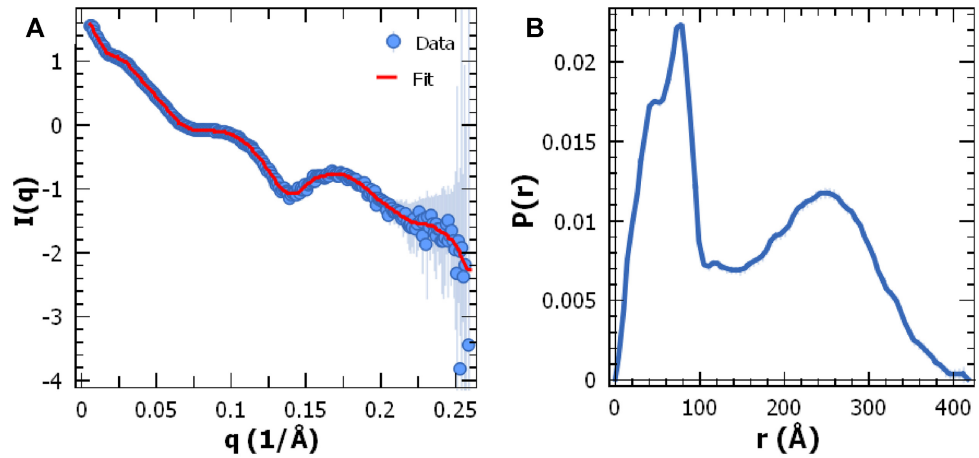

Figure 1. CVSAXS data (**A**) of trinucleosomes and associated distance distribution (**B**). The fit of the CVSAXS data associated with the distance distribution is shown in red. Both y axes (log of I(q) and P(r)) are in arbitrary units, a.u..

One nucleosome models

Previous studies from our lab showed the power of the $P(r)$ approach in revealing unwrapping transitions of single nucleosomes (18). Wrapped nucleosomes show a strong peak at ∼80 Å corresponding to the diameter of the DNA superhelix and a smaller peak ∼40 Å due to the separation of the turns around the nucleosome (Figure 2A). With added salt, the nucleosomal DNA unwraps beginning from the entry/exit points. The extent of the unwrapping manifests through several features of the $P(r)$ curves and this analysis informs some aspects of a $P(r)$ analysis of larger nucleosome arrays. At the buffer conditions employed for the trinucleosome experiments, specifically the low background of salt, we expect that the DNA remains wrapped, meaning that the 80 Å peak should be pronounced. Additionally, as seen with single nucleosomes, the maximum spatial extent of the trinucleosome can be determined from a $P(r)$ plot from the distance at which $P(r)$ goes to zero.

Two nucleosome models

We now increase the complexity of the model by adding a second nucleosome without linker DNA. These simple calculations allow us to understand how features reflecting internucleosome interactions are reflected in the $P(r)$ curves. We specifically consider three degrees of freedom: distance, rotation and stacking.

Modeling the distance between two nucleosomes. We start by examining how changing the separation of the two disconnected nucleosomes affects the simulated distance distribution (Figure 2B). The DNA is isolated from the crystal structure of a nucleosome (1AOI) and imported into Pymol (21). The DNA is then duplicated, and the translate tool is used to move the copied DNA a distance (200, 300 or 400 Å) in the plane of the nucleosome, perpendicular to the dyad axis (the axis passing through the middle base pair of the nucleosomal DNA), and observing how this distance is reflected in the distance distribution computed using CaPP (22). CaPP calculates $P(r)$ directly from a structural model, and as such gives much higher resolution (1 Å) distance distributions than those obtained from an experimental measurement. The distance distribution of these models always contains two peaks, a low distance peak at ∼80 Å that does not change, and a second peak that corresponds to the translation distance. The first peak corresponds to the distances within a single nucleosome (the intranucleosome peak) and appears at the same location as the single nucleosome peak. The second peak appears at the distances between the two nucleosomes (the internucleosome peak). It is accompanied by two satellite peaks that correspond to the distance between near edges of the nucleosomes and far edges of the nucleosomes. It is also interesting to see that the largest distance observed in the model is reflected by the x value at which the curve goes to zero: 300, 400 and 500 Å respectively for the three models shown (Figure 2B). This behavior is as expected; it reflects the translation (additional separation) by 100 Å of the nucleosomes in the three different cases. Similar behavior was also seen for non-planar separations of the same distances.

Modeling the relative orientation of two nucleosomes, 200 Å apart. In a trinucleosome we do not expect the nucleosomes to remain co-planar, so we introduced a rotational difference in the nucleosomes to see how the internucleosome peak would change (Figure 2C). These calculations used nucleosomes separated by 200 Å, with one rotated about the dyad axis. 200 Å was chosen because it is the shortest distance displaying a full separation of features reflecting internucleosome or intranucleosome distance distributions, as shown in Figure 2B. As the second nucleosome is rotated, the internucleosome peak broadens before eventually splitting into two separate peaks as the nucleosomes approach perpendicularity.

Modeling nucleosome stacking. Finally, we attempted to capture the known behavior of systems containing multiple nucleosomes: stacking. Stacked nucleosomes have been implicated in higher order structure formation based on crystal structures (23,24), and attractive effects have been observed between nucleosomes in solutions using SAXS (25).

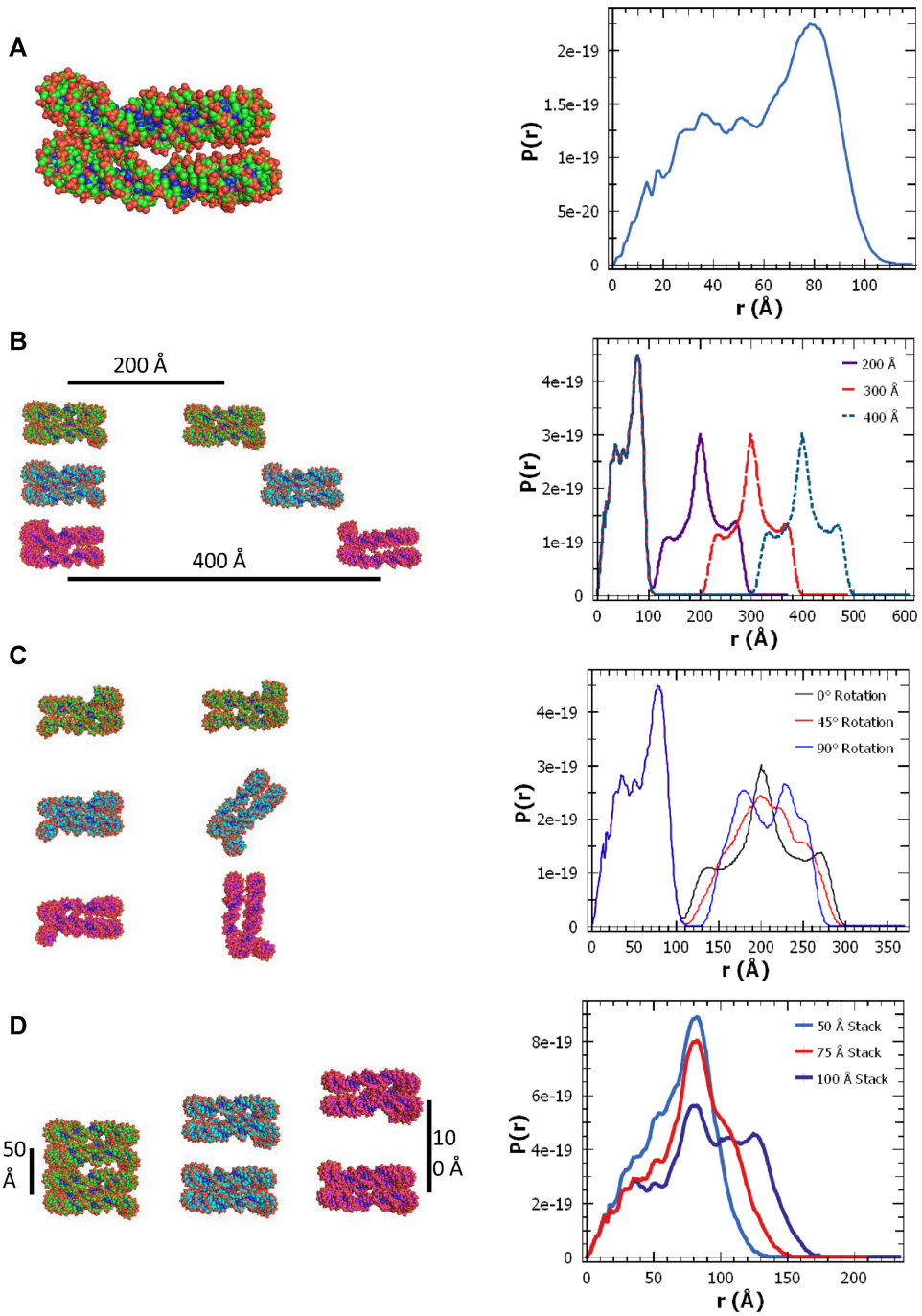

Figure 2. Distance distributions of simple one and two nucleosome models (in arbitrary units, a.u.). For panels with multiple models shown, each is colored differently. (**A**) Distance distribution for nucleosomal DNA from a single nucleosome. (**B**) Distance distributions for coplanar nucleosomal DNA as a function of center-center distance. The second peak in $P(r)$ is located at the center-center distance for each model. (**C**) Distance distributions for nucleosomal DNA separated by 200 Å as one nucleosome is rotated. The sharpest peak occurs when the nucleosomes are coplanar. The peak is then broadened by the rotation, before finally splitting into two distinct peaks when the nucleosomes are perpendicular. (**D**) Distance distributions for stacked nucleosomal DNA as a function of stack spacing. All have roughly the same peak position with a tight stack (50 Å) broadening the peak relative to the mononucleosome. Shoulders appear and expand as the stack is separated.

Two stacked nucleosomes will be separated by short distances, and, in contrast with separations shown in Figure 2B, these displacements are along the cylindrical axis common to the illustrated nucleosomes (Figure 2D). Because the length scales for intra and inter nucleosomes overlap at these shorter distances, the features in the P(r) curve are not as readily separable as described above. These models show that stacking deforms the 'single nucleosome' peak of Figure 2A, reflecting the overlap of large portions of the inter and intra nucleosome length scales. At the minimum stacking distance of ∼50 Å the peak is slightly broadened, and low distance features corresponding to the helix–helix separation in the single nucleosome are washed out. As the stacking distance increases, shoulders in the peak begin to develop and eventually become plateaus at ∼100 Å stacking distance. Beyond this point, the shoulders separate, and a two peak pattern is observed, like that of Figure 2B.

Trinucleosome models

The pool of structures. Although it would seem simple to just add a third nucleosome to model the trinucleosome arrays, we also must consider the constraints imposed by the linkers to fully model the trinucleosome. These 60-bp (∼200 Å) segments of DNA are much shorter than the persistence length of DNA (∼500 Å) and are expected to be relatively (though not perfectly) straight. They therefore dictate the distances and orientations of the flanking nucleosomes relative to the central nucleosome and each other. A pool of 10 000 models was created using the automated approach described in Methods. The scattering profile and P(r) distribution was computed for each model, and our interpretation is informed from our understanding of the simpler one- and two-nucleosome systems discussed above.

Visual inspection of the models in the pool reveals that the trinucleosomes are conformationally flexible (Supplementary Figures S4–S28). A majority of this variation arises from changes in the linker trajectory as base pairs are unwrapped from the center nucleosome. Each base pair unwrapped contributes to a roughly 5° shift in exit angle, which corresponds to a shift of ∼30 Å at the end of the linker. Most models display some amount of bending and twisting of the linker. For the majority of models these variations do not cause large deviations and the linker remains mostly straight, however in some extreme examples the linker has been observed to bend over 90°. All of these factors lead to a high variance in the separation and orientations of the nucleosomes.

Critical structural information is contained in the P(r) curves. Experience gained from the models of one and two nucleosomes allows us to make observations about some general features of the $P(r)$ curves derived from the DNA in trinucleosomes. Each nucleosome should contribute to the intranucleosome peak, and each pair of nucleosomes should give rise to an internucleosome peak. These internucleosome peaks will be located at the center-center distance between the two nucleosomes. The shapes of these peaks depend on the orientation of the nucleosomes, for well separated nucleosomes a larger relative rotation causes a broadening then splitting of the peak. For more closely spaced nucleosomes, stacking plays a bigger role and causes either a smoothing of the ∼70 Å peak for close stackings or the appearance of shoulders on this peak if there is a larger stacking distance.

We now focus on four test models (Figure 3, left) chosen from the pool to connect features in $P(r)$ plots to structural information. While Figure 3 only highlights some of the models in our pool, their distance distributions are representative of the pool as a whole. The Supplementary Information contains a larger library of models with corresponding distance distributions.

Test case 1: a fully extended trinucleosome. We now illustrate how our understanding of the simple two nucleosome models is manifested in the computed profiles of the DNA from trinucleosome arrays. First, we examine an extreme case, illustrated in Figure 3A, which shows a highly extended trinucleosome with asymmetric unwrapping of the central nucleosome. The internucleosome distances between the center and flanking nucleosomes give rise to peaks near 300 and 350 Å, due to the different linker DNA lengths, while the internucleosome distances between the outer nucleosomes give rise to a peak near 620 Å. The maximum distance between atoms in this model can deduced from the curve, as the point where the distance distribution approaches zero: ∼720 Å.

Test case 2: a roughly equilateral triangle. When the connecting linker DNA comes together and the center nucleosome is fully wrapped, it is less intuitively clear what the distance distribution will look like as the distance between the outermost nucleosomes is not well distinguished from the distance from the center to outer. An extreme example of this occurs when the trinucleosome adopts a roughly equilateral orientation, as in Figure 3B. In this case the peaks reflecting the distances between all three pairs of nucleosomes occurs near 250 Å and therefore merge into a single peak.

Test case 3: closer approach of the outer nucleosomes. As the outer nucleosomes come closer together (Figure 3C), the distance between the outer nucleosomes drops enough so that the flanker–flanker peak can be resolved between the intranucleosome peak and the flanker-center peaks (which remain merged).

Test case 4: closest approach of the outer nucleosomes. When the outer nucleosomes come close enough together (Figure 3D), the flanker–flanker peak begins to overlap with the intranucleosome peak, raising and broadening it. A small shoulder also appears, reminiscent of those seen at moderate stacking distances (Figure 2D) and co-planar nucleosomes (Figure 2C).

Interpreting experimental data. We now use the information derived from the model curves to provide some insight into our experimentally derived distance distribution (Figure 1); the insight obtained from the models allows us to draw conclusions regarding the structural features of the experimental ensemble. We review these data from the largest to small length scales shown. First, the $P(r)$ curve of Figure 1 drops to zero ∼400 Å, indicating that under experimental

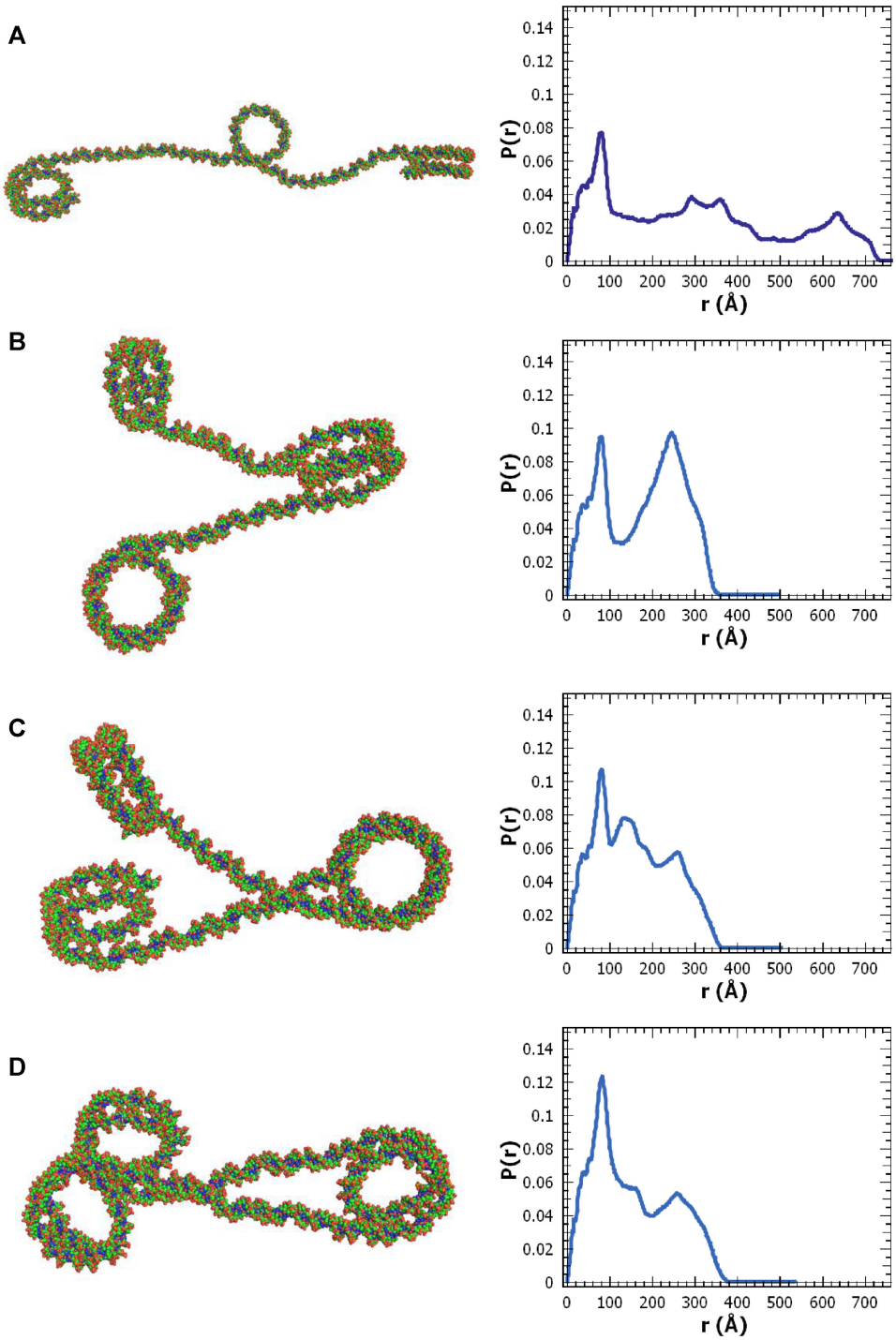

Figure 3. Selected trinucleosome models (left) and associated distance distributions (in a.u.) (right). (**A**) A model with the flanking nucleosomes highly extended and significant unwrapping of the center nucleosome. (**B**) A model with the flanking nucleosomes separated by roughly the length of a DNA linker arm. (**C**) A model with the flanking nucleosomes at an intermediate distance. (**D**) A model with the flanking nucleosomes close together.

conditions, trinucleosomes stay reasonably compact, unlike the model in Figure 3A. This indicates that slightly more expanded arrays than the model in Figure 3B, which drops to zero near 350 Å, are present in the sample. Next, the location of the second peak contains information about the distance between the center and flanking nucleosomes. As this distance is constrained by the linker length, it also reports whether the nucleosomal DNA is fully wrapped. When this is the case, this peak appears near 250 Å (as in test cases 2–4), and as DNA unwraps from the histone octamer this peak shifts to higher distances (as in test case 1). The experimental peak is present at roughly 250 Å, indicating that the experimental ensemble remains largely wrapped. At even shorter length scales, we note the presence of a well-defined minimum at ∼140 Å. To interpret the absence of this length scale we refer to the test models of Figure 3, specifically those that prominently exhibit this length scale (e.g. test case 3). Here, length scales ∼140 Å arise from distances between the flanking nucleosomes that are shorter than the linker lengths, but longer than the lengths represented by individual nucleosomes. The absence of this length scale from the experimental $P(r)$ curve suggests that structures with small distances between the peaks are not present to any significant degree in our experimental system.

Thus, based on the longest dimension measured and the minimum ∼140 Å, we rule out the existence of a substantial population of very extended structures, and structures where the flanking nucleosomes are close to each other (at a distance that is larger than the mononucleosome length scale). However, the roughly equilateral test case (Figure 3B) which satisfies both of these constraints, insufficiently describes the data, as the two peaks have roughly the same height, which does not match the experimental measurement. The next closest model, test case 4, also does not exactly match the data: the $P(r)$ curve has a shoulder where the minimum would be and has a much lower second than first peak. We now consider that the likely scenario that the conformations most likely represent an ensemble of structures, not a single conformation. The distance distribution of an ensemble of structures is simply the average of distance distributions of the members of the ensemble.

Looking back to the two nucleosome stacking models (Figure 3B and D), we surmise that the shoulder in the 100–150 Å range of test case 4 would disappear when the flankers are stacked together. The methodology used to generate the trinucleosome models does not preclude the existence of these models, however the likelihood of having the flanking nucleosomes stack precisely is very low, and none were seen in the model pool. The modeling framework used to generate the structure pool (cgDNA), does not account for energetic gains that would accompany stacking. To model this possible conformation, we took the structures from test case 3 above and manually stacked the flanker nucleosomes (the lower flanker as viewed in Figure 3C was moved). This manually constructed model is shown with its computed $P(r)$ in Figure 4. The relocated flanking nucleosome is no longer connected correctly to its linker, as moving and rotating the entire strand correctly is non-trivial; however, it remains close to the end of the linker to maintain the distance, and we expect that the qualitative features of the $P(r)$ are not impacted by this discontinuity.

We observed that the addition of stacking to the model in test case 4 removes the original shoulder, without changing the peak height ratio. The removal of the shoulder makes this a better candidate for explaining the experimental data, though the peak heights suggests that we still need a combination of stacked and roughly equilateral models to explain our data. The average of test cases 2 and 4 qualitatively recreates the data, as shown in Figure 5. Based on this curve, we expect that these two test cases are good representatives of two populations that explain the experimental data, consisting of trinucleosomes with flanking nucleosomes either stacked, or separated by 200–300 Å.

Previously we derived insight into the structures of single nucleosomes, using a genetic algorithm (EOM) (26) to select structures from the pools whose summed scattering profile matches the experimentally determined curve (12). We attempted this approach with these data as well; however, we were unable to capture the experimental data. This is likely due to the lack of stacked structures in the pool, which our P(r) analysis indicates are a significant portion of the experimental sample. While our goal is not to provide detailed models of intra-nucleosome interactions, which would clearly require inclusion of the histone proteins, the value of our approach arises from the direct 'readout' of the relative positions of the nucleosomes. The P(r) profiles would, in a straightforward way, reveal changes in the relative positions of the nucleosomes, for example, if the central nucleosome slid towards one of the flanking nucleosomes. In this case, the single length scale that represents the equal distance between the central and flanking nucleosomes would be replaced by two, unequal distances between the central nucleosome and the others. Although beyond the scope of the present study, there are opportunities to expand this approach to model the inter-nucleosome interactions, for example by incorporating phenomenological potentials, such as those in (27) that include stacking interactions. While EOM may provide more detailed information about the contents of the ensemble, the main parameters of interest relate to the unwrapping of the nucleosomes and the organization of the flankers, both of which are readily found through P(r).

CONCLUSIONS

Here, we report contrast variation SAXS measurements of nucleosome arrays to obtain information on their relative orientation in solution. This approach has the potential to highlight the different distances, orientations, and angles between nucleosomes in arrays in solution and unconstrained by crystal lattice or EM grid surface. Our measurements are consistent with the presence of two dominant distinct populations, both of which are formed by nucleosomes that are fully wrapped but that differ in their relative orientation. One population consists of trinucleosomes where the flanking nucleosomes are stacked close together (Figure 4), and the other where they are separated by ∼200–300 Å and do not interact (Figure 3). Recently, Takizawa *et al*. observed similar 'stacked' populations in trinucleosomes using cryo-electron microscopy (6), and this configuration was stabilized by $MgCl_2$. These trinucleosomes were connected by shorter linker lengths (22 and 30 bp instead of the 60

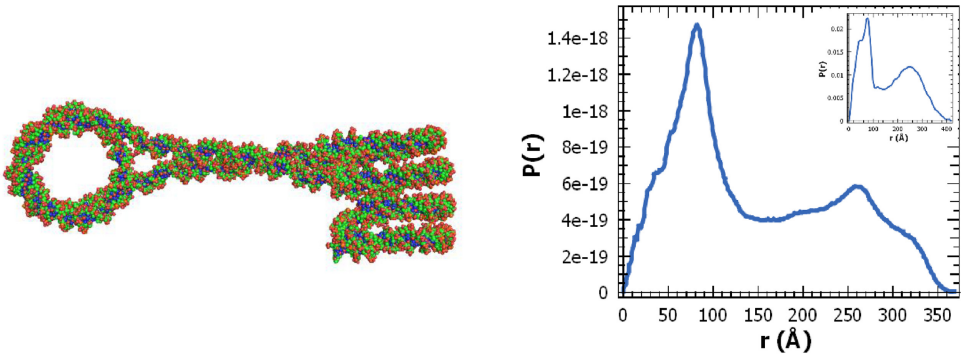

Figure 4. Manually stacked trinucleosome and associated distance distribution (a.u.). Flanking nucleosomes from test model 4 (Figure 3D) were manually stacked by moving the DNA from a flanking nucleosome. This change is reflected in the distribution as a slightly stronger low distance peak, and a loss of the previously seen shoulder. The experimental data is inset for comparison.

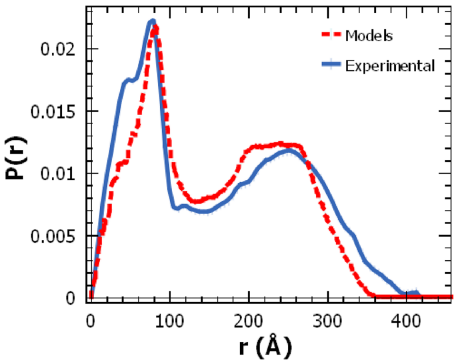

Figure 5. Comparison of experimental distance distribution to two model ensemble (in a.u.). A linear combination of the distance distributions of a roughly equilateral model (Figure 2B) and the manually stacked model (Figure 4) are used to model the experimental distribution.

bp used here), and a difference between these two shorter length confirms the intuitive assumption that longer linker length increases the degrees of freedom.

The sensitivity of CVSAXS to changes in the configuration of nucleosomal DNA makes it an ideal platform for future studies. Physiologically relevant protein partners, such as remodelers or other nucleosome associated proteins, can be added without changing the scattering due to the contrast matching. This allows for direct observation of the changes caused by these proteins and could even be observed as a function of time depending on the time scales associated with the reorganization. Histone variants or mutants can also be easily screened for effects on higher order organization using this system.

DATA AVAILABILITY

SAXS data will be made available through the SASBDB, under accession number *SASDJ96*.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Joshua Tokuda and Yujie Chen for assistance in data acquisition at CHESS.

FUNDING

National Science Foundation [DGE-1650441 to A.M.]; National Institutes of Health (NIH) [R35-GM122514 to L.P.]; CHESS is supported by National Science Foundation (NSF); National Institutes of Health/National Institute of General Medical Sciences (NIH/NIGMS) via NSF award [DMR-0936384 to CHESS]; Howard Hughes Medical Institute (to K.L. and U.M.); NIH-NCI [R01 CA218255]. Funding for open access charge: National Institutes of Health.

Conflict of interest statement. None declared.

REFERENCES

1. Luger,K., Mäder,A.W., Richmond,R.K., Sargent,D.F. and Richmond,T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, **389**, 251–260.
2. Kuznetsova,M.A. and Sheval,E.V. (2016) Chromatin fibers: from classical descriptions to modern interpretation. *Cell Biol. Int.*, **40**, 1140–1151.
3. Hansen,J.C. (2002) Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annu. Rev. Biophys. Biomol. Struct.*, **31**, 361–392.
4. Baldi,S., Korber,P. and Becker,P.B. (2020) Beads on a string—nucleosome array arrangements and folding of the chromatin fiber. *Nat. Struct. Mol. Biol.*, **27**, 109–118.
5. Song,F., Chen,P., Sun,D., Wang,M., Dong,L., Liang,D., Xu,R.M., Zhu,P. and Li,G. (2014) Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science*, **344**, 376–380.
6. Takizawa,Y., Ho,C.H., Tachiwana,H., Matsunami,H., Kobayashi,W., Suzuki,M., Arimura,Y., Hori,T., Fukagawa,T., Ohi,M.D. *et al.* (2020) Cryo-EM structures of centromeric tri-nucleosomes containing a central CENP-A nucleosome. *Structure*, **28**, 44–53.
7. Daban,J.R. (2011) Electron microscopy and atomic force microscopy studies of chromatin and metaphase chromosome structure. *Micron*, **42**, 733–750.
8. Maeshima,K., Rogge,R., Tamura,S., Joti,Y., Hikima,T., Szerlong,H., Krause,C., Herman,J., Seidel,E., DeLuca,J. *et al.* (2016) Nucleosomal arrays self-assemble into supramolecular globular structures lacking 30-nm fibers. *EMBO J.*, **35**, 1115–1132.
9. Poirier,M.G., Oh,E., Tims,H.S. and Widom,J. (2009) Dynamics and function of compact nucleosome arrays. *Nat. Struct. Mol. Biol.*, **16**, 938–944.

10. Kilic,S., Felekyan,S., Doroshenko,O., Boichenko,I., Dimura,M., Vardanyan,H., Bryan,L.C., Arya,G., Seidel,C.A.M. and Fierz,B. (2018) Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1α. *Nat. Commun.*, **9**, 235.
11. Blanchet,C.E. and Svergun,D.I. (2013) Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution. *Annu. Rev. Phys. Chem.*, **64**, 37–54.
12. Mauney,A.W., Tokuda,J.M., Gloss,L.M., Gonzalez,O. and Pollack,L. (2018) Local DNA sequence controls asymmetry of DNA unwrapping from nucleosome core particles. *Biophys. J.*, **115**, 773–781.
13. Dyer,P.N., Edayathumangalam,R.S., White,C.L., Bao,Y., Chakravarthy,S., Muthurajan,U.M. and Luger,K. (2004) Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.*, **375**, 23–44.
14. Muthurajan,U., Mattiroli,F., Bergeron,S., Zhou,K., Gu,Y., Chakravarthy,S., Dyer,P., Irving,T. and Luger,K. (2016) In vitro chromatin assembly: strategies and quality control. In: *Methods in Enzymology*. Academic Press Inc. pp. 3–41.
15. Demeler,B. and Gorbet,G.E. (2016) Analytical ultracentrifugation data analysis with ultrascan-III. In: *Analytical Ultracentrifugation: Instrumentation, Software, and Applications*. Springer, Japan. pp. 119–143.
16. Tokuda,J.M., Pabit,S.A. and Pollack,L. (2016) Protein-DNA and ion-DNA interactions revealed through contrast variation SAXS. *Biophys. Rev.*, **8**, 139–149.
17. Blose,J.M., Pabit,S.A., Meisburger,S.P., Li,L., Jones,C.D. and Pollack,L. (2011) Effects of a protecting osmolyte on the ion atmosphere surrounding DNA duplexes. *Biochemistry*, **50**, 8540–8547.
18. Chen,Y., Tokuda,J.M., Topping,T., Sutton,J.L., Meisburger,S.P., Pabit,S.A., Gloss,L.M. and Pollack,L. (2014) Revealing transient structures of nucleosomes as DNA unwinds. *Nucleic Acids Res.*, **42**, 8767–8776.
19. Skou,S., Gillilan,R.E. and Ando,N. (2014) Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nat. Protoc.*, **9**, 1727–1739.
20. Svergun,D.I (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.*, **25**, 495–503.
21. Schrödinger,L. (2015) The PyMOL Molecular Graphics System, Version 2.0.
22. Larsen,A.H., Arleth,L. and Hansen,S. (2018) Analysis of small-angle scattering data using model fitting and Bayesian regularization. *J. Appl. Crystallogr.*, **51**, 1151–1161.
23. Dorigo,B., Schalch,T., Kulangara,A., Duda,S., Schroeder,R.R. and Richmond,T.J. (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science (80-.)*, **306**, 1571–1573.
24. Schalch,T., Duda,S., Sargent,D.F. and Richmond,T.J. (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature*, **436**, 138–141.
25. Andresen,K., Qiu,X., Pabit,S.A., Lamb,J.S., Park,H.Y., Kwok,L.W. and Pollack,L. (2008) Mono- and trivalent ions around DNA: a small-angle scattering study of competition and interactions. *Biophys. J.*, **95**, 287–295.
26. Tria,G., Mertens,H.D.T., Kachala,M. and Svergun,D.I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ*, **2**, 207–217.
27. Norouzi,D. and Zhurkin,V.B. (2018) Dynamics of chromatin fibers: comparison of Monte Carlo simulations with force spectroscopy. *Biophys. J.*, **115**, 1644–1655.