DNA Flexibility on Short Length Scales Probed by Atomic Force Microscopy

Alexey K. Mazur

UPR9080 CNRS, Université Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique,
13, rue Pierre et Marie Curie, Paris, 75005, France

Mounir Maaloum

Institut Charles Sadron, CNRS–University of Strasbourg, 23 rue du Loess, BP 84087, 67034 Strasbourg cedex 2, France

Unusually high bending flexibility has been recently reported for DNA on short length scales. We use atomic force microscopy (AFM) in solution to obtain a direct estimate of DNA bending statistics for scales down to one helical turn. It appears that DNA behaves as a Gaussian chain and is well described by the wormlike chain model at length scales beyond 3 helical turns (10.5 nm). Below this threshold, the AFM data exhibit growing noise because of experimental limitations. This noise may hide small deviations from the Gaussian behavior, but they can hardly be significant.

Long double stranded DNA behaves as a continuous elastic rod with bending deformations described by the harmonic wormlike chain (WLC) model [1–3]. In many biological processes the DNA flexibility, notably its ability to wrap around proteins, plays a key role, therefore, the bendability of DNA is actively studied [4–6]. Recent experimental data indicate that the WLC model significantly underestimates the probability of strong bends on length scales shorter than the persistence length ($l_b = 50$ nm) [7–12]. This hypothesis is vigorously disputed because it does not agree with all data [13–23] and because the effect is crucial for biology [24–26].

Atomic force microscopy (AFM) has the advantage of directly observing DNA when adsorbed onto supporting surfaces. Earlier studies showed that in mild conditions the DNA molecules equilibrate on the surface by 2D diffusion so that the chain statistics is not perturbed [27,28]. Two groups earlier used this method for studying the statistics of bending in short DNA [9,11]. It was found that for lengths > 30 nm the probability distributions of bend angles and end-to-end distances agree with the WLC model, but for shorter lengths the populations of strongly bent conformations are much higher than the WLC predictions. In contrast to DNA cyclization, where high probabilities of small circles can be due to rare fluctuations like melting bubbles [29–36], the AFM data suggested that the double helix is intrinsically kinkable, that is, it is kinked rather than bent smoothly even for small angles. These results were accounted for by the linear subelastic chain (LSEC) model [9]. According to the model the bending of DNA fragments of finite length $l = 2.5$ nm obeys Boltzmann statistics with an empirical energy function

$$E_{\text{LSEC}}(\theta) = \alpha|\theta|kT, \qquad (1)$$

where θ is the bend angle and kT is the thermal energy. The dimensionless constant α fit to experimental data equals 6.8. Lengths shorter than 2.5 nm are not considered.

In the earlier literature, one aspect of the LSEC model has escaped attention. If the neighboring base-pair steps (bps) are approximately independent, the bending probability distribution for $l = 2.5$ nm represents a convolution of several single-bps distributions. By solving the inverse problem one can derive the effective single-step potential from Eq. (1). A reasonably accurate approximate solution can be easily found by trials and errors. Figure 1 displays the results of Brownian dynamics simulations with such potential. The discrete coarse-grained model of DNA from the earlier report [37] was employed with one bead per base pair. The single-bps bending potential was as follows:

$$U_1(\theta) = \begin{cases} q\theta^2 & \theta < \theta_0 \\ q\theta_0^2 - \frac{q}{k}\theta_0(\theta_0 - \pi)\left[1 - \left(\frac{\theta - \pi}{\theta_0 - \pi}\right)^{2k}\right] & \theta \geq \theta_0, \end{cases}$$
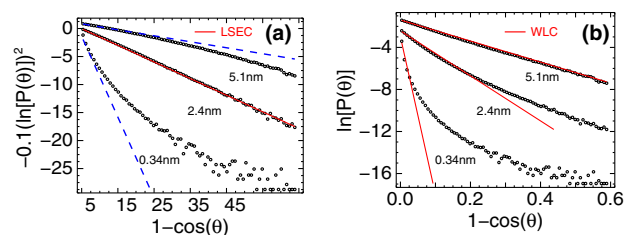

FIG. 1 (color online). Probability distributions of DNA bending obtained by Brownian dynamics simulations of the LSEC model. The DNA lengths are shown near the plots. In panel (a) the solid red line approximates the LSEC postulate. The dashed traces are shown for visual convenience. Solid lines in panel (b) represent WLC predictions for $l_b = 50$ nm.

with $\theta_0 = 2°$, $k = 2$, and $q = 110$ kcal/mole. It is harmonic around zero to avoid singularity, but concave beyond a narrow vicinity, which favors sharp bends. The X and Y axis scales in Figs. 1(a) and 1(b) linearize the bend angle distributions of the LSEC and WLC models, respectively [14]. The results are shown for three DNA lengths. For $l = 0.34$ nm the bend angle distribution is very broad due to easy kinking. For $l = 2.4$ nm it agrees with the LSEC hypothesis [Fig. 1(a)], and diverges from the WLC distribution with large angles [Fig. 1(b)]. However, for $l = 5.1$ nm the distribution is already indistinguishable from the WLC prediction with $l_b = 50$ nm [Fig. 1(b)]. This result demonstrates that Eq. (1) is not sufficient to account for the AFM data because the uniform flexibility of the double helix ensures very rapid convergence to a Gaussian behavior [26]. A much stronger tacit assumption of the LSEC model is that sharp bends are spaced by 2.5 nm intervals of straight DNA [38], which is hard to believe.

To shed light upon the above difficulty we used the AFM method to evaluate DNA bending at short length scales. Linear DNA with a fixed length of 4363 bp was obtained by cutting a PBR322 plasmid with EcoRI restrictase. Experiments were performed in a solution containing a 10 mM tris-HCl buffer, pH 7.5, supplemented with 1 mM MgCl2, to a final DNA concentration of 1 mg/ml. 200 ml of this DNA solution was injected in the AFM liquid cell and DNA molecules adsorbed onto freshly cleaved muscovite mica at room temperature. Images were collected using a Nanoscope 8 (Bruker) operated in the tapping mode in solution, with a pixel size (grid spacing) of 1.95 nm. Ultrasharp noncontact silicon cantilevers Multi75Al (NanoAndMore) were driven at oscillation frequencies in the range of 20–26 kHz. During AFM imaging, the force was reduced in order to avoid dragging of DNA by the tip. The line scan rate was usually 1.4 Hz. The integral gain was adjusted to give sharp images. Images were taken without on-line filtering and were subsequently processed only by flattening to remove the background slope. The AFM images of DNA were transformed into discrete chains under visual control by using a custom implementation of the tracing algorithm by Wiggins *et al.* [9]. This procedure was repeated several times using different link lengths l_0. Five independent sets of contours thus obtained were chosen for further analysis. The corresponding l_0 values were 2.5, 3.5, 7, 10.5, and 14 nm, respectively. The total contour length of DNA observed in the AFM images was ~348 μm (~10^6 bp). With $l_0 = 2.5$ nm this gave about 139000 angles between adjacent links (compared to 98000 in the earlier report [9]). For statistical estimates the contours were divided into fragments so that every measured angle was counted only once.

The AFM data were compared with Monte Carlo (MC) simulations of planar discrete WLC and LSEC models. A phantom chain was considered without the excluded volume effect. The bend angles were sampled directly

from appropriate Boltzmann distributions. To get a feel of statistical errors the DNA length and the volume of sampling were similar to those in experiment. The link length l_0 in the digitized AFM DNA contours is always larger than one bps. For short chains this gives a significant bias with respect to the underlying DNA. To take this into account, MC simulations of the WLC model were performed with one bead per bp, but the resulting chain configurations were resampled by stepping along MC bead positions with fixed strides corresponding to link lengths in AFM data. These new contours were processed in the same way as experimental data to generate reference WLC curves.

Before the beginning of this study we considered two possible origins of short-length deviations of AFM results from the WLC theory. The first of them is related to experimental conditions. The short length deviations from the WLC theory were earlier observed in AFM of dry DNA [9,11]. During drying, strong DNA-ion interactions and other electrostatic effects may change the DNA conformation. In contrast, solution AFM allows direct visualization of DNA in nearly physiological conditions [39]. To check if this experimental difference plays a role we analyzed our AFM images by a procedure identical to that in the original report by Wiggins *et al.* [9]. It turned out that the experimental estimates of DNA bendability obtained on air and in solution agree nearly perfectly. Some representative results are shown in Fig. 2(a). In agreement with the LSEC model, short DNA exhibits excessive flexibility, but for lengths beyond 30 nm everything converges to the WLC theory. The only notable difference from the earlier data is a somewhat larger asymptotic l_b value (56 nm). This may be due to solvent conditions, a systematic bias in the measured DNA length, or the exclusion volume effect [27].

The second possibility we considered was that the deviations from the WLC model might be due to the choice of $l_0 = 2.5$ nm in the tracing algorithm. This length corresponds to 0.7 of a helical turn. As a result, the bending
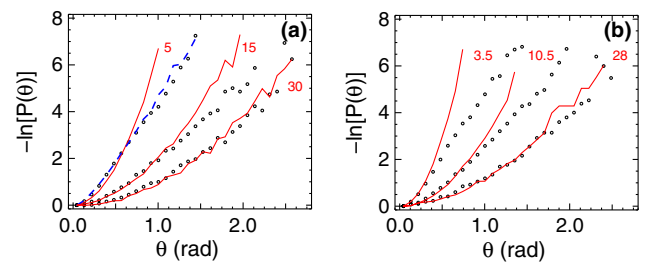

FIG. 2 (color online). Negative logarithm of the probability distribution $P(\theta)$ for the angle θ between tangents separated by different contour lengths indicated in the figure. The AFM images were traced with link lengths $l_0 = 2.5$ nm [panel (a)] and 3.5 nm [panel (b)]. The dots are experimental data. Solid lines: MC evaluation of the same function for the WLC model with $l_b = 56$ nm. The dashed line in panel (a) displays the MC evaluation for LSEC model [Eq. (1)].

is measured only for DNA fragments with nonintegral numbers of turns. It is known from all atom molecular dynamics simulations that measuring bend angles in such fragments is prone to large errors due to rotation of reference bp frames and anisotropy of bending towards DNA grooves [40,41]. The AFM resolution is lower, but this difficulty should persist for any method that tries to probe bending in DNA fragments of a few helical turns. Figure 2(b) reveals, however, that qualitatively similar deviations from the WLC model are evident also when the AFM images are traced with $l_0 = 3.5$ nm corresponding to one helical turn.

Continuing the search, we decided to check the consistency of the results obtained with different link lengths. Figure 3 compares the bend angle distributions for 14 nm DNA in AFM contours traced with $l_0 = 3.5$, 7, and 14 nm (1, 2, and 4 DNA turns, respectively). Strong deviations from the WLC model are observed only with $l_0 = 3.5$ nm. With $l_0 = 7$ nm they are much smaller and disappear completely with $l_0 = 14$ nm. For any smooth contour the measured bend angles depend upon l_0 simply due to discretization. As explained above, this effect is taken into account in the WLC curves in Fig. 3. With l_0 increased, the shape of the WLC probability distribution is preserved, but it is uniformly scaled and corresponds to a higher l_b value (see the three reference WLC curves in Fig. 3 and discussion below). In contrast, the experimental data in Fig. 3 reveal that, with l_0 increased, strong bends are suppressed selectively and so that the distribution approaches the theoretical result of the WLC model.

To get further data, we systematically checked AFM images that contributed high populations of strong bends with $l_0 = 3.5$ nm. For instance, in the molecule in Fig. 4(a) there was a kink of 78.6°. The tracing was repeated 30 times in opposite directions starting from different points. These new contours usually contained a few bends beyond 43° [the upper limit for the left WLC curve in Fig. 2(b)]. However, these bends almost never occurred near the original kink and the new kink locations varied in repeated contours. Figure 4(b) shows a fragment of this DNA with a bundle of 30 contours superimposed. In the AFM images
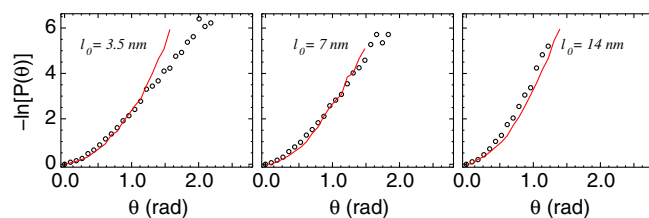
the width of the DNA varied between 4 and 8 pixels [42]. The bundle width in Fig. 4(b) also is not constant and reaches three pixels (5.8 nm). Irreproducible strong bends commonly belonged to zones where the measured DNA width was larger than average. With l_0 increased, the bundle width is reduced. Figure 4(c) shows the results of similar tests with $l_0 = 14$ nm.

The above results demonstrate that the measured population of large bending angles rapidly grows when the l_0 value is reduced to lengths comparable to the DNA width in AFM images. Figures 3 and 4 strongly suggest that this effect represents a limitation of data processing rather than the physical property of DNA. A sensible comparison with the WLC theory is possible only when this effect is small. Based upon Figs. 3 and 4 we concluded that the appropriate link length for our data is $l_0 = 7$ nm (two helical turns). With larger link lengths the contours are increasingly prone to round sharp bends [see Fig. 4(c)]. Some deviations from the WLC model are still evident with $l_0 = 7$ nm, but they are not very significant. The link length $l_0 = 10.5$ nm was also checked, and this gave results similar to $l_0 = 14$ nm, that is, deviations from the WLC model were absent, but the apparent l_b values were slightly overestimated due to cutting sharp bends. Figure 5 shows a more detailed analysis of the shape of the probability distribution functions obtained with $l_0 = 7$ nm in the range of small angles
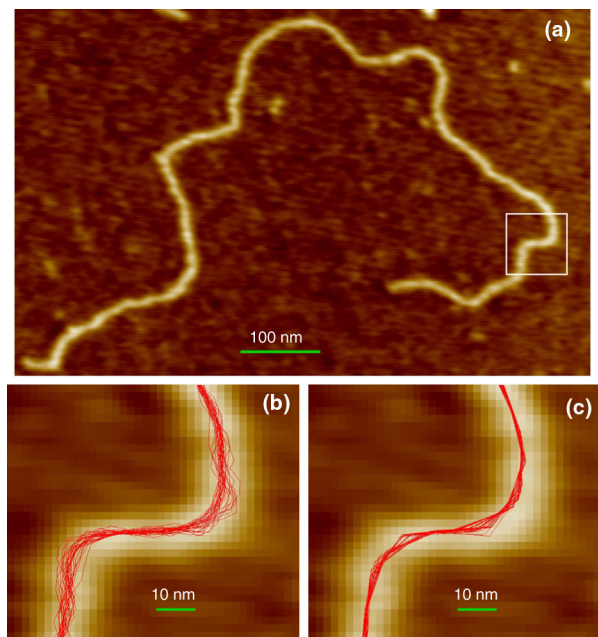

FIG. 4 (color online).   Panel (a) shows a high resolution AFM image of a DNA molecule selected as described in the text. It was traced 30 times using two alternative directions and starting from arbitrary points at opposite ends. Panel (b) shows a zoomed out view of the area highlighted by the white rectangle including a fragment of DNA with the bundle of contours computed with the link length $l_0 = 3.5$ nm. Panel (c) shows similar results for tracing with a larger link length $l_0 = 14$ nm.

FIG. 3 (color online).   Negative logarithm of the probability distribution $P(\theta)$ for the angle θ between tangents separated by a contour length of 14 nm. The AFM images were traced with three different link lengths indicated in the figure. The dots are experimental data. Solid lines: MC evaluation of this function for the WLC model traced with the same link lengths.
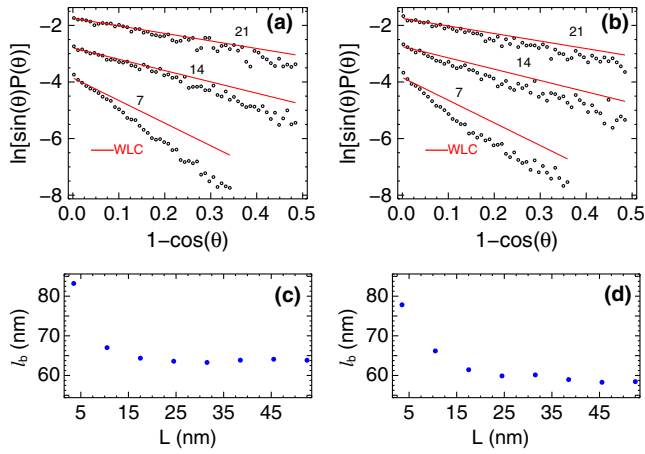
FIG. 5 (color online).   Probability distributions of DNA bending obtained with $l_0 = 7$ nm for MC simulations [panels (a), (c)] and experimental data of AFM [panels (b), (d)]. The plots are labeled by the corresponding DNA lengths (in nm). Straight lines in both panels represent WLC predictions for $l_b = 56$ nm. Panels (c) and (d) display the values of the bending persistence length l_b obtained with Eq. (3)

where the data can be linearized by an appropriate choice of scales.

According to the WLC model, DNA of contour length L equilibrated on a plane is described by the normalized bend angle distribution [27]

$$P[\theta(L)]_{2D} = \sqrt{\frac{l_b}{2\pi L}} \exp\left(-\frac{l_b \theta^2}{2L}\right) \qquad (2)$$

so that

$$\langle \theta^2 \rangle = L/l_b. \qquad (3)$$

For small angles Eq. (2) gives linear plots in coordinates $\ln[\sin(\theta)P(\theta)]$ versus $[1 - \cos(\theta)]$, for instance. Figure 5(a) shows the results of MC simulations of a WLC model with one bead per bp traced with the link length $l_0 = 7$ nm. The distribution functions for the three smallest contour lengths all have linear shapes corresponding to the WLC model. The apparent l_b value is visibly overestimated due to discretization. As seen in Fig. 5(c), with L increased, l_b decreases, but does not reach the value of the underlying WLC model (56 nm) [42]. Figures 5(b) and 5(d) display similar plots for experimental AFM data. As expected, for $L = 7$ nm the shape of the probability distribution slightly deviates from the WLC model, nevertheless, the overall pattern is evidently similar to that in MC simulations.

The non-Gaussian statistics of bending fluctuations in short DNA were observed in two earlier AFM studies [9,11], and also here in Fig. 2. The experimental conditions and parameters of DNA images in these three cases were not identical, but the linear contours were obtained with the same tracing algorithm. Transforming AFM data into linear

contours is not trivial because there is no constructive definition of the centerline of a DNA image. Defining it as a minimum-cost path, for instance, leads to biased contours with underestimated flexibility [43]. The tracing algorithm by Wiggins *et al.* [9] is rapid and simple. It uses a manually set starting point and search direction; therefore, its result is a bundle of contours rather than a single line. We found that the spread of this bundle dramatically grows when the tracing link length is reduced below a certain limit. This effect was responsible for the apparent deviations from the Gaussian behavior in our experiment. When it is negligible, the measured bending statistics agrees with the WLC model, which involves DNA lengths beyond three helical turns. The origin of non-Gaussian effects in the earlier studies [9,11] could be different because they employed AFM in air and parameters of DNA images were different. We believe that AFM in solution can be also used for probing smaller DNA lengths, but this should require additional work on experimental conditions and data processing.

In summary, AFM experiments demonstrate that bending fluctuations in DNA absorbed on a plane in solution are Gaussian and well described by the WLC model at all length scales beyond 3 helical turns (10.5 nm). With DNA lengths reduced below this threshold, the AFM data exhibit growing noise because of experimental limitations. This noise may hide small deviations from the Gaussian behavior, but they can hardly be significant.

———————

[1] S. E. Bresler and Y. I. Frenkel, Zh. Eksp. Teor. Fiz. **9**, 1094 (1939).

[2] L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part 1* (Nauka, Moscow, 1976).

[3] C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry, Part III: The Behavior of Biological Macromolecules* (W. H. Freeman, San Francisco, 1980).

[4] F. H. C. Crick and A. Klug, Nature (London) **255**, 530 (1975).

[5] P. J. Hagerman, Annu. Rev. Biophys. Biophys. Chem. **17**, 265 (1988).

[6] J. P. Peters and L. J. Maher, Q. Rev. Biophys. **43**, 23 (2010).

[7] T. E. Cloutier and J. Widom, Mol. Cell. **14**, 355 (2004).

[8] T. E. Cloutier and J. Widom, Proc. Natl. Acad. Sci. U.S.A. **102**, 3645 (2005).

[9] P. A. Wiggins, T. V. D. Heijden, F. Moreno-Herrero, A. Spakowitz, R. Phillips, J. Widom, C. Dekker, and P. C. Nelson, Nat. Nanotechnol. **1**, 137 (2006).

[10] C. Yuan, H. Chen, X. W. Lou, and L. A. Archer, Phys. Rev. Lett. **100**, 018102 (2008).

[11] H. Chen, H. Fu, Z. Zhou, and J. Yan, Int. J. Mod. Phys. B **24**, 5475 (2010).

[12] R. Vafabakhsh and T. Ha, Science **337**, 1097 (2012).

[13] Q. Du, C. Smith, N. Shiffeldrim, M. Vologodskaia, and A. Vologodskii, Proc. Natl. Acad. Sci. U.S.A. **102**, 5397 (2005).

[14] A. K. Mazur, Phys. Rev. Lett. **98**, 218102 (2007).

[15] H. Shroff, D. Sivak, J. J. Siegel, A. L. McEvoy, M. Siu, A. Spakowitz, P. L. Geissler, and J. Liphardt, Biophys. J. **94**, 2179 (2008).

[16] R. P. Linna and K. Kaski, Phys. Rev. Lett. **100**, 168104 (2008).

[17] D. Demurtas, A. Amzallag, E. J. Rawdon, J. H. Maddocks, J. Dubochet, and A. Stasiak, Nucleic Acids Res. **37**, 2882 (2009).

[18] A. J. Mastroianni, D. A. Sivak, P. L. Geissler, and A. P. Alivisatos, Biophys. J. **97**, 1408 (2009).

[19] R. A. Forties, R. Bundschuh, and M. G. Poirier, Nucleic Acids Res. **37**, 4580 (2009).

[20] V. Ortiz and J. J. de Pablo, Phys. Rev. Lett. **106**, 238107 (2011).

[21] R. Schopflin, H. Brutzer, O. Muller, R. Seidel, and G. Wedemann, Biophys. J. **103**, 323 (2012).

[22] H. You, R. Iino, R. Watanabe, and H. Noji, Nucleic Acids Res. **40**, e151 (2012).

[23] A. Vologodskii, Q. Du, and M. D. Frank-Kamenetskii, Artif. DNA: PNA XNA **4**, 1 (2013).

[24] R. Podgornik, Nat. Nanotechnol. **1**, 100 (2006).

[25] P. C. Nelson, Science **337**, 1045 (2012).

[26] A. Vologodskii and M. D. Frank-Kamenetskii, Nucleic Acids Res. **41**, 6785 (2013).

[27] C. Rivetti, M. Guthold, and C. Bustamante, J. Mol. Biol. **264**, 919 (1996).

[28] F. G. A. Faas, B. Rieger, L. J. van Vliet, and D. I. Cherny, Biophys. J. **97**, 1148 (2009).

[29] G. Altan-Bonnet, A. Libchaber, and O. Krichevsky, Phys. Rev. Lett. **90**, 138101 (2003).

[30] J. Yan and J. F. Marko, Phys. Rev. Lett. **93**, 108108 (2004).

[31] Y. Zeng, A. Montrichok, and G. Zocchi, J. Mol. Biol. **339**, 67 (2004).

[32] J. Yan, R. Kawamura, and J. F. Marko, Phys. Rev. E **71**, 061905 (2005).

[33] P. A. Wiggins, R. Phillips, and P. C. Nelson, Phys. Rev. E **71**, 021909 (2005).

[34] P. Ranjith, P. B. Sunil Kumar, and G. I. Menon, Phys. Rev. Lett. **94**, 138102 (2005).

[35] C. Yuan, E. Rhoades, X. W. Lou, and L. A. Archer, Nucleic Acids Res. **34**, 4554 (2006).

[36] N. Destainville, M. Manghi, and J. Palmeri, Biophys. J. **96**, 4464 (2009).

[37] A. K. Mazur, J. Phys. Chem. B **112**, 4975 (2008).

[38] P. A. Wiggins and P. C. Nelson, Phys. Rev. E **73**, 031906 (2006).

[39] Y. L. Lyubchenko and L. S. Shlyakhtenko, Proc. Natl. Acad. Sci. U.S.A. **94**, 496 (1997).

[40] A. K. Mazur, Biophys. J. **91**, 4507 (2006).

[41] A. K. Mazur, J. Phys. Chem. B **113**, 2077 (2009).

[42] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.112.068104 for additional comments and experimental details.

[43] R. J. van Heekeren, F. G. Fass, and L. J. van Vliet, in *Lecture Notes on Computer Science.*, Vol. 4552, edited by B. K. Ersboll and K. S. Perdersen (Springer Verlag, Berlin, 2007), p. 263.