Optics Letters

Sub-10-nm imaging of nucleic acids using spectroscopic intrinsic-contrast photon-localization optical nanoscopy (SICLON)

Adam Eshein,[1,†] Yue Li,[2,†] Biqin Dong,[1,3] Luay M. Almassalha,[1] John E. Chandler,[1] The-Quyen Nguyen,[1] Karl A. Hujsak,[4] Vinayak P. Dravid,[4] Cheng Sun,[3] Hao F. Zhang,[1] and Vadim Backman[1,*]

[1]Department of Biomedical Engineering, Northwestern University, 2145 Sheridan Rd., Evanston, Illinois 60208, USA
[2]Applied Physics Program, Northwestern University, 2145 Sheridan Rd., Evanston, Illinois 60208, USA
[3]Department of Mechanical Engineering, Northwestern University, 2145 Sheridan Rd., Evanston, Illinois 60208, USA
[4]Department of Material Science and Engineering, Northwestern University, 2145 Sheridan Rd., Evanston, Illinois 60208, USA
*Corresponding author: v-backman@northwestern.edu

Received 25 July 2018; revised 21 September 2018; accepted 3 October 2018; posted 9 October 2018 (Doc. ID 340369); published 29 November 2018

Elucidating chromatin structure *in vitro* requires resolution below 10 nm to visualize the mononucleosome has been an ongoing challenge. In this work, we achieve sub-10-nm imaging of nucleic acids via spectroscopic intrinsic-contrast photon-localization optical nanoscopy (SICLON) without the use of external labels. SICLON leverages two key innovations: using endogenous nucleotides as the emission source and a custom-made imaging system that can simultaneously record the position and optical spectra of emitting molecules. With a novel spectral regression algorithm that identifies the spectroscopic fingerprints of neighboring molecules that were previously indistinguishable, we demonstrate the utility of SICLON by visualizing unlabeled poly-nucleotides and linear single-stranded DNA fibers with a resolution of 6.2 nm.

https://doi.org/10.1364/OL.43.005817

The nanoarchitecture of chromatin underlies and regulates essentially all genetic machinery with a complex organization ranging from individual macromolecules that are a few nanometers in size (e.g., DNA) to macromolecular assemblies that may span tens of nanometers (e.g., nucleosome, chromatin fiber) to micron-scale structures forming topologically associating domains (TAD) and chromatin compartments [1]. The major scientific challenge is to understand this heterogeneous chromatin structure at all length scales in non-perturbed states. Tremendous advances in techniques, such as neutron scattering, small angle x-ray scattering, electron microscopy, super-resolution microscopy, etc., have greatly enriched our knowledge of chromatin nanostructures in the past decades [2–5]. Super-resolution microscopy, specifically photon-localization microscopy (PLM), can resolve chromatin at subdiffractional length scales (<200 nm) with a micron-scale field of view. PLM has been employed to visualize the second-order (e.g., chromatin fiber) and the higher-order chromatin structure (e.g., TAD) [6,7]. However, to visualize even finer details of chromatin, such as a single nucleosome or DNA (length scales <10 nm), using conventional PLM is challenging. The resolution of conventional PLM is limited by two key factors: labeling density and the number of photons from each emission (localization uncertainty). High label density is required to resolve structures according to the Nyquist criterion. Even with sufficiently high labeling density, the precision of fitting a point-spread function (PSF) during image reconstruction in PLM is dependent on the number of photons in each emission event [8]. Specifically, the spatial resolution of the PSF fit after reconstruction is proportional to $1/\sqrt{n}$, where n is the number of emitted photons. Together, these two factors have limited the resolution of conventional PLM techniques to approximately 20 nm. To improve the resolution of PLM, we describe spectroscopic intrinsic-contrast photon-localization optical nanoscopy (SICLON), a super-resolution microscopy technique to overcome both of these obstacles. Using SICLON, we demonstrate the ability to visualize DNA with sub-10-nm resolution.

To date, the methods for imaging chromatin generally require staining, such as the fluorescence dye in super-resolution microscopy [e.g., photo-activated localization microscopy (PALM)/stochastic optical reconstruction microscopy (STORM)] [9] and combined fluorescence dye and heavy metal stain in electron microscopy (e.g., ChromEMT) [4]. However, the use of exogenous dyes has fundamental limitations: the uptake, diffusion, and localization of dyes depends non-linearly on the local environment and could render the chromatin image difficult to interpret, especially for the high label densities required to satisfy the Nyquist criterion for sub-10-nm resolution; and, with spacing smaller than 30 nm, steric hindrance and epitope accessibility for most fluorophore probes becomes a significant issue, as the labels are nearly the size of the molecules of interest. Additionally,

the use of DNA intercalating dyes has the drawbacks of potentially distorting and even damaging DNA structure [10]. An alternative approach of using label-free contrast for direct imaging has considerable advantages.

To tackle the labeling density issue and eliminate the potential artifacts introduced by labels, we utilized a newly discovered phenomenon: DNA can fluoresce under visible light, and the nucleotides of DNA itself are used as the source of photon emission. While DNA had previously been considered "dark" in the visible spectral range, recently it was shown to exhibit photo-switchable autofluorescence when illuminated by visible light using ground-state depletion (GSD) with dark-state shelving and stochastic return [11]. Dong *et al.* previously investigated the photochemical properties of DNA autofluorescence in detail [11,12]. By leveraging GSD, direct nanoscopic imaging of nucleic acids using their intrinsic fluorescence for contrast has been demonstrated. Under visible-light illumination, unmodified DNA has the capacity to stochastically emit photons, allowing for their use in PLM. While nucleic acids have weak absorption within the visible range, they have remarkably high quantum efficiency, a long-lived dark state, and comparably high photon emission counts [11]. Furthermore, as the base unit of chromatin is nucleic acids, their use for PLM completely bypasses the labeling density limits presented by exogenous approaches and the non-linear effects of local macromolecular density would have on molecular mobility and binding affinities. Using the intrinsic stochastic fluorescence of nuclear acids, the capacity to image chromatin without exogenous fluorophores with ~20 nm resolution has been demonstrated [12].

To measure spatial information and spectra of fluorescent molecules (e.g., nucleotides) simultaneously, we built the SICLON instrument shown in Fig. 1, where continuous wave laser illumination is focused onto the back focal plane of a high numerical aperture (NA) objective to allow collimated laser light to exit the objective. The position and angle of the laser can be controlled to achieve total internal reflection fluorescence (TIRF) illumination to excite the fluorescent molecules into long-lived dark states and subsequently recover them by stochastic photo-switching. The emission image is collected by the high NA objective and coupled into a Czerny–Turner monochromator (SP2150, Princeton Instruments) via a matched tube lens. The emission image is split into its zeroth-order and spectrally dispersed first-order image via a blazed dispersive grating (150 grooves per mm) and projected onto a high sensitivity electron-multiplying CCD (EMCCD, ProEM, Princeton Instruments) with a resolution of 0.63 nm. This configuration allows simultaneous collections of spatial (zeroth-order) and spectral (first-order) information from each emission event. It should be noted that only ~1/4 of the light is sent to the zero order, so we see a decrease in localization precision equal to approximately $\sqrt{4} = 2$ times. Experimentally we have measured an average localization uncertainty of approximately 40 nm on average, indicating that our theoretical resolution should be 20 nm, comparable to traditional label-based STORM techniques.

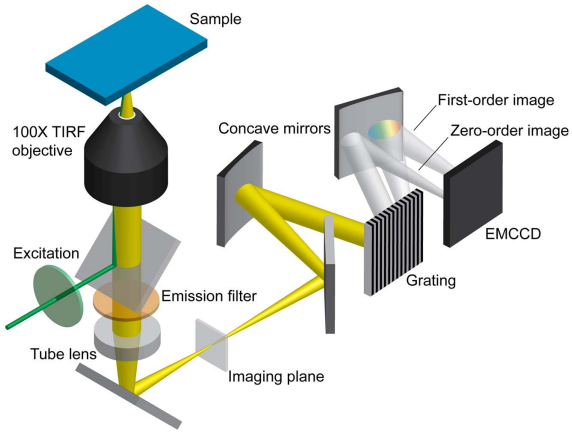

Fig. 1. System schematic. Fluorescent images were collected by a high numerical aperture objective lens and sent to a Czerny–Turner monochromator, where the zero- and first-order images are read separately and simultaneously with an EMCCD.

As mentioned previously, the precision of image reconstruction in most PLM techniques like STORM/PALM relies on fitting the PSF to emission events. The resolution of the localization of any emission event is proportional to $1/\sqrt{n}$. In order to combine emission events from the same emitter and increase the number of photons, conventional reconstruction algorithms consider all emissions occurring in the same resolution limited pixels in consecutive time frames from the "same" molecule [13]. However, this approach has two shortcomings: (1) it might lead to poor localization precision due to incorrect merging of unique emitters, and (2) it does not consider multiple emission events from the same emitter that are not in temporally neighboring frames. To overcome this, we have developed a spectral regression (SR) algorithm to more accurately merge recurrent emission events from a single emitter [8].

The principle of the SR algorithm is that the emission spectra of the same molecule from different blinking events remain consistent compared to the variations of the spectra from other molecules in close proximity [8]. For different nucleotides, we found that the nucleic acid emission has spectroscopic specificity, meaning that the peak position and frequency of the emission spectra depends on the nucleotide residue. Figure 2(a) shows that 20 base pair (bp) DNA molecules (Integrated DNA Technologies, USA) with different molecular makeups (different nucleotides) have unique spectral properties.

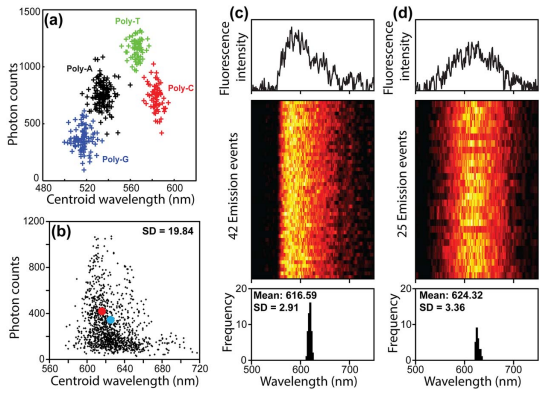

Fig. 2. (a) Photon counts and emission spectra centroids associated with 100 blinking events for unlabeled 20 bp sequences of poly-G, poly-A, poly-T, and poly-C. (b) Spectral heterogeneity of unlabeled poly-DNA molecules. (c) and (d) Emission spectra of individual stochastic localization at two neighboring spots indicated by blue and red dots in (b). The emission events can be separated clearly according to their emission properties, indicating they are from two different emitters.

Figure 2(b) shows the heterogeneity of DNA emission spectra extracted from salmon (Sigma-Aldrich, USA). We were able to differentiate distinct molecules by comparing the average standard deviation (SD) of the spectral centroids of recurrent emissions. For single DNA emitters, we calculated the average SD of the spectral centroid from recurrent blinking evens to be 3.25 nm. For comparison, the SD of the spectral centroid of emissions from unique DNA emitters was 19.84 nm. Thus, the spectra from a given emitter are consistently unique and can be differentiated from the spectra of other molecules.

The probability that a different molecule generates a spectrum with the same spectral centroid as a target molecule (α) can be estimated from the error function (erf) and the SD of spectral centroids from the same molecules (σ_{single}) and different molecules (σ_{multiple}) using the equation

$$\alpha \sim \sqrt{\frac{2}{\pi}} \int_0^{\infty} \left\{ \left[\operatorname{erf}\left(\frac{x + \frac{\sigma_{single}}{\sigma_{multiple}}}{\sqrt{2}} \right) - \operatorname{erf}\left(\frac{x}{\sqrt{2}} \right) \right] \exp\left(-\frac{x^2}{2} \right) \right\} dx.$$

In the particular case in Figs. 2(b)–2(d), the added relative localization uncertainty is negligible at $\sim\alpha/2 \approx 4\%$. Under the key principle, we can safely rely on the centroid wavelength to determine whether two blinking events within the resolution limit of conventional STORM originate from the same molecule. The working principles of the SR algorithm are shown in Fig. 3. To implement the algorithm, the neighborhood around each blinking event within the localization uncertainty in all frames will be searched to yield candidate events that can be categorized as repeat emissions. Based on the position of the spectral centroid, we determine whether multiple emissions belong to the same molecule. All emissions corresponding to the same molecule were merged by calculating the average of their localized positions. This process was repeated until all unique molecules in the dataset have had their repeated emissions combined and merged centroids localized [8]. Molecules with few repeated emission events were filtered out to remove false blinking events and ensure sufficiently high localization precision of all molecules.

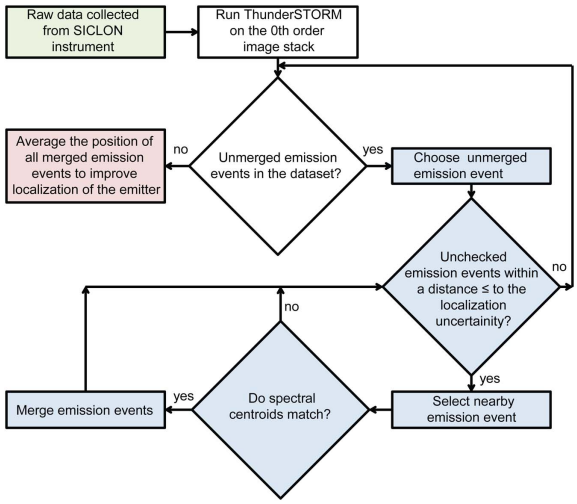

Fig. 3. Flow chart of SR algorithm. Conventional PLM reconstruction was performed in ThuderSTORM. Using the SR algorithm, emission events that belong to the same molecule were added (merged) to improve the accuracy of emitter localization.

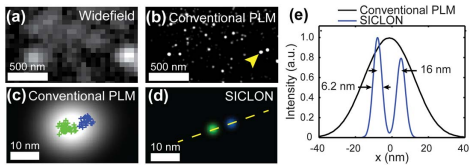

Fig. 4. (a) Wide-field image of 20-mer poly-DNA, (b) STORM reconstruction of 20-mer poly-DNA, (c) enlarged view of two 20-mer poly-DNA molecules, (d) after SR the two 20-mer poly-DNA molecules are separable, and (e) the FWHM of each individual molecule is shown to be 6.2 nm, where the molecules are shown to be spaced 16 nm apart.

To improve imaging resolution, the intrinsic fluorescence contrast from nucleic acids that form chromatin completely bypasses the labeling density limits presented by exogenous approaches for high resolution imaging, and the SR algorithm significantly improves the localization quality by precisely classifying neighboring emissions and identifying recurrent emissions from the same molecule. It is possible to improve the resolution to the sub-10-nm range by taking advantage of both techniques. Using the new nanoscale imaging technology combining time-resolved visible-light stochastic photon localization with simultaneous spectroscopic molecular-signature-carrying intrinsic fluorescence detection, referred to as SICLON, we achieve a lateral resolution of 6.2 nm in both linear single-stranded DNA (ssDNA) fibers and unmodified poly-nucleotide sequences. The resolution of SICLON is confirmed by atomic force microscopy (AFM) imaging of similar linear ssDNA fibers prepared in parallel.

As demonstrated in Fig. 4(a), isolated 20-mer poly-DNA complexes deposited on glass coverslips have weak, but observable fluorescence under wide-field illumination. By using SICLON and analyzing the zero-order emissions, resolution can be improved to ≤40 nm [Figs. 4(b) and 4(c)]. As the zero-order source is composed of distinct emission spectra owning to conformational and molecular variations, each molecule can be identified utilizing SR [Fig. 4(c)]. Significantly, even for molecules that are less than 20 nm apart (theoretical resolution limit), SR allows clear delineation of molecular features down to 6.2 nm, as measured by the full-width half-maximum (FWHM) of the nucleotide source. With this resolution, SICLON can fully differentiate two molecules that are 16 nm apart [Fig. 4(d)].

To test the improved resolution for unmodified nucleotide samples, we imaged isolated salmon sperm ssDNA (Fig. 5). The ssDNA was spin-coated on a glass coverslip to produce linear DNA fibers [14]. To confirm that the molecular origin of these fibers was nucleic acids, we utilized Hoechst 33342 wide-field fluorescent imaging using a 405 nm source [Fig. 5(a)]. Next, we performed SICLON imaging of the fibers using 532 nm excitation to produce a conventional PLM image (zeroth-order image) of the DNA [Figs. 5(b) and 5(d)]. We observed co-localization between the PLM image and wide-field fluorescent Hoechst 33342 imaging, thus confirming the DNA origin of the structure. Next, we utilized SICLON collection of the zeroth-order emission and the first-order spectra from the fiber to apply SR and emission integration [Figs. 5(c) and 5(e)]. Using SICLON, we observe an average of approximately 2100 photons recorded from each emitter merged in the zero-order

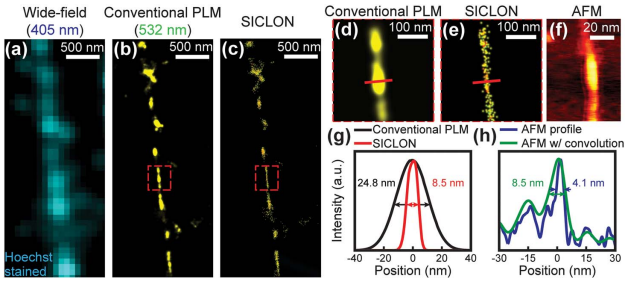

Fig. 5. (a) Wide-field fluorescence image of Hoechst stained DNA fiber, (b) STORM reconstruction of DNA fiber using 532 nm illumination, (c) STORM reconstruction of DNA fiber after applying SR algorithm, (d) enlarged view of selection shown in (b), (e) enlarged view of selection shown in (c), (f) an image acquired by AFM of a separate single DNA fiber, (g) comparison of resolution between traditional PLM and SICLON with SR shows a four times resolution enhancement, and (h) AFM reveals the ground truth width of the fiber is estimated to be 4.1 nm.

image, resulting in a minimal FWHM of the DNA fiber samples of 8.5 nm [Fig. 5(g)]. To compare this resolution with the ground truth of the fiber, we performed AFM of DNA samples spin-coated in parallel on freshly cleaved mica at 0.98 nm resolution [Fig. 5(f)]. With AFM, we observe a minimal FWHM of 4.1 nm of fibers with similar geometry to those observed with SICLON in Figs. 5(a)–5(e). Through convolution of the minimal FWHM of 4.1 nm observed with AFM and the 6.2 nm resolution observed in SICLON (Fig. 4), a fiber with an observable FWHM of 8.5 nm is yielded, which matches the SICLON result, as shown in Figs. 5(g) and 5(h). This further confirms the estimated lateral resolution of SICLON as 6.2 nm.

In this work, we present a nanoscopic imaging technique, SICLON, to directly image nucleic acids with the resolution required to fully image the organizational structure of chromatin down to the single nucleosomal level without the need for extrinsic dyes. Using the intrinsic contrast produced by the stochastic emission of nucleic acids and SR, we demonstrate in isolated nucleotides and ssDNA a lateral resolution of 6.2 nm. This resolution surpasses the capabilities presented *ex-vivo* with DNA point accumulation for imaging in nanoscale topography (DNA-PAINT) [15] or dual-objective STORM [16]. SICLON can also be paired with these enabling technologies to further increase resolution and provide molecular-specific information about the *in-vitro* structure of chromatin. We envision that subsequent work with SICLON will provide direct mapping of nucleosomal occupancy in fixed cells using complementary hybridization to demarcate genes of interest. Importantly, as we observe small but detectable differences in the spectral emission from each nucleotide, it may be possible in the future to utilize SR to both identify the molecular composition (molecular fingerprint) and structure of chromatin without exogenous labels. We believe the potential of SICLON to fully map the structure and chemistry of chromatin *in-vitro* will greatly expand our understanding of how chromatin topology influences essentially all genetic machinery.

Funding. National Cancer Institute (NCI) (R01CA200064, U54CA193419); National Institute of Biomedical Imaging and Bioengineering (NIBIB) (F31EB022414); National Science Foundation (NSF) (CBET-1240416, CBET-1706642); LUNGevity Foundation; Soft and Hybrid Nanotechnology Experimental (SHyNE) Resource (NSF ECCS-1542205); MRSEC program (NSF DMR-1720139); International Institute for Nanotechnology (IIN); Keck Foundation; State of Illinois, through the IIN, and the Air Force Office of Scientific Research (AFOSR) (FA9550-17-1-0348).

Acknowledgment. This work made use of the SPID facility of Northwestern University's NUANCE Center, which has received support from the SHyNE; MRSEC at the Materials Research Center; IIN; Keck Foundation; and the State of Illinois, through the IIN, and the Air Force Office of Scientific Research.

[†]These authors contributed equally to this work.

REFERENCES

1. R. J. Ellis and A. P. Minton, Nature **425**, 27 (2003).
2. K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, Nature **389**, 251 (1997).
3. J. M. Harp, B. L. Hanson, D. E. Timm, and G. J. Bunick, Acta Crystallogr. Sect. D **56**, 1513 (2000).
4. H. D. Ou, S. Phan, T. J. Deerinck, A. Thor, M. H. Ellisman, and C. C. O'Shea, Science **357**, eaag0025 (2017).
5. B. J. Beliveau, E. F. Joyce, N. Apostolopoulos, F. Yilmaz, C. Y. Fonseka, R. B. McCole, Y. Chang, J. B. Li, T. N. Senaratne, and B. R. Williams, Proc. Natl. Acad. Sci. USA **109**, 21301 (2012).
6. P. J. Fabre, A. Benke, S. Manley, and D. Duboule, *Cold Spring Harbor Symposia on Quantitative Biology* (Cold Spring Harbor Laboratory, 2015), pp. 9–16.
7. S. Wang, J.-H. Su, B. J. Beliveau, B. Bintu, J. R. Moffitt, C.-T. Wu, and X. Zhuang, Science **353**, 598 (2016).
8. B. Dong, L. Almassalha, B. E. Urban, S. Khuon, T.-L. Chew, V. Backman, C. Sun, and H. F. Zhang, Nat. Commun. **7**, 12290 (2016).
9. R. Henriques, C. Griffiths, E. Hesper Rego, and M. M. Mhlanga, Biopolymers **95**, 322 (2011).
10. E. Richard, S. Causse, C. Spriet, N. Fourré, D. Trinel, X. Darzacq, B. Vandenbunder, and L. Heliot, Photochem. Photobiol. **87**, 256 (2011).
11. B. Dong, L. M. Almassalha, B. T. Soetikno, J. E. Chandler, B. E. Urban, C. Sun, H. F. Zhang, and V. Backman, Opt. Express **25**, 7929 (2017).
12. B. Dong, L. M. Almassalha, Y. Stypula-Cyrus, B. E. Urban, J. E. Chandler, C. Sun, H. F. Zhang, and V. Backman, Proc. Natl. Acad. Sci. USA **113**, 9716 (2016).
13. M. Ovesný, P. Křížek, J. Borkovec, Z. Švindrych, and G. M. Hagen, Bioinformatics **30**, 2389 (2014).
14. H. Yokota, J. Sunwoo, M. Sarikaya, G. van den Engh, and R. Aebersold, Anal. Chem. **71**, 4418 (1999).
15. J. Schnitzbauer, M. T. Strauss, T. Schlichthaerle, F. Schueder, and R. Jungmann, Nat. Protoc. **12**, 1198 (2017).
16. K. Xu, H. P. Babcock, and X. Zhuang, Nat. Methods **9**, 185 (2012).