Supporting Information

Super-resolution imaging of self-assembled nanocarriers using quantitative spectroscopic analysis for cluster extraction

Janel L. Davis ^a, Yang Zhang ^a, Sijia Yi ^a, Fanfan Du ^a, Ki-Hee Song ^a, Evan A. Scott ^a, Cheng Sun ^b and Hao F. Zhang ^{a, *}

- a. Department of Biomedical Engineering, Northwestern University, 2145 Sheridan Road, Evanston, Illinois, 60208, United States
- b. Department of Mechanical Engineering, Northwestern University, 2145 Sheridan Road, Evanston, Illinois, 60208, United States

* Corresponding Author

Corresponding Author Email: hfzhang@northwestern.edu

Supporting Information Summary: 11 Pages, 15 Figures, & 3 Notes		
Page No.	Supporting Information Table of Contents	
S2	Figure S1	Bulk and single-molecule spectroscopic measurements of Nile Red in polymersomes and immobilization reagents
S3	Figure S2	sPAINT reconstructions before and after applying the spectral window for selecting polymersomes
S3	Figure S3	Comparison of the accuracy of quantifying polymersomes at different localization densities
S4	Supporting Note 1	Development of quantitative sPAINT
S5	Figure S4	Histograms of the spatial photon counts, spectral photon counts and spectral precisions
S5	Figure S5	Selection of the optimum parameters for density-based cluster analysis
S 6	Figure S6	Selection of the optimum parameters for quantitative sPAINT
S 6	Figure S7	Selection of the optimum parameters for counting polymersomes
S7	Figure S8	Super-resolution reconstruction of the polymersome sample using unutilized localizations
S7	Figure S9	The qSPACE workflow
S8	Supporting Note 2	Comparison of size measurements of polymersomes using sPAINT and qSPACE
S8	Figure S10	The percent variance in size measurements and localization density for clusters with increasing localizations per cluster
S9	Supporting Note 3	Optimization of qSPACE for counting and sizing polymersomes
S9	Figure S11	Threshold selection for detecting potential qSPACE clusters
S10	Figure S12	Selection of the optimum spectral window for qSPACE
S10	Figure S13	CryoTEM image of the polymersome sample
S11	Figure S14	Comparison of the size distributions of artifacts identified in control samples for sPAINT and qSPACE
S11	Figure S15	Bulk spectroscopic measurements of Nile red in bovine albumin serum and neutravidin



Figure S1: (a) Comparison of the average emission spectra of Nile Red in solutions containing polymersomes (PS) and sample preparation reagents (poly-L-lysine (PLL) and phosphate buffered saline (PBS)) measured using a fluorimeter and sPAINT. (b) Representative super-resolution reconstructions of the PLL coated glass used as the control sample. (Scale bar: 1 μ m). (c-e) Histograms of the emission maxima (λ_{max}) of NR interactions in ROIs containing PS from three different images (six clusters were selected in each image). The selected spectral window (SW, 595 – 625 nm) is highlighted in red. (f) Reproducibility test showing the histogram of λ_{max} from PS acquired on a different day with similar experimental conditions. (g) Comparison of the identification percent of localizations from the PS λ_{max} histogram and identification percent of localizations from the SW increases (peak position of SW = 610 nm).



Figure S2: sPAINT reconstructions before applying the spectral window for selecting PS for (a) example 1 and (b) example 2. Overlay of the reconstructions with the localizations designated as PS in cyan and localizations designated from nonspecific binding in red from (c) example 1 and (d) example 2. sPAINT reconstructions of the PS channel for (e) example 1 and (f) example 2. sPAINT reconstructions of the nonspecific binding channels for (g) example 1 and (h) example 2 (Scale bar: 500 nm)



Figure S3: (a) Representative super-resolution reconstruction showing individual clusters with varying sizes identified by the optimized density-based spatial clustering for applications with noise (DBSCAN) algorithm marked by the red plus signs (Scale bar 500 nm). (b) Comparison of the differences in the cluster counting accuracy using DBSCAN for trials using data with different cluster densities (number of nearest neighbors or number of NN), the red bar shows data selected based on the spectroscopic and spatial information while the blue bar shows data selected for analysis based on spatial information alone.

Supporting Note 1

We compared the performance of DBSCAN alone versus DBSCAN using localizations within the selected spectral window (SW). We used localizations with at least 450 photons in representative PS and control images for both trials. The photon distributions in the spatial and spectral domains are included in Figures S4a and S4b. Additionally, the distribution of the spectral precision (Figure S4c) shows that the spectral precision of a typical localization used in this study would be sufficient to be separated from the nonspecific binding with emission maxima populations ~30 nm away from the emission maximum of the sample. We selected SW (610±3 nm) based on the histograms from the PS shown in Figures 2 and Figures S1c-e. We then selected a 30 nm SW where the nonspecific binding to the PLL coated glass was minimized in the control sample (PLL coated glass with no PS) while maximizing the size of the window to allow as much NR interactions with PS to be included (Figure S1g). The selected SW rejected 83% of the localizations from the control sample while 60% of localizations associated with the PS were accepted. Changing the SW window size would affect the percentage localizations from both the nonspecific binding and the PS sample. For example, using a window size of 10 nm, the percentage of localizations from nonspecific binding would be reduced to 5% and the percentage of localizations from the PS sample would be reduced to 23%. This change would reduce the number of localizations per cluster in both cases causing inaccuracy in the size and morphology of the clusters using DBSCAN.

We used the selected cluster radius (40 nm) to determine the threshold on the minimum number of NN (min-NN) for each trial. For clarity, the min-NN threshold set based on DBSCAN alone is referenced by NN_C. Meanwhile, nearest neighbor thresholds based on sPAINT will be designated as NN_V. Figures S5a-c and S5d-f show the comparison between the histograms of NN for localizations in the control and PS samples with and without application of the SW. For the control and PS images the number of localizations was ~10⁵ before the SW filter and ~10⁴ the SW filter. By selecting localizations within the SW, the average number of NN reduced from 13 to 7 in the PS sample. Alternatively, in the control image the average number of NN reduced 5-fold from 10 to 2. This indicates that though the number of localizations was reduced for both groups, the application of the spectra-based criteria preferentially selected localizations which were members of clusters.

Based on this finding, a range of min-NN thresholds were tested for the defining clusters in the PS and control sample. For sPAINT a range of min-NN_V from 5 to 20 were tested (Figure S5g). Meanwhile, for analysis based on DBSCAN alone, a range of min-NN_C from 15 to 50 was tested (Figure S6a). A threshold was selected to achieve the minimal number of clusters in the control without causing morphology changes in the PS image (Figure S6b). Using this threshold, localizations with too few neighbors were rejected as noise. For sPAINT, a min- NN_V of 10 was selected while a min-NN_C of 20 was selected for DBSCAN alone. Figure S7 shows the application of the selected thresholds on an additional dataset with 42 and 31 identified before and after the SW filter respectively.



Figure S4: Histograms of the (a) spatial photon counts, (b) spectral photon counts and (c) spectral precision used for sSMLM analysis from a representative image.



Figure S5: Histograms showing the comparison between the number of NN of localizations with 450 photons in the spatial domain before $(n=\sim10^5)$ and after application of the selected spectral window $(n=\sim10^4)$ of (a-c) 3 control images and (d-f) 3 PS images. (g) Plots showing the number of clusters using localizations within the SW as the min-NN increases in representative PS and control images. The black dashed line shows the threshold used in this study.



Figure S6: Plots showing the number of clusters with localizations with at least 450 photons in the spatial domain as the min-NN increases in the (a) PS and control images. The black dashed line shows the threshold used to compare spatial clustering without considering the PS window. Representative super-resolution reconstructions of the (b) PS sample (white arrows indicate changes in PS morphology) and (c) the control sample after application of the clustering thresholds (min-NN_C =25 and Int=450). (Scale bar 500 nm)



Figure S7: (a) Clusters extracted from a representative image of the PS sample based on DBSCAN alone; (b) Clusters extracted from the PS sample after the SW (Scale bar: 500 nm)



Figure S8: Localizations excluded from spectroscopic analysis based on photon number (Average localization uncertainty ~25 nm, Photon budget: 300-450 in the spatial domain)



qSPACE Workflow

Figure S9: The detailed qSPACE workflow outlines how spectroscopic information is used to recover localizations from clusters from the PS sample while rejecting artifacts from nonspecific binding.

Supporting Note 2

We measured the effect of the reduced number of localizations per cluster on size measurements by selecting 71 sPAINT clusters with at least 40 localizations. We calculated the percent variance $(P_{\Delta SIZE} = \frac{|\overline{D} - D_n|}{\overline{D}})$ where \overline{D} is the final diameter of the cluster and the D_n is the estimated diameter for the cluster with n localizations) for each cluster as the number of localizations sequentially increased. The average percent variance in cluster size reduced from 48% to <5% as the number of localizations increased (blue line Figure S10). This trend was further investigated by grouping the analyzed clusters based on their varying sizes. We found that the majority of clusters smaller than 100 nm could be accurately sized with ~25 localizations per cluster. Meanwhile, clusters larger than 200 nm required 38 localizations per cluster for accurate sizing.

We accounted for the variation in the size requirements for clusters of different sizes by using a threshold based on the localization density ($LD = \frac{N}{\pi r^2}$ where N is the number of localizations in the cluster and r is the radius of the cluster). The red line in Figure S10 represents the LD for an average PS (with 113 nm diameter) as the number of localizations increased. Using the results from the size variation measurements we set a LD threshold of 3.5×10^{-3} nm⁻² for accurate sizing.



Figure S10: The average percent variance in size measurements ($P_{\Delta SIZE}$) and localization density for clusters as the number of localizations per cluster increases. The black line shows where $P_{\Delta SIZE} = 5\%$. The dashed blue and red lines indicate the min-NN of 35 and LD of 3.5×10^{-3} nm⁻² required for accurate sizing.

Supporting Note 3

We tested the performance of qSPACE to reject artifacts from nonspecific binding and identify PS clusters. We used the clusters identified using the sPAINT criteria previously described (SW = [595-625], min-NN_v = 10) as a validation map. Potential PS clusters were identified using localizations using a min-NN_c of 35 to limit variations in size measurements to 5% (Figures S11 & S12). While traditional density-based algorithms would rely on stricter clustering threshold to remove potential artifacts, qSPACE uses a separate channel for sample validation. Using the validation channel, the specificity can be tuned by adjusting the min-NN_v independently (Figure S12). The spatial coordinates of localizations in the validation map are compared to the spatial coordinates of localizations which formed potential clusters. Localizations which did not have neighbors in the validation map were rejected. All other localizations can be retained as members of true clusters. After validation, the min-NN_c (35) for selecting the potential nanocarriers threshold set was used for cluster assignment.



Figure S11: Plots showing the number of clusters at least 300 photons in the spatial domain as the minimum number of NN increases in the (a) PS and control images. The dashed black line represents the threshold used for detecting potential clusters for qSPACE. Representative super-resolution reconstructions of the (b) PS sample and (c) the control sample after application of the clustering thresholds (min-NN_C=35 and Int=300), (d) PS sample and (e) the control sample after application of the clustering thresholds (min-NN_C=45 and Int=300). The white arrows indicate morphology changes in d. (Scale bar: 500 nm)



Figure S12: Plots showing the tuning of the min-NNv for noise removal in the validation channel using qSPACE with a SW window of 595-625 nm and min-NN_c of 35 for potential cluster assignment. The dashed black line shows the selected min-NN_v.



Figure S13: CryoTEM image of PS sample.



Figure S14: (a) Histograms showing the size distribution of artifacts identified from 4 reconstructions of the control sample (809 clusters with average size120±4 nm) and clusters identified as artifacts in the LC (504 clusters with average size 111±5 nm) and HC (501 clusters with average size 90±7 nm) PS samples by qSPACE. (b) Histograms showing the size distribution of artifacts identified in 4 reconstructions of the control sample (229 clusters with average size 199±2 nm) and clusters identified as artifacts in the LC (273 clusters with average size 220±11 nm) and HC (186 clusters with average size 211±8 nm) PS samples by sPAINT (clusters formed from localizations with λ_{max} outside the SW filter).



Figure S15: The average emission spectra of Nile Red in solutions containing PS and sample preparation reagents (bovine serum albumin (BSA) and Neutravidin (NEU)) measured using a fluorimeter.