CORRESPONDENCE

Identification of clustering artifacts in photoactivated localization microscopy

To the Editor: Fluorescent proteins are known to display in certain cases an on-off blinking and switching behavior[1]. We could not yet find a report investigating the impact of this phenomenon on super-resolution techniques that are based on the sequential photoactivation and bleaching of individual emitters such as photoactivated localization microscopy (PALM)[2]. In our hands, even monomeric (m)Eos2, one of the most promising photoconvertible fluorescent proteins reported in this journal[3], had non-negligible light-induced fluorescence recovery[4] (**Supplementary Figs. 1** and **2**). This is particularly important for imaging membrane receptors, for which phenomena such as oligomerization and clustering of proteins can be properly identified only if the number of their constituents is correctly estimated.

Here we compared a negative and a positive clustering control by using plasma membrane–bound proteins labeled with mEos2 to determine how photoblinking or reactivation-induced artifacts may be erroneously interpreted as biological clusters. A PALM image of a fixed HeLa cell expressing, as a negative clustering control, the 15-amino-acid N terminus of the protein Src (SrcN15; **Supplementary Methods**) showed that the localized single molecules formed clusters over a length scale of 100 nanometers (**Fig. 1a**). Using a color code for the instant of localization of each molecule (**Fig. 1b**), we found that most of the observed clusters were largely monochromatic, indicative of a suspicious con-

comitant spatial and temporal clustering, which we confirmed in the cluster kymographs (**Fig. 1c**). If we allowed for longer molecular dark times (up to 10 s) in regrouping the fluorescence traces (**Supplementary Methods** and **Supplementary Fig. 1**), temporal clusters disappeared (**Fig. 1d**), yielding a pattern consistent with what we observed upon labeling SrcN15 with the photoswitchable fluorescent protein, PS-CFP2 (ref. 5) (**Fig. 1e**).

As a positive control for physical clustering, we imaged fixed cells expressing the prototypical G protein–coupled receptor β2-AR a few minutes after stimulation with the agonist isoproterenol. Under these conditions β2-AR internalizes in endosomes on the cell membrane that are even visible in diffraction-limited total internal reflection fluorescence (TIRF) microscopy images[6]. In PALM images (**Fig. 1f**), we detected endosomes as clusters even when we allowed for long fluorescence dark times (**Fig. 1g**). In addition, the kymographs of the clusters (**Fig. 1h**) revealed that the temporal distribution of localizations of the molecules comprising clusters was homogeneous over the entire duration of data acquisition, consistent with the stochastic photoactivation mechanism at the base of PALM measurements.
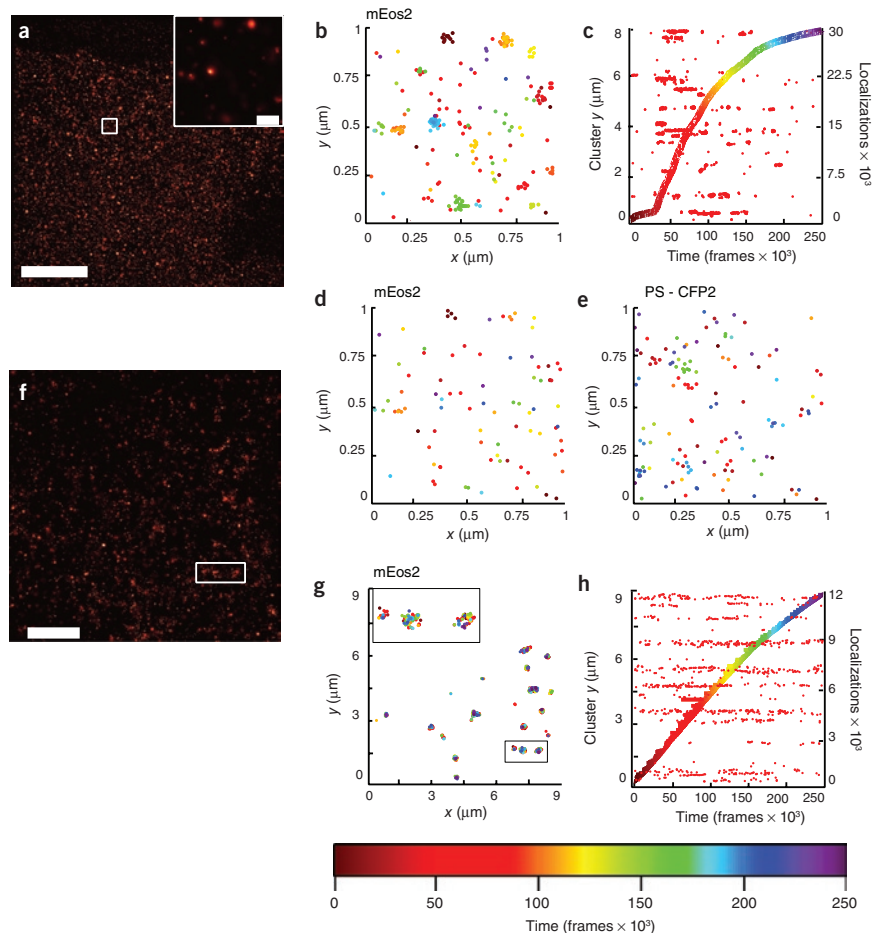

Figure 1 | Protein clustering on plasma membrane. (**a**) TIRF-PALM image of a plasma membrane region of a fixed HeLa cell expressing SrcN15-mEos2. (**b**) Map of the localized emission centers for each SrcN15-mEos2 molecule in the inset in **a**; the color scale identifies the temporal instant of molecular localization; dark time (t_d) = 2 s. (**c**) Representative kymographs of molecules shown in **a** belonging to spatial clusters; superimposed color-coded curve shows localized molecules as a function of time. (**d**) Map for same data as in **b** but with t_d = 10 s. (**e**) Map of the localized emission centers for each SrcN15–PS-CFP2 molecule; t_d = 2 s. (**f**) TIRF-PALM image of a fixed HeLa cell expressing β2-AR, labeled with mEos2, during endocytosis. (**g**) Map of the localized emission centers for each β2-AR–mEos2 molecule; t_d = 10 s. (**h**) Representative kymographs of molecules shown in **f** belonging to spatial clusters. Scale bars, 5 μm (**a**), 200 nm (inset in **a**) and 2 μm (**f**).

Therefore, depending on the fluorophore used, scanning for temporal clusters is an effective strategy to identify photoblinking-induced artifacts in super-resolution images of small membrane-protein aggregates. It appears that allowing for longer single-molecule fluorescence dark times in the data analysis allows one to obtain largely artifact-free images while retaining the essential information about physical clusters (**Supplementary Fig. 3**), although more specific application-targeted methods to identify and remove temporal clustering may be envisioned.

Note: Supplementary information is available on the Nature Methods website.

Paolo Annibale[1,3], Stefano Vanni[2,3], Marco Scarselli[1,3], Ursula Rothlisberger[2] & Aleksandra Radenovic[1]

[1]Laboratory of Nanoscale Biology, Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. [2]Laboratory of Computational Chemistry and Biochemistry, Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. [3]These authors contributed equally to this work.
e-mail: aleksandra.radenovic@epfl.ch

1. Dickson, R.M., Cubitt, A.B., Tsien, R.Y. & Moerner, W.E. *Nature* **388**, 355–358 (1997).
2. Betzig, E. *et al. Science* **313**, 1642–1645 (2006).
3. McKinney, S.A., Murphy, C.S., Hazelwood, K.L., Davidson, M.W. & Looger, L.L. *Nat. Methods* **6**, 131–133 (2009).
4. Annibale, P., Scarselli, M., Kodiyan, A. & Radenovic, A. *J. Phys. Chem. Lett.* **1**, 1506–1510 (2010).
5. Chudakov, D.M. *et al. Nat. Biotechnol.* **22**, 1435–1439 (2004).
6. Puthenveedu, M.A. & von Zastrow, M. *Cell* **127**, 113–124 (2006).

PSICQUIC and PSISCORE: accessing and scoring molecular interactions

To the Editor: To study proteins in the context of a cellular system, it is essential that the molecules with which a protein interacts are identified and the functional consequence of each interaction is understood. A plethora of resources now exist to capture molecular interaction data from the many laboratories generating such information, but whereas such databases are rich in information, the sheer number and variability of such databases constitutes a substantial challenge in both data access and quality assessment to the researchers interested in a specific biological domain.

Integrating data from these disparate resources remained a challenge until 2004, when the Human Proteome Organization Proteomics Standards Initiative (HUPO-PSI) released the PSI molecular interaction (MI) XML format, a community standard for the representation of molecular-interaction data. To concomitantly standardize annotation across the different databases, they also developed a controlled vocabulary enabling a detailed but consistent description of molecular interactions[1]. A simplified, standardized format for interaction data, the Molecular Interaction Tabular format (MITAB), is also available[2]. PSI-MI formats are now broadly accepted and widely implemented by over 30 databases and supported by key software tools.

The PSI-MI formats facilitate the integration of molecular interaction data from multiple sources, both by the user community and by dedicated software tools. However, users must still first collect data from each of the individual databases, which typically involves different queries at multiple websites or downloading data files from different web servers. Additionally, the retrieved data has then to be kept up to date with each release of the originating database. This challenge has led to the development of the PSI common query interface (PSICQUIC), a community standard for computational access to molecular-interaction data resources.

All data sources implementing PSICQUIC can be queried in the exact same way. Formulating the query once is sufficient to retrieve the relevant data from many interaction data sources. Independently published observations of an experimental system, curated by independent databases, are then integrated in response to a user query (**Fig. 1**). A PSICQUIC query can be a simple protein identifier or a complex construct using the syntax defined by the molecular interaction query language (MIQL) (**Supplementary Note 1**).

The existence of an open-source reference implementation for PSICQUIC allows the rapid setup of a local server for interaction data with limited effort. The PSICQUIC project site (http://psicquic.googlecode.com/) offers open-source client libraries and code examples,
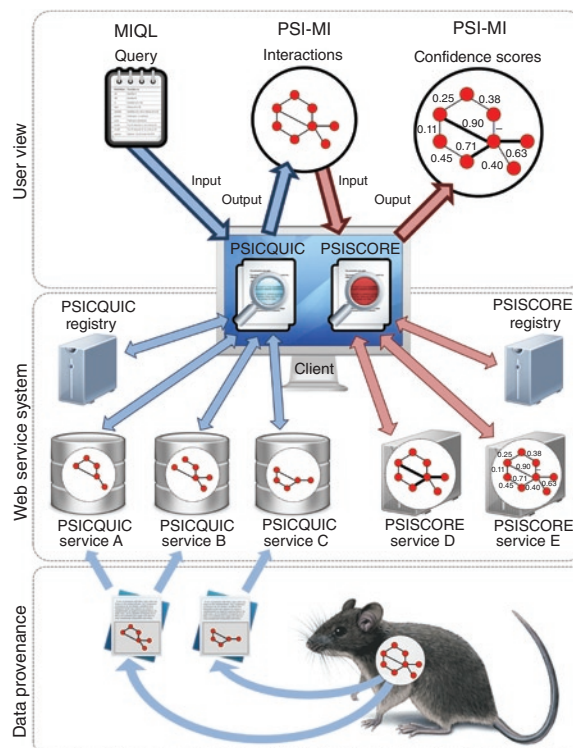

Figure 1 | PSICQUIC and PSISCORE architecture. A given biological system (sample) is observed by different experimental technologies, resulting in different publications reporting different, potentially partial, observations. A publication is potentially curated by more than one database. A PSICQUIC application sends a user query formulated in MIQL to all currently available PSICQUIC servers. Responses in unified PSI-MI format allow the PSICQUIC application to assemble a complete network view of the originally observed system. A given interaction network can be scored by multiple PSISCORE servers, each of them implementing one or more scoring methods (here symbolized by different line thickness (PSISCORE service D) and numbers (PSISCORE service E). The PSISCORE client application then presents the combined results to the user.