Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Single molecule fluorescence:

A tool to study dynamics and structure in single cells

Ivo Marten Glück

aus

Münster, Deutschland

2021

Abstract

Cells are the smallest, indepedently-functioning unit inside the human body. The interplay of different cell types makes up organs and organs form the organism. In order to understand the physiology of the body and how it is altered under disease conditions, it is necessary to understand how a single cell functions. Partially due to their small size, experiments on single cells are challenging. Over time scientists developed a large repertoire of sophisticated methods enabling these investigations. An important goal is to understand how molecules interact to build up the structure of the cell. Gaining insight into this question involved the development of microscopy as a very important tool since it enabled the visualization of processes inside the body and inside cells that are not observable with the naked eye. Another important aspect for understanding cell function is the dynamics taking place inside the cell. Although it is generally possible to directly image dynamics by microscopy, this is only possible when the dynamics are not too fast and if the spatial resolution of the microscope is sufficiently high to observe the process of interest. Correlation spectroscopy is an alternative approach, which enables insight into dynamics down to the single-molecule level in cases in which the complexes cannot be imaged directly.

Both methods, microscopy and spectroscopy, depend on a high contrast or signal-to-noise ratio to distinguish specific signal from background. Generally speaking, the intrinsic fluorescence inside cells is low which limits the application of the before mentioned methods in unmodified cells. It is the combination with fluorescence that turns microscopy and spectroscopy into important tools for the understanding of the cell. Nowadays, both fluorescence microscopy and fluorescence spectroscopy, are well-established techniques in cell research.

This thesis applies both approaches to study four different biological questions. Chapter 1 briefly introduces why these tools are important for biological research. Chapter 2 presents the basics of the methods. As such, it deals with the fundamentals of fluorescence as well the different types of fluorescence microscopy and spectroscopy applied in the experiments. Chapter 3 features the investigations of the dynamics of the foamy virus fusion process by spinning-disk confocal microscopy. Chapter 4 deals with the application of fluorescence super-resolution microscopy to study the structure of the ASC speck inside cells. The diffusion of the ASC protein inside cells measured by fluorescence spectroscopy is additionally discussed in the appendix A. Chapter 5 finally applies fluorescence spectroscopy to study the interaction of two proteins inside the nucleus of mouse embryonic stem cells. The thesis concludes with a summary of the obtained results and an outlook in chapter 6.

Contents

1 Introduction

Illnesses have always caused trouble for mankind. They have always been a strong driving force for humans to understand the processes occuring within the human body. In order to be able to counteract diseases, a fundermental step is to gain understanding of the processes happening in a healthy body. Generally speaking, research in this direction is denoted as fundermental research. Another step is to gain knowledge about the processes underlying diseases. Often, both levels of understanding go hand-in-hand and cannot be strictly separated. Research aiming to apply the gained knowledge to find concrete ways of fighting diseases is called applied or translational research.

This thesis provides contributions into both directions: fundamental as well as more application-driven research. The experiments conducted in the area of inflammasome research and the epigenetic regulation of pluripotency during embryo developement are classical contributions to fundamental research. Although the experiments concerning the fusion process of the foamy virus also have a strong fundamental aspect, the driving question behind them belongs to the area of translational or applied research. The common denominator of all conducted experiments is that they apply very specialized physical methods to address detailed biological questions.

Since both the biological/medical research questions and the methods to address them have become more and more specialized, modern biological research strongly benefits from collaboration between experts from both research areas in order to make progress. The research presented in this thesis is a typical example for such a collaborative work. It addresses biological questions raised by experts from the field with the latest advanced microscopy techniques developed and implemented by experts in the field of optics. These collaborations are especially productive if the researchers have an understanding of each others research areas. As I regard this approach of collaborative research as very fruitful, it was my aim to gain knowledge in both areas. Having obtained a biochemistry-focused background during my Bachelor's and Master's degree, I decided to do a method-focused PhD. Since microscopy always fascinated me, I chose this as the method to work with.

Microscopy has developed into a central tool in the biological and medical sciences as it enables the perception of features in a sample that would remain hidden otherwise. A central characteristic of microscopes is their ability to magnify. However, contrast is also needed in order to be able to distinguish features in an image. As good contrast is not given for every sample, different techniques to enhance the contrast of the image have been developed since the invention of optical microscopy. Important techniques include dark field microscopy, specific staining of features in a sample with dyes, and phase contrast as well as differential interference contrast (DIC) imaging. Among the different methods invented for microscopy, fluorescence is by far the most commonly used and today the term "Fluorescence Microscopy" denotes a large variety of microscopy techniques, and is a research area on its own. It is characterized by high contrast, specificity, sensitivity and selectivity combined with multiplexing capabilities and live cell compatibility. Enabling steps in the developement of this technique were the discovery of fluorescent dyes and proteins (section 2.2). In addition to imaging, the contrast enhancement by fluorescence has also been exploited by techniques that analyze the fluctuations in a fluorescent signal, which are summarized under the term "Fluorescence Fluctuation Spectroscopy" [1, 2]. The work presented in this thesis is based on the detection of fluorescence and exploits both modalities, imaging and spectroscopy (chapter 2), to study four biological problems on a single-cell level. As mentioned at the beginning of this chapter these problems are all linked to the aim of understanding the processes inside the human body and the aim of treating diseases. Chapter 3 studies the fusion of the foamy virus by

single virus tracking. The motivation for those investigations arises from the fact that the foamy virus is a potential vector for gene therapy. Being a retrovirus it integrates its genome into the host cell. This in combination with the fact that it causes apathogenic infections in humans renders it a potential vector for gene therapy. The ASC protein, which is addressed in chapter 4 and the appendix A of this thesis, plays an important role in the innate immune system and more specifically in the inflammasome pathway. This pathway plays a pivotal role in human health and disease. In order to be able to specifically interfere with this pathway a detailed understanding of how the involved proteins interact is crucial. The experiments presented as part of this thesis contribute to this aim by applying super-resolution fluorescence microscopy as well as fluorescence correlation spectroscopy. A detailed understanding of embryo developement is important for the understanding of what is altered under disease conditions. Chapter 5 contributes to this aim by investigating the interplay of two proteins in the nucleus of mouse embryonic stem cells using fluorescence correlation spectroscopy.

2 Theory and Methods

2.1 Fundamentals of fluorescence

When molecules exposed to light absorb a photon, electrons become excited into an electronically excited state. Relaxation of the molecule into the electronic ground state can be accompanied by emission of a photon. When this transition happens between two singlet states, the emitted light is called "fluorescence" as coined by Sir George Gabriel Stokes in 1852 [3] and the molecule is called a fluorophore. When the relaxation occurs between the triplet state T_1 and the singlet state S_0, the emitted light is called "phosphorescence".

Figure 2.1 depicts the typical cycle of fluorescence and phosphorescence in a Jablonski diagram proposed in 1933 by Erwin Jablonski [4]. It illustrates the possible electronic transitions within a molecule after it has absorbed a photon. Thick lines indicate electronic states and the space in between them is filled with vibrational and rotational states, indicated by thin lines. At room temperature, a molecule usually populates the lowest vibrational level of the ground state, S_0. Upon excitation with light of a specific wavelength, the electron is lifted from the electronic ground state into an electronically excited state, which is usually S_1 or S_2. Only a specific range of wavelengths lead to such a transition for a specific fluorophore and the transition occurs within femtoseconds (10^{-15}s). During this time, the relatively heavy nuclei do not move (Born-Oppenheimer approximation [5]), which leads to the fact that the electron ends up in a higher vibrational level of the electronically excited state (Franck-Condon principle). Transitions into different vibrational and rotational levels are possible whereby the energy of the absorpted photon corresponds to the energy difference associated with the respective transition. Transitions into the different energy levels are visible as absorption bands in the absorption spectrum, which are further broadend by thermal motion and interactions with the solution into a continuous spectrum. In the absorption or excitation spectrum (strictly speaking these are not the same) of a fluorophore, the relative absorption or the probability for absorption is plotted against the wavelength of the exciting light (Figure 2.2, dark gray spectrum). This plot can be obtained on a spectrophotometer by scanning through the excitation of a sample of the dissolved fluorophore and detecting the transmission at a fixed wavelength. Within picoseconds (10^{-12}s), the electron relaxes back to the lowest vibrational level of the electronically excited state releasing the corresponding energy in form of heat. This process is know as internal conversion or vibrational relaxation [6]. This is followed by a transition of the electron to a vibrationally excited state of the electronic ground state (here the Franck-Condon principle and the Born-Oppenheimer approximation again apply) within 1 - 10 ns (10^{-9}s), which can be accompanied by the emission of fluorescence photon. The average time a molecule spends in the excited state before emitting a fluorescence photon is called fluorescence lifetime. Fluorescence emission usually occurs from the vibrational ground state of the first excited state and hence is independent of the excitation wavelength (Kasha's rule, [6]). The electron again undergoes internal conversion by quickly relaxing to the vibrational ground state of S_0 and releasing the energy as heat. Since relaxation into all kinds vibrational and rotational levels of the electronic ground state are possible and thermal motion again leads to a further broadening, fluorescence emission, analogously to absorption, also forms a broad spectrum plotted as relative fluorescence or probability of fluorescence emission against emission wavelength. This spectrum can be measured on a spectrophotometer by keeping the excitation wavelength of a fluorophore solution fixed and scanning the detection wavelength (Figure 2.2, light gray spectrum). For many fluorophores, the spacing between the vibrational levels of the ground state is similar to those of the first excited state resulting in the absorption and emission spectrum being mirror images of one another. This completes the fluorescence cycle and the electron can be excited again. A fluorophore can undergo several hundered of thousands of

excitation cycles before it photobleaches, a transition that leads to alterations of covalent bonds within the molecule and renders it permanently unable to fluoresce. Since during the process of internal conversion, some of the exciting energy is converted into heat, through collisions of the excited molecule with the surounding solvent molecules, fluorescence is less energetic and therefore of longer wavelength, compared to the excitation light. This phenomenon was first described by Sir George Gabriel Stokes and is accordingly denoted as the Stokes shift [3] (Figure 2.2). There is also the possibility of a non-radiative decay from S_1 to S_0 in which case the energy is lost in collisions with surounding solvent molecules and no fluorescence is emitted. Alternatively, the electron can undergo a transition from the ground state of S_1 into the triplet state T_1, which is of slightly lower energy, a process known as intersystem crossing occuring within ~ 10 ns. This transition is theoretically forbidden and associated with spin inversion. Relaxation from the lowest energy level of the triplet state T_1 into a vibrationally excited state of the singlet state S_0 can also be accompanied by the emission of a photon. In this case, the radiation is called phosphorescence and occurs within ms to mins to up to hours since this process is also theoretically forbidden and also requires spin inversion and therefore is very unlikely. The following section will introduce the types of fluorophores used in this thesis, namely fluorescent proteins and fluorescent dyes.
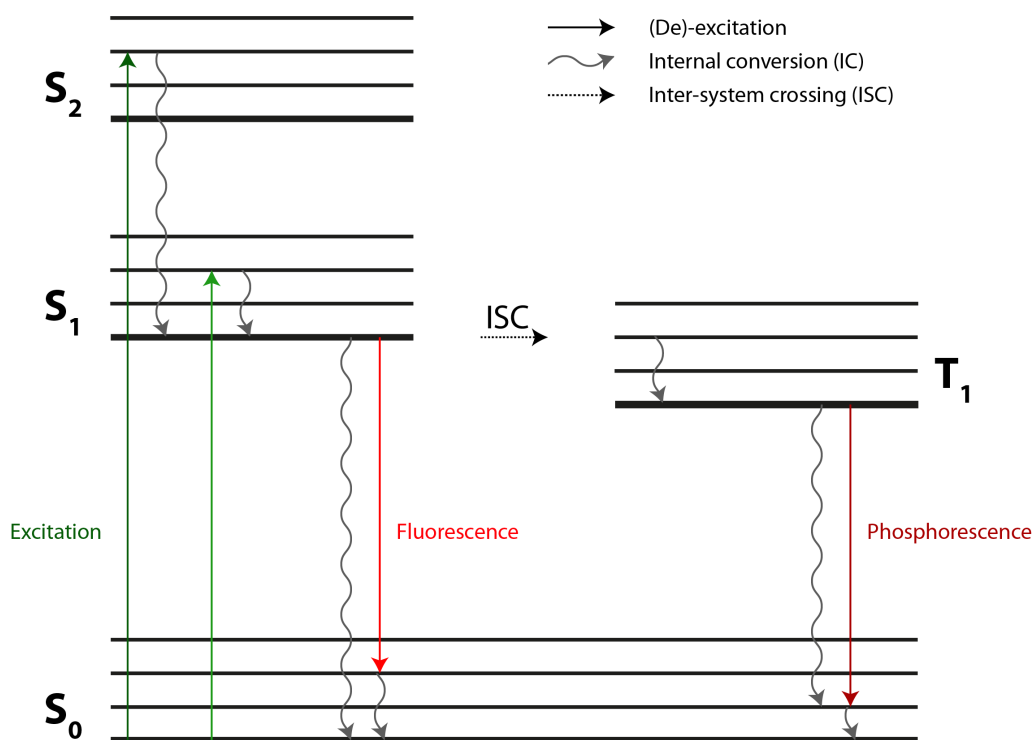
Figure 2.1: **Electronic transitions of a molecule exposed to light illustrated in a Jablonski diagram**
At room temperature, a molecule usually populates the singlet ground state S_0. Exposure to light can lead to its transition into a vibrational state (thin horizontal lines) of a singlet excited state, which is, in most cases, S_1 or S_2 (green arrows). From there, an excited molecule will first relax into the lowest vibrational level of the first electronically excited state releasing the associated energy as heat. From here, the molecule can transition into a vibrationally excited state of the electronic ground state under emission of a fluorescence photon (light red arrow) followed by relaxation into the electronic ground state again associated with the emission of heat. Alternatively, the molecule can undergo intersystem crossing (ISC) into a vibrationally excited triplet state T_1, from where it can relax first into the electronic ground state of T_1. From there, it undergoes a transition into a vibrationally excited state of the singlet ground state S_0 under emission of a phosphorescence photon (dark red arrow). From there, it relaxes down to the electronic ground state. The molecule can also relax from the ground state of S_1 and T_1 to S_0 without the emission of photons releasing the associated energy as heat. Gray wavy arrows illustrate internal conversion transitions which are associated with the emission of the corresponding energy as heat. Adapted from [4].
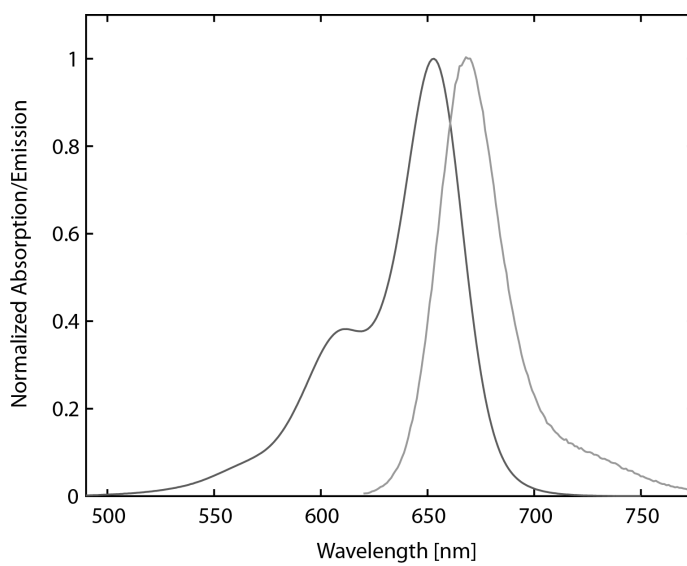
Figure 2.2: **Absorption and emission spectrum of Alexa Fluor 647** [®]
Generally, the absorption (dark gray) and emission (light gray) spectra of fluorophores are composed of distinct absorption and emission bands that are further broadend by thermal motion into a continuous spectrum.

2.2 Fluorescent dyes and fluorescent proteins

There are different types of fluorophores such as organic dyes, nanoparticles (e.g. quantum dots) and fluorescent proteins. Important characteristics of a fluorophore are its excitation and emission spectrum, fluorescence lifetime, fluorescence quantum yield, extinction coefficient and brightness. The term fluorescence lifetime describes the average time after excitation of a fluorophore at which a photon is emitted from the fluorophore and usually occurs on the nanosecond timescale. The quantum yield is the ratio of the number of emitted photons to the number of absorbed photons. The extinction coefficient describes the probability of a fluorophore to absorb light of a specific wavelength, is usually given for the wavelength of maximum absorption and has the units of $cm^{-1}M^{-1}$. The molecular brightness is the product of the fluorescence quantum yield and the extinction coefficient. It describes the number of fluorescence photons emitted per molecule per second. In the experiments done for this thesis, fluorescent proteins (FPs) and fluorescent organic dyes were used. Hence these fluorophores will be addressed in greater detail below.

The first fluorescent protein, the green fluorescent protein (GFP), was discovered and extracted from the jellyfish *Aequorea victoria* by Shimomura and Johnson in 1962 [7]. Martin Chalfie first showed that GFP can be utilized as a fluorescence tag in living cells [8] and Roger Y. Tsien developed GFP-like proteins with different spectral properties [9, 10]. The three scientists were awarded the Nobel Prize in Chemistry in 2008 for their achievements. Today, hundreds of FPs in different colors exist [11]. GFP-like proteins are generally 238 amino acids in length and characterized by a β-barrel made up of 11 β-sheets with an α-helix running down the center of the protein (Figure 2.3A). The α-helix is disrupted by the fluorescent chromophore located at the center of the protein and is made from three amino acids. Changes in this chromophore lead to the characteristic variants of the protein with altered excitation and emission properties. The diameter of fluorescent proteins is typically ~ 3 nm. Next to the above mentioned general properties of fluorophores, four other important parameters of fluorescent proteins are their oligomeric state, their folding efficiency, their maturation time at 37°C and their photostability. Most commonly FPs are expressed inside cells as a fusion protein with a protein of interest to follow its localization and dynamics in living cells. Alternatively, the fluorescent protein can be exploited as a tag which can be specifically recognized by a protein binder such as a nanobody. Both approaches were utilized in the experiments conducted for this thesis and the following fluorescent proteins were deployed: enhanced green fluorescent protein (eGFP) [12], an optimized version of GFP, TagRFP [13], which was engineered on the basis of a GFP-like protein from the sea anemone *Entacmaea quadricolor* and mScarlet [14], a synthetic fluorescent protein.

Organic dyes are mostly planar molecules of about 0.5 - 1 nm in size that exhibit a coordinated π-electron system. Fluorescein, which was first synthesized in 1871 by Adolf von Baeyer, is a well-known fluorescent dye and von Baeyer got awarded the Nobel Prize in Chemistry in 1905 for his achievements. The larger the delocalized π-electron system, the higher the absorption and emission wavelength of an organic dye. Based on the core scaffold, different dye families are distinguished which each constitute a number of dyes within a certain wavelength range. Important families, sorted according to their central wavelength from low to high, are: coumarins, BODIPY dyes, fluoresceins, rhodamines and cyanines. Today there is a large variety of fluorescent organic dyes available [15]. Compared to fluorescent proteins, fluorescent dyes generally have a higher quantum yield and extinction coefficient resulting in a higher molecular brightness and exhibit a much higher photostability. A common application of organic dyes is their conjugation to protein binders such as antibodies or nanobodies for applications in immunofluorescence stainings. During this thesis, the dye Alexa Fluor 647®, a cyanine dye, was used (Figure 2.3B).
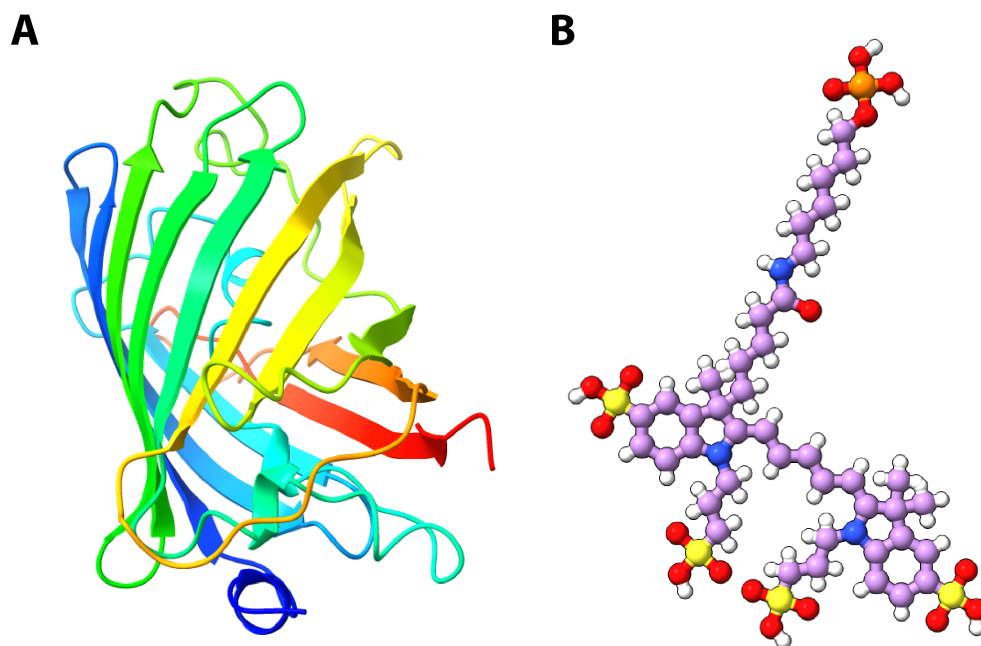
Figure 2.3: **The structure of the two types of fluorophores used in the experiments described in this thesis: a fluorescent protein and a fluorescent dye**
A) Crystal structure of the S65T variant of the green fluorescent protein (as published in [16]; PDB entry: 1EMA)
B) Chemical structure of Alexa Fluor 647®; purple: carbon; white: hydrogen, red: oxygen, yellow: sulfur, orange: phosphorus, blue: nitrogen; (Source: National Center for Biotechnology Information (2021). PubChem Compound Summary for CID 102227060, Alexa Fluor 647. Retrieved April 19, 2021 from https://pubchem.ncbi.nlm.nih.gov/compound/Alexa-Fluor-647)
The structures are not shown to scale.

2.3 Fluorescence microscopy

After the the pioneering work on fluorescence by John Frederick William Herschel in 1845 [17] and George Gabriel Stokes in 1852 [3] the potential of fluorescence for optical microscopy was also realized by Hermann Helmholtz [18] and others. It was August Köhler who, in 1904, for the first time, observed fluorescence with a light microscope and while mainly regarding it as a disadvantage in his experiments, he also anticipated potential benefits for optical microscopy [19]. The first fluorescence microscopes were built between 1911 and 1913 by Otto Heimstädt, Carl Reichert [20], Heinrich Lehmann [21] and Carl Zeiss. In 1929, Eillinger and Hirt invented the intravital microscope, which was based on fluorescence [22]. Another enabling step in the history of fluorescence microscopy was the invention of a procedure to conjugate antibodies with a fluorophore and its application for immunofluorescence by Albert Hewett Coons in 1942 [23]. Over time, fluorescence microscopy was refined and optimized and the following section will introduce the main building blocks of a modern fluorescence microscope. Furthermore, the diffraction-limited resolution in a classical optical microscope is addressed.

2.3.1 General building blocks of a fluorescence microscope

Generally two designs of optical microscopes can be distinguished: the finite and the infinity-corrected optical microscope (Figure 2.4). The finite design was the original way in which microscopes were constructed. Here the specimen is placed slightly outside the front focal plane of the objective (which is typically approximated by a thin lens), which leads to the formation of an image at a fixed distance on the opposite side of the objective. Modern microscopes are, however, constructed as an infinity-corrected optical system. Here, the sample is placed in the front focal plane of the objective and consequently the image is focused at infinity i.e. the light rays from the object are parallel upon leaving the objective. A tube lens is used to image the specimen onto the detector. This design has three advantages: First, the fact that the light rays are parallel behind the objective (infinity space), in principle, allows the tube lens to be placed at an arbitrary distance from the objective and basically an unlimited amount of flat optical elements can be inserted into the beam path without distorting the image. Secondly, since the final magnification of the system depends on the focal length of the objective and the tube lens (equation 1), an additional magnification can be achieved by changing the tube lens. Third, focussing can be achieved by moving the objective. All microscope setups used in this thesis are constructed as an infinity optical system. They all share some basic building blocks, which will be described in the following paragraphs on the basis of an inverted epi-illumination microscope, the geometry in which all microscope setups applied are constructed (Figure 2.5).

A

Objective   Tube lens

B

$f_{obj.}$ Infinity space f_{tl}

Front focal plane   Primary image plane
of objective
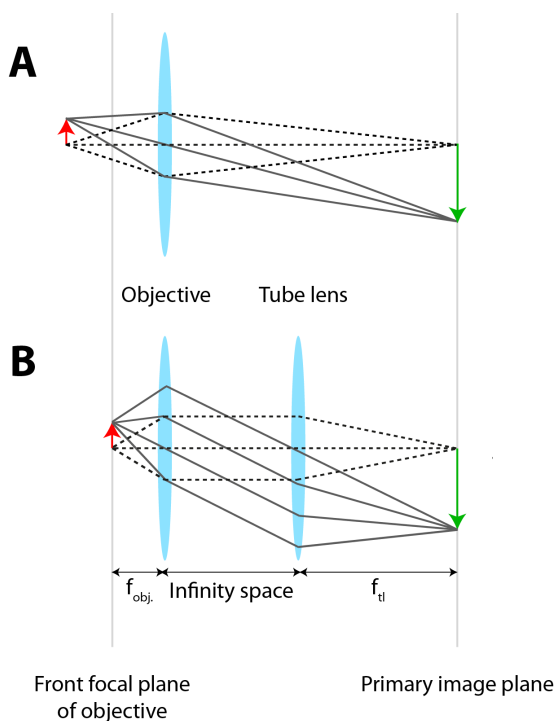
Figure 2.4: **Schematic of a finite and an infinity-corrected optical microscope**
A) In the finite optical microscope, the object (red arrow) is placed at a fixed distance away from the front focal plane of the objective leading to the formation of a magnified image (green arrow) at a fixed distance behind the objective.
B) In the infinity-corrected optical microscope, the object (red arrow) is placed in the front focal plane of the objective leading to the light being focused at infinity (collimated light rays) behind the objective. A tube lens refocuses those rays into an image (green arrow).

In an inverted epi-illumination microscope, the objective is located below the specimen and fluorescence emission is collected by the same objective used for excitation. Light emitted by the sample in the front focal plane of the objective will be collected by the objective and the rays emitting from the sample will be parallel behind the objective. They are then focused by the tube lens into an image on the detector in the primary image plane. The magnification of this system is determined by:

$$M = \frac{f_{tl}}{f_{obj}} \tag{1}$$

with f_{tl} and f_{obj} being the focal distance of the tube lens and objective, respectively. Note that M is the magnification specified on an objective and, therefore, is always given in combination with a tube lens of certain focal length. Although the tube lens can, in essence, be located at an arbitrary distance from the objective, in practice, it is most commonly placed so that the backfocal plane of the objective and the focal plane of the tube lens coincide.

As explained in section 2.1, a molecule needs to be excited by light in order for fluorescence to occur. As such, every fluorescence microscope is equipped with a light source. Different light sources are used in modern fluorescence microscopy such as lamps (e.g. Mercury arc or Xenon lamps), light emitting diodes (LEDs) and lasers. Lamps have the general disadvantages of needing excitation filters to select a certain excitation wavelength, the emitted light intensity being strongly dependent on

the selected wavelength and being generally unstable. LEDs are a new type of light source in fluorescence microscopy. They are characterized by a more stable emission intensity compared to lamps and are cheaper than lasers. However, the emitted spectra are broader than those of lasers making the use of excitation filters again necessary when sharp excitation spectra are needed. The microscope setups applied in this thesis use lasers as excitation source, which emit collimated light of (in theory) a single wavelength. This is generally beneficial for separating the excitation light from the fluorescence emission and allows a larger proportion of the emission spectrum to be captured without residual excitation light being reflected onto the detector. The installed lasers can be continuous wave or pulsed. A general characteristic of pulsed lasers is a very high peak output power and the spectrum they emit is usually broader than the one of continuous-wave lasers. All microscope setups applied during the course of this thesis used continuous wave lasers as the excitation source except for the laser scanning microscope with which time-correlated single photon counting was performed. In this setup pulsed lasers are necessary.

When multiple excitation sources are used in a setup, their emission is overlayed into a single beam using the appropriate dichroic mirrors and, in some cases, an optical fibre. Using an optical fibre has the advantage that the excitation can in essence be transfered to an arbitrary location within the optical system. A single mode fibre additionally offers the advantage of acting as a spatial filter and a Gaussian beam profile is emitted from the fibre. A drawback of fibre coupling is the, in general, quiet significant loss in laser intensity with the efficiency of the coupling being dependent on the profile of the laser beam.

The laser emission is reflected by a dichroic mirror into the objective above which the fluorescent sample is located. The emitted fluorescence is collected by the same objective (i.e. epi-) from which it is collimated. As fluorescence is red-shifted (i.e. Stokes shift) it can be separated from the excitation light by the dichroic mirror. Afterwards, it is focused onto the detector by the tube lens. The ability of the dichroic mirror to separate excitation and fluorescence is decisive since only a fraction of the exciting laser light leads to the emission of a fluorescence photon. Hence, fluorescence is much weaker in intensity than the excitation light. Optical elements might be inserted into the beam path in front of the tube lens, e.g. the emission might be further separated according wavelength and/or polarization before being focused onto the detector. Residual stray or reflected laser light is mostly blocked by an emission filter in front of the detector which only transmits a wavelength range which is usually centered around the wavelength of maximum fluorescence emission of the fluorophore to be detected. It may still happen that excitation light passes through all filters and reaches the detector. There are two main reasons for this: First, the light emitted by a laser is never of a single wavelength and secondly the optical density of the filters at the laser wavelength might be insufficient to fully block all the reflected laser light. In these cases, either a clean-up filter, which transmits only a range of ~ 10 nm around the central emission wavelength of the laser, can be installed in front of the laser or the residual excitation light can be blocked from reaching the detector by installation of a notch filter in the emission path which transmits all wavelengths except for the one emitted by the laser.

There are generally two types of detectors used in fluorescence microscopy: single point and array detectors of which by far the mostly applied ones are cameras. Important characteristics of a detector are its quantum efficiency, which describes the percentage of incident photons that are converted into photoelectrons, their response time, meaning how fast incident photons can be registered, and the amount of background counts. The most important single point detectors are avalanche photodiodes (APDs) and photomultiplier tubes (PMTs). Cameras are generally characterized by higher quantum efficiency while single point detectors have a shorter response time. Figure 2.5 depicts a scheme of the basic fluorescence microscope setup described.
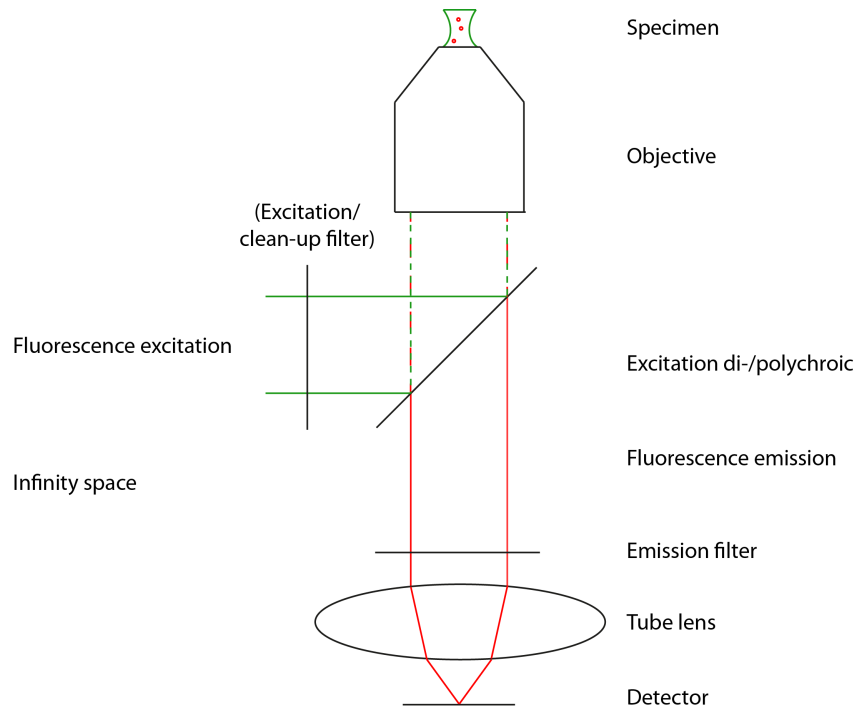
Figure 2.5: **Schematic of the main building blocks of a basic modern epi-fluorescence microscope**
The excitation light (green) is reflected into the objective by an excitation di-/polychroic mirror. The spectrum of the excitation source can be narrowed by installation of an excitation/clean-up bandpass filter when required. The objective delivers the excitiation light to the sample and collects the fluorescence light emitted by the sample (red). The fluorescence is transmitted by the excitation di-/polychroic mirror and focused onto the detector by the tube lens. Residual stray or reflected excitation light are blocked from reaching the detector by an emission bandpass filter installed in the emission path.

2.3.2 The diffraction limit of light microscopes

An important characteristic of a microscope is its resolving power. It describes the minimum distance between two points necessary to distinguish them in the image produced by the microscope. Intuitively, one would assume that higher magnification results in increased resolution; however, there is a limit in the achievable resolution of a microscope, which cannot be overcome by increasing the magnification. In 1873, Ernst Abbe postulated that the lateral resolution d_{xy} of a light microscope is limited to

$$d_{xy} = \frac{\lambda}{2\,NA} \tag{2}$$

where λ denotes the wavelength of the light coming from the object and *NA* stands for the numerical aperture of the objective defined as

$$NA = n\,sin\,\alpha \tag{3}$$

where *n* denotes the refractive index of the medium between the objective and the sample and α is the half angle of the front

aperture of the objective [24].

Already in 1835, Georg Airy made the observation that an infinitesimally small point source is depicted in a microscope as a bright central peak sourrounded by rings gradually decreasing in intensity [25]. This pattern is known as the "Airy disk". Based on this observation, Lord Rayleigh, in 1879, refined Abbe's resolution limit as the distance d_{xy} between the central peak and the first minimum of the Airy disk [26], which is given by:

$$d_{xy} = \frac{0.61\,\lambda}{NA} \tag{4}$$

The central peak of the Airy disk can be approximated with a Gaussian function and the distance d_{xy} is approximately equal to the full width of the Gaussian at half the maximum intensity.

For two points in the image plane this means that the peak maximum of the first point needs at least to coincide with the first minimum of the second point in order for the two points to be distinguishable (Figure 2.6). When imaging GFP, which has an emission maximum of 510 nm, with a 1.4 NA oil immersion objective, this results in a maximum lateral resolution of ~ 222 nm. This is, in some sense, an arbitrary definition because two points might be distinguishable even when they are located in closer proximity. Other definitions of resolution exist, but, to date, the Rayleigh criterion is the most commenly used definition to describe the resolution of a light microscope.
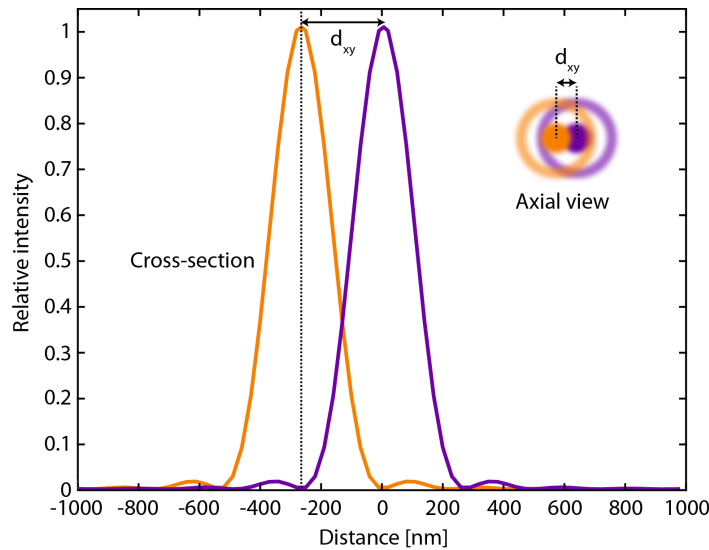
Figure 2.6: **Illustration of the Rayleigh criterion for the resolution limit of a light microscope**
A conventional fluorescence microscope images a point source, such as a fluorescent molecule, as an Airy disk pattern, a central bright peak sourounded by concentric rings of decreasing maximum intensity. A schematic of the Airy disk pattern in 1D (left) and 2D (right) is shown. According to the Rayleigh criterion, the minimum distance two point sources have to be separated in the sample plane in order to be resolvable in the image plane corresponds to the distance d_{xy} between the central maximum and the first minimum of the Airy disk pattern.

However, the question of what causes a point source to appear as an Airy disk in the image plane remains. The Airy disk pattern and, therefore, the resolution limit arises from diffraction and is best explained using Huygens wavelets, a concept introduced by Christiaan Huygens (1629-1695). Huygens postulated that light travels as waves characterized by a frequency,

wavelength and amplitude. He distinguished plane waves, for which the wavefront can be represented as a straight plane propagating in space, and spherical waves, which have a spherical wavefront propagating in space, and originate from or converge into a single point. Huygens postulated that the elemental units of these waves are small point sources of light sending out light in all directions, which he called wavelets. Figure 2.7A depicts how wavelets make up plane and spherical waves.
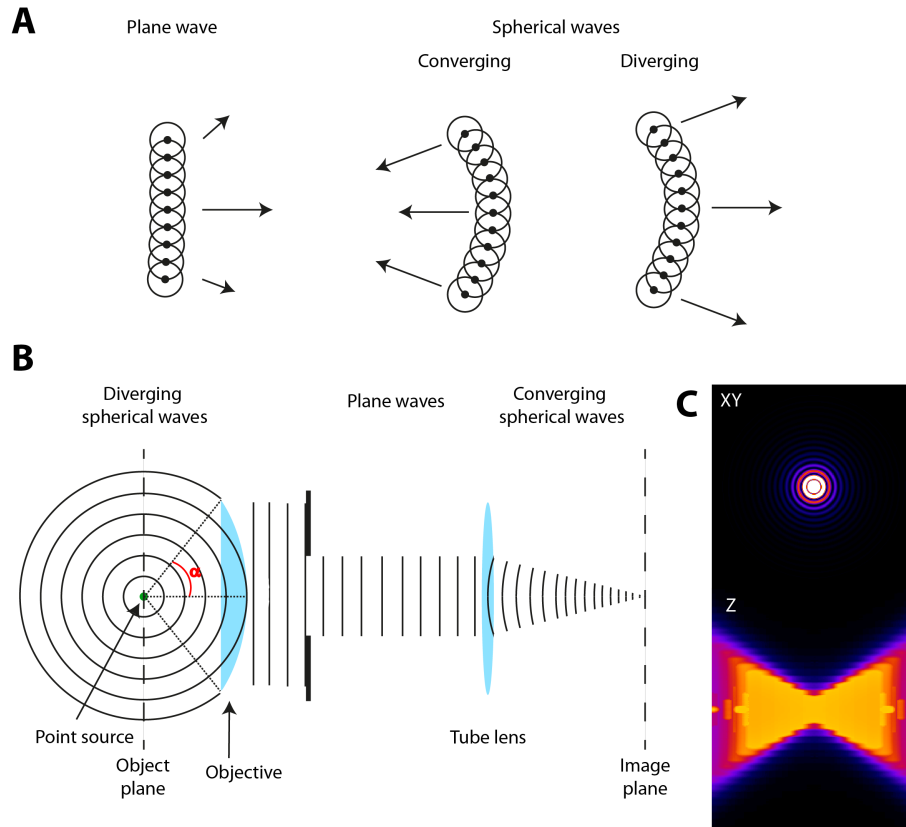
Figure 2.7: **Waveoptics-based graphical illustration of the diffraction limit of light microscopes**
A) An illustration of how wavelets (small point sources of light sending out light in all directions) make up plane waves and converging or diverging spherical waves is shown. The light send out by the composing wavelets interferes in such a way that the overall wavefront propagates as a plane (left) or a spherical wave which is converging (center) or diverging (left).
B) An illustration of how a point source is imaged by an infinity-corrected optical microscope. The point source in the object plane sends out a spherical wave of light, which is collected by the objective. The collected proportion depends on the collection half angle α of the objective. The objective converts the light into a plane wave. The plane wave leaving the back aperture of the objective is collected by the tube lens, which converts it into a converging spherical wave. Interference of the wavelets making up the wave results in the characteristic point spread function intensity pattern in the image plane
C) Simulated point spread function in the *xy*-plane (Airy disk) (top) and in the *z*-dimension (bottom); Generated using the "PSF Generator" Fiji plugin [27]

Figure 2.7B graphically illustrates how this concept can be used to explain the Airy disk pattern observed in a light microscope. A point-like object observed under a light microscope, for example, a fluorescent molecule, can be approximated as a point source of light sending out a spherical wave made up of wavelets. In a microscope, a proportion of this wave will be collected by the objective, with a high numerical aperture objective collecting a higher proportion than a low numerical aperture objective. The characteristic of an objective to collect light emitted from a point source and collimating it is nothing other than converting a diverging spherical wave into a plane wave. This plane wave will travel towards and will be collected by the tube lens which refocuses it i.e. converts it into a converging spherical wave. The composing wavelets will interfere in the image plane on the detector, which registers the interference pattern as intensity. The intensity of a wave is obtained by calculating the square of its amplitude over at least a single wavelength. Areas of constructive and destructive interference result in areas of high and low intensity, respectively. Constructing the intensity pattern formed by all wavelets emitted from a single source in 2D results in the Airy pattern (Figure 2.7C, top).

Interference is not limited to two dimensions but also occurs along the optical axis. The resulting intensity pattern can be derived analogously to the two dimensional case. It results in a central hour glass-shaped cone of high intensity surrounded by increasingly dark areas when looking at the pattern in 3D. The axial resolution, denoted d_z, is analogously to the above mentioned lateral resolution given as the distance from the center of the pattern to the first minimum and defined as

$$d_z = \frac{2n\lambda}{NA^2} \tag{5}$$

in which n denotes the refractive index of the medium between the objective and the sample, λ the wavelength of the light emitted by the sample and NA the numerical aperture of the objective. Comparing equations 4 and 5 it immediately becomes obvious that the resolution of a microscope along the z-axis is much worse compared to the xy-plane. Figure 2.7C bottom shows the diffraction pattern along z. Analogously to the above described lateral case, this results in a maximum axial resolution of ~ 780 nm when imaging GFP.

The fact that a high NA objective collects a larger cone of light compared to a low NA objective leads to a narrower diffraction pattern and thereby to higher resolution. Since this diffraction pattern shows how a point source is spread out in the image plane, it is called the "point spread function" (PSF) of the microscope.

According to the Nyquist-Shannon sampling theorem [28], one needs to sample a signal with at least twice the frequency of the finest detail one wants to resolve. For the resolution in microscopy, this means that, to reach a theoretical resolution of 220 nm the pixel size in the image should be 110 nm or smaller. For example if a 100x objective is used in combination with a camera, its physical pixel size should be 11 μm or smaller. However to image structures within a sample with a certain resolution also the labeling density of the structure of interest plays a role. In this context the term "structural resolution" has been coined. Applying the Nyquist-Shannon sampling theorem here means that the structure of interest needs to be decorated by a fluorophore at least every 110 nm to obain a structural resolution of 220 nm. The structural resolution, i.e. the labeling density of the sample, becomes especially important when imaging beyond the diffraction-limited resolution, as will be addressed in section 2.5 which covers a microscope technique capable of reaching such resolution.

Next to the resolution, it is often the contrast or signal-to-noise which limits the distinguishable details in a microscopy image. It was with this in mind that different ways of illuminating the sample and detecting the signal based on the original widefield implementation were developed over time. For the work presented in this thesis, different techniques were applied: namely widefield (section 2.4.1), total internal reflection (TIRF) microscopy and its adaptation highly inclined and laminated optical

(HILO) sheet microscopy (section 2.4.2), laser scanning confocal (section 2.4.3) as well as its further development spinning-disk confocal microscopy (section 2.4.4). The following sections will describe the basics of these microscopy techniques and address, where applicable, the particularities of the applied microscope setups.

2.4   Diffraction-limited fluorescence microscopy modalities

2.4.1   Widefield microscopy

Most modern widefield microscopes use continuous wave lasers as excitation sources. After being expanded, the laser beams are combined by the appropriate dichroic mirrors and reflected into the objective by a di- or polychroic mirror. A lens introduced into the beam path focuses the light onto the backfocal plane of the objective resulting in collimated light leaving the objective and illuminating the sample. Being built in epi-geometry, fluorescence emitted by the sample is collected by the same objective, transmitted by the di- or polychroic mirror and focused onto a camera for detection. The image captured by the camera is diffraction-limited according to the principles described in section 2.3.2. Figure 2.8 depicts a scheme of a classical modern widefield setup.
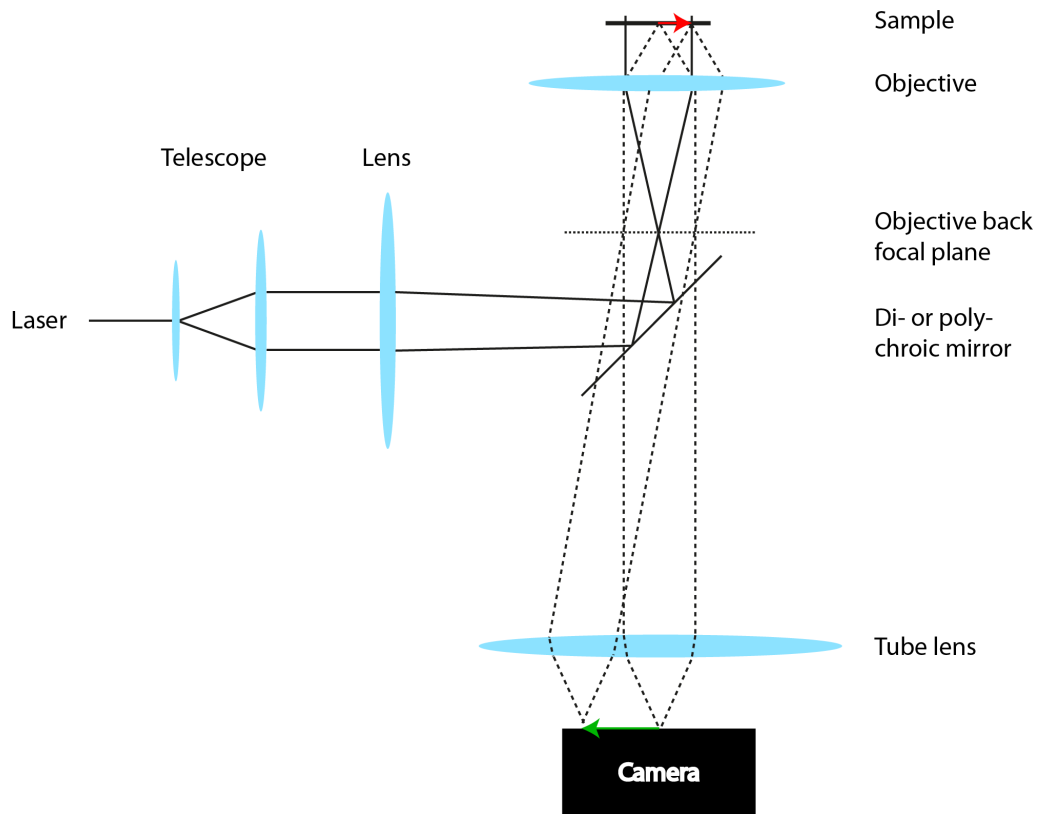
Figure 2.8: **Schematic of a classical widefield setup**
The laser beam used for excitation (solid line) is expanded by a telescope and focused onto the center of the back aperture of the objective resulting in a collimated beam of light exiting the objective and illuminating the sample. Fluorescence emitted by the sample (dashed line) is collected by the objective and imaged onto a camera by the tube lens. The magnification of the image compared to the object depends on the focal lengths of the objective and the tube lens.

Due to the Gaussian profile of the laser beam, the illumination profile reaching the sample in such a setup is also Gaussian shaped. This can be disadvantagous for techniques strongly relying on the power density reaching the sample such as direct stochastic optical reconstruction microscopy (dSTORM) applied in this thesis and described in section 2.5.1. For this reason, the widefield illumination setup applied for the experiments reported here was optimized for homogenous illumination [29] (Figure 2.9).

The setup uses continuous-wave lasers of 405, 642 and 750 nm wavelength as the excitation source. The illumination is optimized compared to the beam path described above through beam-shaping of the excitation lasers by introducing two microlens arrays (MLA) and a rotational diffuser located in between the lenses of a telescope into the excitation beam path. This results in a homogenous flat-field illumination in the sample plane for which the inventors coined the term flat illumination for field-independent imaging (FIFI). This design enables recordings of evenly illuminated images 100 μm x 100 μm in size. The focus stability is maintained by a home-built perfect focus system based on a 850 nm laser and the fluorescence signal is recorded on a scientific complementary metal–oxide–semiconductor (sCMOS) camera.
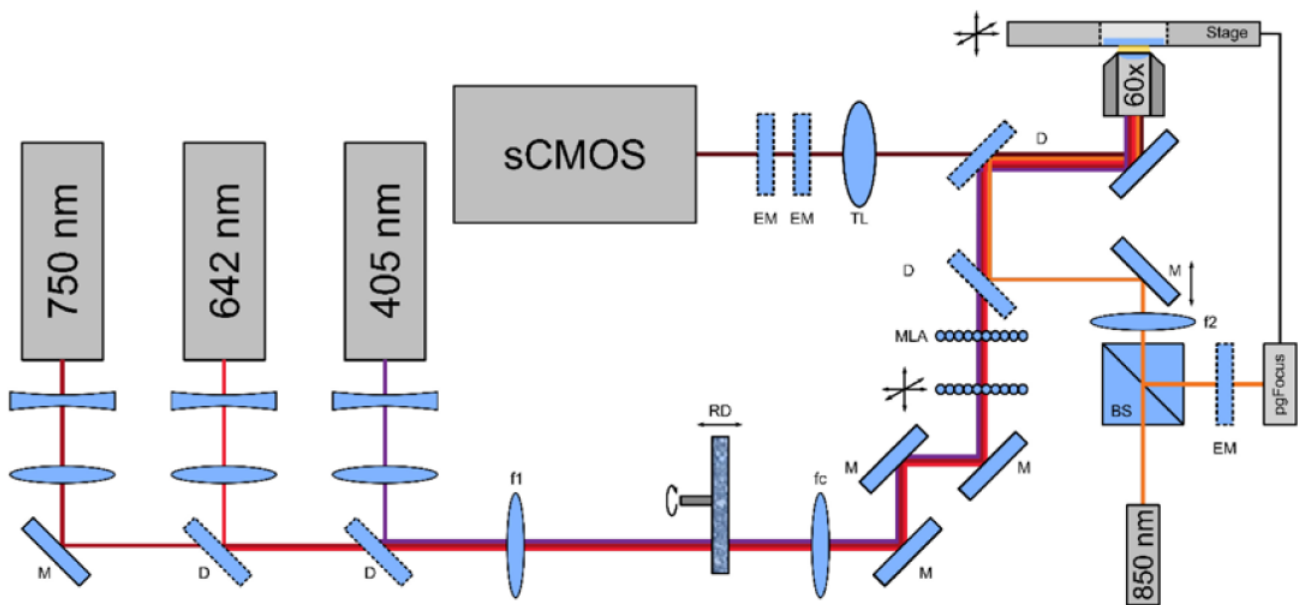
Figure 2.9: **Simplified schematic of the applied flat-field-optimized widefield setup**
Continuous-wave lasers of 405 nm, 642 nm and 750 nm wavelength are expanded by a beam expander and subsequently ovelayed using a mirror (M) and dichroic mirrors (D). The laser light is then introduced into a telescope. In between the lenses of the telescope (f1, fc) a rotating diffus (RD) is located. After the telescope the beam is directed onto the two microlens arrays (MLA) by mirrors. The optimized beam is then introduced into the objective. The sample is illuminated by a flat-field-optimized laser beam. Fluorescence is collected by the objective and imaged onto the sCMOS camera of the system by a tube lens (TL). In front of the camera emission filters (EM) are mounted. Axial drift during the measurement is counteracted by a home-built perfect focus system operating with a 850 nm continuous-wave laser. f2: lens; BS: beam splitter.
From [29]. Reprinted with permission from Nature Publishing Group.

2.4.2   Total internal reflection and highly inclined and laminated optical sheet microscopy

The above described illumination modality of widefield microscopy harbours the major disadvantage that all fluorophores located within the illumination cone above and below the image plane are excited and their out-of-focus fluorescence is also detected on the camera, which leads to a high background in the resulting image. This drawback was addressed by E. J. Ambrose in 1961 [30] and later by Daniel Axelrod [31, 32], who invented and developed total internal reflection fluorescence (TIRF) mircroscopy. Here, the incoming laser beam hits the interface between the microscope slide and the aqueous solution of the sample under such an angle that it is totally internally reflected. The phenomenon of total internal reflection can be explained by Snell's law which describes the bending of light, an electromagnetic wave, at the interface of two media with different refractive indices. Snell's law is given as:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \tag{6}$$

in which n_1 is the refractive index of the medium in which the light is approaching the interface and n_2 the refractive index of the medium on the opposite side of the interface. θ_1 and θ_2 denote the angles between the incoming and the outgoing beam and the normal on the interface, respectively. In the case of TIRF microscopy it is the glass-water interface located at the upper side of the microscope slide where the light is bent (Figure 2.10). In this case n_1 is given for borosilicate glass (n_1 = 1.52) and n_2 for water (n_2 = 1.33) (which is a reasonable approximation e.g. for the cytoplasm of a cell located on the glass slide). The angle θ_1 for which θ_2 reaches 90° is called the critical angle θ_C. If θ_1 is below this angle the light is refracted into the medium on the opposite site of the interface (Figure 2.10, red line).
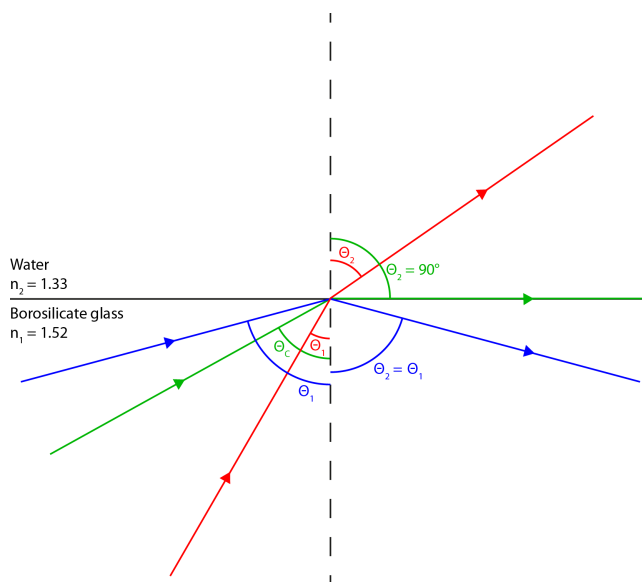
Figure 2.10: **An illustration of Snell's law**
An illustration of how light rays hitting an interface between borosilicate glass and water are bent. If a beam of light hits the interface at an angle smaller than the critical angle it is refracted into the water (red line). If a beam of light hits the interface at the critical angle it is refracted along the interface (green line). If a beam of light hits the interface at an angle above the critical angle the light gets reflected into the glass at the interface (blue line).

If θ_1 takes the value of the critical angle the following relationship follows from Snell's law:

$$\arcsin(\theta_C) = \frac{n_2}{n_1} \tag{7}$$

For the above mentioned interface between borosilicate glass and water the critical angle is 61.04° (Figure 2.10, green line). If θ_1 takes values above the critical angle Snell's law cannot be fullfilled anymore and the light is reflected back into the glass (Figure 2.10, blue line) which is denoted as total internal reflection (TIRF). TIRF leads to an evanescent wave of light emerging on the opposite site of the interface and the intensity of the emerging light decays exponentially with the distance from the interface.

In a TIRF microscope collimated laser light hits the upper interface of the microscope glass silde at an angle above the critical angle. The resulting evanescent wave of laser excitation above the microscope slide reaches only about 100 nm into the sample. The exact penetration depth, d, can be calculated as follows:

$$d = \frac{\lambda}{4\pi \sqrt{n_1^2 \sin^2(\theta) - n_2^2}} \tag{8}$$

where λ is the wavelength of the exciting laser light. For example, an excitation wavelength of 561 nm hitting the interface at an angle of 65° would result in a penetration depth of 124.4 nm. Only fluorophores located within roughly that distance are excited since the laser intensity decays exponentially as mentioned above. Fluorescence is collected in the same way as in widefield microscopy but the resulting image has higher contrast and signal-to-noise since the restricted excitation reduces the background fluorescence. Different ways to achieve total internal reflection illumination have been implemented. By far, the most applied modality, and also the one which has been exploited for this thesis, is objective-type TIRF. Here, the laser beam is focused onto the backfocal plane of the objective, as in widefield microscopy, but is offset from the optical axis and hits the objective at the side of the back aperture (Figure 2.11). The offset results in the excitation beam leaving the objective under an oblique angle. In order to reach total internal reflection, a high numerical aperture objective needs to be applied and there are special objectives available for TIRF illumination. It should be noted that the excitation beam is reflected off the surface and collected by the objective and needs to be prevented from hitting the detector by using appropriate emission filters. Due to the strongly limited penetration depth of the excitation light, this illumination scheme is mostly used to image cellular membranes or molecules immobilized on glass surfaces.

In 2008, Tokunaga et al. published an illumination approach to overcome the limitation of TIRF microscopy to the excitation of fluorophores near the surface. Their approach, which they called highly inclined thin illumination (HILO) [33] is based on similar principles as TIRF microscopy. The difference is that the incoming laser beam is offset to a lesser extend from the optical axis such that the angle under which the illumination leaves the objective is less steep and the laser penetrates into the sample as a thin collimated sheet of light. This approach keeps the benefit of TIRF, that only a thin region of the sample is illuminated, but it is not limited to areas near the microscopy slide. However the sheet of light penetrating the sample in HILO is not as thin as the evanscent field in TIRF microscopy. During the course of this thesis HILO illumination was applied for the recordings of DNA point accumulation for imaging in nanoscale topography (DNA-PAINT) data of the cytosolically located ASC speck, a supramolecular protein complex formed in certain cell types as part of the innate immune response (section 4). DNA-PAINT, in its original implementation exploited for this thesis, relies on optical sectioning as will be explained in

section 2.5.1.
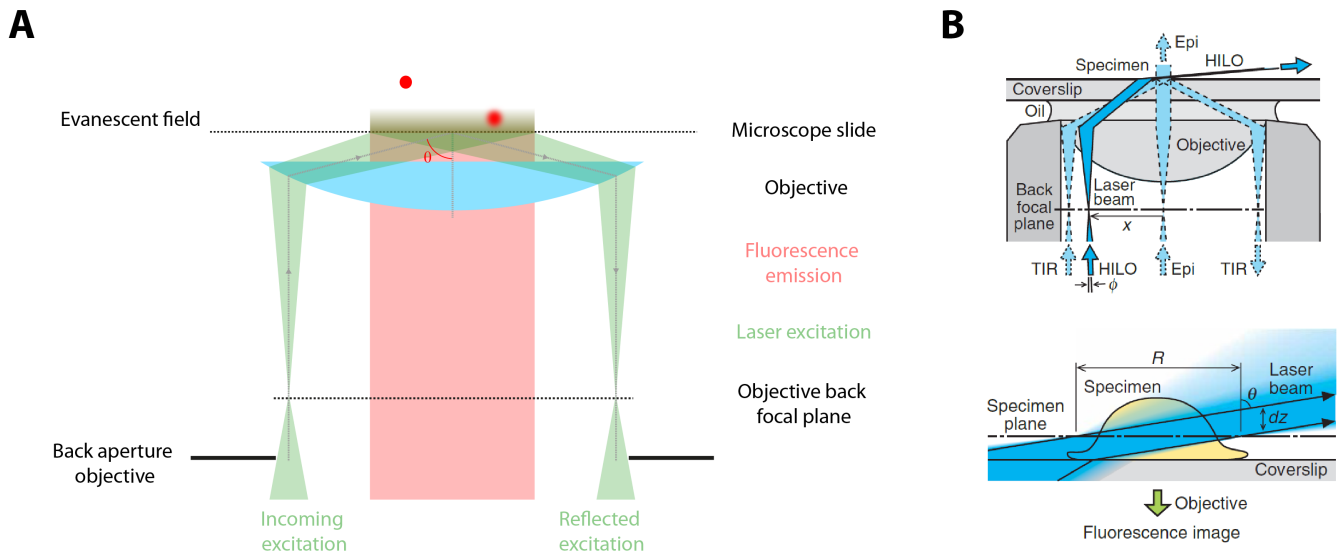
Figure 2.11: **Schematic of sample illumination and fluorescence collection for an objective-type total internal reflection and highly inclined thin illumination microscope**

A) In total internal reflection (TIRF) microscopy, the laser beam used for excitation is entering the objective at the periphery of the back aperture of the objective and focused on the back focal plane of the objective resulting in collimated light leaving the objective under an angle θ. This angle is higher than the critical angle, θ_c (equation 7) and hence the laser light is totally internally reflected at the upper glass interface of the microscope slide leading to an evanescent wave of laser light emerging above the slide. Fluorophores located within ~ 100 nm above the microscope slide are excited by the laser light.

B) In highly inclined thin illumination (HILO) microscopy, the offset of the excitation beam from the center of the objective is reduced compared to TIRF (upper panel) resulting in a thin sheet of excitation light penetrating into the sample (lower panel). The resulting fluorescence emission, in both approaches, is collected by the objective and imaged as described for widefield microscopy.

B) from [33]. Reprinted with permission from the Nature Publishing Group.

2.4.3   Confocal microscopy

Around the same time when E. J. Ambrose invented TIRF illumination, Marvin Minsky introduced another approach to overcome the problem of out-of-focus bluriness, the confocal imaging microscope [34] based on a concept invented in 1951 by Hiroto Naora as a non-imaging confocal microscope for spectroscopic studies of nucleic acids [35].

In a modern implementation of confocal microscopy the excitation laser light reaching the back focal plane of the objective is collimated, which results in the light being focused on the specimen into a diffraction-limited excitation volume. Although the highest laser power density is reaching the specimen at the focus of the objective, the entire sample region located within the volume is excited and hence also emits fluorescence. This fluorescence is collected by the objective, transmitted through a dichroic mirror and focused by the tube lens onto a pinhole introduced in the emission path. Due to its small size and its location in an image plane, almost exclusively photons originating from the center of the excitation volume pass through the focus, whereas photons from above, below or from a laterally shifted location in the focal plane of the objective are blocked by the pinhole. After the pinhole, the fluorescence signal is collected and focused onto the detector, most commonly a single point detector, by a second lens. Typical detectors used in confocal microscopy are photomultiplier tubes (PMT) or avalanche photodiodes (APD). By blocking the majority of out-of-focus fluorescence by the pinhole, the signal-to-noise in the detected signal is increased in comparison to conventional widefield microscopy (2.4.1). Figure 2.12 depicts the described beam path of a confocal microscope, the resulting excitation and the detection volume.

The effective NA (NA_{eff}) of an objective, and therefore the size of the excitation volume, depends, up to a certain extend, on the diameter of the incoming laser beam with

$$NA_{eff} \approx \frac{\phi}{2f} \tag{9}$$

where f denotes the focal length of the objective and ϕ the diameter of the incoming beam. Only the back aperture of the objective is fully filled meaning that the diameter of the incoming beam is at least as large as the back aperture of the objective, the full NA of the objective is reached. Since the intensity profile of the incoming beam is Gaussian shaped and therefore decreases away from the maximum, a fill factor (ratio between laser beam diameter and objective back aperture diameter) of at least two is often used in confocal microscopy (the objective is overfilled). By overfilling the back aperture, the smallest possible excitation volume size can be achieved. On the other hand, the inverse proportionality between the size of the excitation volume and the diameter of the incoming beam described in equation 9 can be exploited to tune the volume size, for example, by intentionally underfilling the back aperture of the objective and thereby increasing the size of the volume. This is for example done in single-molecule experiments in solution when one intends to collect the maximum amount of photons from a single molecule as it diffuses through the confocal volume.

In contrast to widefield, TIRF and HILO microscopy, in confocal microscopy only a small region of the sample is measured at a time, typically a diffraction-limited spot, and not an area. Consequently, the excitation volume needs to be scanned across the sample to generate an image. This can either be implemented by moving the stage and hence the sample with respect to the excitation beam or by scanning the excitation beam across the specimen, an approach introduced to confocal microscopy by White et al. in 1987 [36]. Scanning the stage has the advantage that no modifications are needed in the excitation path avoiding complicated alignment as well as light loss and optical abberations from additional optical elements. However, scanning is slow compared to laser scanning approaches for which reason the scanning confocal setup used during the course

of this thesis is based on laser scanning. The image obtained on a confocal microscope has higher contrast compared to an image acquired of the same sample on a widefield microscope, which itself might lead to an increase in the real resolution. In addition, the PSF of a confocal microscope is smaller compared to the one detected by the microscopy modalities described before (sections 2.4.1 and 2.4.2). The PSF size and hence the resolution of a confocal microscope depends on the size of the pinhole, which is normally given in Airy units (AU). One AU corresponds to the distance between the first minima of the Airy pattern (which equals $2\frac{0.61\lambda}{NA}$ according to equation 4) multiplied by the total magnification of the microscope. For a pinhole of one AU, the lateral resolution is given by

$$d_{xy\,confocal\,pinhole=1\,AU} = \frac{0.43\,\lambda}{NA} \tag{10}$$

The axial resolution, in this case is given by

$$d_{z\,confocal\,pinhole=1\,AU} = \frac{0.88\,\lambda}{n - \sqrt{n^2 - NA^2}} \tag{11}$$

where n denotes the refractive index of the medium between the objective and the sample, NA is the numerical aperture of the objective and λ is the wavelength of the fluorescence emitted by the sample. It is especially the increased axial resolution and hence optical sectioning capability of confocal microscopy that is important and leads to an increase in signal-to-noise. This effect is illustrated in Figure 2.13, which depicts an example of a fluorescently labeled cell recorded during the course of this thesis with a widefield microscope and the confocal microscope described below. The images were included in the book "Fluorescence Microscopy - From Principles to Biological Applications" [37]. Since the detected signal decreases with decreasing pinhole size, in practice one needs to find a compromise between resolution and signal intensity, and usually a pinhole diameter between 0.8 and 1 AU is chosen.
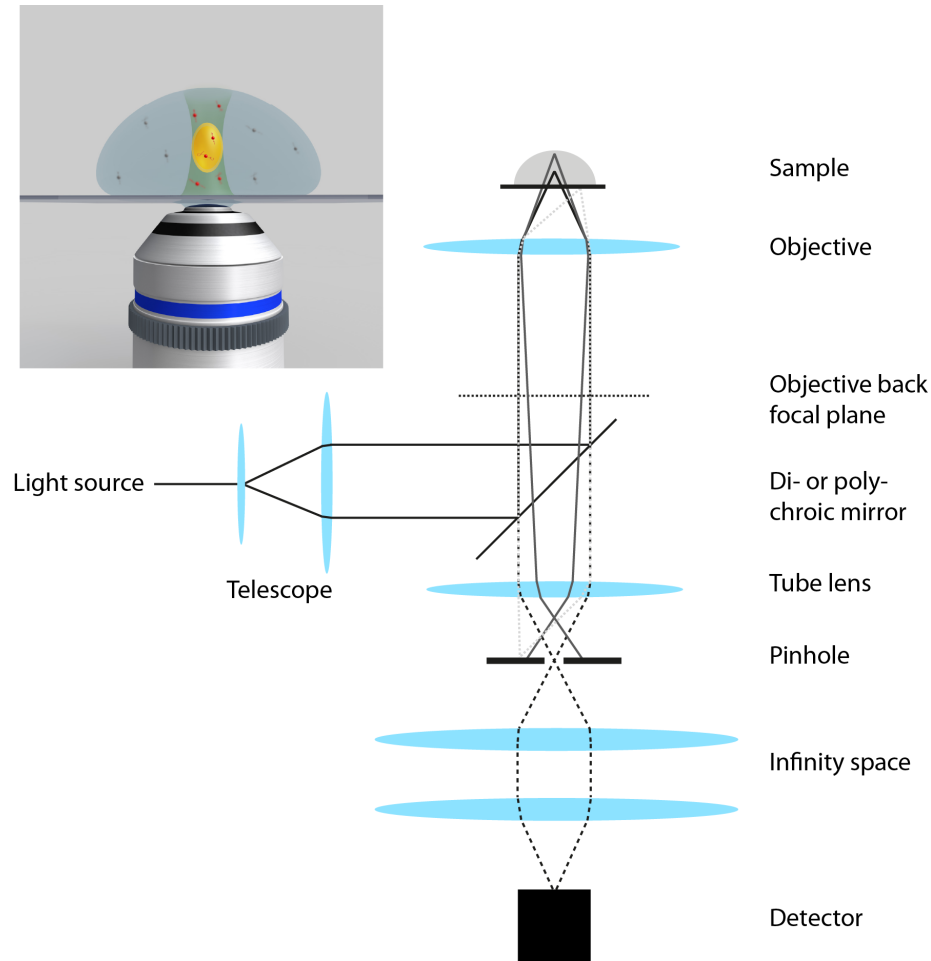
Figure 2.12: **Schematic of a non-scanning confocal microscope setup and its excitation and detection volume**
The laser beam used for excitation (solid line) is expanded by a telescope and enters the back aperture of the objective collimated resulting in a focused excitation. Fluorescence emitted by the sample is collected by the same objective, leaves the objective collimated, is transmitted through the di-/polychroic mirror and focused onto a pinhole by the tube lens. Only fluorescence originating from the center of the excitation volume (black dashed line) passes through the pinhole introduced into a conjugated image plane and reaches the detector. Fluorescence originating from out-of-focus regions of the sample (dark gray solid line) or a location laterally shifted from the center of the excitation focus (light gray dashed line) does not come to a focus at the center of the pinhole and is thereby blocked from reaching the detector. By inserting a lens pair behind the pinhole, an infinity space can be created where flat optical elements can be inserted without distortion of the beam.
The inset at the top left illustrates the excitation (green) and detection (yellow) volumes of a confocal microscope through which fluorescent molecules (depicted in red) diffuse.
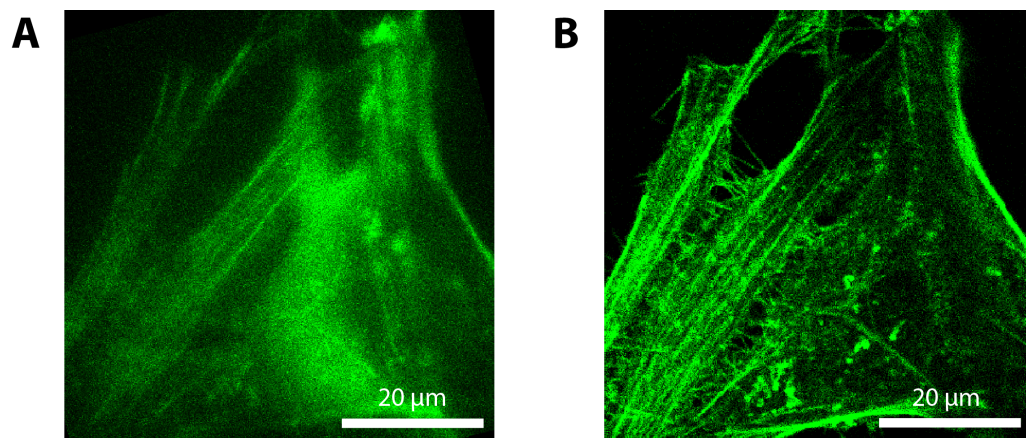
Figure 2.13: **Widefield versus confocal microscopy of actin filaments stained with Alexa Fluor 647®-phalloidin in HeLa cells**
A HeLa cell stained with Alexa Fluor 647®-phalloidin and imaged by (A) widefield microscopy and (B) confocal microscopy. Note the increased contrast in the confocal image due to background reduction compared to the widefield image enabling discrimination of more details inside the cell.

In the following, the main building blocks of the laser scanning confocal microscope utilized in the course of this thesis are described (Figure 2.14). As excitation sources, pulsed lasers at 475 nm (emits picosecond pulses) and 561 nm (emits femtosecond pulses) are utilized. Generally, continuous wave as well as pulsed lasers can be used as the excitation source in confocal microscopy. The usage of sub-nanosecond pulsed lasers in combination with time-correlated single photon counting (TCSPC) [38] generally enables the measurement of the fluorescence lifetime. For experiments requiring excitation with multiple lasers of different wavelengths, it furthermore enables switching between the excitation lasers on the nanosecond time scale for which the term pulsed-interleaved excitation (PIE) has been coined [39, 40]. This allows cross-talk free imaging and fluctuation measurements with quasi-simultanous excitation. During the course of this thesis, TCSPC in combination with PIE was applied for the cross-correlation measurements featured in chapter 5 and will be explained in detail in section 2.6.

As shown in Figure 2.14, the lasers are combined by the appropriate dichroics and introduced into a single mode fibre to ensure proper overlay of the lasers and a Gaussian intensity profile. Neutral density filter wheels allow adjustment of the power of every laser before they are coupled into the fibre. Afterwards, the laser light is reflected onto the first of two closely-spaced galvanometric mirrors which are electronically controlled to move the laser beam in two perpendicular dimensions, thereby scanning the laser beam across the sample. A telescope is introduced into the beam path after the galvanometric mirrors to keep the position of the laser light in the backfocal plane of the objective constant during the scanning process so that only the angle under which the laser hits the objective is changing. This change in the angle of the incoming beam translates into a scanning motion of the confocal volume across the sample. In addition to its function as relay optics, the telescope expands the beam to slightly overfill the back aperture of the objective.

Since the movement of the galvanometric mirrors is slow compared to fluorescence emission, the emitted fluorescence travels back along the same optical path (it is descanned by the galvanometric mirrors), is transmitted through the dichroic mirror separating excitation and emission light, and focused onto the pinhole (80 μm diameter) by the tube lens. Behind the pinhole, the beam is recollimated and introduced into the detection unit of the microscope. Here, the fluorescence is split by two dichroic mirrors into two spectral windows and focused onto APDs in front of which appropriate emission filters are mounted. APDs were chosen as detectors in this confocal setup due to their superior quantum efficiency compared to PMTs. During the scanning process, each detector records a constant photon stream, which is binned into pixels and thereby rendered into an image by the microscope control software. In the software, the size of the scanned area, the pixel size as well as the pixel dwell time can be chosen. The scanning speed and scanning range executed by the galvanometric mirrors are adjusted accordingly. The setup is equipped with a home-built perfect focus system to counteract axial movement of the sample during long term measurements. A widefield path with two continuous wave lasers of 488 nm and 561 nm wavelength facilitates identification of target cells in the sample through the eye piece or via an charge-coupled device (CCD) prior to confocal measurements.

In addition to its main application in imaging, the small observation volume (< 1 fL) and enhanced sensitivity of a confocal setup can also be exploited to analyse the movement of fluorescence species through the confocal volume via fluorescence fluctuation spectroscopy as will be explained in section 2.6. During the course of this thesis, this modality was exploited for the experiments described in chapter 4 and 5 as well as in the appendix A. These chapters feature applications of single-point fluorescence correlation spectroscopy (FCS) and raster image correlation spectroscopy (RICS) measurements, which were performed on a non-scanning confocal microscope and the described laser scanning confocal microscope, respectively.
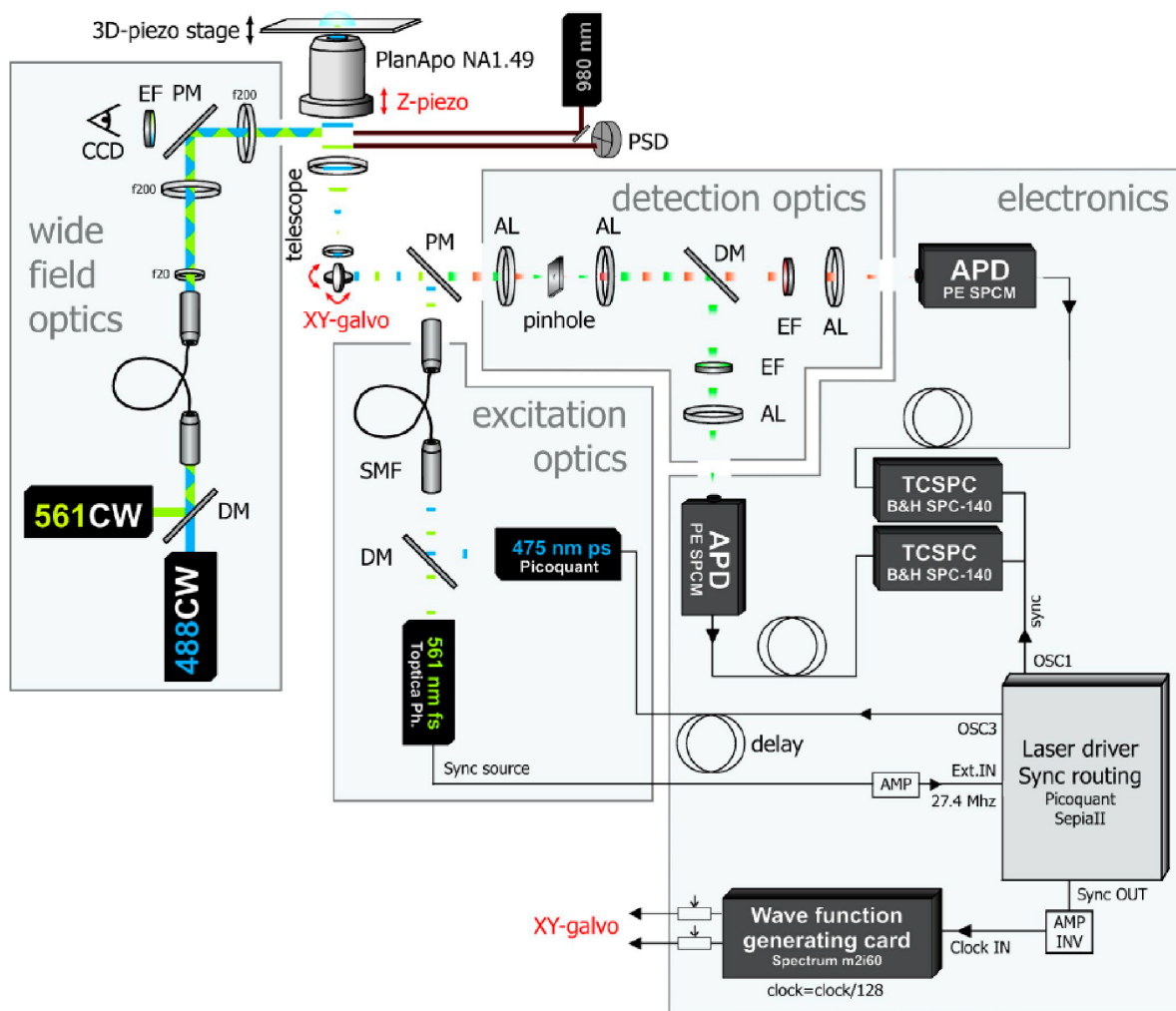
Figure 2.14: **Schematic of the applied laser scanning confocal microscope**
The laser scanning confocal microscope setup is equipped with a confocal excitation and emission path as well as a widefield excitation and emission path. The widefield excitation path contains continuous-wave lasers at 561 and 488 nm. For confocal excitation in combiantion with pulsed-interleaved excitation and time-correlated single photon counting detection, pulsed lasers at 561 and 475 nm can be used. Fluorescence in the confocal emission path is collected on two avalanche photodiodes. From [41]. Reprinted with permission from Rockefeller University Press

2.4.4 Spinning-disk confocal microscopy

Despite considerable improvements over the years, the maximum imaging rate of laser scanning confocal microscopes (LSCM) is in the range of a few frames per second with the limitation being the speed at which the galvanometric mirrors can be moved. This is substantially slower than the camera-based techniques described before and many biological processes happen on a much faster timescale. To enable faster image acquisition while still maintaining the benefits of a confocal microscope, Egger and Petráň, in 1967, came up with an idea to scan multiple confocal excitation volumes across the specimen simultanously and image their emission through multiple pinholes onto a camera. For this, they introduced a so called Nipkow disk into the beam path on which a large number of pinholes is arranged in a Archimedean spiral [42]. Spinning of the disk scans the sample in the *xy*-plane. The pinholes are arranged in such a way that, during one revolution of the disk the entire sample is scanned exactly once (or an integer number of times) and the exposure time is constant for every part of the sample. Since most of the excitation laser light is blocked by the spinning disk, usually only 5% of the laser light is transmitted, denoted as the transmission factor, the Yokogawa company introduced an additional disk into the excitation beam path on which microlenses are arranged in the same pattern as the pinholes. The collimated laser light is send onto the microlens disk behind which it is split up into multiple beams which are focused onto the pinholes of the Nipkow disk. This design improves the transmission factor to up to 40%. The diverging laser light emerging from the pinholes is collimated by a tube lens, introduced into the objective and focused into multiple confocal volumes exciting the specimen at different locations. Fluorescence is collected by the objective and focused onto the same pinholes by the tube lens. As in the classical confocal microscope the pinholes only transmit fluorescence originating from the center of the exciting confocal volumes. The dichroic mirror located in between the disks reflects the fluorescence which is imaged onto a camera by two relay lenses. An emission filter is installed in the emission path blocks light of different wavelength than the fluorescence from reaching the detector. Figure 2.15 depicts the beam path of such a confocal spinning-disk unit.

Modern spinning-disk units allow up to 10,000 revolutions per minute of the spinning disk and thereby exposure times as short as about 1 ms per frame. Furthermore, the higher quantum efficiency of cameras compared to single point detectors can be exploited. As for conventional LSCM, mounting the sample or the objective onto a piezo stage allows acquisition of *z*-stacks. All of these advantages have been exploited for the experiments performed on a spinning-disk confocal microscope (SDCM) during the course of this thesis. However, SDCM also has some disadvantages compared to LSCM. It requires high power lasers for excitation since the transmission factor of the Nipkow disk is still comparably low even when a microlens disk is implemented. Furthermore, the pinhole size cannot be easily adjusted and the *z*-resolution is lower due to emission light from out-of-focus planes leaking through adjacent pinholes, a phenomenon refered to as pinhole cross-talk.

The optimal pinhole diameter in a SDCM depends on the magnification of the objective used. It can be calculated according to the equation:

$$D = 1.2 M \frac{\lambda}{NA} \tag{12}$$

in which M and NA denote the magnification and the numerical aperture of the objective, respectively, and λ the wavelength of the emitted light. As such, the pinhole size is usually optimized for a certain magnification. For a 100x objective with a numerical aperture of 1.4, it equals 44 µm whereas, for a 60x objective with the same NA, 26 µm are optimal. Common pinhole sizes for spinning disks are 25 and 50 µm.
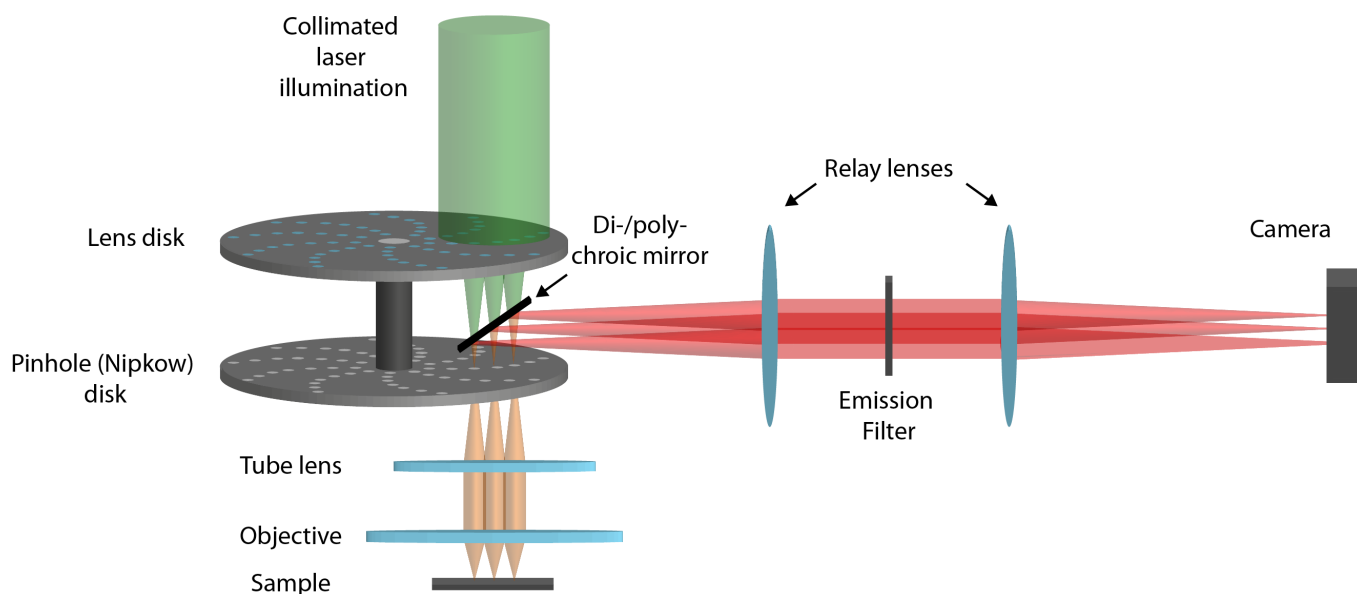
Figure 2.15: **Schematic of a spinning-disk confocal microscope**
Collimated laser light used for fluorescence excitation is focused by microlenses mounted on a lens disk onto the pinholes of a Nipkow disk. In between the disks, a di-/polychroic mirror is mounted which transmits the excitation light. The laser light emerging from the pinholes is recollimated by a tube lens and focused onto the sample by the objective. Fluorescence emitted by the sample is collected by the objective and focused onto the pinholes of the Nipkow disk by the tube lens. Fluorescence not originating from the center of the exciting confocal volumes is blocked by the pinholes. The fluorescence transmitted by the pinholes is reflected by the di-/polychroic mirror and imaged onto the camera by two relay lenses. An emission filter mounted in the emission path blocks light of different wavelength than the fluorescence from reaching the detector.

In the experiments described in chapter 3, a SDCM was used to track fluorescently labeled foamy virus particles in three dimensions in living cells. For this, a commercial Andor Revolution Spinning disk system (Andor Technology Ltd, Belfast, Northern Ireland) was modified. Figure 2.16 depicts a schematic of the system. Addition of the lasers and cameras, as well as the microscope control software were implemented by my coworker Dr. Fabian Wehnekamp. Separation of the channels by dichroic mirrors and filters as well as characterization of the setup was done with the help of my coworker Chen Qian. All filters listed in Table 2.1 and 2.2 were purchased from AHF Analysentechnik AG, Tübingen, Germany. The commercial part contains a laser combiner (Andor Laser Combiner (ALC) 401) equipped with four continuous wave lasers of wavelength 405 nm (Coherent Cube, 100 mW), 488 nm (Coherent Sapphire, 50 mW), 561 nm (Cobolt Jive, 50 mW) and 640 nm (Coherent Cube, 100 mW), which are overlayed into a single beam using a mirror and the appropriate dichroic mirrors (D1 - D3) and send through an acousto-optic tunable filter (AOTF), which allows selection of the laser wavelength and tuning of the laser power. The AOTF is connected via the Andor Precision Control Unit (PCU) to a field programmable gate array (FPGA) (National Instruments, Austin, Texas, USA) and the computer (PC) used to operate the microscope. After the AOTF, the beam is coupled into a single mode fibre (OZ Optics, Ontario, Canada). Four more laser lines of 445 nm (Coherent OBIS LX, 75 mW), 514 nm (Coherent OBIS LX, 40 mW), 594 nm (Coherent OBIS LS, 60 mW) and 685 nm (Coherent OBIS LX, 40 mW) wavelength have been added to the system in a home-built laser box where they are also combined using dichroic mirrors (D4: zt 633 RDC, D5: zt 543 rdc, D6: zt488RDC). Laser emission was narrowed down by installing laser clean-up

filters in the laser emission (C1: ZET445/10x, C2: ZET514/10x, C3: ZET594/10x). Both overlayed excitation beams are subsequently coupled into the same single mode fibre (OZ Optics) via a polychroic mirror (P1: zt405/488/561/640rpc) whose other ending is connected to a commercial spinning disk unit (CSU-10) from Yokogawa (Yokogawa Electric Corporation, Musashino, Japan). Inside the CSU-10 unit, the collimated laser beam is directed onto the microlens disk by mirrors and focused onto the 50 µm pinholes of the Nipkow Disk. The excitation lasers are transmitted by the di-/polychroic mirror used to separate excitation from emission light located in between the microlens and the pinhole disk. Depending on the laser lines used different di- /polychroic mirrors can be installed in between the disks (P2) (listed in Table 2.1). Afterwards, the light exits the spinning-disk unit and enters the microscope body (Eclipse TE2000-E, Nikon, Minato, Tokio, Japan) where it gets reflected by a prism, recollimated by the tube lens and transmitted through the objective (Apo TIRF, 100x, 1.49 NA, Nikon), which focuses it onto the specimen. A motorized xy-stage (Pro Scan II, Prior Scientific, Cambridge, UK) enables movement of the sample in *xy* and a motorized z-piezo stage (NanoScanZ, Prior Scientific, Cambridge, UK) the recording of *z*-stacks. To minimize *z*-drift, the microscope body is equipped with a perfect focus system operating with a 770 nm laser. Fluorescence emission is collected by the same objective and focused onto the same pinholes by the tube lens (rotation of the spinning disks is slow compared to the speed of light), is reflected by the dichroic mirror between the disks and directed out of the spinning-disk unit by mirrors. A relay lens located inside the spinning-disk unit behind the di-/polychroic mirror generates an image plane just outside the exit of the CSU-10 unit. This is because in an unmodified commercial setup a camera captures the image at this location. A combination of two 100 mm focal length lenses (the first 100 mm away from the image plane outside the spinning disk unit and a second 100 mm away from each camera) translates the image plane and thereby enables splitting of the fluorescence emission onto three electron-multiplying charge-coupled device (EMCCD) cameras (iXon (X-1731, X-1799), iXon-Ultra (X-10351), Andor Technology Ltd, Belfast, UK) by two dichroic mirrors (D7 and D8). Emission filters are mounted in filter wheels directly behind the emission dichroics (E1-E3). Depending on the fluorophores used, different lasers and filters are applied and Table 2.1 list the currently available filters (D7 + D8 and E1-E3) on the system. Table 2.2 gives an example of how the lasers and filters are to be combined to detect a certain fluorophore combination. The transmission illumination path of the system allows recording of differential interference contrast (DIC) images. For live-cell measurements, the system is equipped with a heatable stage inset (ibidi GmbH, Gräfelfing, Germany) in which the sample is mounted. To minimize drift during measurements performed at 37°C due to temperature differences between the objective and the sample, a heating collar (PeCon GmbH, Erbach, Germany) is used to warm the objective. A heatable lid ensures equal temperature around the sample and avoids condensation of cell medium on the lid of the cell slide. Lasers, cameras, z-piezo movement and the perfect focus system are controlled by a custom-written software written in LabView by Dr. Fabian Wehnekamp running on an FPGA (National Instruments, Austin, Texas, USA). For laser safety reasons and to avoid stray room light from reaching into the cameras, both the laser excitation and the emission path are shielded by custom-made boxes.

Table 2.1: **Fluorescence filters mounted in the home-modified spinning-disk confocal setup**

Emission dichroics	Excitation di-/polychroics	Emission filters	Laser clean-up filters
H 507 LPXR	ZT440-445/514/640tpc-YOKO	ET480/40m	ZET445/10x
568 LPXR ultraflat	ZT488/594tpc	525/50	ZET514/10x
ZT633rdc-UF1		ET535/70m	ZET594/10x
H 643 LPXR superflat		ET555/55m	
		560/40	
		ET605/70m	
		617/73	
		645/75 ET	
		ET650/100m	
		670/30	
		690/70	
		731/137	

Table 2.2: **Example of laser and filter settings for parallel excitation of eGFP and mCherry on the home-modified spinning-disk confocal setup**

Fluorophore	Laser(s) [nm]	Excitation dichroic	Emission dichroic 1	Emission dichroic 2	Emission filter 1	Emission filter 2	Emission filter 3
eGFP + mCherry	488 + 594	zt488/594tpc	H 568 LPXR superflat	-	ET 535/70 M	ET 650/100	-
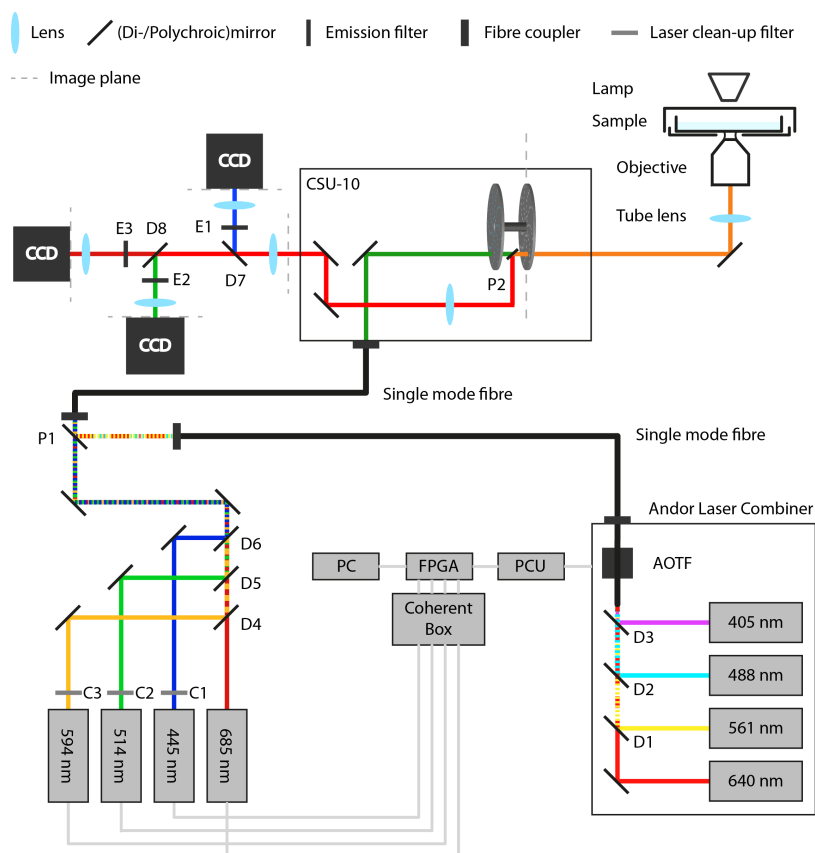
Figure 2.16: **Schematic of the applied, home-modified eight-color spinning-disk confocal microscope**
The commercial part of the system includes four lasers of 405 nm, 488 nm, 561 nm and 640 nm wavelength which are overlayed in an Andor Laser Combiner and coupled into a single mode fibre. An AOTF allows selection of the wavelength as well as the laser power. Four lasers of 445 nm, 514 nm, 594 nm and 685 nm wavelength were added to the system. All lasers are coupled into a single mode fibre, connected to the commercial spinning-disk unit of the system and can be selected for excitation of the sample. Fluorescence emitted by the sample can be imaged onto three EMCCD cameras. The system is additionally equipped with a lamp enabling the recording of DIC images, a heatable sample holder allowing live-cell measurements, a perfect focus system and heating collar for the objective, to minimze *z*-drift. All components of the system are controlled via a custom-written LabView software running on an FPGA.

Characterization of the system   Before starting measurements on the SDCM it was characterized with regards to: camera counts-to-photon conversion of the cameras, the pixel size, the alignment of the cameras in all spatial dimensions, the power of all laser lines reaching the objective and the cross-excitation of mCherry by the 488 nm laser as well as the cross-emission between the eGFP and mCherry channels. In the following, some of these characterizations will be explained.

The pixel size was measured using a Multi Grid Standard (58607) from Edmund Optics, Barrington, New Jersey, USA. For 100x magnification, as was applied for the experiments conducted during this thesis, the pixel size on all three cameras of the system was found to be 142 nm.

Since the majority of the planned measurements were to be performed using parallel excitation of eGFP and mCherry, the amount of cross-excitation and cross-emission on the system between these fluorescent proteins using the filters listed in Table 2.2 was characterized. Figure 2.17 depicts the eGFP and mCherry spectra overlayed with the respective filters.
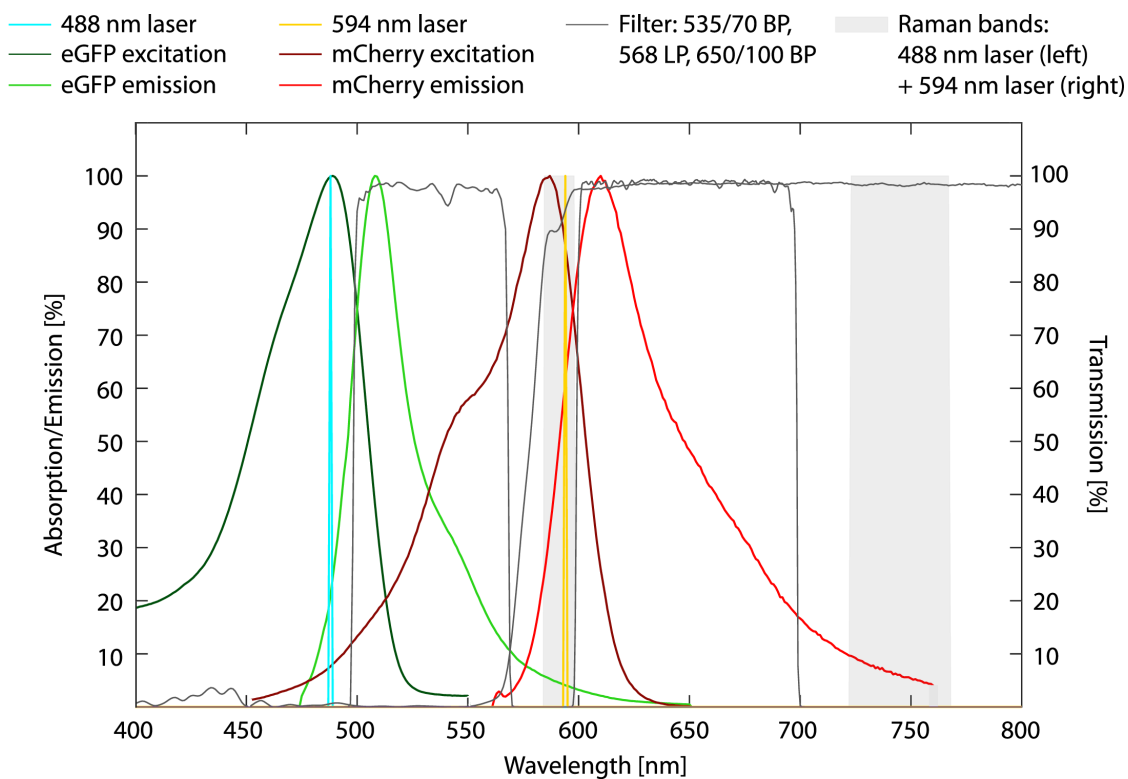
Figure 2.17: **Excitation and emission of eGFP and mCherry in the home-modified spinning-disk confocal microscope**
The excitation and emission spectra of eGFP (dark green and light green) and mCherry (dark red and light red) are shown. Additionally the 488 nm laser (cyan) and the 594 nm laser (yellow) used for excitation of the fluorescent proteins as well as their Raman bands for water (light gray shaded) are depicted. The spectra are overlayed with the transmission spectrum of the dichroic mirror installed in the emission path to separate the fluorescence from both fluorophores and the transmission spectra of the used emission filters (gray lines). BP: bandpass, LP: longpass.

First the amount of cross-excitation of mCherry by the 488 nm laser, commonly referred to as direct excitation, was quantified. HEK cells were transfected with mCherry and imaged on the mCherry camera after 488 nm and 594 nm laser excitation. The settings of the camera and lasers were set to the values chosen for the planned single virus tracking measurements.

The obtained data was analysed using a self-written Fiji macro [43] (Figure 2.18). In the first step, it segmented individual cells in the images recorded after 594 nm excitation by convolving the image with a Gaussian blur filter using a standard deviation of five pixels and subsequently intensity thresholding the image intensity using the IsoData algorithm [44]. The segmentation of the background was obtained by enlarging the segmentation of the cell by 10 pixels in all directions and selecting the image region outside this enlarged cell segmentation. The enlargement was performed to exclude pixels in close proximity to the cells which in some cases exhibited counts above the background. The mean background-corrected pixel intensity within the cell was calculated by subtracting the mean pixel intensity within the background segmentation from the mean pixel intensity within the cell segmentation. The obtained cell and background segmentations were subsequently applied to the images recorded of the same field of view after 488 nm excitation and the mean background-corrected pixel intensity within the cell was calculated analogously to what was done after 594 nm excitation. In total, 16 cells were analyzed. Taking the ratio of the obtained values revealed that the intensity obtained after 488 nm excitation was on average $4.2 \pm 0.97\%$ of the value obtained after 594 nm excitation. Consequently, when using parallel excitation, ~ 4% of the pixel counts obtained on the mCherry camera would be due to cross-excitation of mCherry by the 488 nm laser. For the planned measurements cross-excitation is irrelevant but in cases in which one only wants to measure mCherry signal after 594 nm excitation one would need to subtract these 4% from the counts on the red camera. Figure 2.18 illustrates the described analysis procedure and Table 2.3 summarizes the results.

Figure 2.18: **Image data underlying the calculation of cross-excitation of mCherry by the 488 nm laser**
A) mCherry-transfected HEK cells recorded on the mCherry camera after excitation with the 594 nm laser; top: signal segmentations, bottom: background segmentations
B) mCherry-transfected HEK cells recorded on the mCherry camera after excitation with the 488 nm laser; top: signal segmentations, bottom: background segmentations

Table 2.3: **Results of the cross-excitation measurements of mCherry by the 488 nm laser**; SD: standard deviation, SEM: standard error of the mean

Cell	1	2	3	4	5	6	7	8	9	10
Cross-excitation [%]	4.89	3.68	6.31	4.60	4.32	5.49	4.88	4.70	2.87	3.68

Cell	11	12	13	14	15	16	Mean	SD	SEM
Cross-excitation [%]	3.17	3.43	5.31	3.37	3.59	3.67	4.2	0.97	0.24

Next, the amount of cross-emission, commonly also called cross-talk, from eGPF into the mCherry channel was measured. For this purpose, HEK cells were transfected with eGFP and the signal of individual cells was recorded on the eGFP and the mCherry camera after 488 nm excitation. The settings of the camera and lasers were set to the values chosen for the planned single virus tracking measurements. The image analysis was performed analogously to the above described procedure for calculation of the cross-excitation with the following exceptions: the cells and background were segmented on the data recorded on the eGFP camera after 488 nm excitation, a Gaussian blur filter with standard deviation of two pixels was applied and thresholding was done by applying the Huang algorithm [45]. Subsequently, the mean background-corrected pixel intensity within the cell was calculated for both channels. The analysis revealed that, on average, 16.54 ± 1.70 % of the signal on the eGFP camera shows up on the mCherry camera (when excluding cells 4, 5 and 9 from the analysis since, for them, artifically high values were obtained due to low eGFP expression). Using parallel excitation, one should subtract 16.54% of the pixel counts measured for each pixel on the eGFP camera from the counts measured in the respective pixels on the red camera. Figure 2.19 illustrates the described analysis procedure and Table 2.4 summarizes the results.

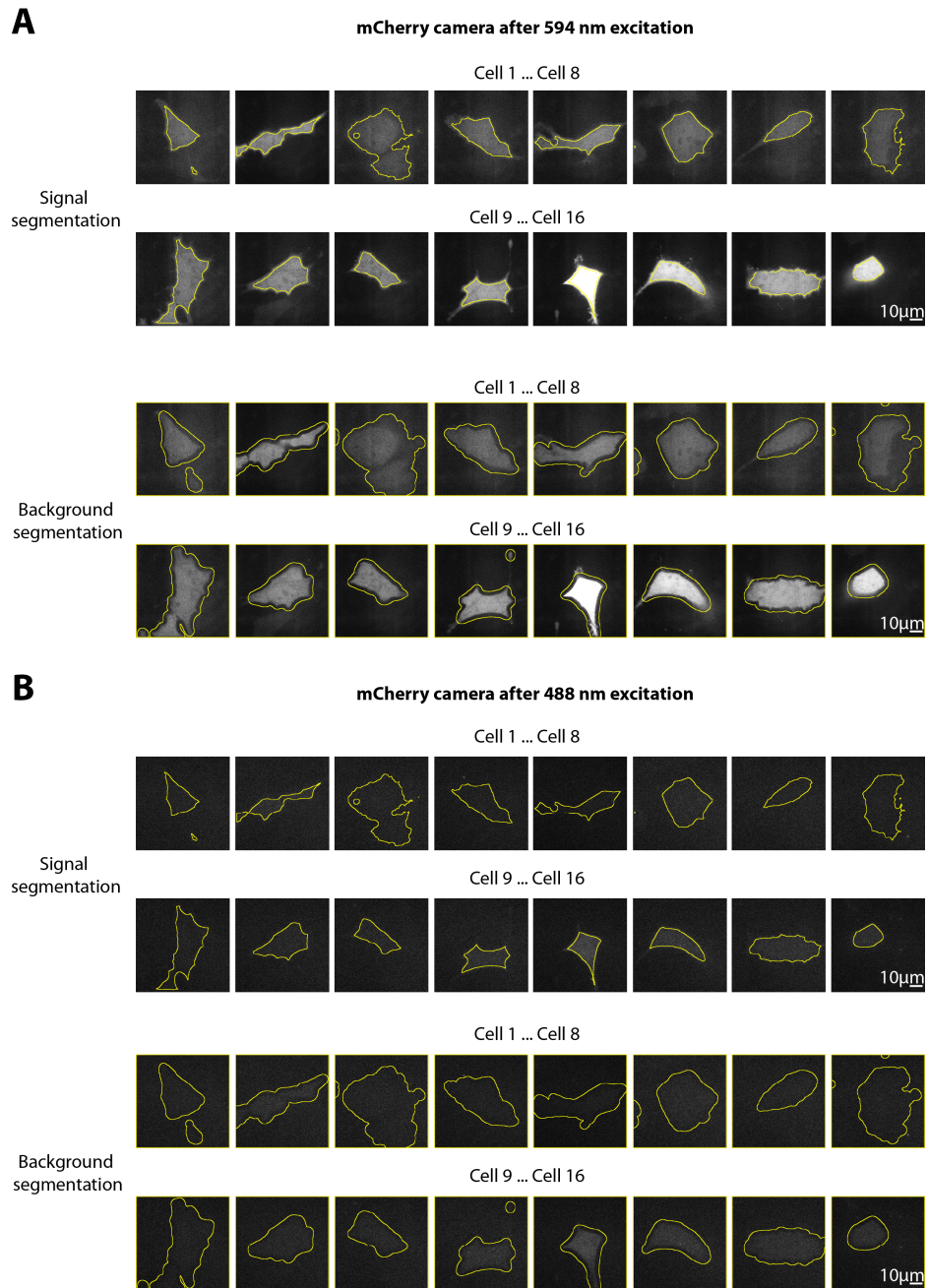**A**   **eGFP camera after 488 nm excitation**

Cell 1 ... Cell 6

Signal
segmentation

Cell 7 ... Cell 12

10µm

Cell 1 ... Cell 6

Background
segmentation

Cell 7 ... Cell 12

10µm

B   **mCherry camera after 488 nm excitation**

Cell 1 ... Cell 6

Signal
segmentation

Cell 7 ... Cell 12

10µm

Cell 1 ... Cell 6

Background
segmentation

Cell 7 ... Cell 12

10µm

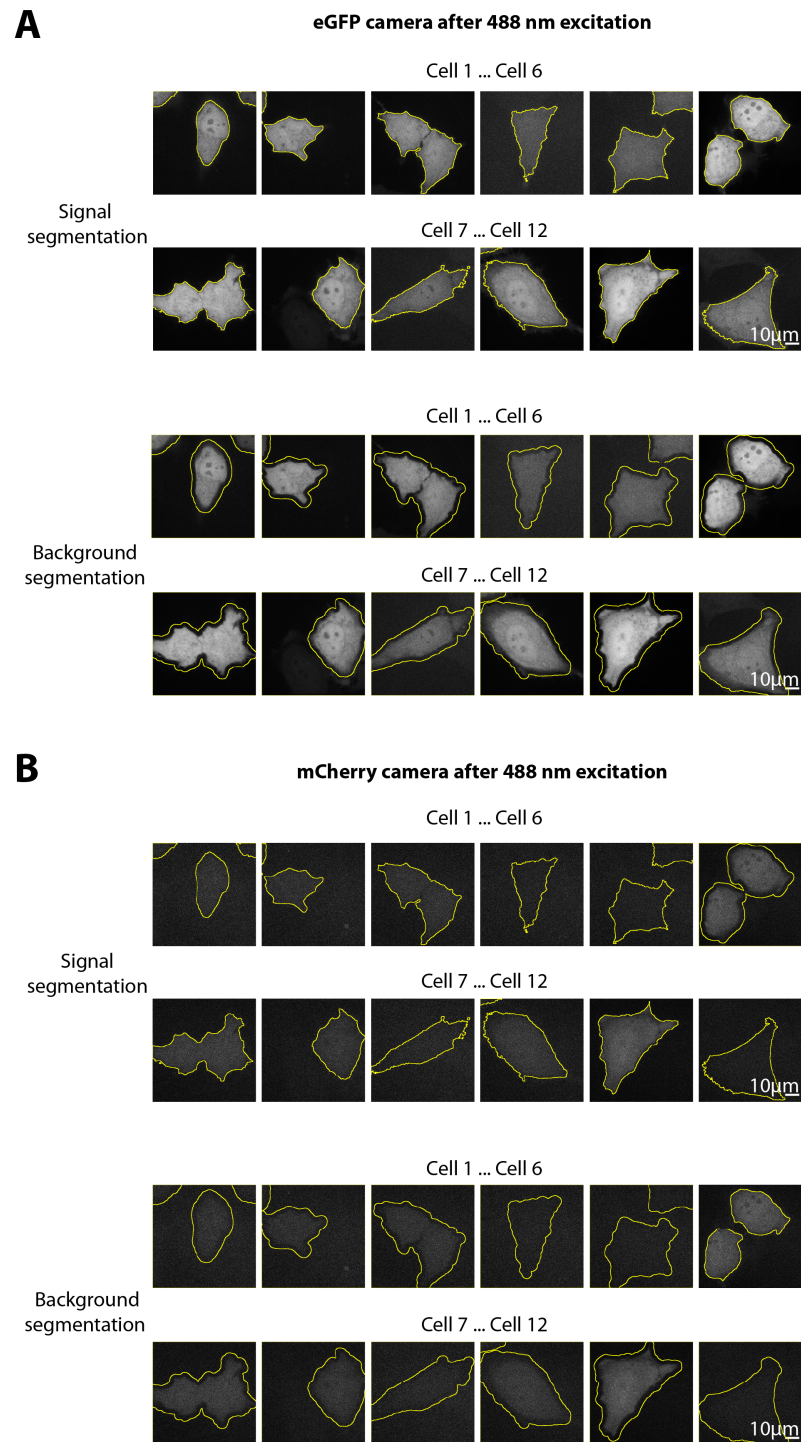Figure 2.19: **Image data underlying the calculation of cross-emission of eGFP into the mCherry channel**
A) eGFP-transfected HEK cells recorded on the eGFP camera after excitation with the 488 nm laser; top: signal segmenta-
tions, bottom: background segmentations
B) eGFP-transfected HEK cells recorded on the mCherry camera after excitation with the 488 nm laser; top: signal segmen-
tations, bottom: background segmentations

Table 2.4: **Results of cross-emission measurements of eGFP into the mCherry channel**; values in brackets were excluded due to low eGFP expression of the corresponding cells. SD: standard deviation, SEM: standard error of the mean

Cell	1	2	3	4	5	6	7	8
Cross-emission [%]	16.98	15.60	19.60	(32.80)	(23.72)	13.46	15.86	17.65

Cell	9	10	11	12	Mean	SD	SEM
Cross-emission [%]	(37.39)	17.12	14.91	17.66	16.54	1.8	0.60

During the measurements to monitor the uptake of foamy virus, the imaging slide was cooled to ~ 10°C and then mounted on the setup, which was pre-warmed to 37°C. This led to severe axial drift. To counteract this problem, the above mentioned perfect focus system was implemented into the microscope control software. To verify the effectiveness of this system, 3 μm fluorescent beads were adhered to an 8-well chambered cover glass, each chamber of the slide was filled with 300 μl water and the slide was equilibrated for 8 mins at 10°C or room temperature. For every temperature condition consecutive z-stacks of the same field of view were recorded on the bead sample over a total period of time of ~ 44 mins. The planes of the z-stack were thereby chosen 300 nm apart. These measurements were conducted with and without the perfect focus system being active after every stack. For every measurement the mean pixel intensity within a squared region of interest (ROI) selected around a single bead was calculated for the entire image series. The sub-plane precise maximum mean intensity within that ROI was subsequently calculated by fitting the calculated intensity values for every recorded z-stack with a Gaussian and determining the maximum of the Gaussian (Figure 2.20A). As a measure of focus stability the obtained sub-plane precise values were plotted against the number of the recorded z-stack. In an ideal situation without drift in the z-direction, the extracted z-position would stay constant between stacks. Figure 2.20B-D depicts an overview of the results. It clearly shows that the perfect focus system is needed to minimizes the axial drift in a situation where the sample is at 10°C and the microscope at 37°C. That the drift is stronger for the depicted measurement with the slide equilibrated at room temperature and the perfect focus system enabled (Figure 2.20C) compared to the slide equilibrated at 10°C and the perfect focus system enabled (Figure 2.20D) is due to the fact that the effectiveness of the system varies between measurements.
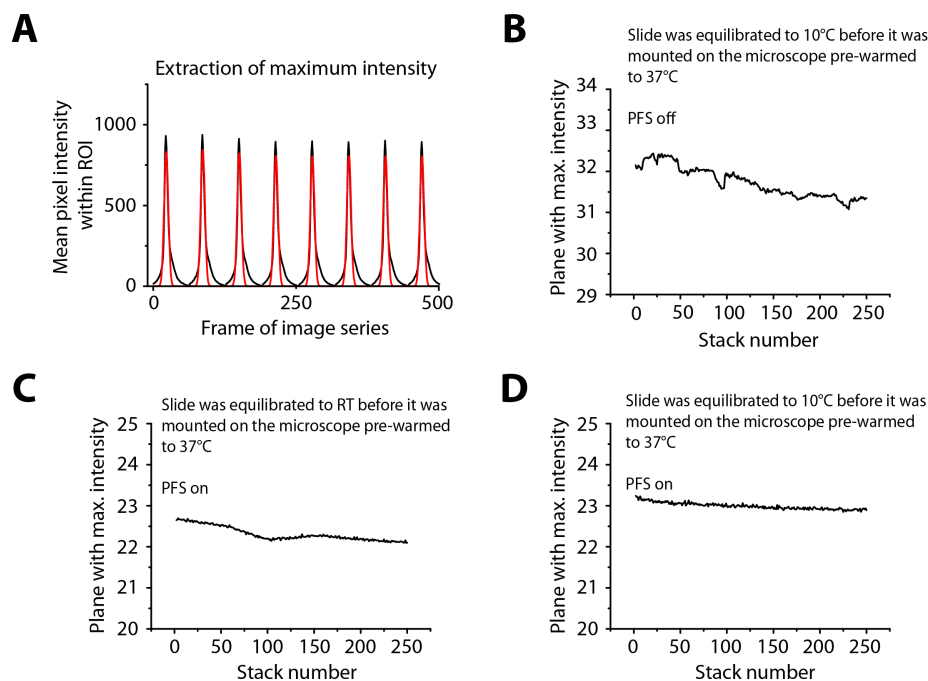
Figure 2.20: **Characterization of the perfect focus system of the home-modified spinning-disk confocal microscope**
To test the stability of the perfect focus system (PFS) of the SDCM, the imaging slide was equilibrated at 10°C or room temperature, placed on the SDCM pre-warmed to 37°C and multiple *z*-stacks of surface-adherent fluorescent beads were recorded over time. The measurements were performed with and without the PFS enabled. The mean pixel intensity within a squared region of interest (ROI) selected around a single bead was measured for every frame of the image series.
A) The average pixel intensity within a ROI plotted against the frame of the image series. Overlayed are Gaussian fits (red) to the extracted intensity values. From the fits, the z-position of the maximum intensity was determined with sub-plane precision for each stack.
B) - D) Plots of the extracted *z*-position of maximum intensity plotted against the number of the recorded stack. In an ideal situation without drift in the *z*-direction, the extracted *z*-position would stay constant between stacks. Hence, the extracted *z*-position served as a measure of focus stability for different temperatures of the slide and the microscope measured with and without the PFS enabled.
B) Data obtained for a slide equilibrated at 10°C measured on a microscope system warmed to 37°C without the PFS enabled. A clear shift in the extracted values is observable with increasing stack number.
C) Data obtained for a slide equilibrated at room temperature (RT) measured on a microscope system warmed to 37°C with the PFS enabled. A minor shift in the extracted values is observable with increasing stack number.
D) Data obtained for a slide at 10°C measured on a microscope system warmed to 37°C with the PFS enabled. The extracted value stays more or less constant with increasing stack number.
The measurements show that the PFS is necessary and suffficient to correct for axial drift if the slide is at 10°C and the objective at 37°C. The total number of recorded stacks correspond to ~44 mins recording time.

As an example for the capabilities of the home-modified spinning-disk confocal microscope Figure 2.21 depicts a live cell recorded on the system in 4-channels. It shows a live HeLa cell stained with CellTracker Deep Red Dye for intracellular structures and infected with foamy virus particles labeled with mCerulean at the Env protein and YFP at the Gag protein. Fluorescence excitation was performed using parallel excitation with the 445 nm, 514 nm and 642 nm lasers of the system. Additionally, a DIC image of the cell was recorded.
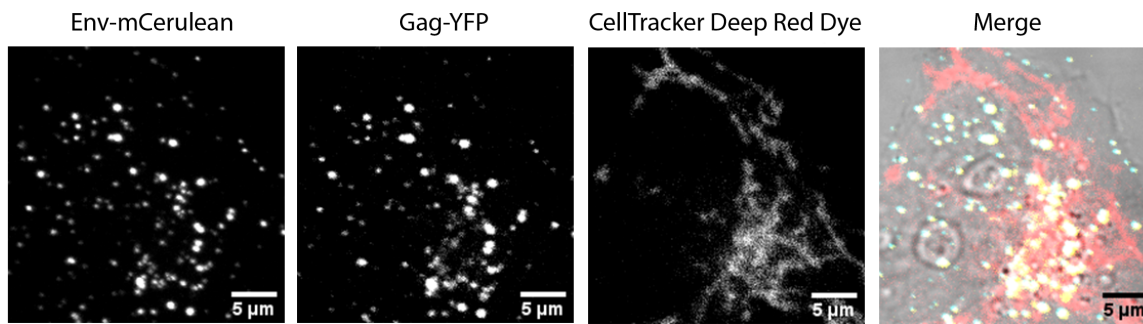

Figure 2.21: **4-channel image recorded on the home-modified spinning-disk confocal microscope**
A HeLa cell stained with CellTracker Deep Red Dye and infected with Env-mCerulean/Gag-YFP-labeled foamy virus particles is shown. The individual fluorescent channels as well as a merge showing an overlay of all three channels with a differential interference contrast image of the cell are depicted.

2.4.5 Single virus tracking

In addition to high contrast, fluorescence microscopy is characterized by high sensitivity. After the establishment of fluorescence microscopy-based techniques allowing detection of single molecules, these capabilities were exploited in multiple ways beyond classical imaging to tackle biological problems. A very prominent example of such an application is single virus tracking (SVT), which was introduced in 2001 by Seisenberger et al. [46]. Although in this first study the virions were only labeled with a single fluorophore this is not a requirement for SVT and many studies which applied this approach in fact labeled the virion with multiple fluorophores. Labeling single virus particles with a single or multiple fluorophores enables their detection by fluorescence microscopy. Their movement can then be captured by recording a movie of the signal they emit over time. By this approach, viral egress from as well as virus entry into living cells can be visualized. When combined with optical sectioning and the recording of multiple *z*-planes, single virus particles can be followed in all three spatial dimensions. Determining the center-of-mass of their emission for every frame of the movie and connecting the resulting coordinates results in a track or trace visualizing the movement of the particle, hence the term "single virus tracking/tracing" [47]. Besides the location of the virion, other parameters such as the fluorescence intensity, lifetime, the instantaneous velocity and interaction with cellular components can be determined along the track. Since its invention, it has been combined with widefield, classical confocal, spinning-disk confocal, orbital tracking as well as light sheet microscopy. Groundbreaking experiments showed several stages of motion, each characterized by distinct diffusion characteristics in the infection pathway of the adeno-associated virus [46]. Using an in vitro system, the kinetics of hemifusion and content mixing of the influenza virus were investigated [48]. Further studies characterized the infection process of the influenza virus in living cells [49, 50]. Several studies applied SVT to elucidate details of the fusion process of simian virus 40 [51, 52, 53] as well as the one of the human polyoma virus

[54] and the echovirus I [55].

As part of this thesis SVT was applied to study the fusion process of the foamy virus using the above described confocal spinning-disk microscope. The foamy virus particles contained an eGFP-labeled capsid and a mCherry-labeled envelope. By recording *z*-stacks of single virus infected cells the virus particles could be followed in all three spatial dimensions as they infected the cell. The data was analysed by a self-written software [56] which tracks individual foamy virus virions. The key component of this software is to detect single virus fusion events based on image cross-correlation (section 2.6.2) and hence it was named "Tracking Image Correlation" (TrIC). Next to the image correlation-based analysis the software extracts the following parameters for every virus particle over the time period the particle can be tracked: the sub-voxel precise position of the particle in 3D, the instantaneous velocity and the background-corrected intensity of the capsid as well as the envelope. Details on how the experiments and the the analysis were performed as well as the obtained results are described in chapter 3.

2.5   Sub-diffraction-limited microscopy modalities

The resolution of conventional optical light microscopes is limited by diffraction (section 2.3.2). It was at the beginning of the 1990s when the first methods to achieve images with resolutions beyond this limit in far-field microscopy were published. As described in section 2.3.2, single molecules are detected as the point spread function (PSF) by the detector of a light microscope, a structure much larger than the underlying molecule. For an isolated molecule, one can approximate the projection of this pattern onto the *xy*-plane, i.e. Airy disk, with a 2D Gaussian and estimate the position of the underlying molecule by finding its center-of-mass. The precision with which the center-of-mass can be determined depends on the number of detected photons, the standard deviation of the 2D Gaussian fit of the Airy pattern as well as the background and the pixel size of the image [57, 58, 59]. However if two features are closer together than defined by the Rayleigh criterion (equation 4) in a diffraction-limited image their centers-of-mass cannot be estimated with high precision since their PSFs overlap. Hence the PSFs need to be separated. This can either be achieved by shrinking the size of the effective PSF or by separating neighbouring molecules on the basis of a parameter that distinguishes them (e.g. spectrally or temporally). Both approaches have been realized to achieve resolution beyond the diffraction limit with a light microscope and the corresponding microscopy modalities are summarized under the term super-resolution microscopy. Stimulated emission depletion (STED) [60, 61] and structured illumination microscopy (SIM) [62, 63] are based on the approach of shrinking the effective PSF and single molecule localization microscopy (SMLM) [64] on the approach of separating molecules based on a feature that distinguishes them. For their pioneering work Stefan Hell, Eric Betzig and William E. Moerner, who was the first one to perform single molecule spectroscopy [65], were awarded the Nobel Prize for Chemistry in 2014.

The central objective of this thesis was to visualize the nanoscale organization of the endogenous ASC speck, a micrometer-sized cytosolic protein complex, inside cells by fluorescence microscopy (chapter 4). Due to the high density of this structure, maximal spatial resolution was needed to resolve structural characteristics of the complex. Among the above mentioned imaging approaches, SMLM offers the highest resolution and therefore was the method of choice. Furthermore, the experiments were designed to image a large number of specks since they are very heterogenous in nature and general characteristics can only be revealed if sufficient statistics are recorded.

Generally speaking, super-resolution methods, and especially SMLM techniques, are low in throughput due to their long data acquisition times. However, this disadvantage has been increasingly addressed in recent years by enlarging the field of view as well as by parallelization and automatization of data acquisition [66]. Pioneering work in this direction was done by the group of Professor Suliana Manley at the École Polytechnique Fédérale de Lausanne (EPFL) who implemented a flat field-optimized high-throughput direct stochastic optical reconstruction microscopy (dSTORM) setup, one particular form of SMLM introduced in section 2.5.1 [29]. For this reason, it was decided to collaborate with her group to address the above mentioned biological question. Furthermore, it was planned to apply labeling probes of different sizes on the same sample to investigate potential density differences within the structure, which made two-color SMLM imaging necessary. Fittingly, the group of Professor Manley also implemented this possibility into their workflow [67].

A different implementation of SMLM offering higher resolution than all previously mentioned approaches, namely DNA point accumulation for imaging in nanoscale topography (DNA-PAINT) [68], became increasingly popular when the experiments on this project were started. Having the above mentioned density of the ASC speck in mind, a collaboration with Sebastian Strauss from the group of, Professor Ralf Jungmann, the inventor of this technique, from the Max-Planck-Institut in Martinsried, was established. The following paragraphs explain the general working principle of SMLM as well as the particularities of

dSTORM and DNA-PAINT.

2.5.1   Principle of super-resolution single molecule localization microscopy

It was in 1995 that Eric Betzig published an idea on how to achieve resolution beyond the Abbe diffraction limit using a light microscope [64]. He claimed that, if it would be possible to separate the signal of individual emitting fluorophores in a diffraction-limited image, one could localize them and, by plotting the localizations of all molecules, achieve an image of subdiffraction-limited resolution (Figure 2.22).
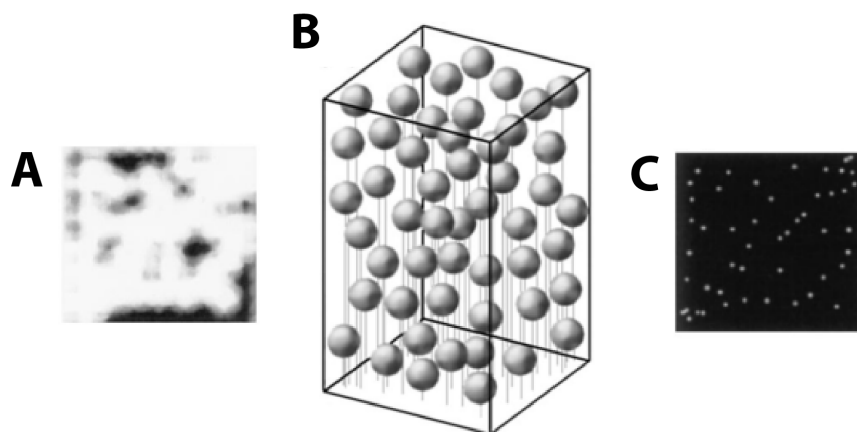

Figure 2.22: **Illustration of the principle of single molecule localization microscopy as published in the paper proposing the approach**
A) Simulated diffraction-limited image in which individual emitters cannot be distinguished due to overlapping point spread functions
B) Illustration of seperation of individual emitters based on a characteristic that distinguishes them
C) Simulated image of subdiffraction-limited resolution reconstructed from the localizations of the emitters depicted in panel A)
From [64]; Reprinted with permission from The Optical Society.

It took until the developement of the first photoactivatable protein, by George H. Patterson and Jennifer Lippincott-Schwartz in 2002 [69], for this strategy to become possible to implement in a method called photoactivated localization microscopy (PALM) [70]. Photoactivatable proteins can be switched from a non-fluorescent state into a fluorescent state by irradiation with light. For example, photoactivatable green fluorescent protein (paGFP) is weakly fluorescent in its native state and, upon irradiation with light of 413 nm wavelength, is switched to a stable state with absorption and emission maxima at 492 nm and 517 nm, respectively. This switched form of the molecule exhibits strong fluorescence after excitation with e.g. a 488 nm laser. By tuning the power of the activating 413 nm laser and recording a time series of the sample, it was possible to stochastically activate less than one single molecule per diffraction-limited area per frame and thereby separate the emission signals of individual molecules in time. Localizing the isolated single molecules and summing up all the single molecule localizations allowed them to obtain a super-resolved image. Separating individual molecules in time subsequently turned into an established strategy to localize individual molecules and different ways to achieve this were implemented. Since they all rely on the same basic principle, they are summarized under the term single molecule localization microscopy (SMLM). After recording of the blinking data, there are several processing steps required in order to obtain the final image for which

different software solutions have been developed [71]. Key aspects of the analysis procedure as well as the extension of the method for imaging in three dimensions will be described in the following. Afterwards the optical resolution achieved by these methods will be addressed.
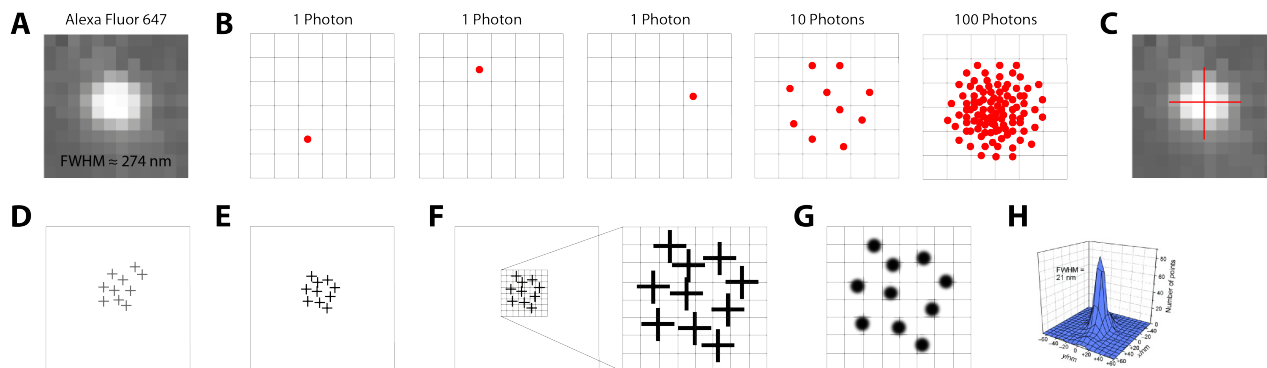

Figure 2.23: **Graphical illustration of the steps involved in single molecule localization microscopy**
A) Diffraction-limited image of a single emitting Alexa Fluor 647® molecule imaged with a objective having a numerical aperture of 1.33 on a sCMOS camera.
B) An illustration of where photons emitted by a single molecule are collected on the pixels of a camera chip to make up the point spread function (PSF).
C) An illustation of the localization of the molecule based on the shape of its PSF. Note that the larger the number of photons collected from the molecule on the camera, the more precisely its location can be determined.
D) An illustration of the spread of the localizations obtained from multiple blinking events of a single molecule. Lateral drift of the sample during the data recording might introduce a shift of the localizations. The localizations need to be corrected for this shift by a drift correction step.
E) An illustration of the drift-corrected localizations of a single molecule.
F) An illustration of how the localization distribution needs to be binned into smaller pixels in order to fullfill the Nyquist sampling criterion.
G) An illustration of how the localizations are convoluted with a blurring filter to represent the uncertainty with which they can be determined.
H) 2D Gaussian fit of all blurred localizations obtained from a single molecule; the full width at half maximum (FWHM) of the fit illustrates the increased resolution compared the diffraction-limited case in A); G) from [72]; Reprinted with permission from Wiley-VCH.

First, the signals of individual molecules have to be identified and, secondly the positions of the underlying molecules need to be determined e.g. by fitting with a 2D Gaussian or a model derived from the PSF of the microscope [73]. For the experiments described here an algorithm specific for sCMOS cameras was used [74]. Next, localizations with low precision, which can e.g. be caused by low number of photons for a given single molecule blinking event or by a large standard deviation of the corresponding fit of the PSF signal, are filtered out (provided there are sufficient localizations remaining) as they negatively affect the resolution of the resulting image. Since recording of the raw data can take anything from several minutes to several hours, drift of the sample often becomes a serious issue. In the *z*-direction, it is usually counteracted by a perfect focus device. *xy*-drift, however, needs to be corrected by processing the obtained data. Two common strategies of doing this are tracking fiducial markers such as gold beads, visible in every frame of the recorded image series, and image cross-correlation [75]. When exploiting the first strategy, one corrects the positions of the localizations by the amount and direction the fiducial

shifted in the corresponding frame from its original position at the beginning of the data acquisition.

As will be explained in section 2.6.2, the *xy*-shift between two otherwise identical images can be quantified by calculating the image cross-correlation. Hereby, the position of the cross-correlation maximum with respect to the center, the 0,0-lag corresponds to the amount of shift between the underlying images. For drift correction, this is exploited by cross-correlating images reconstructed from subsections of the raw blinking data. This approach was applied on the data shown in this thesis (chapter 4). The final step of the analysis is to render the localizations into an image. For that they are binned into pixels which results in a 2D histogram of localizations. Often the histogram is convolved with a Gaussian probability density function with volume and standard deviation equal to one (one pixel blur) as it was also done for the data presented in this thesis. More precisely, localizations can be depicted as a Gaussian of volume one and the standard deviation equal to the median localization precision of all localizations within the image or with the standard deviation equal to the localization precision of each individual localization. The latter one gives the best impression of how well a structure is resolved. Figure 2.23 illustrates the steps involved in single-color 2D SMLM.

When recording two-color SMLM images, an additional processing step becomes necessary when the data was recorded in spectrally separated channels. The two-color images presented in this thesis were recorded sequentially on the same camera using different fluorophores, Alexa Fluor 647® and DyLight 755®. Differences and aberrations (spherical and chromatic, magnification and rotation) between the two recordings were corrected by affine transformation. To this end an image of immobilized fluorescent beads was recorded in both channels and used to register the localizations in the DyLight 755® channel to the ones in the Alexa Fluor 647® channel.

Generally speaking, the axial position of single fluorophores cannot be determined from data obtained under standard widefield illumination due to the symmetry of the PSF along the *z*-axis, which initially limited SMLM to imaging in 2D. However, in 2007, Huang et al. adopted an approach that was introduced by Kao et al. in 1994 for single particle tracking [76] to enable SMLM in three dimensions [77]. By introducing a cylindrical lens into the detection beam path after the objective, the light becomes asymmetrically focussed in the *x*- and *y*-direction depending on the *z*-position of the emitter relative to the objective. This results in an asymmetrically elongated PSF in the *xy*-plane (Figure 2.24). The dependence of the PSF shape on the axial position can be calibrated by recording a *z*-stack of a point source, such as a fluorescent bead, fitting its image for every recorded plane with a Gaussian in *x* and in *y*, and determining the standard deviation of the Gaussians. From the obtained data a calibration curve can be generated. It visualizes the dependence of the width of the Gaussians in the *x*- and in the *y*-direction on the axial position of the emitter relative to the focal plane of the objective (Figure 2.24). For subsequent localization measurements, the 3D position of single emitters can then be accurately determined by fitting its signal with a Gaussian in the *x*- and in the *y*-direction and looking up the corresponding *z*-position in the calibration curve. The 3D SMLM data presented in this thesis were obtained using this approach. Over time, alternative approaches to obtain 3D information in SMLM were developed [78, 79].
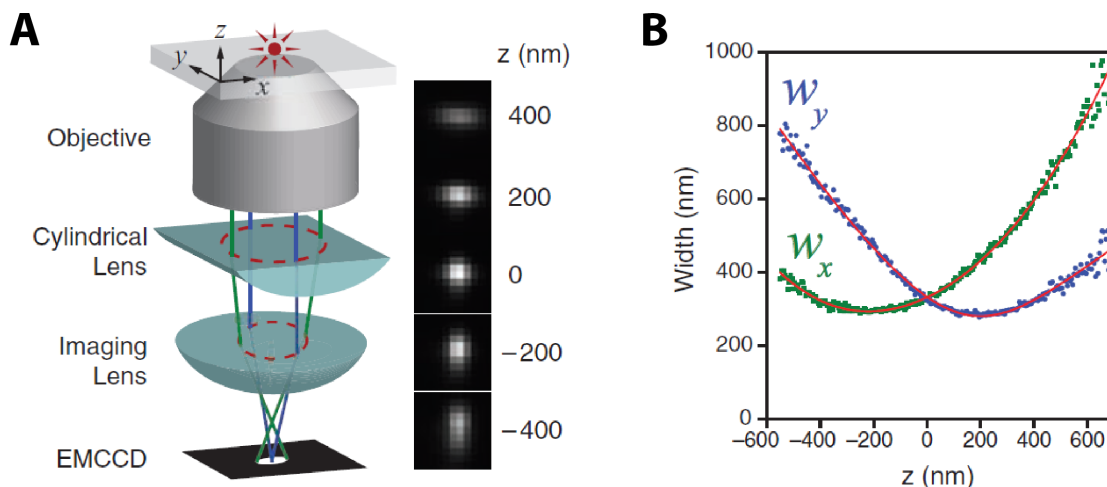
Figure 2.24: **Illustration of how to achieve three dimensional images in single molecule localization microscopy by introducing a cylindrical lens into the emission beam path**
A) A schematic illustrating how the point spread function is shaped as a function of the axial position of a fluorophore relative to the focus plane of the objective by introducing a cylindrical lens into the emission path.
B) Calibration curve obtained from single Alexa Fluor 647® molecules immobilized on a glass slide. Each data point represents the average of six molecules. It visualizes how the width of the PSF in the *x*- and *y*-dimension depend on the axial position of the molecules relative to the focal plane of the objective.
From [77]; Reprinted with permission from AAAS

The optical resolution achieved by SMLM depends mainly on three parameters. The first two are the precision and the accuracy with which individual molecules can be localized and the third is the molecular sampling rate, meaning the fraction of molecules making up the structure that are successfully localized. These three parameters will be addressed in the following. Omitting the influence of background and pixel size the localization precision for a single blinking event is defined as:

$$\sigma_{Super-resolution} = \frac{\sigma_{PSF}}{\sqrt{N}} \tag{13}$$

Where σ_{PSF} denotes the standard deviation of the 2D Gaussian fit to the PSF of the single molecule and N the number of photons detected from the single molecule. In SMLM, a single molecule is usually blinking multiple times and the localization precision of a single molecule can be determined by fitting the distribution of its localizations with a Gaussian and measuring the full width at half maximum (FWHM) of the fit. This is often also taken as an estimation of the resolution analogue to the diffraction-limited case in which the FWHM of the PSF is determined (compare Figure 2.23). Alternatively, the distribution of the relative pairwise distances of the localizations of a single molecule or the distribution of nearest neighbor positions in adjacent frames can be used as an estimate for the global localization precision [80].
The localization accuracy depends on the size of the labeling probe used which determines the offset of the detected fluorophore from the target protein. Since current SMLM techniques are approaching a resolution in the single digit nanometer regime, the size of the labeling probe is gaining increasing importance. Furthermore, a small probe size usually improves labeling density and hence the molecular sampling rate. In recent years, smaller probes such as nanobodies have been intro-

duced. For the experiments conducted for this thesis, the combination of primary + secondary antibody (~17.5 nm), primary antibody + secondary F_{ab} fragment (~11 nm) and nanobody (~4 nm) have been applied. Figure 2.25 illustrates the difference between localization precision and localization accuracy.
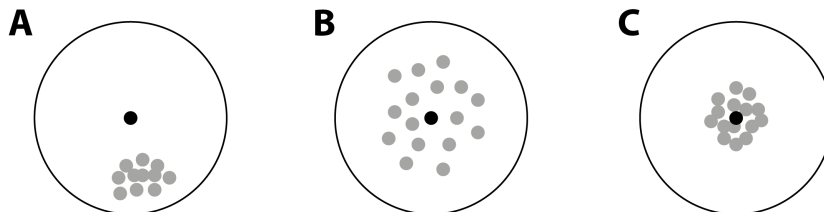

Figure 2.25: **Schematic illustration of localization precision versus localization accuracy**
The black dot indicates the position of the target protein, gray dots indicate individual localizations
A) An illustration for an event with high localization precision but low localization accuracy
B) An illustration for an event with low localization precision but high localization accuracy
C) An illustration for an event with both high localization precision and high localization accuracy;
Note that the size of the labeling probe puts a lower limit on the achievable localization accuracy and the achieved localization precision depends mainly on the number of photons emitted by the emitting molecule, the background and the pixel size of the camera.

The molecular sampling rate largely depends on the labeling density of the structure. As previously mentioned, smaller labeling probes generally increase this parameter. By applying the above mentioned Nyquist-Shannon sampling theorem (section 2.3.2) to the molecular sampling rate of a structure in SMLM, the distance between successfully localized molecules forming the target structure needs to be at least two-times smaller than the optical resolution to be obtained. For example, to obtain an optical resolution of 10 nm, a localization at least every 5 nm is needed.

Different ways of measuring the resolution in SMLM can be found in the literature. One approach is to measure the FWHM or the standard deviation from the distribution of all blinking events of a single molecule. A second frequently found approach is to measure the peak-to-peak distance between two neighbouring resolved single molecules. Since these approaches only quantify the resolution in a particular area of the image, methods to globally quantify the optical resolution of an image were introduced recently [81, 82].

Important forms of SMLM implemented over time include: the above mentioned photoactivated localization microscopy (PALM) [70], (direct) stochastic localization microscopy ((d)STORM) [83, 72], ground-state depletion with individual molecule return (GSDIM) [84] and (DNA-) point accumulation for imaging in nanoscale topography ((DNA-)PAINT) [85, 68, 86] For the experiments described in this thesis dSTORM and DNA-PAINT were applied. The following paragraphs will briefly summarize their working principle and address some important aspects which need to be considered when working with them.

direct Stochastic Optical Reconstruction Microscopy (dSTORM)   dSTORM, introduced in 2008 [72], uses standard car-
bocyanine dyes, such as Alexa Fluor 647®, to achieve stochastic emitting of single molecules. In contrast to the above
introduced PALM, the fluorophore is fluorescent at the beginning of the experiment and needs to be efficiently transfered into
a non-fluorescent state. For this, the sample is illuminated with high laser power density (for the experiments presented in this
thesis, 10 kW/cm² of 642 nm were used) in the presence of a reducing agent-containing buffer. For the described experiments,
a buffer containing the reducing agents mercaptoethylamine and β-mercaptoethanol was utilized. In addition the oxygen scav-
enger system protocatechuic acid (PCA)/protocatechuic dioxygenase (PCD) to deplete oxygen and reduce photobleaching as
well as cyclooctatetraene (COT) to stabilize the dye were used. Under these conditions, the majority of molecules is driven
into a long-lived non-fluorescent dark state. However, at the same time, a sparse subset of molecules returns into the fluores-
cent state, emitting fluorescence and can be localized. After a while, the emitting molecules are photobleached. At this point,
the sample is additionally excited with 405 nm laser light to reactivate molecules from the long-lived dark state by pushing
them into the ground state of S_0 from where they can be reexcited and subsequently emit fluorescence (Figure 2.26). The
power of the 405 nm laser light hereby determines the amount of molecules returned into the ground state and consequently
the density of emitting molecules. In practice, the UV laser power is gradually increased, e.g. varying from 1 - 10 mW laser
output power, while ensuring that the density of emitting molecules allows for single molecule fitting. The number of blinking
events per molecule is limited (for Alexa Fluor 647® to ~10) and afterwards, the molecule is permanently photobleached. This
limits the number of times a single molecule can be localized.

dSTORM data usually is recorded using widefield illumination. The data recorded for this thesis were obtained on a the flat-
field-optimized widefield microscope in the laboratory of Professor Suliana Manley which offers homogenous illumination
over a 100 µm x 100 µm large field-of-view as described described in section 2.4.1.
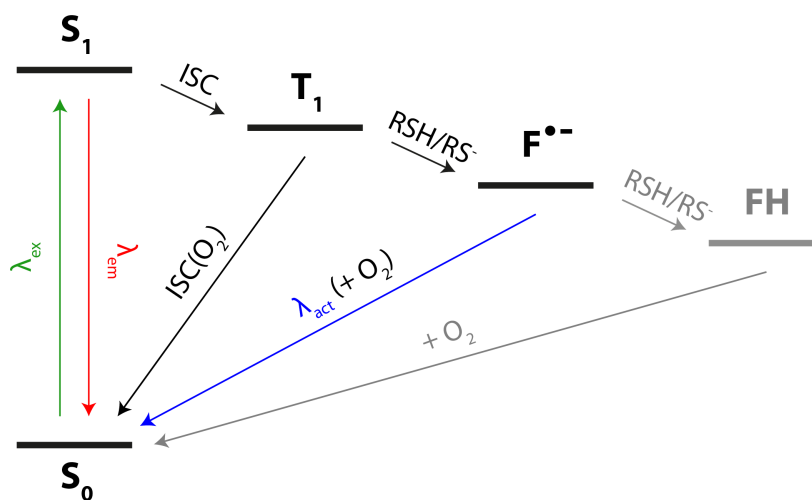
Figure 2.26: **Simplified Jablonski diagram of the electronic transitions underlying the switching mechanism exploited by dSTORM**
The dye is excited from the ground state S_0 into the first excited singlet state $S_1(\lambda_{ex})$. From here it can emit fluorescence (λ_{em}) or undergo intersystem crossing (ISC) into the triplet state T_1. From T_1 it can react with oxygen and relax back into the ground state. Alternatively it can react with a thiol to form a long-lived, non-fluorescent, semireduced dye radical (F^{-}). Reacting with oxygen (O_2) or irradiation with UV light pushes it back to the electronic ground state. Some fluorophores can transition into the fully reduced leuco-form (FH) from where they can also transition back into the ground state by reacting with oxygen.

DNA Point accumulation for imaging in nanoscale topography (DNA-PAINT)   In 2006, Alexey Sharonov and Robin M. Hochstrasser suggested a variation of SMLM in which individual molecules are identified through transient binding with fluorophore-labeled molecules in solution, an approach which they called point accumulation for imaging in nanoscale topography (PAINT) [85]. In 2010, Jungmann et al. introduced an implementation of this approach, named DNA-PAINT, in which the target molecule is decorated with a short oligonucleotide (~7-10 nucleotides), the docking strand, usually by application of a conjugated protein binder. This docking strand transiently interacts with a complementary fluorophore-labeled oligonucleotide, the imager strand, in solution [68]. When imaging these interactions by TIRF or HILO, which both decrease background, single hybridization events are detected as diffraction-limited signals on the camera. The duration of a single hybridization event is long enough to enable capturing of sufficient photons on the camera during a single frame for the event to appear as a diffraction-limited signal. Fluorophores in solution, in contrast, diffuse at such speed that the photons they emit only contribute to background signal on the camera.

This approach has several advantages over the other existing SMLM implementations: First, the duration of a single binding event can be adjusted such that it allows detection of all photons from the conjugated fluorophore, which directly translates into a high localization precision according to equation 13. Secondly, the essentially unlimited amount of labeled imager strands in solution eliminates the problem of photobleaching that limits the amount of available photons per molecule e.g. in dSTORM. Third, since the specificity of the detected signal for a particular protein is encoded in the sequence of the interacting oligonucleotides, a very large number of target molecules can be imaged sequentially with the number of targets being in theory only limited by the number of orthogonal sequences and specific protein binders available. This translates directly into

a fourth advantage, namely that the same fluorophore can be used for different targets, removing the difficulties arising from chromatic abberations in spectrally multiplexed fluorescence imaging. A fifth advantage comes with the programmability of the binding kinetics, which allows one to infer the number of underlying molecules from the blinking kinetics even though the molecules are not specially resolvable, an approach named qPAINT [87].
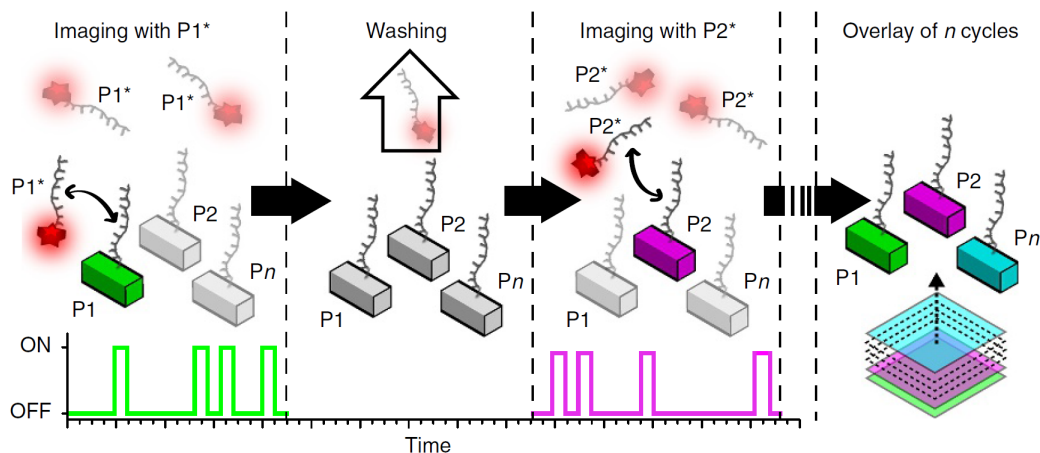

Figure 2.27: **Illustration of the principle of Exchange-(DNA-)PAINT**
The target molecule is decorated with a short DNA oligonucleotide (docking strand) P1 which transiently interacts with a fluorophore-labeled complementary DNA oligonucleotide in solution (imager strand) P1*. The binding and dissociation creates an ON-OFF blinking pattern on the camera (lower part of the figure) that is used for localizing the molecule. After a washing step, a second target molecule, decorated with a docking strand of a different sequence P2 can be imaged by applying the complementary imager strand P2* in solution. Repetition of this process allows one to image a larger number of molecules within the same sample, which, in principle is only limited by the amount of orthogonal oligonucleotide sequences available (Pn). Overlaying the obtained images allows one to receive a multicolor super-resolution image (right part of the figure) From [86]; Reprinted with permission from Nature Publishing Group.

Despite the numerous advantages, there are two main disadvantages of the original DNA-PAINT approach: First, the solution in which the sample is immersed contains dye-conjugated imager strands that, under widefield illumination, lead to an increased background and thereby decrease the signal-to-noise ratio and the resulting localization precision. This problem can be minimized by applying the above mentioned optical sectioning techniques (TIRF and HILO), but these are limited to imaging at the surface or within a thin section through the cell. To address this limitation, sample-based strategies have been suggested such as coupling a FRET acceptor dye to the docking strand and a donor dye to the imager strand [88], or the use of imager strands that are quenched in solution and only emit light upon binding to the docking strand [89]. The second major disadvantage is that DNA-PAINT measurements are time consuming. This is due to the single blinking/binding events being long and that the frequency of binding cannot be simply increased by higher imager strand concentration, since that would increase the background when using the original DNA-PAINT approach. This restricts the throughput of the method and practically limits it to the recording of only a few field of views within a single imaging session. The above mentioned approach based on a FRET pair at the docking and the imager strand or a quenched dye at the imager strand however allow to increase the imager concentration in solution without an increase in background. Since higher imager concentration increases the frequency of the binding events this ultimately increases the imaging speed and the throughput of the method.

Furthermore, the low imaging speed has been recently adressed by optimizing the sequence of the used oligonucleotides and the buffer conditions [90, 91].

An additional problem worth mentioning is that the docking strands get damaged over time mainly due to the formation of reactive oxygen species. In practice, this limits the maximal recording time and thereby the sampling of the binding sites and the multiplexing capabilities. However, by addition of oxygen scavenging reagents to the imaging buffer in combination with a triplet quencher, this problem can be efficiently suppressed [92].

2.6 Fluorescence correlation spectroscopy

Data recorded on fluorescence microscopes can also be analysed in terms of intensity fluctuations within the signal. The related methods are summarized under the term "fluorescence fluctuation spectroscopy" (FFS). In the following the three FFS methods applied during this thesis are described.

2.6.1 Single-point fluorescence correlation spectroscopy

The methodology of fluorescence correlation spectroscopy (FCS) was introduced in 1972 by Magde, Elson and Webb [93, 94, 95]. It extracts information from fluctuations in a fluorescence signal. Typically one observes the movement of fluorescent particles through an observation volume. Since these fluctuations are most pronounced when looking at a small number of particles, it took until the combination with the confocal microscope, characterized by its detection volume of about 1 femtoliter, for the method to be widely applied [96]. Nowadays, the idiom single-point fluorescence correlation spectroscopy refers to the method of correlating the temporal fluorescence fluctuations in a photon stream detected from a stationary confocal volume (Figure 2.28A). The method is most sensitive in the concentration range from ~ 1 pM to ~ 200 nM. To extract information about the processes underlying the fluctuations, the temporal autocorrelation function $G(\tau)$ of the photon stream is calculated according to equation 14 in which $I(t)$ is the photon count rate at time point t and $I(t + \tau)$ is the photon count rate at time point t plus the lag time τ, typically given in seconds. The angle brackets describe the average across all time points t.

$$G(\tau) = \frac{\langle I(t)\, I(t+\tau)\rangle_t}{\langle I(t)\rangle_t^2} - 1 \tag{14}$$

$G(\tau)$ can also be defined in terms of the fluctuations of the detected fluorescence signal $\delta I(t)$ which are defined as:

$$\delta I(t) = I(t) - \langle I(t)\rangle_t \tag{15}$$

$G(\tau)$ is then expressed as:

$$G(\tau) = \frac{\langle \delta I(t)\, \delta I(t+\tau)\rangle_t}{\langle I(t)\rangle_t^2} \tag{16}$$

Plotting $G(\tau)$ against τ on a logarithmic scale results in a characteristic FCS curve, which, usually approaches zero for increasing lag times. Intuitively explained the correlation function describes the self-similarity of the fluctuating signal on different time scales. From this perspective, it makes sense that it has a maximum value for short lag times and drops to zero with increasing lag time, if the motion underlying the fluctuations is random. There are a large number of processes that can lead to fluctuations in the underlying photon stream which, if correlated, all show up in the correlation curve. The most common application though is to analyse the fluctuations arising from a fluorescent species diffusing through the confocal volume. Considering only the fluctuations in the detected signal arising from translational diffusion, the maximum amplitude of the autocorrelation function is inversely proportional to the average number of independent fluorescent species N in the detection volume. When the shape factor of the detection volume γ is known, N can be calculated according to:

$$N = \frac{\gamma}{G(0)} \tag{17}$$

The detection volume of a confocal microscope can be well approximated by a 3D Gaussian. For this case γ equals $2^{-\frac{2}{3}}$. When the volume of the confocal volume is known, this number can be converted into a concentration.

The lag time for which $G(\tau)$ drops to half of the maximum amplitude corresponds approximately to the average lateral diffusion time, τ_D, of the diffusing species (when only correlations due to translational diffusion are considered). The radial size ω_r of the confocal detection volume is defined as the distance at which the excitation intensity decreases to e^{-2} relative to the intensity at the center of the confocal volume. When ω_r is known, the diffusion coefficient D of the diffusing species can be calculated according to equation 18.

$$D = \frac{\omega_r^2}{4\tau_D} \tag{18}$$

For free 3D translational diffusion of a single fluorescent species through the confocal volume, approximated by a 3D Gaussian, $G(\tau)$ can be expressed according to equation 19:

$$G(\tau) = \frac{\gamma}{N} \left(1 + \frac{4D\tau}{\omega_r^2}\right)^{-1} \left(1 + \frac{4D\tau}{\omega_z^2}\right)^{-\frac{1}{2}} \tag{19}$$

where ω_z denotes the distance in the axial direction at which the intensity decreases to e^{-2} relative to the intensity at the center of the confocal volume.

Figure 2.28 depicts a typical FCS curve obtained for Atto 565 in water diffusing through the confocal detection volume at 37°C.
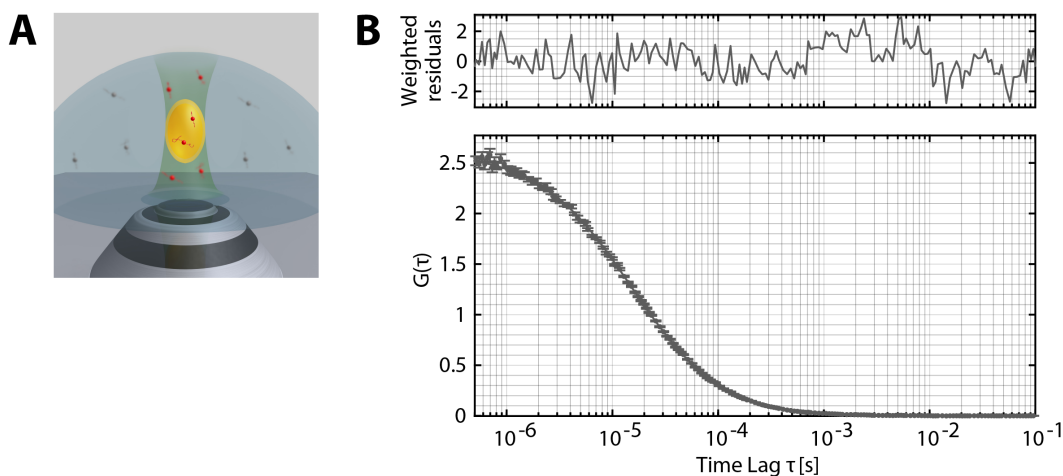
Figure 2.28: **Illustration of single-point fluorescence correlation spectroscopy**
A) Graphical illustration of the diffusion of a dye (red) through the confocal volume. The excitation volume is shown in green and the detection volume in yellow.
B) Autocorrelation function of a single-point fluorescence correlation spectroscopy (FCS) measurement of Atto 565[®] diffusing in water. The dots represent the measured data and the solid line is a fit with the function for free 3D diffusion (equation 19). The upper graph shows the weighted residuals of the fit. Error bars indicate the standard deviation of the measured data points derived from FCS curves calculated for the individual time bins of the underlying photon stream.

When the fluctuations are not due to simple free 3D diffusion of a single species through the confocal volume, the above equation will not fit well to the data and a different model function needs to be applied. At time lags shorter than approximately 2×10^{-6} seconds, additional processes typically contribute to the measured fluctuations, namely fluorophore blinking due to intersystem crossing into a dark triplet state, molecular rotation and antibunching. Those processes can also be analysed on the basis of the information contained in the FCS curve. In general, FCS cannot access processes happening below ~ 10 ns due to a lack of available photons and the maximum timescale is limited by the time a species spends in the confocal volume. During the course of this thesis, single-point FCS was used to verify the correct alignment of the laser scanning confocal microscope before recording RICS data (chapter 5 and appendix A) and to check whether there was free dye in a solution containing a self-labeled nanobody (chapter 4).

The ability to monitor the diffusion coefficient of a species by FCS can also be exploited to investigate interaction of molecules in solution. If two species that are very different in size interact and the diffusion of the smaller species is monitored, this interaction can be observed as an increase of the measured diffusion coefficient. However the interaction of two molecules which are similar in size cannot be precisely monitored by this approach since a doubling of the molecular weight of the diffusing species only leads to a 26% change of its diffusion coefficient. In these cases single-point fluorescence cross-correlation spectroscopy (FCCS) is a much more sensitve approach. It cross-correlates the signal of two differently fluorescent species diffusing through the confocal volume. The corresponding cross-correlation curve only shows an amplitude if the two fluorescent species exhibit correlated movement through the detection volume which is an indication for an interaction of the molecules.

2.6.2   Image correlation spectroscopy

While in single-point FCS the temporal correlation of fluorescence fluctuations is calculated, image correlation spectroscopy (ICS) correlates fluctuations of pixel intensity values within an image in space. For an ergodic process, for which the time average is equal to the ensemble average, single-point FCS and ICS contain the same information. ICS was invented by Petersen et al. in 1993 [97]. The 2D image correlation function G_S is given by the equation:

$$G_S(\xi, \psi) = \frac{\langle I_i(x,y) I_j(x+\xi, y+\psi)\rangle_{x,y}}{\langle I_i(x,y)\rangle_{x,y}\langle I_j(x,y)\rangle_{x,y}} - 1 \tag{20}$$

It depends on the spatial lag variables ξ and ψ along the x and y dimension of the image $I(x,y)$, respectively. The angled brackets indicate the average over all spatial positions. When image i (I_i) and j (I_j) are the same, G_S is denoted as the image auto-correlation function. Otherwise, G_S represents the image cross-correlation function [98]. A schematic representation of how the correlation is calculated is depicted in Figure 2.29.
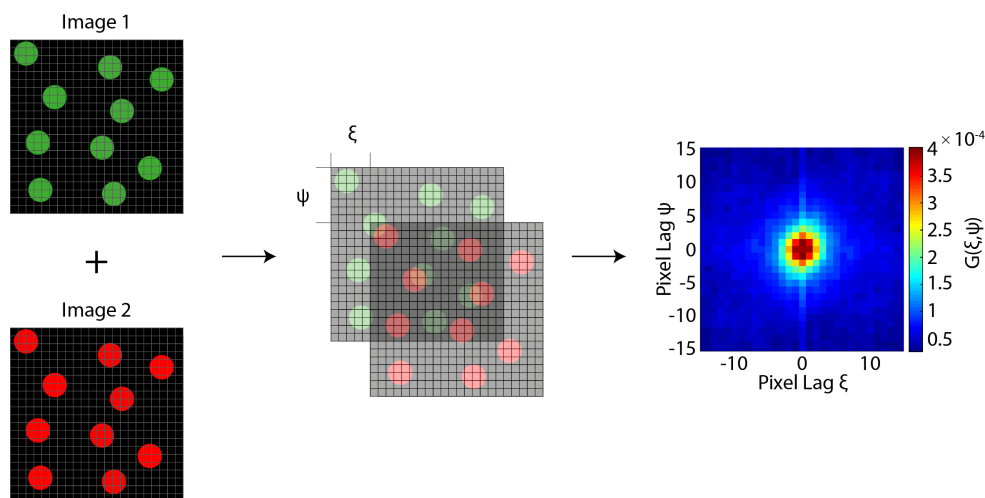
Figure 2.29: **Schematic illustration of image correlation spectroscopy**
Calculation of the image correlation between two images results in a correlation function from which spatial similarities between the images as well as the average number of independent particles within the PSF of the microscope can be derived. ξ and ψ denote the spatial shift variables along the x- and the y-dimension of the image, respectively. The SACF is color-coded according to the correlation value with blue representing low correlation and red representing high correlation. As an example the image autocorrelation function of single-molecule blinking events recorded on a resting THP-1 cell stained against the ASC protein with a primary anti-ASC antibody and a secondary Alexa Fluor 647®-cojugated F_{ab} fragment is shown. The central line observable in the correlation function is due to a correlation of pixels on the sCMOS camera used for data recording.

The amplitude of the autocorrelation function at zero spatial lag in both dimensions is inversely proportional to the average number of independent fluorescently-labeled particles N in the observation volume of the microscope on which the image was taken. The shape of the correlation function represents the convolution of the average size and shape of these species with the

PSF of the microscope. Consequently, information about the size and shape of the particles can be obtained by deconvolving the autocorrelation function with the PSF. However in practice a deconvolution only makes sense for particles whose size is on the order of the size of the PSF, since for particles larger than the PSF, the shape of the correlation function already represents the average size and shape of the particles and for particles much smaller than the PSF the deconvolution does not give reliable results. For randomly distributed particles much smaller than the PSF, the ICS function, G_S, can be approximated with a 2D Gaussian according to equation 21.

$$G(\xi, \psi) = \frac{\gamma}{N} \exp\left(-\frac{\delta r^2 \left(\xi^2 + \psi^2\right)}{\omega_r^2}\right) \tag{21}$$

As for single-point FCS, γ here denotes the shape factor of the PSF which is $2^{-\frac{2}{3}}$ for a 3D Gaussian and 2^{-1} for a 2D Gaussian (which would apply, for example, for measurements on a cell membrane). N is the average number of independent fluorescently-labeled objects in the observation volume, δ_r the pixel size, ω_r the radial size of the observation volume measured as the radius at which the intensity decreases to e^{-2} relative to the intensity at the center of the excitation volume, and ξ and ψ are the spatial lags along the *x*- and *y*-dimension of the image, respectively. The amplitude of the 2D Gaussian is equal to γ over the average number of independent fluorescently-labeled objects in the observation volume N. From the obtained values for N and ω_r the average (cluster) density CD of independent fluorescently-labeled objects within the observation volume can be determined according to:

$$CD = \frac{N}{\pi \omega_r^2} \tag{22}$$

When the image analysed by the classical ICS algorithm described here was obtained on a laser scanning microscope, there should be no movement of the fluorescent particles within the image or it has to be very slow. An algorithm to extract information regarding moving particles from an image obtained on a laser scanning confocal setup will be described in the subsequent section.

A characteristic of the image cross-correlation function is that a spatial shift between two otherwise identical images shows up as a shift of the correlation maximum. The distance the maximum is shifted from the center equals the amount of shift between the images while the direction of the deviation shows the direction in which the images are shifted. A prerequisite for the described method to work is that the PSF of the microscope is sufficiently sampled by pixels.

Provided that the sampling is sufficient the cross-correlation between two images only shows an amplitude when there are spatial similarities within the input images. In practice, image cross-correlation is frequently applied to evaluate colocalization of signals in two images of the same field of view detected in different fluorescence channels. During the course of this thesis, image cross-correlation spectroscopy was applied to evaluate the colocalization between spectrally separated images recorded on foamy virus particles and to correct for the drift occuring during dSTORM and DNA-PAINT data acquisition, as mentioned in section 2.5, and commonly done for SMLM data.

2.6.3 Raster image correlation spectroscopy

The method of single-point FCS described in section 2.6.1 provides high temporal resolution, but lacks spatial information. ICS, described in section 2.6.2, on the other hand, yields spatial but no temporal information. Raster image correlation spectroscopy (RICS), which is an extension of ICS, fills this information gap [99, 100]. It allows determination of the concentration and mobility of fluorescently-tagged molecules inside single living cells. Figure 2.30 graphically illustrates the recording, correlation and fitting of RICS data. Being performed on image data obtained by a laser scanning confocal microscope, such as the home-built setup described in section 2.4.3, it takes advantage of the fact that the image is acquired by scanning a laser focus over the sample. This introduces a time structure into the image since different parts of the image are acquired at different times. More precisely, the image is built up by continuously scanning the excitation beam across the sample until the end of the line is reached at which point the excitation beam is quickly retracted, shifted by one pixel in the perpendicular direction and scanned along the next line (Figure 2.30A). Generating the image in such way (i.e. raster scanning) introduces a fast and a slow time axis that are determined by the pixel time τ_p and the line time τ_l. For molecules moving on the timescale of the scanning process, the diffusion coefficient as well as the average number of molecules inside the confocal volume can be determined. To extract this information, a spatial correlation function is calculated according to the previously mentioned formula for ICS (equation 20). The results are commonly plotted in a 3-dimensional graph in which the *x*-axis describes the direction of the line scan, adjacent lines fall on the *y*-axis and the correlation amplitude at a given pixel *x,y* is denoted in *z*. Generally speaking the diffusion coefficient is encoded in the shape of the correlation function and the amplitude is equal to γ over the average number of molecules inside the focus of the microscope. The speed at which the fluorescent species in the sample moves with respect to the scanning speed is reflected in the shape of the correlation function along the *x*- and the *y*-axis. For example a species moving fast, relative to the movement of the beam in the *x*-direction, leads to a steep decay of the function along the *x*-axis but might show up as a tail of the function in this dimension for higher lag-values. The pixel dwell time is usually chosen in the microsecond range, the time between adjacent lines (line time plus retrace time), in the millisecond range and the frame time in the range of seconds. In theory, a single image frame is sufficient to calculate this function. In pratice, however, multiple frames are recorded and the average correlation function is calculated to increase statistics (Figure 2.30C, D).

RICS is used to extract information about diffusing species in a sample. If applied to extract the movement of diffusing species within a cell, the static or very slow moving cell components correlate and overlay with the correlations due to the movement of the fluorescent species. To extract just the correlations which are due to the movement of the fluorescent species a moving average of the acquired image frames is subtracted from each image frame before the spatial correlation is calculated. This filters out species not moving within the time span of the moving average window. Taking a moving average instead of a global average is necessary to account for the slow moving cell components. After subtracting a moving average the correlation function tends to fluctuate a lot since the denominator is close to zero. Hence, a constant, usually the total mean of all frames and pixels, is added. Of course, after this correction, the amplitude of the correlation is no longer equal to γ over the average number of molecules inside the focus and a correction factor needs to be introduced to extract this information.

For free 3D diffusion and approximating the confocal excitation volume with a 3D Gaussian, the correlation can be fitted with the function $G(\xi, \psi)$ (equation 23) from which the diffusion coefficient D and, in case no moving average correction is applied, the average number of fluorescent molecules inside the confocal volume N can be derived.

$$G(\xi, \psi) = \frac{\gamma}{N} \left(1 + \frac{4D\left(\tau_p \xi + \tau_l \psi\right)}{\omega_r^2} \right)^{-1} \left(1 + \frac{4D\left(\tau_p \xi + \tau_l \psi\right)}{\omega_z^2} \right)^{-\frac{1}{2}} \exp\left(-\frac{\delta r^2 \left(\xi^2 + \psi^2\right)}{\omega_r^2 + \left(4D\left(\tau_p \xi + \tau_l \psi\right)\right)} \right) \tag{23}$$

RICS is especially suitable to quantify diffusion processes on the subsecond timescale.

With the recently introduced extension Arbitrary-region RICS (ARICS) [101], it is now possible to measure these parameters in freely selectable regions of the image (Figure 2.30B). This allows, for example, comparison of the behavior of tagged molecules within different regions inside a cell.

When applied to cells in which two molecules are labeled with different fluorophores, as in one of the projects described in this thesis (chapter 5), a spatial cross-correlation function can be calculated to probe for interaction between the two species (cross-correlation (cc)RICS) [102]. The wavelength-dependent shift between the focus volumes δx and δy [103] is taken into account in the spatial term when fitting of the cross-correlation. For free 3D diffusion, the cross-correlation can then be described by:

$$G(\xi, \psi) = \frac{\gamma}{N} \left(1 + \frac{4D\left(\tau_p \xi + \tau_l \psi\right)}{\omega_r^2} \right)^{-1} \left(1 + \frac{4D\left(\tau_p \xi + \tau_l \psi\right)}{\omega_z^2} \right)^{-\frac{1}{2}} \exp\left(-\frac{\delta r^2 \left((\xi - \delta x)^2 + (\psi - \delta y)^2\right)}{\omega_r^2 + \left(4D\left(\tau_p \xi + \tau_l \psi\right)\right)} \right) \tag{24}$$

where the variables are defined as denoted previously. As mentioned for single-point FCCS, this approach allows for very sensitive detection of interactions between two fluorescent species since the correlation function only shows significant signal when there is an interaction between the cross-correlated molecules.

In the course of this thesis, RICS was applied to address two different biological questions, which are described in chapter 5 and the appendix A, respectively.
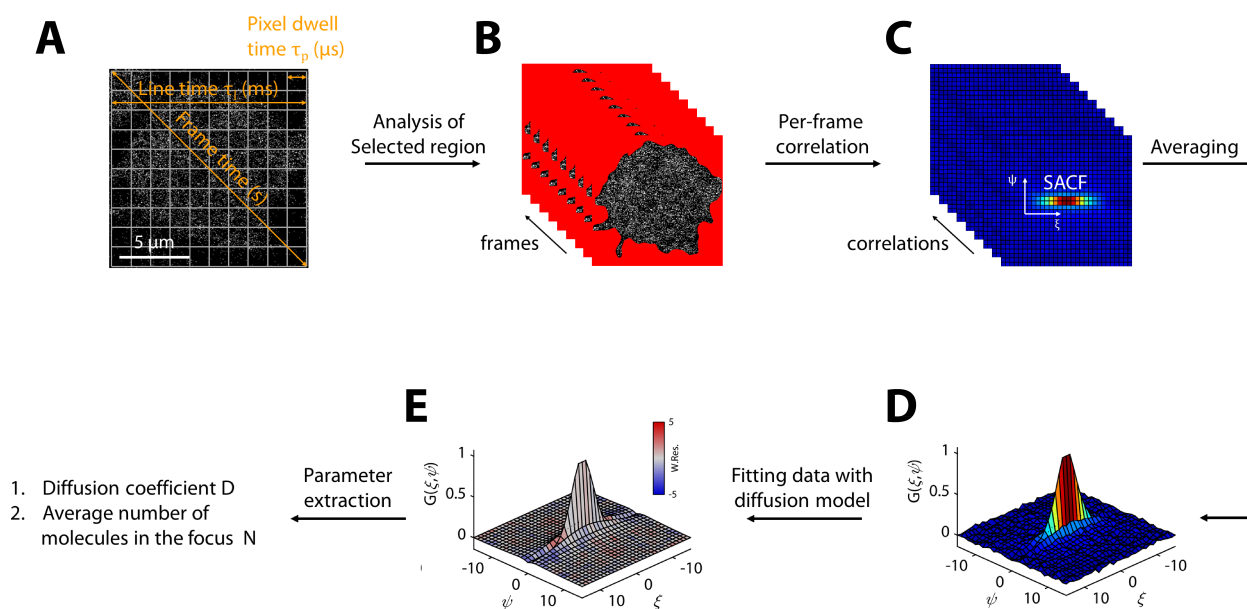
Figure 2.30: **Illustration of the steps involved in raster image correlation spectroscopy**
As an example the analysis of the diffusion of the protein ASC-TagRFP in the cytoplasm of BlaER1 cells is depicted (see also appendix A).
A) Schematic drawing of the raster image correlation spectroscopy (RICS) data acquisition process. A cell in which a fluorescently-tagged species is diffusing is raster-scanned on a confocal microscope setup introducing a time structure into the data. The commonly associated time scales are indicated.
B) For the analysis illustrated here the ARICS approach was chosen. The region to analyse was obtained by manually segmenting the cell outline in the obtained image series and pixels overlayed in red were excluded from the analysis. Note that this step is not essential for a RICS analysis and could be omitted.
C) An illustration of the obtained spatial autocorrelation function (SACF) determined for each frame of the image series; the spatial lag variables along the *x*- and *y*-axis, ξ and ψ, are indicated. The SACF is color-coded according to the correlation value with blue representing low correlation and red representing high correlation.
D) The average autocorrelation obtained from the individual correlations of the entire image stack. The colors encode the same as in C).
E) An illustration of the fit of the spatial autocorrelation with a diffusion model. Weighted residuals of the fit are color-coded from negtive values in blue to positive values in red. RICS allows determination of the average diffusion coefficient *D* and the average number of fluorescently-labeled molecules within the focus *N*.

2.6.4 Time-correlated single photon counting with pulsed-interleaved excitation

The sensitivity of FCCS using parallel laser excitation can be decreased by the spectral cross-talk between the fluorophores used. It denotes the leakage of emission photons from the shorter wavelength fluorophore into emission channel of the longer wavelength fluorophore (Figure 2.31A). When measuring the cross-correlation between two fluorescent species, cross-talk leads to a cross-correlation which is not due to processes in the sample but has technical reasons. Although it is possible to correct for this by factoring the fraction of cross-talk into the calculation of the cross-correlation function, a low number of cross-correlating molecules might be missed by this operation. Leakage of photons originating from the longer wavelength fluorophore into the channel of the shorter wavelength is normally not a problem since it can usually be avoided by the use of appropriate emission filters. For data recorded on a confocal microscope equipped with the suitable hardware, cross-talk can be circumvented using a method called pulsed interleaved excitation (PIE) [39, 40] in combination with time-correlated single photon counting (TCSPC) detection [38]. In PIE, subnanosecond pulsed lasers used for excitation are delayed with respect to each other by 12-20 nanoseconds, which is equivalent to about 4-5 times the fluorescence lifetime of the fluorophores used. Generally this is achieved by the first laser serving as the master clock and triggering the second laser, which is electronically delayed. Consequently, the respective fluorophores present in the sample are excited at different time points and their emitted photons also arrive at the detector in different time windows (Figure 2.31B). The detector registers the arrival of single photons and converts them into an electrical signal. These pulses are registered by the TCSPC electronics, which records the photon arrival time. The master clock starts sending clock ticks at a constant frequency at the beginning of the experiment which are registered by a TCSPC card. For each photon, the card records the macro- and the microtime relative to the clock ticks. The macrotime is given by the number of ticks registered by the card since the beginning of the experiment and gives information about the photon arrival on the 10 - 100 ns time scale. The more precise microtime (with accuracies of a few picoseconds) is the time between the photon arrival and the previous or the following clock tick. The photon arrival time is the sum of the macro- and the microtime. Figure 2.31C shows a scheme depicting the determination of the photon arrival time. In the analysis software, the number of registered photons is plotted in a histogram against the microtime or photon arrival time for each detector (Figure 2.31D). This excitation and detection scheme allows one to separate emission photons not only based on their wavelength by means of an emission dichroic but also by the exciting laser. For the setup used in this thesis, four PIE channels, assuming a hypothetical measurement with a green and a red fluorophore, can be distinguished: Green photons after green excitation, red photons after red excitation, red photons after green excitation and green photons after red excitation. The latter can normally be omitted since the red laser typically does not excite the green dye and there is no significant overlap between the red emission spectrum and the green emission filter, as mentioned above. Cross-correlating only green photons after green excitation and red photons after red excitation avoids the artificial cross-correlation signal due to cross-talk and thereby enables a more sensitive cross-correlation analysis. The cross-correlation RICS measurements performed during the course of this thesis and described in chapter 5 illustrate this benefit well.
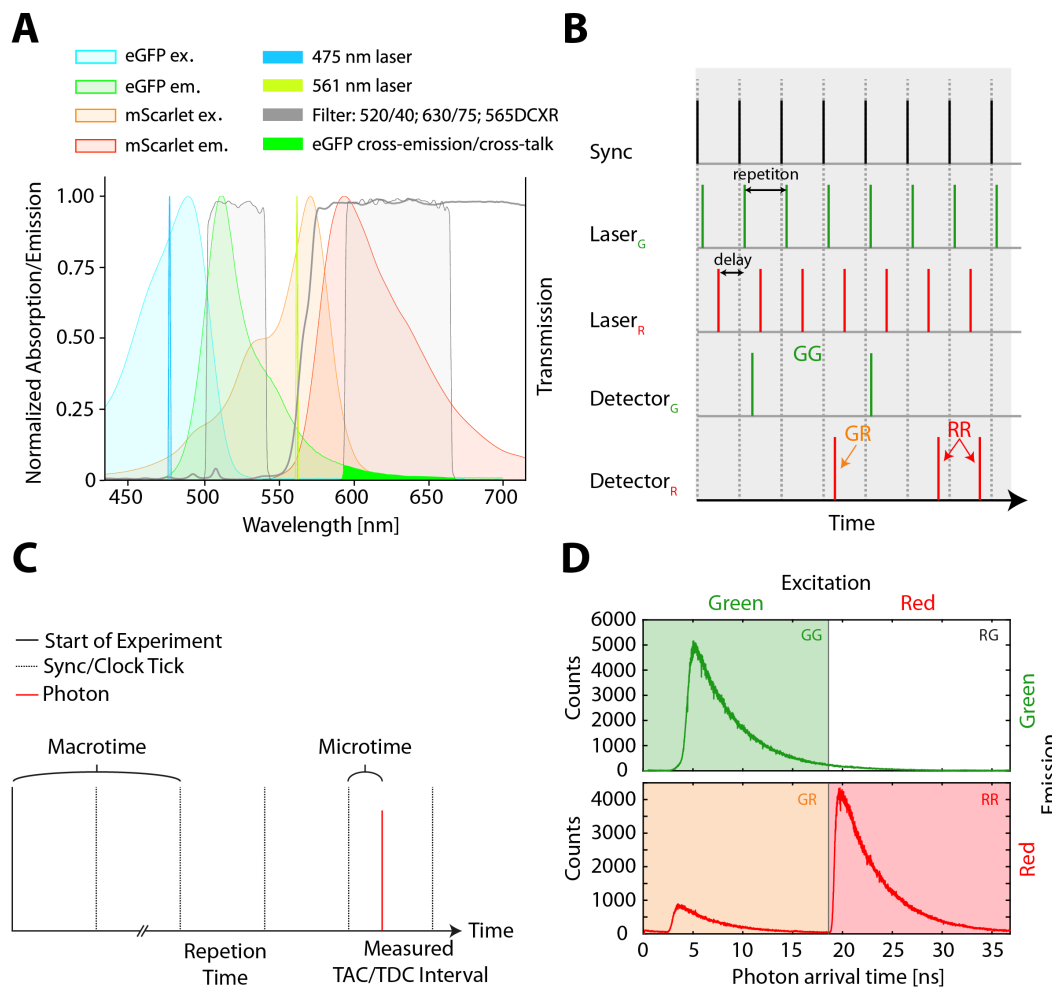
Figure 2.31: **Principle of time-correlated single photon counting with pulsed-interleaved excitation for cross-talk free imaging of eGFP and mScarlet**

A) Excitation and emission spectra of eGFP (blue + green) and mScarlet (orange + red) overlayed with the spectra of the applied filters (gray) and the lasers used for excitation of the fluorescent proteins; area plotted in solid neon green illustrates cross-talk from eGFP into the mScarlet emission channel

B) Schematic representation of the timing in a pulsed-interleaved excitation (PIE) experiment. A sync pulse which sets the repetition rate, determines the triggering of two lasers (green and red) with a time delay of 12 - 20 nanoseconds. The fluorescence emission is spectrally split between two detectors (a detecor that registers photons in the green part of the electromagnetic spectrum and a detecor that registers photons in the red part of the electromagnetic spectrum). Registering the photon arrival time on both detectors with respect to the exciting laser pulses allows to define four PIE channels: Green photons after green excitation (GG), green photons after red excitation (RG), red photons after green excitation (GR) and red photons after red excitation (RR).

C) Schematic representation of the photon arrival time determination using time-correlated single photon counting (TCSPC) detection. For each photon, the detector registers a macro- and a microtime. The macrotime is determined as the number of sync ticks since the beginning of the experiment and the microtime is the time passed in between the photon registration and the previous sync tick; TAC: time-to-amplitude converter; TDC: time-to-digital converter

Continued Figure 2.31: D) Histograms of photon arrival times obtained by measuring a mixture of Atto 488® and Atto 565® in water on the confocal microscope setup described in section 2.4.3 using TCSPC in combination with PIE. Overlayed onto the histograms are the respective PIE channels as defined in B).

3 Tracking image correlation of foamy virus

Publication underlying this chapter (to be found in appendix B, section B.1):

3.1 Biological background

Foamy viruses belong to the family of *Retroviridae* in which they fall into the subfamily of *Spumaretrovirinae* [104]. They are thought to be the oldest retroviruses and were originally observed in animals. As all retroviruses, they are enveloped viruses and contain a single-stranded RNA genome of positive polarity. A hallmark of foamy viruses is their broad cell tropism. The exact mechanism by which the foamy virus particles enter host cells remains to be elucidated. It has been shown that they attach to heparan sulfate [105] and the engagement of a currently unknown receptor on the cell surface is assumed. The virus particles are mostly taken up by receptor-mediated endocytosis. Release of the capsid into the cytoplasm is thought to be mediated by a pH-dependent mechanism involving the viral glycoprotein Env [106] which is incorporated into the envelope of the virus. Once released, the capsid is transported via dynein on microtubules towards the microtubule-organizing center (MTOC) where viral capsids accumulate [107]. Disassembly of the capsid through viral and cellular proteases seems to be cell cycle-dependent and nuclear import of the viral preintegration complex presumably requires break down of the nuclear membrane [108, 109, 110]. After entering the proviral stage, viral messenger RNA (mRNA) is produced and exported into the cytoplasm where translation into the viral proteins takes place. An exception are the viral glycoproteins whose mRNA is translated at the rough endoplasmatic reticulum. These proteins enter the secretory pathway [111]. Capsid preassembly takes place at the MTOC and reverse transcription of the viral mRNA mostly occurs within the formed capsid rendering those particles infectious [112, 113, 114]. Budding of the particles mostly occurs from intracellular membranes and release of viral particles is dependent on the Env protein. However, subviral particles without capsid are observed [115, 116]. Virus release is mediated via the cellular machinery of vacuolar protein sorting (Vps) [117, 118].

The only foamy virus species known to infect humans is the prototype foamy virus (PFV), which was previously classified as human foamy virus (HFV). Later on it was recognized to orignate from a chimpanzee that zoonotically infected a human [119, 120, 121, 122, 123]. One characteristic of foamy viruses is the apparent apathogenicity in their original host as well as zoonotically infected humans, which along with their ability to integrate into the host cell genome, makes them an interesting potential vector for gene therapy. However, a much better understanding of the viral replication cycle is needed for such an application to be put into practise.

The work conducted during the course of this thesis contributed to this aim by addressing a step in foamy virus fusion using single virus tracking (section 2.4.5). It was carried out in collaboration with the research group of Professor Lindemann at the Technische Universität Dresden, Germany.

3.2 Additional Materials and Methods

This section provides experimental details, that go beyond what is given in the attached paper.

3.2.1 Sample preparation and data recording

To study the infection process of single virus particles into cells, we exploited the high sensitivity and speed of fluorescence spinning-disk microscopy (section 2.4.4). To mimic physiological virus infection and, at the same time, enable virus particle detection on a fluorescence microscope Professor Dirk Lindemann and coworkers generated foamy virus particles containing PFV-Gag (the Gag protein is part of the viral capsid) and Env from either PFV or macaque simian foamy virus (SFVmac) origin, which were labeled with different fluorescent proteins at the Gag and Env protein. We decided to utilize HeLa cells to study the fusion process since previous studies showed that this cell line is susceptible to foamy virus infections. The following briefly describes the experimental procedure applied. Details of the procedure were adjusted during the experiments and hence the protocol presented here slightly varies from the one presented in the manuscript underlying this chapter of the thesis. If not stated differently, all reagents were purchased from Thermo Fisher Scientific, Massachusetts, USA.

HeLa cells were cultivated at 37°C in 5% CO_2 atmosphere in Dulbecco's Modified Eagle Medium (DMEM-Glutamax (high glucose, + pyruvate, 31966021), supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS) (10500064) and split every 2-3 days. One day prior to the experiment, 2 x 10^4 cells/well were seeded in 8-well microscopy slides (Nunc LabTek II, 154534PK) coated with collagen A-solution (C4243, Sigma-Aldrich, St. Louis, MO, USA) according to the manufacturer's protocol to improve cell adhesion. On the day of the experiment, cells were kept for 5 minutes at 10°C. Cooling of the cells to 10°C was performed on a ThermoMixer C (Eppendorf, Hamburg, Germany) to ensure reproducible temperature without harming the cells. Subsequently, the cells were washed once with Leibovitz's L15 medium without phenol red (21083027) and without FBS at 10°C, and diluted foamy virus particles were added. For these experiments, frozen cell supernatants containing virus particles were thawn on ice, centrifuged at 1200 rpm at 4 °C for 5 minutes to separate off virus aggregates and 5 μl virus supernatant, pipetted from the top of the liquid, was mixed with 200 μl L15 medium without phenol red and without FBS at 10°C. Initial experiments applied concentrated particles in medium containing FBS. Since both, the concentration and the FBS, lead to aggregation of the particles this was adjusted in later experiments of the study. To allow for virus attachment to the cell membrane while preventing virus uptake, the sample was immediately incubated at 10°C for 8 mins. Initially cells were cooled on ice to approximately 4 °C. Since cooling of the cells to 10 °C turned out to be sufficient to prevent virus uptake and lowered the risk of harming the cells this temperature was chosen in later experiments. To remove floating, non-attached particles, the cells were washed once with L15 medium of 10°C and 300 μl fresh L15 medium was added. The sample was directly transferred to the microscope stage, which had been pre-warmed to 37°C beforehand. A cell to with a suitable amount of virions bound, usually 40 - 80 particles, was selected for imaging data recording was started immediately.

Experiments were performed on the home-built spinning disk setup described in section 2.4.4. Virus particles were excited in parallel with lasers matching the excitation spectrum of the fluorescent proteins attached to the Gag and Env protein, respectively. The laser power was optimized such that photobleaching was minimized to allow maximum observation time of the virus particles while still maintaining sufficient signal-to-noise for virus particle tracking. To follow virus fusion in all three spatial dimensions as well as over time, *z*-stacks of single virus-decorated cells were recorded for up to 20 mins. The *z*-planes were spaced by 300 nm to ensure sufficient oversampling to achieve the maximum resolution determined by the optics of the microscope while minimizing sample photobleaching. The number of recorded *z*-planes was adjusted to cover

the entire cell volume and was usually between 20 and 30 planes. For the microscope setup, a frame transfer period of 36 ms was needed after every camera frame. To keep measurements comparable with previously conducted experiments, the camera exposure time was set to 130 ms resulting in a interframe time of 166 ms. Axial drift during data acquisiton was counteracted by the perfect focus system of the setup controlled by the microscope control software, which automatically readjusted the focus after a certain number of acquired *z*-stacks. The frequency of focus adjustment was determined empirically and usually a readjustment was performed every ten stacks. During focus readjustment, data acquisition was paused. Since the time needed by the microscope mechanics to readjust the focus was not regular, the microscope control software recorded timestamps for every camera frame. These timestamps were saved in a file along with the data. This file was loaded into the custom software used for the analysis of the obtained data. The data analysis will be explained in the following section.

3.2.2   Data analysis and the TrIC software

In the recorded movies, single virus particles were visible as diffraction-limited spherical signals, and for fully assembled virus particles the capsid and envelope signals colocalized in the two recorded channels. The custom-written image analysis applied to the data was based on tracking the signal of individual viruses in one channel while following the intensity in both channels for the single virus particles as they infect the recorded cell. In the data, such an event was visible as a transition from the Gag and Env fluorescence signals colocalizing while the virus particle is complete to both signals separating as the viral envelope fuses with a cellular membrane. Since we were aiming to investigate also the virus infection steps subsequent to membrane fusion, we tracked the capsid signal to follow single particles.

The applied image analysis algorithms are part of a MATLAB-software developed in our research group called Tracking Image Correlation (TrIC). It was invented by Dupont et al. [56] and the main underlying analysis algorithms were implemented by Dr. Aurélie Dupont (currently affiliated with the Université Grenoble Alpes, Grenoble, France). During the course of this thesis, my coworker Chen Qian and I revised and extended the software. The following briefly describes the workflow of a TrIC analysis, the underlying algorithms and the extensions that were implemented during the preparation of this thesis.

After starting the software, the user is presented with a graphical user interface (GUI) showing user controls on the right of the window and several data depiction panels on the left (Figure 3.1). The main inputs into the software are the acquired raw image data and the corresponding 2D tracks of the virus particles. At the beginning of my PhD, virus particles were tracked manually and the tracks were loaded into and analyzed by the TrIC software, one at a time. In order to increase statistics, and to avoid the time consuming, tedious step of manual tracking, I extended the software with the capability to load tracks obtained by the TrackMate ImageJ plugin [124]. Furthermore, a batch processing option enabling the loading of tracked data obtained on multiple cells was added.

To begin the analysis, the user has to enter the following information about the acquired data and the type of analysis to be done into the GUI (Figure 3.1, (1)):

1. Whether to run a 2D or 3D analysis

2. Whether to run the analysis on a single track or multiple tracks obtained on one cell or multiple tracks obtained on multiple cells

3. Whether to remove the first stack and the first frame of each stack (this option was implemented since an error in the microscope control software leads to an incorrect recording of those frames)

4. The number of *z*-slices recorded per *z*-stack (when chosing the multicell batch analysis option the user is asked for this information for every loaded cell)

5. The time interval between frames in seconds or the information that the data was recorded using the perfect focus system of the microscope (when the second option is chosen, the user will be askes to load a separate file containing the timestamps of every frame)

6. The channel in which the particle was tracked (generally, the capsid channel was chosen as explained above)

7. The *z*-step size in µm

8. The pixel size in µm

9. The minimum length a track needs to have to be included into the analysis in frames

Typically, the next step in the workflow is to generate a so called transformation matrix (Figure 3.1, (2)). Imperfect alignment of optical elements and optical aberrations lead to an unavoidable shift between the two recorded spectral channels. To still evaluate colocalization between the two channels with high accuracy, the TrIC analysis applies a geometrical transformation, stored in the form of a transformation matrix, to correct for the shift by transforming the coordinates obtained on the tracked channel to match the non-tracked channel. Of note, the raw image data is not altered in this step. Upon clicking the button "Generate transformation matrix", the user is asked to select a calibration image recorded in both channels. Usually, an image of the regular pattern of bright spots of the non-rotating Nipkow disk, which was recorded prior to data acquisition, is used for this purpose. Alternatively, an image of beads could be used. The underlying algorithm, implemented by my coworker Chen Qian, identifies bright features in the image based on intensity and area, and then determines their centroid position. Using a k-nearest neighbors search and an additional step of image cross-correlation spectroscopy, the corresponding peaks in both channels are identified and the geometric transformation needed to match the features is calculated.

Next, the user clicks the button "Select files for analysis" (Figure 3.1, (3)) and is asked to select:

1. The image data obtained on the microscope in the tracked and the untracked channel

2. A file containing the *xy*-coordinates of the individual tracks for each frame of the recorded image series obtained by manual tracking or the TrackMate plugin (these coordinates were obtained by calculating the maximum projection of the *z*-stacks in the recorded data and performing 2D tracking on the resulting image series.)

3. In case the data were recorded using the perfect focus system: a file containing the time stamps

4. The transformation matrix generated to correct for the *xy*-shift between the detection channels

5. The directory where the analysis results are to be saved

Loading of the files is started by clicking the button "Load files for analysis" (Figure 3.1, (4)), which might take a while, depending on the size of the image data set. The analysis is started by clicking "Run analysis" (Figure 3.1, (5)). Afterwards,

the software automatically executes a number of algorithms for every loaded track. These algorithms are briefly explained below. Analysis steps evaluating the data in 2D are generally performed on the maximum projection of every recorded z-stack. When the 2D analysis option is selected for the analysis, steps involving evaluation of the z-stack are skipped.

In the first step, the xy-coordinates of the track loaded into the software are refind by a least-squares fitting of the pixel intensity values within a 9 x 9 pixel box around the initial coordinate with a 2D Gaussian given by:

$$c_1 + c_2 \exp\left(-\left(\frac{(x-c_3)}{c_4}\right)^2 - \left(\frac{(y-c_5)}{c_6}\right)^2\right) \tag{25}$$

with:

c_1: offset

c_2: amplitude

c_3: x-coordinate of the Gaussain maximum

c_4: standard deviation in x

c_5: y-coordinate of the Gaussain maximum

c_6: standard deviation in y

The x- and y-coordinate of the Gaussian maximum are taken as the refined subpixel-precise location of the particle.

The following step takes a box of 7x7 pixels around the xy-coordinate loaded into the software and determines the z-plane of the z-stack with the highest mean pixel intensity within that box. A box in z of four planes around the determined z-plane (two planes above and two below) is taken and the mean pixel intensity values per plane (obtained from the 7x7 pixels large xy-box around the xy-coordinate loaded into the software) are fitted with a 1D Gaussian given by:

$$A \exp\left(\frac{-(x-m)^2}{2\,s^2}\right) \tag{26}$$

with:

A: amplitude

m: position of the maximum

s: standard deviation

The obtained value for *m* is taken as the precise z-coordinate of the track with sub-plane resolution. The precision with which the position of the particle can be determined in x, y and z (tracking precision) follows equation 13 and, hence, is dependent on the number of detected photons, the standard deviation of the detected signal, the background and the pixel size. The precision was estimated to be 30 nm in xy and 40 nm in z.

After refinement of the track in all three spatial dimensions, more specific characteristics of the track are determined. Since the different stages in virus infection are often characterized by their type of motion and speed, the software next calculates the instantaneous velocity of the particle by dividing the distance the particle moved between frames by the corresponding time interval.

The following steps constitute the core functionality of the analysis and are mainly aimed at determining the time point of virus fusion. The analysis monitors whether the particle is fully assembled along the track by determining the degree of colocalization between the fluorescence signals originating from the viral capsid and the viral envelope. Since colocalization analyses, based on intensity, require a high signal-to-noise ratio in order to give reliable results, the TrIC software evaluates

colocalization based on image cross-correlation. Here, the noise level plays less of a role since it does not correlate rendering this approach more sensitve. This is especially valuable for single particle tracking where data is often recorded with a low signal-to-noise ratio as the number of fluorophores per particle is comparably low and/or one aims to obtain signal from the particle for as long as possible and hence a low excitation power is applied to minimize photobleaching. The colocalization analysis starts by calculating the 3D image cross-correlation between both channels within a 21 x 21 pixels large region around the previously determined *xy*-particle coordinate for each *z*-stack of the data set. Subsequently, the maximum correlation amplitude for the entire stack is determined and its subvoxel accurate *xyz*-position is calculated by Gaussian fitting. The *xy*-coordinate is determined by fitting the *z*-plane containing the maximum correlation with a 2D Gaussian and the *z*-coordinate is determined by fitting the maximum correlation values of each plane with a 1D Gaussian. All correlation values are normalized to 1. In the analysis, the maximum amplitude of the image correlation as well as its shift from the center, the 0,0,0-lag are evaluated for every stack of the recorded movie. If the maximum of the cross-correlation amplitude is located near the origin, this indicates that the capsid and envelope signal are colocalizing and hence indicates that the virus is fully assembled (Note that this assumption does not always hold true. For example for viruses with an elongated envelope the particle might still be fully assembled although the capsid and envelope signals are not colocalizing. On the other hand a virion might not be fully assembled but the capsid and envelope might still colocalize.). A shift of the cross-correlation maximum from the origin, however, represents a shift between the fluorescent signal of the underlying images. The amount of the shift of the cross-correlation function corresponds to the amount of shift between the images. During the TrIC analysis, this effect becomes apparent when the capsid and envelope signal separate while still both being located within the observation box and it allows one to determine the relative distance between the two signals. The accuracy with which the distance between the signals can be determined was evaluated by tracking a fusion incompetend virus mutant. The average measured distance between the capsid and envelope signal over a 10 mins track was 46 nm with a standard deviation of 20 nm resulting in a 95% confidence boundary at 86 nm. Based on this a relative distance, a threshold, of 100 nm was set for signal separation. The direction in which the maximum is shifted additionally gives information about the direction in which the signals are shifted in the image.

To determine the level of non-specific correlation, the *z*-stack in the tracked channel is additionally correlated with a *z*-stack containing the randomly redistributed pixel values of the *z*-stack in the non-tracked channel.

Virus fusion within the TrIC analysis was defined as the time point at which the relative distance between the capsid and the envelope signal was larger than 400 nm (our significance level for signal seperation plus 300 nm). The time point at which the envelope signal moves out of the observation box of 21x21 pixels can be identified from the cross-correlation amplitude which shows a clear drop at this time point. To clearly define a time point, the TrIC analysis takes the maximum correlation amplitude obtained for non-specific correlation plus three times the standard deviation (determined from all non-specific maximum correlation values of the track) as the threshold below which the cross-correlation amplitude has to drop.

As the last step of the analysis, the background-corrected intensity of the particle along the track is determined in both channels. This information can, for example, be helpful to evaluate the amount of photobleaching during the track. Furthermore, for a fusion event, a clear drop in the intensity measured for the envelope should be observable as the fluorescent protein is diluted in the cellular membrane. For background correction, a 9 x 9 pixels large region around the subpixel-accurate *xy*-position of the particle is evaluated, the average pixel intensity within a circle of radius two pixels around the center is calculated and corrected by subtracting the average pixel intensity of the remaining pixels inside the cutout which is taken as the background. After completing the analysis the program saves the results as a table in a .txt file.

As a large number of tracks is generated by the above outlined analysis, I implemented several options to review the analysed tracks in the software. Clicking the button "Open analysis" (Figure 3.1, (6)) enables the user to load the results of a previous analysis session and the program now offers the possibility to load multiple tracks into a list (Figure 3.1, (7)). By clicking on the list of loaded tracks, the user can chose which track to review and the corresponding analysis results are shown in the data windows of the software. To facilitate looking at a particular time point of the track in all data windows simultaneously, an indicator was implemented into the software which can be moved using a slider bar at the bottom of the GUI (Figure 3.1, (8)). A clipout from the raw movie showing the tracked virus in both channels as well as a merge allows the user to visually review the track (Figure 3.1, (9)). For presentation purposes, this cutout of the entire virus track can also be exported as a movie sequence in the TIFF and AVI format (Figure 3.1, (10)). Above the clipout, a window depicts the complete field of view obtained during data acquisition highlighting where the currently reviewed particle is located with respect to the entire cell (Figure 3.1, (11)). It enables the user to visually control whether the particle was correctly tracked. An overview movie showing the tracks of all tracked particles can also be exported (12). The button "Show full plot" opens a separate window showing a summary of the analysed track (Figure 3.1, (13)).

From the TrIC analysis of the foamy virus data sets, Dr. Aurélie Dupont and coworkers were able to identify an intermediate step in foamy virus fusion which is characterized by the capsid and envelope signal being separated by 100 - 400 nm while still being tethered together. The corresponding publication is attached to this thesis [125] and summarized in the following section.
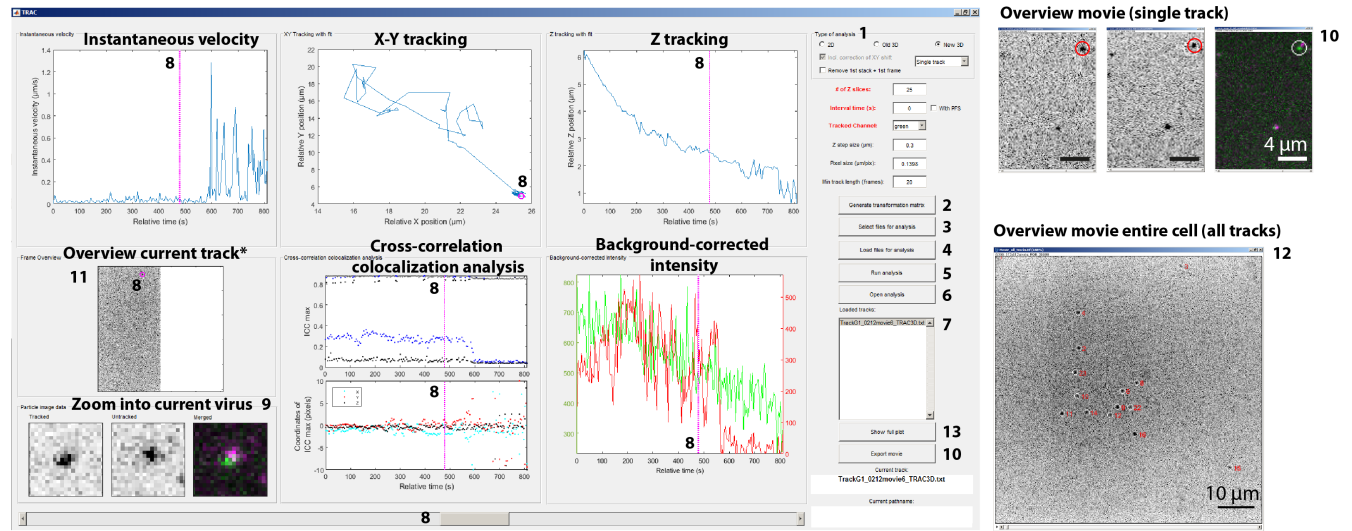
* The version of the software in September 2021 allows loading of 256 x 256 as well as 512 x 512 pixel movies

Figure 3.1: **Graphical user interface of the updated Tracking Image Correlation (TrIC) software**
The user input to the software are located on the right of the interface. In the center and on the left of the screen, the analysis results are visualized and can be reviewed by the user. To start the analysis the user has to enter several pieces of information into the software (1). First, the analysis type needs to be selected. The analysis can be performed in 2D and in 3D. It can be performed on a single track, multiple tracks from the same cell or multiple tracks from multiple cells. Furthermore, the following parameters need to be entered into the software: the number of *z*-slices recorded per stack (in case multiple cells are analyzed this information needs to be entered for every cell), the interval time between the frames of the recorded image series, whether the data was recorded with the perfect focus system enabled, whether the particle was recorded in the capsid or the envelope channel, the distance between the planes of the recorded *z*-stacks, the pixel size within the recorded image series and the minimum length a track should have in order to be included in the analysis. A transformation matrix correcting for the spatial shift between the channels can be generated by clicking the button "Generate transformation matrix" (2). The transformation matrix enables mapping of the coordinates in the tracked channel to the corresponding positions in the untracked channel. The data files needed for the analysis are selected by clicking the button "Select files for analysis" (3) and the data is loaded by clicking the button "Load files for analysis" (4). Clicking the button "Run analysis" starts an analysis run on the loaded data (5). The results of an analysis can be loaded by clicking the button "Open analysis" (6). In the box (7) below the button "Open analysis" the user can switch between the currently loaded tracks and the corresponding data is visualized by the software. With the slider bar at the bottom of the GUI the user can move on indicator (8) in every data depiction panel and inspect the corresponding frames of the raw data in panel 11 and 9. Panel 11 depicts the location of the current particle with respect to the entire recorded field-of-view; Panel 9 shows clipouts from the image series depicting the particle in the tracked and the untracked channel as well as an overlay of both channels; The "Export movie" button allows the user to export a movie of the selected track in which the corresponding particle is encircled (10). The panels in the center and on the left of the interface visualize the following particle parameters: top left: the instantaneous velocity against the time since the beginning of the track; top center: the subpixel precise *xy*-track; the subvoxel precise *z*-position versus the time since the beginning of the track; left center: the corresponding frame of the image series with the particle encircled; bottom center: the amplitude and the location of the cross-correlation function versus the time since the beginning of the track; bottom right: the background-corrected intensity in the tracked (green) and in the untracked channel (red). An overview figure depicting the corresponding results can be generated by clicking the "Show full plot" button (13). A movie in which all analyzed tracks are encircled and numbered (12) is saved in addition to the other results of the analysis at the end of each analysis run.

3.3 Summary of the paper

Paper to be found in appendix B, section B.1

In the following the main results of the paper underlying this chapter of the thesis are summarized. The figure references given reference the figures in the paper which can be found in the appendix A, section B.1. The study identifies an intermediate step in the fusion process of the foamy virus by single virus tracking. During this intermediate step, which lasts for several minutes, the capsid and the envelope of the virus particle are separated by a distance between 100 nm and 400 nm while still being tethered together.

The understanding of the fusion process of the foamy virus with its host cell remains incomplete until today. The study reported by Dupont et al. investigates the fusion process by following single fluorescently-labeled particles as they infect the cell. For this an experimental approach combining spinning-disk confocal microscopy and single-particle software tracking is taken.

In the first part of the paper the applied virus particles are characterized. In addition to fusion at the plasma membrane it has been shown that foamy virus particles can also be taken up via endocytosis [126, 106] where they encounter a decrease in pH within the endocytic pathway from early endosomes (pH ~ 6.5), to late endosomes (pH ~ 5.5), to lysosomes (pH ~ 4.5). The PFV and SFVmac particles used in this study were labeled with eGFP at the Gag protein, which is part of the foamy virus capsid, and mCherry at the Env protein, which is part of the virus envelope (Figure 1 a) - e)). As it is known that eGFP fluorescence is quenched at acidic pH values [127], the stability of the fluorescent proteins incorporated into the virus particles under these conditions was tested (Appendix Figure A6). By monitoring the fluorescence intensity of particles containing both Gag-eGFP and mCherry-Env at different pH values, it was observed that eGFP shows strong quenching at pH 5.5 whereas the mCherry intensity stayed largely unaffected. In order to minimize the influence of quenching on the fluorescence signals the Env protein was labeled with mCherry since it might be directly in contact with the acidic environment of the endosome/lysosome during virus uptake whereas the eGFP label was placed at the capsid where it is shielded from the potentially acidic environment during virus uptake. Since the study aims at investigating the fusion process of the virus particles, the infectivity of the particles was addressed next. A previous study showed that 100% labeling of Gag with eGFP reduces the infectivity of the particles by an order of magnitude but that the infectivity can be recovered to almost wildtype levels by mixing wildtype Gag with Gag-eGFP at a ratio of 3:1 [126]. Consequently, these particles were chosen for the experiments performed. In order to be able to detect virus fusion on a single particle level, a high labeling efficiency of the virus particles is needed. For this reason, the percentage of green particles colocalizing with red and the percentage of red particles colocalizing with green was determined. It was found that 93 +/- 1% of the particles with a green labeled capsid also contained an Env-mCherry signal and about 50% of the Env-mCherry labeled particles showed an Gag-eGFP signal [126]. It is known that the foamy virus cannot bud without the envelope, [128, 116] whereas subviral particles without capsid exist, [129] hence these observations were not unexpected.

Investigations of foamy virus fusion on a single particle level comprise the main part of the study and use the characterized virus particles. These experiments were conducted and the data was analyzed as described in sections 3.2.1 and 3.2.2, respectively and illustrated in the paper in Figure A1 - A3. For PFV Env-mediated fusion, an exemplary fusion event most likely taking place at the plasma membrane is presented (Figure 2). In this context, the study reports an approach to extract

the location of the cell surface in the *z*-dimension from the *z*-positions of all particles in the beginning of the recorded image series (Figure A5). The depicted close ups from the recorded image series of the fusion event, in combination with the image cross-correlation data, show that the particle is intact at the beginning. Subsequently, color seperation is observed in the image series accompanied with a decrease in the cross-correlation amplitude. An increase in the relative distance between the fluorescent signals and a sudden drop in the measured capsid fluorescence suggest fusion of the particle. These observations coincided with a sudden increase in the instantanous velocity of the capsid indicating active transport of the capsid inside the cell. Interestingly, an intermediate step in the fusion process was observed. It is characterized by an increase in the relative distance between the envelope and the capsid of approximately 300 nm and lasts for about 5.5 mins before the capsid and the envelope fully separate. Plotting the relative position of the Env signal with respect to the Gag signal shows that the envelope signal is moving around the capsid during the intermediate step.

As a comparison to the observations made for PFV Env-mediated fusion, the study presents a fusion event mediated by SFVmac Env (Figure 3). Here, the particle is taken up into the cell before fusion takes place as indicated by directional motion towards the nucleus before seperation of the Gag and Env signal. The time point of the fusion event can be clearly identified based on the depicted close ups from the recorded image series in combination with the TrIC analysis and the fluorescence intensity of the particle in both channels. Strikingly, also for this event, an intermediate step with similar characteristics to the fusion events for PFV is observed.

Comparing all observed fusion events (Figure 4) it can be concluded that PFV Env-mediated fusion takes significantly longer at the plasma membrane than when it occurs from the endosome. For SFVmac Env-containing particles only fusion events from endosomes were observed. The observation that PFV Env-mediated fusion but not SFVmac Env-mediated fusion was observed at the plasma membrane is in agreement with previous observations showing that PFV Env can already fuse at neutral pH [126, 106] whereas SFVmac Env-mediated fusion strongly depends on low pH [126]. A comparision of all endosomal fusion events suggests that endosomal fusion takes longer if it mediated by SFVmac than if it is mediated by PFV Env.

Using a mean-squared displacement analysis on the the relative trajectory of envelope with respect to capsid and on the absolute trajectories of the Gag-eGFP and mCherry-Env signal reveals that the two virion components seem to be tethered during the intermediate phase (Figure 5). The study reports on experiments aimed at elucidating what could tether the viral components together. To this end, viral RNA was investigated as a possible tether however the experiments suggested that viral RNA is not the molecule tethering the components.

3.4   Outlook

Future experiments should investigate what causes the observed intermediate step. Here different viral as well as cellular components have to be investigated for which multicolor fluorescence microscopy seems to be the method of choice. To get insight into the processes at higher resolution one could imagine to stop the fusion process during the intermediate step and image the sample by correlative light and electron microscopy [130]. Furthermore future experiments will elucidate when the intermediate step takes place with respect to the other known steps in foamy virus fusion. For these experiments, the viral membrane could be, for example, stained with DiD to pin point the time point of lipid mixing as it was previously done for the influenza virus [49]. Content mixing could be visualized by incorporation of a content marker into the virion. Those investigations would need data acquisition and analysis in three fluorescent channels, which are possible on the described home-modified spinning-disk system and the described software would be easily extenable to three channels. Furthermore,

the above described data acquisition and analysis workflow could be applied to study other viruses in the future.

4 Super-resolution single molecule localization microscopy of ASC in the classical NLRP3 inflammasome pathway

Publication underlying this chapter (to be found in appendix B, section B.2):

Nanoscale organization of the endogenous ASC speck by Ivo M. Glück, Grusha Primal Mathias, Sebastian Strauss, Thomas S. Ebert, Che Stafford, Ganesh Agam, Suliana Manley, Veit Hornung, Ralf Jungmann, Christian Sieben, Don C. Lamb. *Manuscript submitted to Nature Cell Biology.*

4.1 Biological background

The innate immune system is the first line of defense of the body. Important molecules of the innate immune system are so called pattern recognition receptors (PRRs) which upon engagement trigger central signalling cascades of the immune response. Molecular entities triggering the activation of the innate immune response by binding to PRRs are summarized under the terms pathogen-associated molecular patterns (PAMPs) and danger-associated molecular patterns (DAMPs). Inflammasomes are multiprotein complexes that assemble e.g. in macrophages, monocytes and dendritic cells, after activation of particular intracellularly-located PRRs [131]. Key components of the majority of inflammasomes are a sensor protein that becomes activated by PAMPs or DAMPs, the adaptor protein apoptosis-associated speck-like protein containing a caspase activation and recruitment domain (ASC) [132, 133] and the inflammatory caspase 1 [134, 135, 136]. Inflammasome activation triggers the recruitment of these proteins into a single micrometer-sized complex per cell [137]. The associated proximity of caspase-1 molecules drives their activation by self-cleavage and activated caspase-1 results in release of the proinflammatory cytokines interleukin (IL)-1β and IL-18, and additionally leads to pyroptosis, a proinflammatory form of cell death. This process has been shown to be an all-or-nothing response meaning that, once activated, it always runs until completion leading ultimately to cell death.

The assembly of ASC molecules was first described in 1999 by Masumoto et al. who coined the term "ASC speck" [133]. Interaction of the named proteins within the inflammasome complex is mediated via homotypic domain interaction of the caspase activation and recruitment domains (CARDs) [138] and Pyrin domains (PYD) [139], respectively. Due to their complementarity in structure and charge these have a high tendency to bind to each other [140, 141, 142, 143]. How these proteins are organized within the supramolecular inflammasome complex has been the subject of intensive research since their discovery. The study underlying this chapter investigates the ASC speck. ASC is a 22 kDa protein composed of 195 amino acids [133, 132]. The solution crystal structure of the full-length, human protein, solved in 2009, showed that the C-terminal CARD and the N-terminal PYD domain are folded into the typical structure of a death fold and are connected via a 23 amino acid long semi-flexible linker [144]. Death folds are tertiary structure motifs found in proteins involved in signaling cascades leading to cell death and inflammation [145]. These motifs have the above mentioned tendency to interact via homotypic domain interaction due to there complementarity in structure and charge. At low pH or in chaotropic solution, ASC is soluble. However, at physiological pH, the protein can oligomerize into filaments as has been shown by in vitro experiments applying solid-state NMR spectroscopy and electron microscopy (EM) [146, 147, 148, 149, 150, 151]. Experiments performed by purifying endogenous ASC from inflammasome-activated cells revealed star-shaped assemblies [137], isolated filaments or amorphous aggregates potentially composed of intertwined filaments ([149, 152]). To investigate

and visualize the interactions within the speck inside cells, numerous studies applied overexpression of ASC tagged with a fluorescent protein. The resulting speck structure is usually much larger than 1 μm and, occasionally, filaments protruding from the edge of the structure were observed [150, 152, 153, 154]. Introducing mutations at particular residues into the interacting domains of the protein disassembled this structure into filaments [153, 155, 140, 150].

The endogenous ASC speck inside cells is frequently visualized by immunofluorescence as a single perinuclear spot of about 1 μm in diameter [156, 157]. Expression of a GFP-labeled nanobody directed against the CARD domain of ASC stabilized a filamentous speck intermediate inside inflammasome-activated cells [158]. To date, there are only a few studies that visualized the endogenous, unperturbed structure inside cells at higher resolution either by electron microscopy [159] or super-resolution fluorescence microscopy [160]. They mostly resolve the speck as an amorphous aggregate potentially made up of intertwined filaments and a model of the inflammasome as an overall amorphous assembly has been proposed [149, 158, 161]. However, in contrast to the previously mentioned model numerous studies depict and describe the ASC speck as a hollow, ring-shaped complex including the study describing the ASC speck for the first time [133, 136, 153, 162, 163, 164, 165, 166, 157, 167, 168, 169]. Based on this observation alternative models for the ASC speck and the entire inflammasome complex have been proposed [166, 157, 170, 171, 172, 173]. Despite structural investigations at various levels of resolution and complexity, the supramolecular structure of the endogenous speck has not yet been systematically addressed. Due to the high tendency of the involved proteins to interact, it still remains to be confirmed whether the endogenous structure differs from the results obtained by *in vitro* and overexpression experiments.

The experiments underlying this chapter investigate the endogenous ASC speck formed after NOD-, LRR- and pyrin domain-containing protein 3 (NLRP3) inflammasome activation with lipopolysaccharide (LPS) [174] and the bacterial toxin nigericin [175]. They were performed in two different cell types: THP-1 cells, which is a monocytic leukemia cell line, that can be differentiated into macrophage-like cells [176, 177], and BlaER1 cells, which is a B cell-derived cell line that can be transd-ifferentiated into macrophage/monocyte-like cells [178]. Both cell lines express the inflammasome proteins endogenously. In the applied BlaER1 cells ASC was endogenously tagged with TagRFP. By systematically imaging a large number of endogenous specks by widefield, confocal and super-resolution localization microscopy the experiments provide new insights into the above mentioned open questions. In THP-1 cells two different labeling approaches were taken to label endogenous ASC: a primary monoclonal ASC antibody in combination with a secondary F_{ab} fragment, and a nanobody directed against ASC were applied. In BlaER1 cells endogenous ASC was labeled by a primary monoclonal ASC antibody in combination with a secondary antibody, as well as by a nanobody directed against TagRFP. Imaging by dSTORM and DNA-PAINT revealed that the speck is composed of a dense core and a less dense periphery. Steric exclusion of the larger antibody from the dense center of the structure thereby can lead to the appearance of the speck as an overall ring-like assembly. By sorting the specks into a quasi-temporal sequence the experiments furthermore suggest that the endogenous speck mainly becomes denser but only marginally larger during its formation.

4.2   Additional results

This section provides experimental results, that are not given in the attached paper.

4.2.1   Evaluation of the anti-ASC staining procedure using a primary antibody and a secondary F_{ab} fragment applied on THP-1 cells

In the experiments underlying this chapter ASC speck formation in cells was induced by activating the NLRP3 inflamma-some. Subsequently the cells were fixed and stained against ASC by immunofluorescence. As previously mentioned NLRP3 inflammasome activation leads to caspase-1-driven cell death in the form of pyroptosis. Pyroptosis was expected to reduce the number of ASC specks in the sample since cell death leads to a detachment of the cells from the microscopy slide. For this reason caspase-1 was knocked out to prevent the initiation of pyroptosis by the enzyme in the THP-1 cells applied for the experiments. To assess the quality of the staining of endogenous ASC in THP-1 cells several control stainings were included in every repetition of the experiments. These controls also included THP-1 cells in which ASC was knocked out. Both knock-out cell lines, THP-1 caspase-1 KO and THP-1 ASC KO, were generated by Dr. Che Stafford from the group of Professor Veit Hornung in the Biochemistry Department at the Ludwig-Maximilians-Universität, München, Germany. The quality of the staining was evaluated by recording snap shots of the cell preparations on a widefield microscope at 10x magnification. To quantify the results, the background-corrected integrated widefield intensity of individual cells was extracted using a self-written pipline for Cell Profiler [179]. Figure 4.1 shows the results of this analysis. Of note, these results have to be evaluated with care since there were some technical problems leading to uneven illumination of the sample during the recording of the data.

Unstimulated cells and cells stimulated only with LPS showed an elongated morphology typical for the macrophage-like phenotype of differentiated THP-1 cells. The extracted integrated intensity in the Alexa Fluor 647® channel is constant between these two treatments which suggests that also the ASC content in the cells is constant. Cells treated with LPS and nigericin showed a rather round phenotype as can be clearly seen from the recorded brightfield images. These cells exhibited a reduced integrated intensity in the Alexa Fluor 647® channel compared to unstimulated cells or cells only stimulated with LPS. This might indicate that the ASC content in these cells is reduced. In a subset of the cells an ASC speck was observed. Cells containing an ASC speck exhibited reduced ASC staining in the cytoplasm which is indicative of the redistribution of the protein into the speck. There is no difference observable in the extracted mean integrated intensity per cell regardless of whether cells with speck are included in the analysis or not. This might indicate that speck formation does not lead to a loss of the ASC protein from the cell. However due to the limited number of cells showing a speck it is questionable whether a difference would be observable in the mean integrated intensity per cell evaluated over all measured cells. A separated analysis of the cells exhibiting a speck and the cells without speck would be more informative in this case.

Control preparations to which only the secondary antibody was applied or which were left unstained showed no signal in the Alexa Fluor 647® channel. This shows that the level of unspecific binding of the secondary antibody as well as the level of autofluorescence is very low after application of the staining procedure. To further control the specificity of the staining it was applied to THP-1 ASC KO cells. Evaluating the signal in the Alexa Fluor 647® channel on unstimulated cells, a low level of intracellular signal was observed. A more pronounced signal was observed after treatment of the cells with LPS and nigericin. If no staining was applied to these cells no intracellular staining was detected. These results indicate that the applied staining procedure leads to a low level of binding of the applied antibodies in THP-1 ASC KO cells which is stronger in cells treated

with LPS and nigericin. Since this signal was not detected to such a large extent when imaging the cells at 60x magnification the observations made here are most likely due to the above mentioned technical issues during the recording of the data.
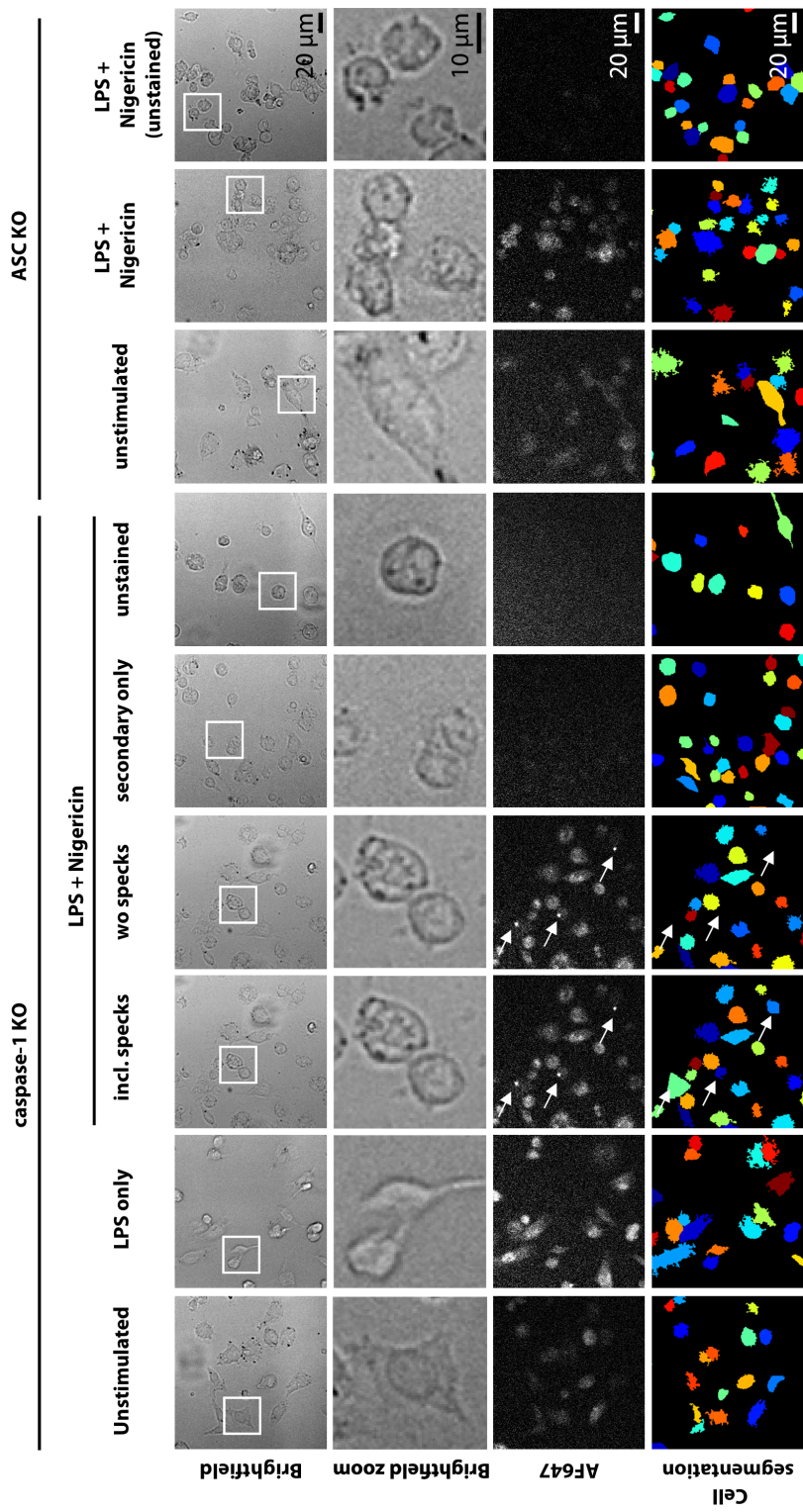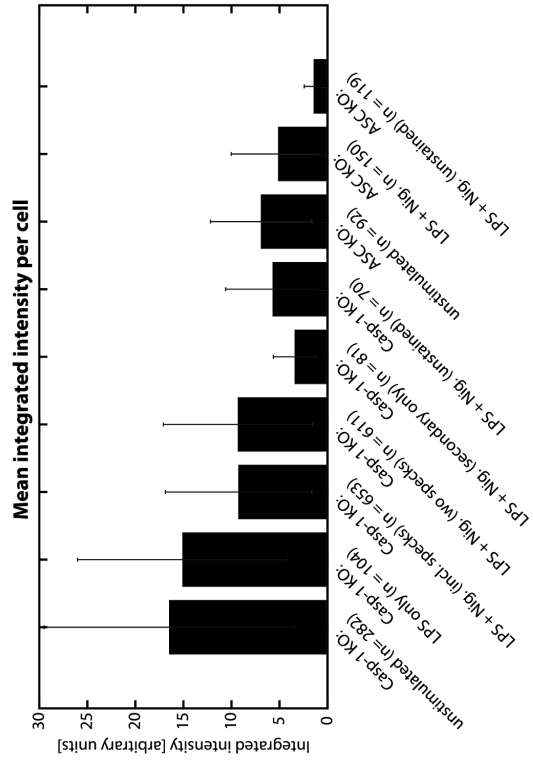
Figure 4.1: **Quantification of the anti-ASC staining procedure using a primary antibody and a secondary F$_{ab}$ fragment applied on THP-1 cells**

81

Continued Figure 4.1: A) Brightfield (first and second row) and anti-ASC Alexa Fluor 647® fluorescence (third row) snap shots of the THP-1 cell samples prepared during the experiments and the corresponding cell segmentations (fourth row).

The recorded brightfield images allow assessment of the cell morphology. The marked areas in the first row correspond to the zooms in the second row. Unstimulated cells or cells only stimulated with LPS show an elongated phenotype whereas cells treated with nigericin have a rather round morphology. Application of the anti-ASC staining procedure to caspase-1 KO cells results in cytoplasmic staining as observable in the depicted snap shots recorded in the Alexa Fluor 647® fluorescence channel. Treatment of the cells with LPS and nigericin leads to the formation of ASC specks in a subset of the cells (arrows). Controls in which only the secondary antibody or no antibody were applied exhibit no detectable signal in the Alexa Fluor 647® fluorescence channel.

ASC KO cells to which the staining procedure was applied show an intracellular signal in the Alexa Fluor 647® fluorescence channel. Since this signal was not observed to such an extent on the same preparations recorded at 60x magnification it is most likely due to technical problems during the recording of the data rather than staining of cell components different from ASC.

The depicted segmentations of individual cells were applied to the snap shots recorded in the fluorescence channel and are color-coded arbitrarily. For LPS + nigericin-treated caspase-1 KO cells the mean integrated intensity per cell was evaluated including the cells exhibiting a speck (column "incl. specks") and excluding cells exhibiting a speck (column "wo specks").

B) Corresponding quantification of anti-ASC staining as integrated intensity measured per cell; error bars indicate the standard deviation calculated from the integrated intensity of individual cells.

4.2.2   Imaging the endogenous ASC speck in THP-1 cells by single- and dual-color dSTORM

After verifying the specificity of the established anti-ASC staining by diffraction-limited microscopy at low magnification the samples were imaged by dSTORM at super-resolution. Before recording of the blinking data a diffraction-limited snap shot of the cells was recorded. Figure 4.2 below depicts a subset of the ASC specks imaged in THP-1 cells and shown in the attached manuscript. Endogenous ASC was stained with primary ASC antibody + secondary Alexa Fluor 647$^{®}$-labeled F_{ab} fragment. Of note, some of those specks exhibit a hollow center. To evaluate whether this hollow center might be due to the labeling probe being sterically excluded from the dense center of the structure, we combined the antibody staining with a much smaller anti-ASC nanobody. For that we performed dual-color immunofluorescence labeling against ASC using the primary anti-ASC antibody in combination with the secondary Alexa Fluor 647-labeled F_{ab} fragment and the nanobody conjugated with DyLight 755$^{®}$ on the same sample. After verifying the specificity of the labeling by diffraction-limited fluorescence microscopy we performed dual-color dSTORM. Before recording of the blinking data a diffraction-limited snap shot of the cells in the Alexa Fluor 647 channel was recorded. A subset of the recorded specks is depicted in Figure 4.3. The obtained data confirmed the observation that some of the specks shows a less dense center after antibody staining. Furthermore it revealed that the nanobody is located at the center of the speck but is very rarely localized at the periphery of the structure. These results support the hypothesis that the larger antibody is sterically excluded from the dense center of the complex whereas the much smaller nanobody can penetrate into these regions. The lower labeling density of the nanobody at the periphery of the structure can be explained by the low binding affinity of the nanobody (~ 159 nM) due to which it is washed away from the less dense periphery of the speck [158].
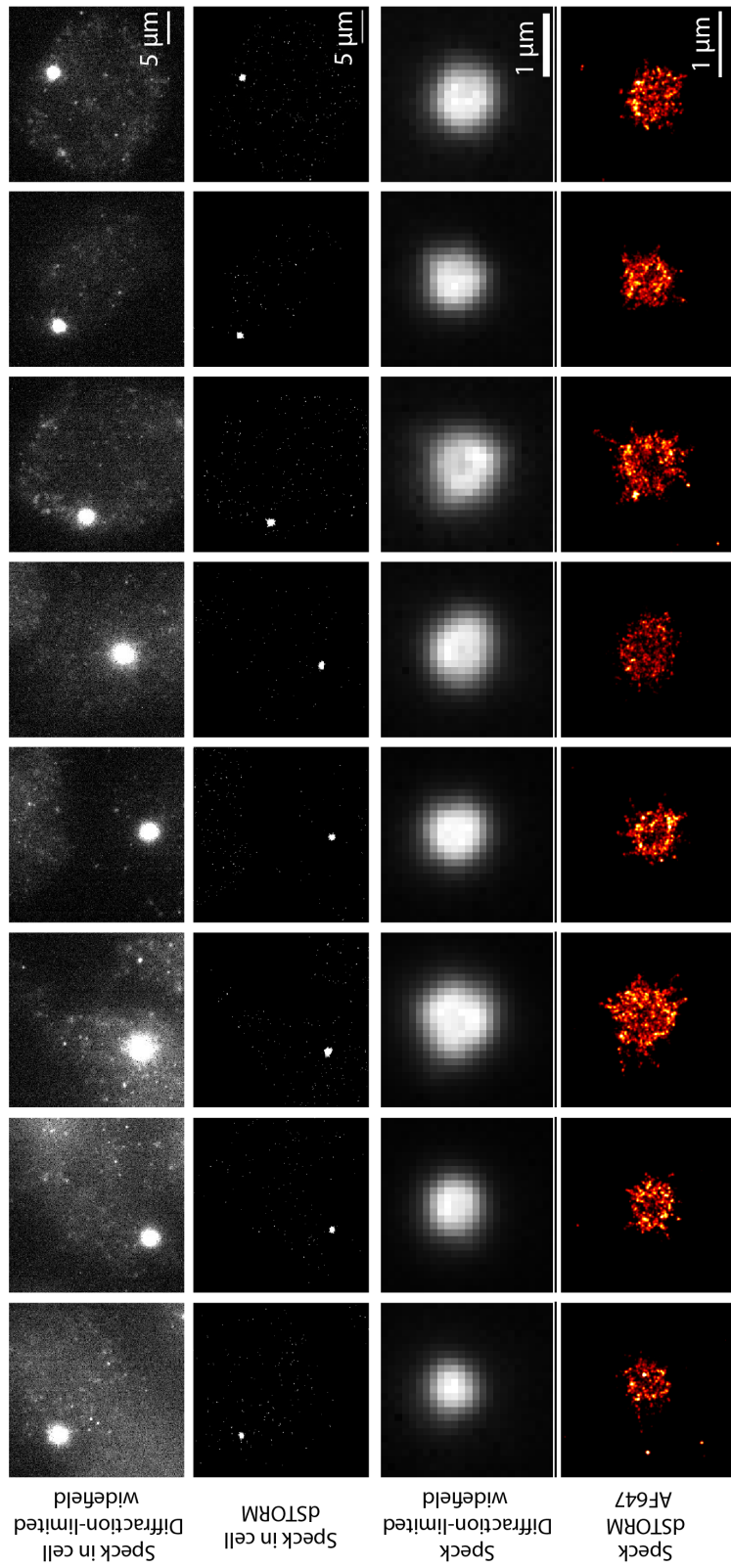
Figure 4.2: ASC specks in THP-1 cells stained with primary anti-ASC antibody and secondary Alexa Fluor 647®-labeled F_ab fragment

A subset of the recorded specks in THP-1 cells stained with primary monoclonal anti-ASC antibody and secondary Alexa Fluor 647-labeled F_ab fragment is shown. Different representations of the same speck are shown in each column. First row: Diffraction-limited widefield images of THP-1 cells containing a speck. Second row: 2D dSTORM reconstruction of the same field-of-view shown in the first row. The reduced background compared to the diffraction-limited images is clearly visible. Third row: Diffraction-limited close up of the speck. Fourth row: 2D dSTORM reconstruction of the speck. The speck exhibits an amorphous structure characterized by a dense core from which occasionally filamentious extensions reach out into the periphery. Some of the specks exhibit a less dense center.
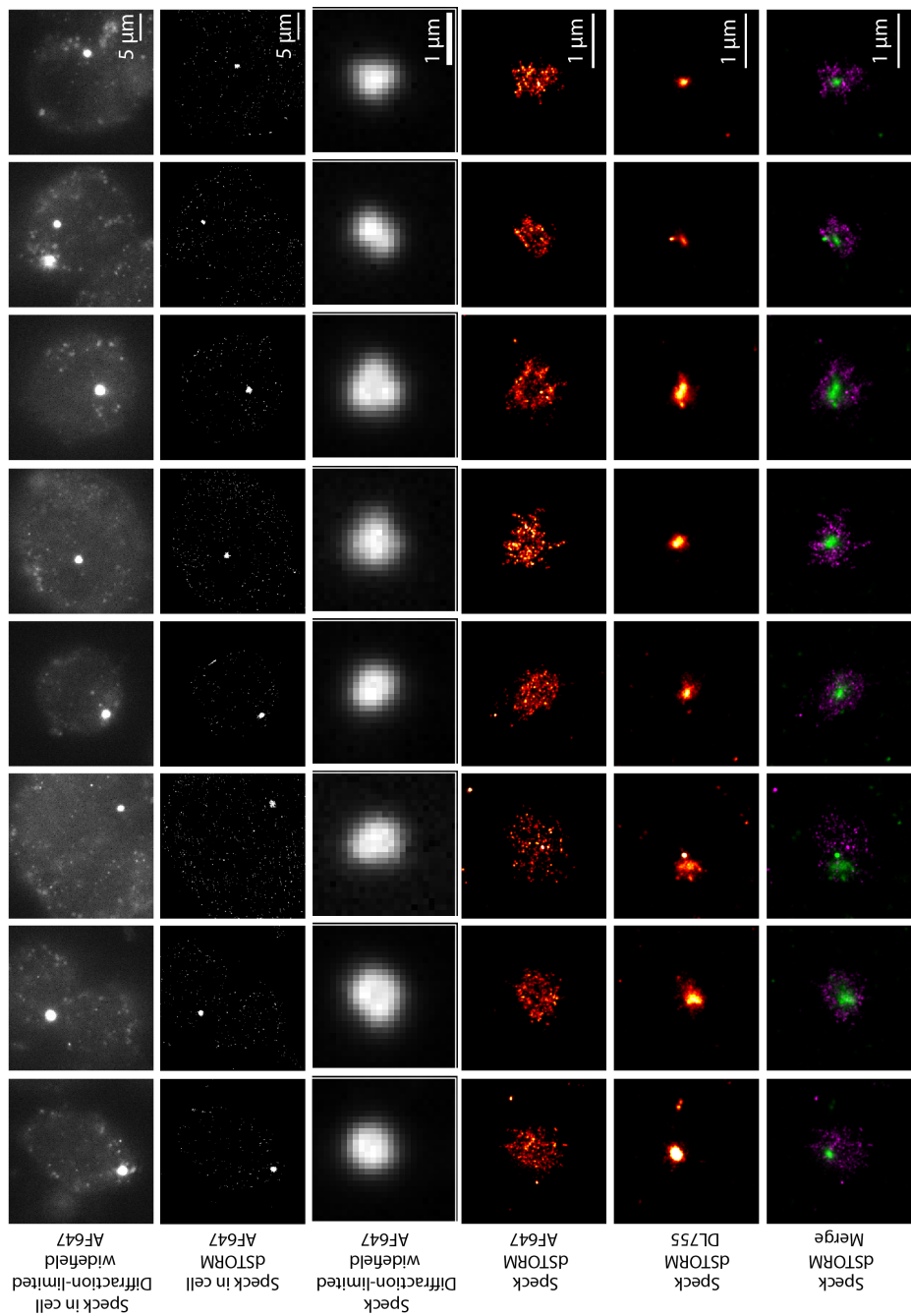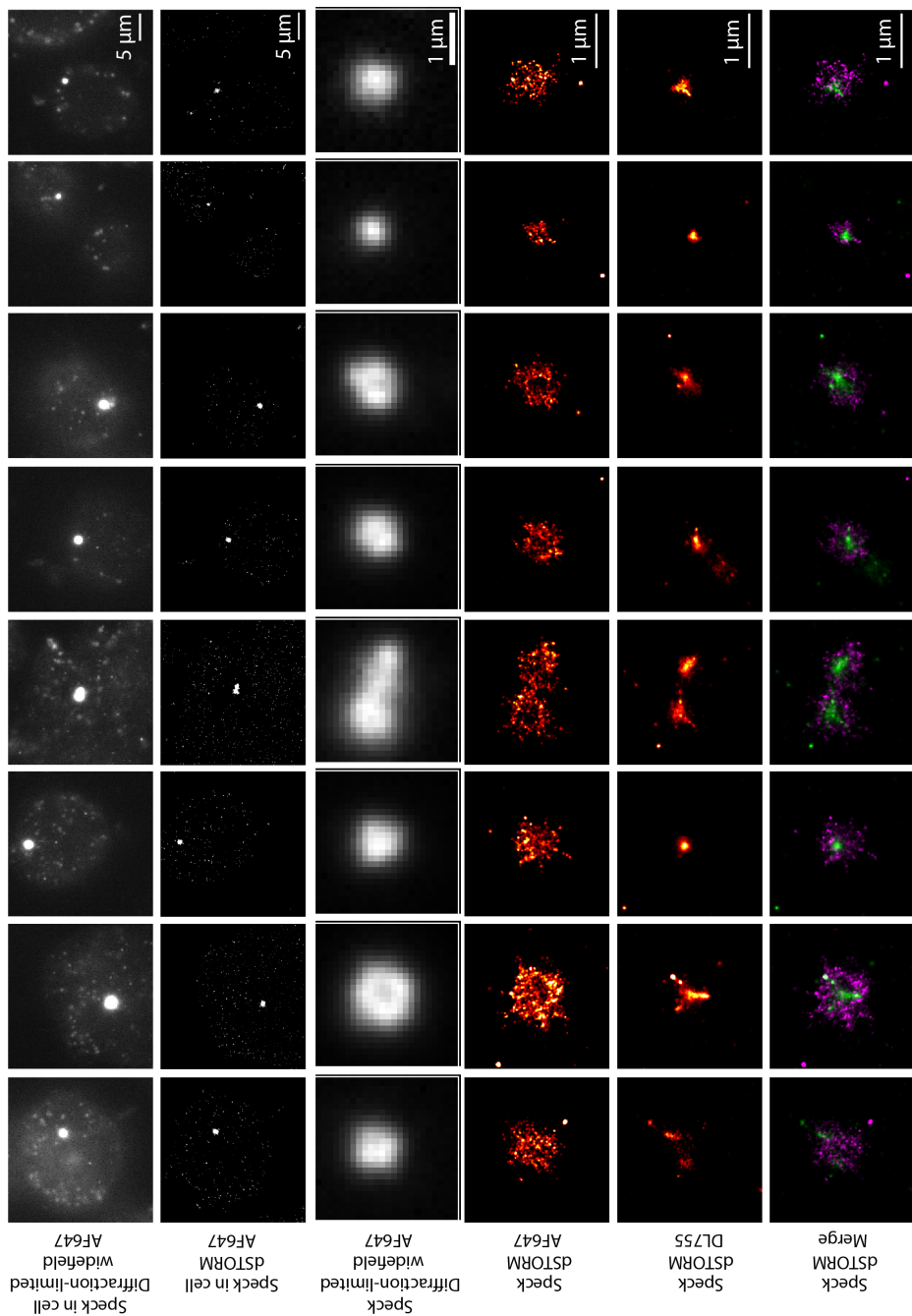
AF647: Alexa Fluor 647®

Figure 4.3: ASC specks in THP-1 cells stained with primary anti-ASC antibody and secondary Alexa Fluor 647®-labeled Fab fragment and with DyLight 755®-labeled anti-ASC nanobody

A subset of the recorded specks in THP-1 cells stained with primary monoclonal anti-ASC antibody and secondary Alexa Fluor 647®-labeled Fab fragment and DyLight 755®-labeled anti-ASC nanobody is shown. Different representations of the same speck are shown in each column. First row: Diffraction-limited widefield images of the cells recorded in the Alexa Fluor 647® channel corresponding to the antibody staining. Second row: 2D Alexa Fluor 647® dSTORM reconstruction corresponding to the antibody staining of the same field-of-view shown in the first row. The reduced background compared to the diffraction-limited images is clearly visible. Third row: Diffraction-limited close up of the speck recorded in the Alexa Fluor 647® channel corresponding to the antibody staining. Fourth row: 2D Alexa Fluor 647® dSTORM reconstruction, corresponding to the antibody staining, of the speck. The speck exhibits an amorphous structure characterized by a dense core from which occasionally filamentious extensions reach out into the periphery. Some of the specks exhibit a less dense center. Fifth row: 2D DyLight 755® dSTORM reconstruction, corresponding to the nanobody staining, The smaller size of the dSTORM reconstruction compared to the signal obtained in the Alexa Fluor 647® channel is clearly visible. It results from the lower binding affinity of the nanobody which leads to incomplete labeling of the low-density periphery of the speck. Furthermore decreased resolution of the DyLight 755® reconstructions resulting from the lower number of photons emitted by the dye is clearly visible. Sixth row: Merge of the dSTORM reconstructions obtained from the Alexa Fluor 647® and the DyLight 755® signal. AF647: Alexa Fluor 647®; DL755: DyLight755®

85

Continued Figure 4.3: ASC specks in THP-1 cells stained with primary anti-ASC antibody and secondary Alexa Fluor 647®-labeled F$_{ab}$ fragment and with DyLight 755®-labeled anti-ASC nanobody

A subset of the recorded specks in THP-1 cells stained with primary monoclonal anti-ASC antibody and secondary Alexa Fluor 647®-labeled F$_{ab}$ fragment and DyLight 755®-labeled anti-ASC nanobody is shown. Different representations of the same speck are shown in each column. First row: Diffraction-limited widefield images of the cells recorded in the Alexa Fluor 647® channel corresponding to the antibody staining. Second row: 2D Alexa Fluor 647® dSTORM reconstruction corresponding to the antibody staining of the same field-of-view shown in the first row. The reduced background compared to the diffraction-limited images is clearly visible. Third row: Diffraction-limited close up of the speck recorded in the Alexa Fluor 647® channel corresponding to the antibody staining. Fourth row: 2D Alexa Fluor 647® dSTORM reconstruction, corresponding to the antibody staining, of the speck. The speck exhibits an amorphous structure characterized by a dense core from which occasionally filamentious extensions reach out into the periphery. Some of the specks exhibit a less dense center. Fifth row: 2D DyLight 755® dSTORM reconstruction, corresponding to the nanobody staining, The smaller size of the dSTORM reconstruction compared to the signal obtained in the Alexa Fluor 647® channel is clearly visible. It results from the lower binding affinity of the nanobody which leads to incomplete labeling of the low-density periphery of the speck. Furthermore decreased resolution of the DyLight 755® reconstructions resulting from the lower number of photons emitted by the dye is clearly visible. Sixth row: Merge of the dSTORM reconstructions obtained from the Alexa Fluor 647® and the DyLight 755® signal.
AF647: Alexa Fluor 647®; DL755: DyLight755®

4.2.3   Staining of endogenous ASC in BlaER1 cells using an anti-TagRFP nanobody

In the attached manuscript three approaches to stain for endogenously expressed ASC are exploited. In THP-1 cells a primary anti-ASC antibody in combination with a secondary F_{ab} fragment and an anti-ASC nanobody are applied. In BlaER1 cells the primary anti-ASC antibody is combined with a secondary antibody. All of these approaches have disadvantages. On the one hand, the antibody-F_{ab} fragment staining approach and the primary-secondary antibody approach result in a complex of about 11 nm and 17 nm in size, respectively [180]. This large size compared to alternative protein binders bears several disadvantages: First, it leads to an offset, linkage error, between the target protein and the fluorescent dye of about 11 nm and 17 nm, respectively and sets an ultimate lower limit on the achievable localization accuracy, and thereby on the achievable resolution (Figure 2.25). Secondly, a large size of the labeling probe often negatively affects the labeling density of the target structure and thereby the molecular sampling rate. This can be due to steric hindrances between neighbouring epitopes or, for dense target structures, the probe might not be able to access all binding sites since it is sterically excluded from dense regions of the structure. This might lead to incomplete structures or structures of lower resolution in cases where the Nyquist criterion is not fullfilled. The smaller size of the ASC nanobody of about 1.5×2.5 nm [180] helps to cirumvent the before mentioned problems. However, it suffers from low affinity (~ 159 nM). All these disadvantages were recognized to play a crucial role during the experiments on the ASC speck. The relatively large size of the antibody-based approaches limited the achievable resolution on the structure and occasionally led to incomplete labeling of the dense center of the structure. Application of the nanobody, on the other hand, led to incomplete labeling in the low-density periphery of the ASC speck.

Due to these disadvantages, I established another staining approach during the preparation of this thesis. It applies an Alexa Fluor 647®-labeled anti-TagRFP nanobody on the BlaER1 monocytes expressing ASC-TagRFP under the control of the endogenous ASC promotor [181]. Staining of a target protein via a nanobody directed against a fluorescent protein fused to the target protein has been described for other super-resolution applications [182, 183, 184, 185]. The applied nanobody was purchased from NanoTag Biotechnologies GmbH, Göttingen, Germany and has a higher affinity than the ASC nanobody (personal communication, NanoTag Biotechnologies GmbH). The linkage error for this approach can be estimated to ~ 7 nm (4.2 nm TagRFP + 2.5 nm nanobody) and is thus lower than that for the antibody-based approaches described above. However, only one ASC allel of the cells is tagged with TagRFP. Therefore, only half of the expressed ASC proteins harbour a TagRFP tag restricting the molecular sampling rate. On the other hand, only having every second molecule labeled might also be beneficial for the achievable resolution since it would lower the probability of overlapping PSFs, which is high due to the density of the speck. The samples were prepared as described in section 4.3.2. Figures 4.4 and 4.5 depict the results of the established staining procedure. Unfortunately an evaluation of those samples by dSTORM was not possible due to difficulties with the safety regulations. Due to those difficulties planned DNA-PAINT measurements using the TagRFP nanobody conjugated to a DNA-PAINT docking strand were not conducted.
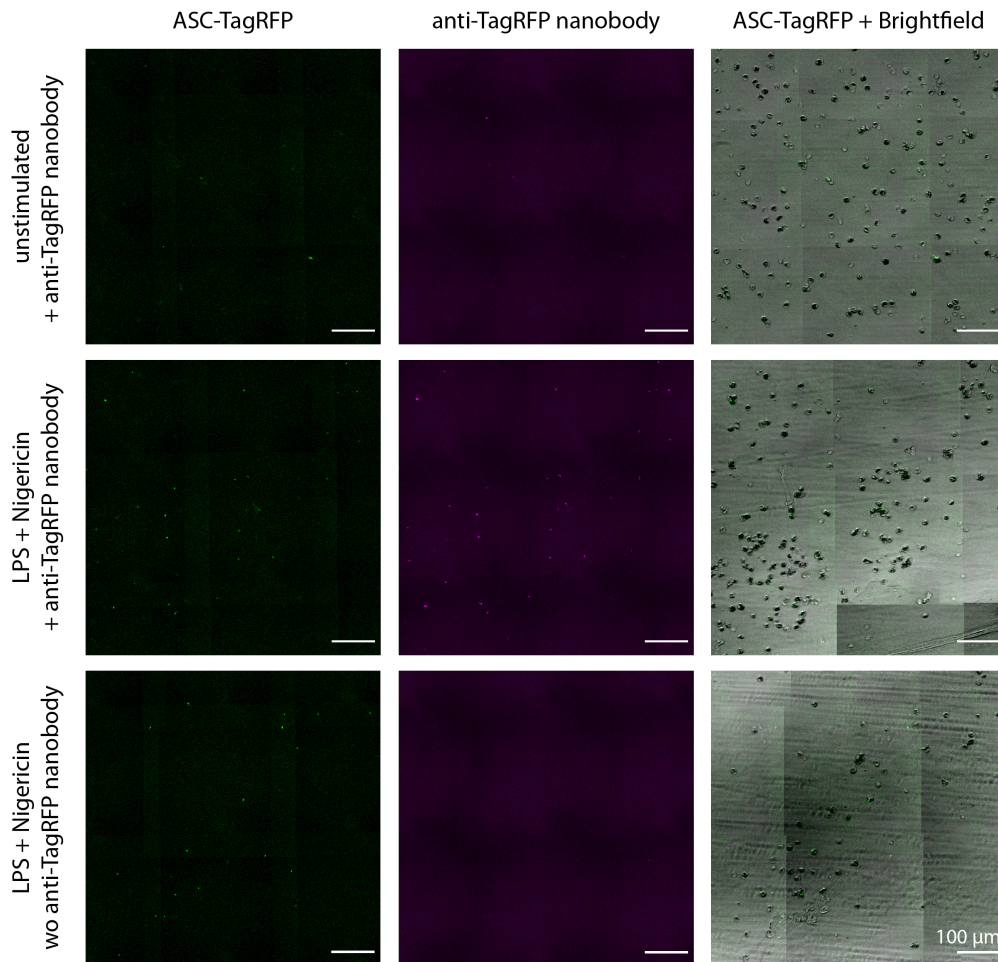
Figure 4.4: **Immunofluorescence of endogenously expressed ASC-TagRFP in BlaER1 cells via anti-TagRFP nanobody**
BlaER1 cells expressing ASC-TagRFP under the endogenous ASC promotor were transdifferentiated into macrophage-like cells and left unstimulated (top row) or were incubated with LPS and nigericin to activate the NLRP3 inflammasome (central and bottom row). Cells were subsequently stained by immunofluorescence using an anti-TagRFP nanobody (top and central row) or left unstained (bottom row). The cells were imaged by exciting TagRFP (left column) and Alexa Fluor 647® (central column). Additionally a brightfield image was recorded and overlayed with the signal recorded in the TagRFP channel (right column).
Treatment of the cells with LPS + nigericin specifically leads to formation of ASC specks inside the cells as visible in the TagRFP and the merged channel (left and right column). Those specks can be specifically stained with the anti-TagRFP nanobody conjugated with Alexa Fluor 647® as can be seen in the Alexa Fluor 647® channel (central column). To increase the number of cells evaluated for each condition large field of views are shown which were obtained by stitching together multiple field of views recorded at 10x magnification on a spinning-disk confocal microscope. In the fluorescence channels maximum projections of *z*-stacks across the entire cell volume are depicted.
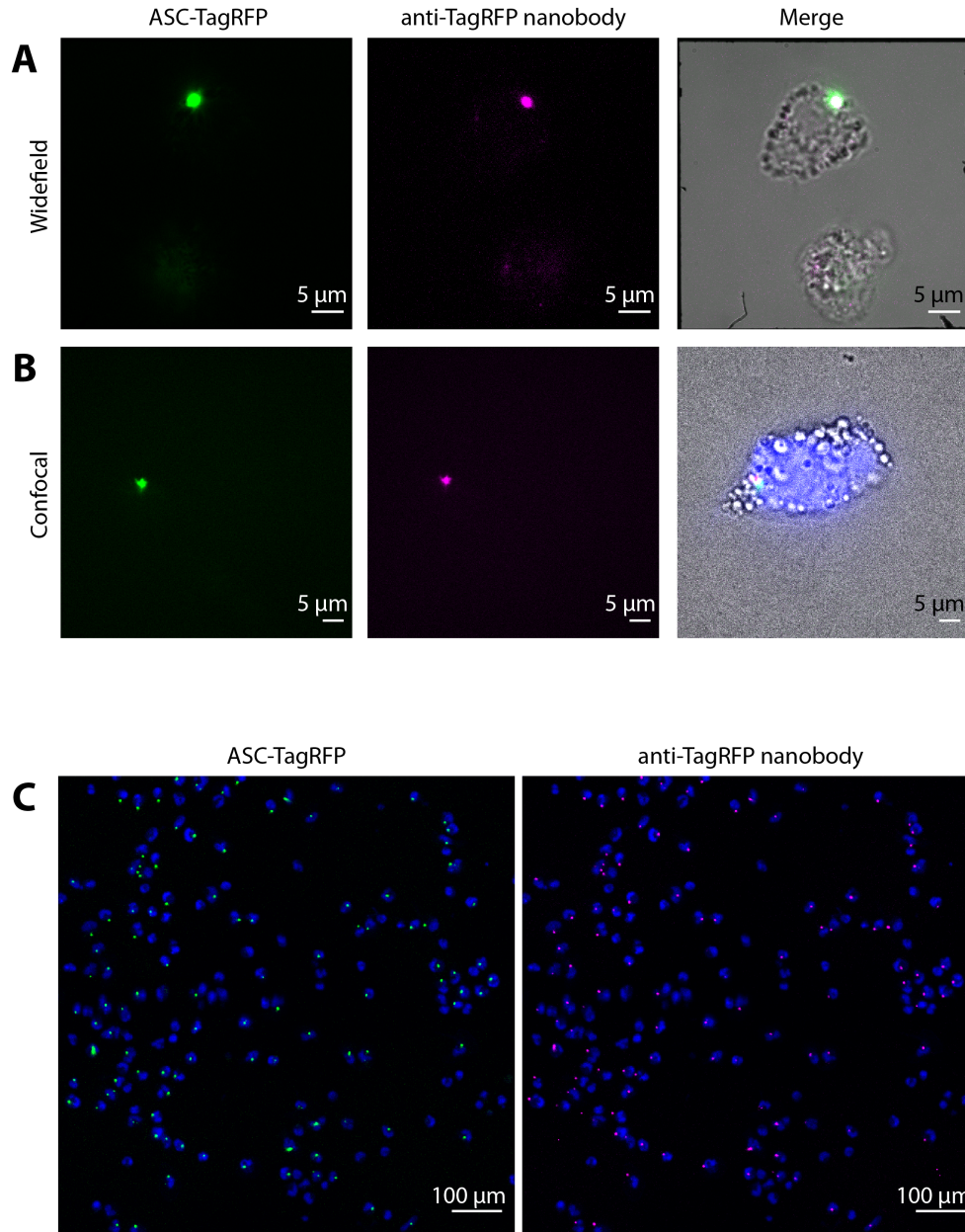
Figure 4.5: **Specific staining of ASC specks formed by endogenously expressed ASC-TagRFP in BlaER1 cells using the anti-TagRFP nanobody**
BlaER1 cells transdifferentiated into macrophage-like cells are depicted. ASC speck formation was induced by incubation of the cells with LPS and nigericin. The cells were stained using the Alexa Fluor 647®-labeled anti-TagRFP nanobody and imaged by exciting TagRFP and Alexa Fluor 647®.
A) Intracellular ASC speck imaged using widefield microscopy at 100x magnification. The TagRFP signal (left), the anti-TagRFP Alexa Fluor 647® signal (center) and a merge of the fluorescence channels with a brightfield image (right) are shown.

Continued Figure 4.5: B) Intracellular ASC speck imaged using spinning-disk confocal microscopy at 100x magnification. The TagRFP signal (left), the anti-TagRFP Alexa Fluor 647® signal (center) and a merge of the TagRFP signal, the Alexa Fluor 647 ® signal, a counterstaining of the DNA by DAPI and a brightfield recording of the cell are depicted.

The data presented in panel A) + B) confirms the specificity of the anti-TagRFP staining as well as the low level of background staining detectable in the cells even if they are imaged under widefield illumination.

C) Overview images obtained by stitching together multiple fields-of-view obtained on a spinning-disk confocal microscope at 10x magnifcation. The TagRFP signal is shown on the left and the anti-TagRFP Alexa Fluor 647® signal is shown on the right. For both channels maximum projections of *z*-stacks across the entire cell volume are depicted. The DNA was counterstained with DAPI. Note that almost all cells exhibit a perinuclear speck and almost all specks are stained by the anti-TagRFP nanobody.

4.3   Additional Materials and Methods

This section provides information about the performed experiments and analyses, that are not given in the attached paper.

4.3.1   ASC nanobody cloning and labeling

A plasmid encoding the cDNA of the ASC nanobody [158] used was provided by Professor Dario Alessi, University of Dundee via the MRC - Protein Phosphorylation and Ubiquitylation Unit, University of Dundee, Dundee, Scotland (DU Number 54832). It expresses the fusion protein FLAG-Von Hippel Lindau protein (VHL)-5xGLY-anti-ASC nanobody, which was originally used in the Affinity-directed PROtein Missile (AdPROM) system [186]. The region coding for the nanobody plus a C-terminal cysteine or sortase tag was amplified via a polymerase chain reaction (PCR) including 5'- and 3'-overhangs for insertion into the pCoofy2 expression plasmid. Figure 4.6 depicts the results of the sequence amplification for the cysteine construct. This construct was applied in subsequent experiments. The cDNA sequence encoding the nanobody including the C-terminal tag (Figure 4.6A) and the utilized primers (Figure 4.6B) are depicted. Correct amplification and length of the PCR product was verified by agarose gel electrophoresis on a 2% agarose gel (Figure 4.6C). The product was purified using the Monarch® PCR & DNA Cleanup Kit (New England Biolabs, Frankfurt am Main, Germany) following manufacturer's instructions and its concentration and purity was assessed by spectrophotometry on a Nanodrop^{TM} device (ThermoFisher Scientific, Waltham, Massachusetts, USA). The DNA concentration was calculated based the absorbance at 260 nm and the absorbance ratios at 260/280 nm and 260/230 nm were used to measure the purity of the DNA product (Figure 4.6D). The PCR product was ligated into the pCoofy2 expression plasmid [187] and expressed and purified at the Protein Production Core Facility at the Max Planck Institute for Biochemistry in Martinsried, Germany by Dr. Sabine Suppmann and coworkers.

A fraction of the nanobody, which has a C-terminal cysteine, was labeled by me with the help of my coworker, Ganesh Agam, using Alexa Fluor 647® C2 maleimid (A20347, ThermoFisher Scientific, Waltham, Massachusetts, USA). The rest was labeled with Alexa Fluor 647® and DyLight 755® by NanoTag Biotechnologies GmbH, Göttingen, Germany. The labeling protocol was adapted from Pleiner et al. [188]. Maleimid labeling buffer (100 mM potassium phosphate pH 6.4, 150 mM NaCl, 1 mM EDTA, 250 mM sucrose) was adjusted to pH 6.4 with HCl or NaOH, filtered through a 0.2 µm filter, degased with vacuum and purged with nitrogen. The C-terminal cysteine of the nanobody was freshly reduced by 10 mins incubation in 15 mM Tris(2-carboxyethyl)phosphin (TCEP) on ice. The buffer was exchanged to maleimid labling buffer on PD-10 Sephadex G-25 M desalting columns (17-0851-01, GE Healthcare, Chicago, Illinois, USA) previously equilibrated with maleimid labeling buffer. The nanobody was concentrated at 6°C using Vivaspin size exclusion columns (VS01910, Vivaspin 500, 3.000 MWCO, PES, Satorius, Göttingen, Germany). Concentrated nanobody was mixed with Alexa Fluor 647® C2 maleimid at a ratio of 1.2:1 dye:nanobody and incubated for 1.5 hrs on ice. Labeled nanobody was separated from free dye on a PD-10 column using maleimid labeling buffer supplemented with 1 mM TCEP and 1 mM DTT, and the nanobody was concentrated on a size exclusion column as described before. The degree of labeling of the nanobody which was labeled in house with Alexa Fluor 647® was estimated by recording a UV-Vis spectrum on a nanodrop spectrophotometer and determination of the protein-to-dye ratio was based on the absorbance at 280 and 650 nm. The integrity as well as proper labeling of the nanobody was verified using moving boundary electrophoresis, which showed a single band at ~ 13 kDa (Figure 4.7A). Exposing the gel to UV light confirmed labeling of the nanobody while no free dye was detected (Figure 4.7B). Since the presence of free dye was expected to have a strong negative effect on the immunofluorescence experiments for which the nanobody was generated, the presence of free dye was further evaluated using single-point FCS measurements. The obtained FCS curve could be fit well to

a model for free 3D diffusion of a single species with a diffusion coefficient of 115 μm^2/s (equation 19) indicating that there was no detectable amount of free dye in the sample (Figure 4.7 C). The derived diffusion coefficient also agreed well with the expected value for the labeled nanobody with a molecular weight of 13.4 kDa. The labeling efficiency of the DyLight 755$^{®}$ -labeled nanobody was evaluated by NanoTag Biotechnologies GmbH, Göttingen, Germany.

A

ATGGCTCAGGTCCAGCTGCAAGAGTCTGGCGGAGGACTT
GTTCAGCCTGGCGGAAGCCTGAAGCTGTCTTGTGCCGCC
AGCGGCTTCACCTTCAGCAGATACGCCATGAGCTGGTACA
GACAGGCCCCTGGCAAAGAAAGGGAAAGCGTGGCCAGA
ATCAGCAGCGGCGGAGGCACAATCTACTACGCCGATAGC
GTGAAGGGCAGATTCACCATCTCCAGAGAGGACGCCAAG
AACACCGTGTACCTGCAGATGAACTCCCTGAAGCCTGAGG
ACACCGCCGTGTACTACTGTTACGTCGGCGGCTTTTGGGG
CCAGGGCACACAAGTGACAGTGTCTAGCGGAGGCTGCTAG

Stop codon
Cysteine-coding codon

B

pCoofy2-ASC-nanobody-cysteine-tag-5'

5'-AAGTTCTGTTCCAGGGGCCCATGGC...
        ...TCAGGTCCAGCTG-3'

pCoofy2-ASC-nanobody-cysteine-tag-3'

5'-CCCCAGAACATCAGGTTAATGGCGC...
        ...TAGCAGCCTCCGCTAGACACTGTC-3'

Length PCR product: 398 nucleotides

C

D
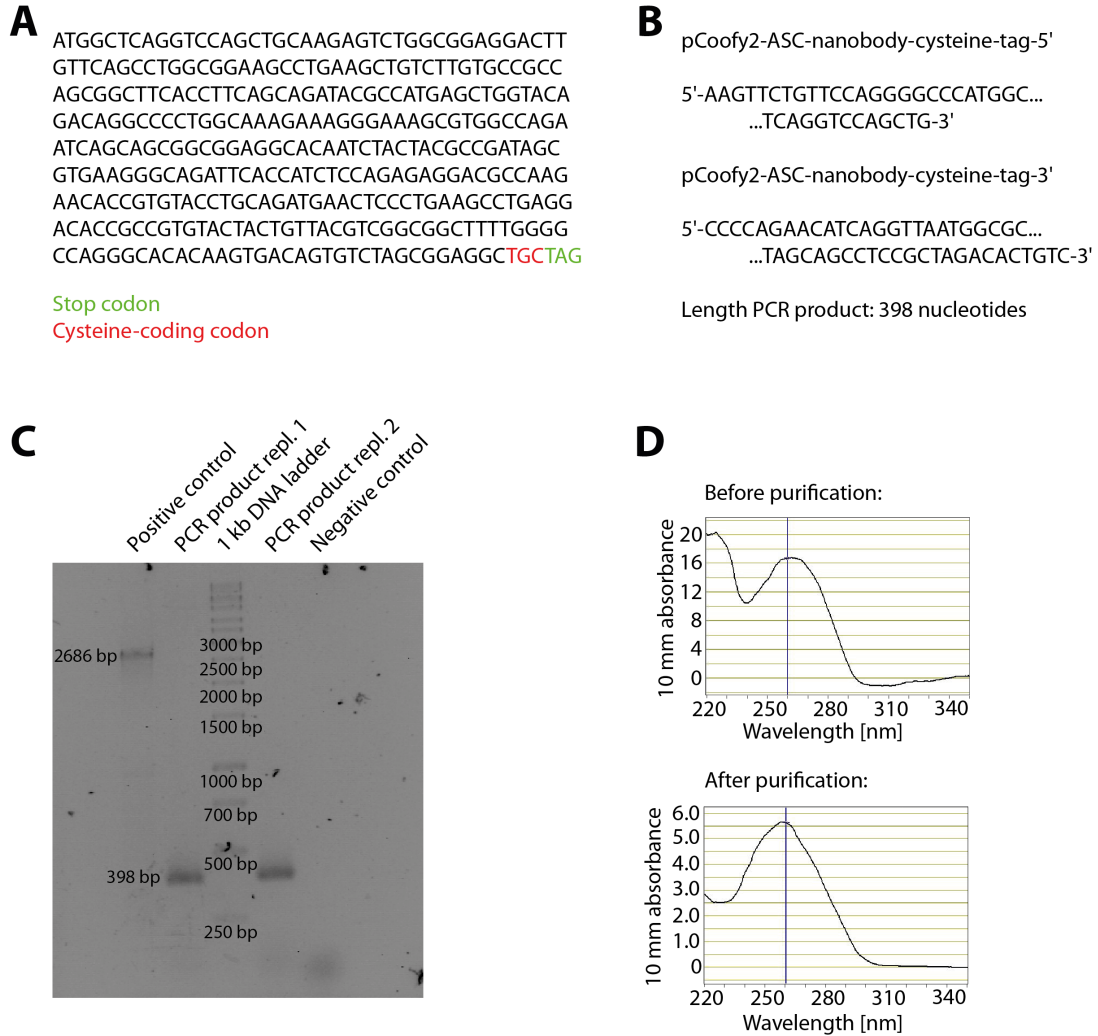

Figure 4.6: **Amplification of the ASC nanobody-encoding sequence including a C-terminal cysteine by PCR**
A) cDNA sequence of the ASC nanobody including a C-terminal cysteine tag
B) PCR primers utilized to amplify the sequence shown in A) including overhangs for ligation into the pCoofy2 expression plasmid
C) Agarose gel verifying the correct length of the PCR products; repl.: replicate
D) Absorbance spectra of the PCR product before (top) and after (bottom) purification verifying successful purification as quantified by the 260 nm / 280 nm and 260 nm / 230 nm ratio.
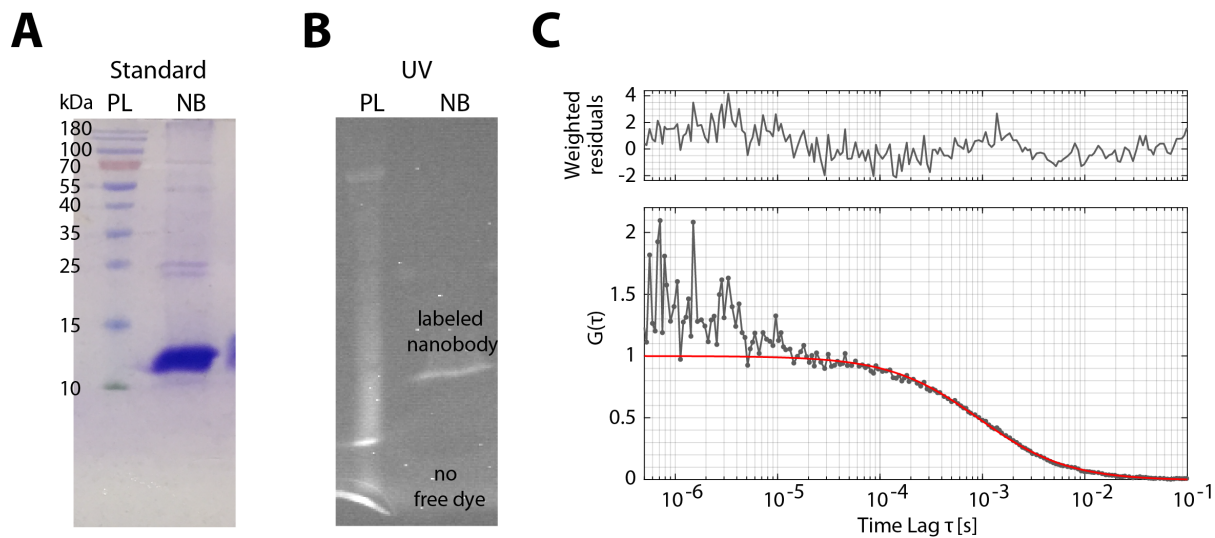
Figure 4.7: **Verification of nanobody labeling**
A) Coomassie-stained polyacrylamide gel as obtained by SDS-PAGE of the Alexa Fluor 647®-labeled nanobody. A protein ladder (PL) was applied to the gel on the left and the labeled nanobody (NB) on the right. The characteristic band at ~ 13 kDa and the absence of additional lower molecular weight bands confirm that the nanobody did not degrade during the labeling procedure
B) Gel shown in A) after exposure to ultraviolett light (UV) to excite Alexa Fluor 647®. An Alexa Fluor 647® band corresponding to the nanobody band is visible confirming successful labeling of the nanobody. No free dye was observed in the gel.
C) The autocorrelation function from a single-point FCS measurement of the labeled nanobody (dots depict data points) and the corresponding single component fit of the data (red line). The autocorrelation function is fit well by a single component 3D free diffusion model and the derived diffusion coefficient of 115 μm^2/s supports the observation that the nanobody is labeled and no free dye is detected in the sample. Note, the additional correlation at lag times shorter than 10^{-5} seconds is due to triplet blinking of Alexa Fluor 647®

4.3.2   Immunofluorescence staining of ASC-TagRFP in BlaER1 cells using an anti-TagRFP nanobody

BlaER1 B cells expressing ASC-TagRFP were cultivated at 37°C, 5% CO_2 in culture medium consisting of Roswell Park Memorial Institute (RPMI) 1640 medium (21875034) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS) (10500064), 1 mM sodium pyruvate (11360039) and 100 U/ml Penicillin/Streptomycin (15140122), and maintained at a density between $1x10^5$ and $1x10^6$ cells per milliliter. Cells were transdifferentiated into BlaER1 monocytes for a week by seeding them at 60,000 cells/cm^2 in culture medium supplemented with 50 ng/ml IL-3 (200-03 B, Peprotech, New Jersey, USA), 50 ng/ml M-CSF (300-25 B, Peprotech) and 500 nM β-Estradiol (E8875, Sigma-Aldrich) on cell culture-treated 8-well plastic bottom slides (80826, ibidi, Gräfelfing, Germany). On the seventh day, cells were trypsinized and transferred to 8-well glass bottom slides (0030742001, Eppendorf, Hamburg, Germany) coated with 100 μg/ml poly-l-ornithine (A-004-C, Merck-Millipore, Massachusetts, USA) at the same density in culture medium without growth factors and incubated overnight at 37°C and 5% CO_2. Differentiation of the cells into monocytes was confirmed by a change in morphology from completely round cells into cells with cellular extrusions. Cells were primed by incubation with 200 ng/ml LPS-EB (tlrl-3pelps, Invivogen) for 14 hours and caspase activity was inhibited by addition of Z-VAD FMK at 20 μM final concentration for 60 mins. The NLRP3 inflammasome was activated by the addition of nigericin to a final concentration of 6.5 μM. Formation of ASC specks was monitored by fluorescence spinning-disk confocal microscopy and nigericin incubation was stopped once a sufficient number of ASC specks were observed. In an ideal preparation almost 100% of the cells formed a speck. After the medium was removed from the cells, they were washed once with PBS and subsequently fixed by incubation with 4% paraformaldehyde (PFA) (E15710-S, Electron Microscopy Sciences, Pennsylvania, USA) in PBS. Subsequently, the sample was rinsed once with PBS and then quenched by rinsing once with 0.1 M NH_4Cl (254134, Sigma-Aldrich) in PBS followed by a 15 min incubation with 0.1 M NH_4Cl in PBS. Permeabilization and blocking was done for 30 mins in 10% normal goat serum (NGS) (16201), 0.5% Triton X-100 (T8787, Sigma-Aldrich) in PBS followed by a 2 mins incubation step in PBS and 30 mins incubation in Image-iTFX Signal enhancer (I36933). After two additional PBS washing steps, cells were incubated for 60 mins with 200 μl/8-well of anti-TagRFP nanobody conjugated with Alexa Fluor 647® (N0510, NanoTag Biotechnologies GmbH, Göttingen, Germany) diluted to a final concentration of 5 nM in 3% normal goat serum, 0.1% Triton X-100 in PBS. Subsequently, cells were washed three times with 10% normal goat serum, 0.1% Triton X-100 in PBS for 2, 3, and 5 mins, respectively. In the following the sample was incubated with 0.1% Triton X-100 in PBS for 2, 3, and 5 mins, respectively. Finally, the sample was incubated twice for 5 mins in PBS. Optionally, DNA was counterstained by incubation of the sample for 5 - 10 mins in DAPI solution followed by a washing step with PBS. For longterm storage, the cells were postfixed by incubation in 3% PFA in PBS.

4.3.3   Clustering of non-speck-associated ASC

In order to investigate whether the ASC protein forms smaller complexes before associating into the ASC speck we decided to perform a clustering analysis on the cytoplamic, non-speck associated ASC localizations in THP-1 cells stained with primary and secondary F$_{ab}$ fragment. For this purpose we chose the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [189] as it is know to give reliable results even when performed on noisy data. Figure 4.8 depicts an illustration of the DBSCAN algorithm. The algorithm has two input parameters: the search radius *n* and the minimum number of neighbouring points a point must have in order to be called a core point of a cluster *m*. Based on these input parameters all points fed into the algorithm are classified as core points, boarder points or outliers. For our analysis a search radius of 70 nm

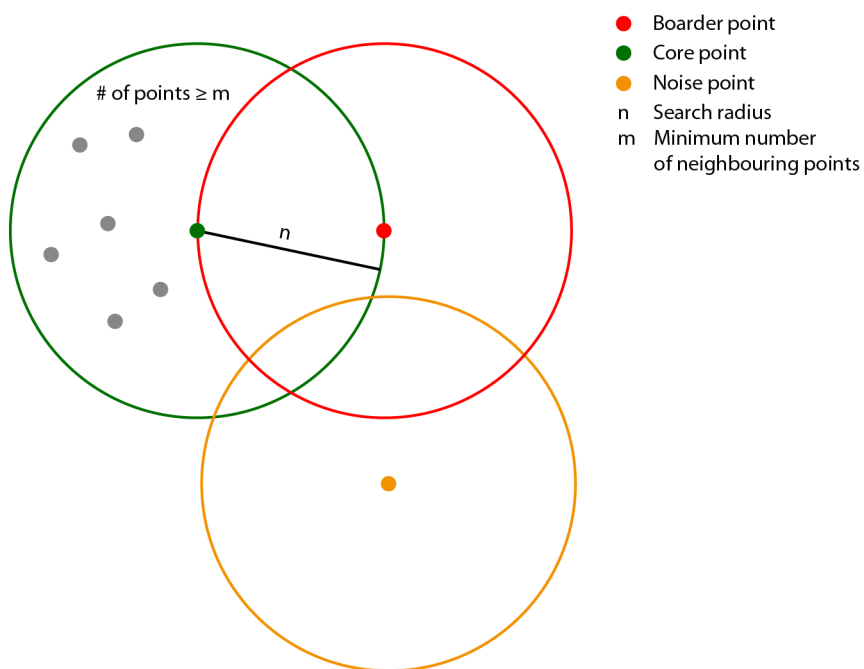and a number of points per cluster of $10 - 300$ was chosen.

Figure 4.8: **Illustration of the density-based spatial clustering of applications with noise (DBSCAN) algorithm**
The DBSCAN algorithm has two inputs: 1) the search radius denoted as n and 2) the minimum number of neighbouring points a points needs to have in order to be called a core point of a cluster m. For every data point the algorithm evaluates, whether there are at least m data points within the search radius n. If this is the case, the point is classified as a core point (green). If a point has less than m neighbouring points within the search radius but can be reached by the cluster, the point is called a boarder point of the cluster (red). If a point has less than m neighbouring points and cannot be reached by a cluster, it is called an outlier (yellow) and not classified as part of a cluster.

4.4   Summary of the paper

Paper to be found in appendix B, section B.2

In the following the main results of the paper underlying this chapter of the thesis are summarized. The figure references given reference to the figures in the paper which can be found in the appendix B, section: B.2. In the manuscript underlying this chapter, different fluorescence microscopy imaging techniques were applied to provide insight into the supramolecular structure of the endogenous ASC speck. Specifically widefield, laser-scanning confocal microscopy, spinning-disk confocal microscopy, 2D and 3D dual-color dSTORM [72, 29], and 2D and 3D DNA-PAINT [68, 86] super-resolution microscopy were utilized.
The experiments investigated the ASC speck formed after NLRP3 inflammasome activation by stimulating the cells with LPS and the bacterial toxin nigericin. Low magnification confocal microscopy images show a homogenous distribution of ASC in unstimulated cells and the formation of a single, perinuclear speck in inflammasome activated cells. Quantification of the

ASC content per speck from 60x magnification widefield images reveals that the amount of incorporated ASC per speck varies by almost one order of magnitude (Figure 1). Next, whole cell dSTORM imaging was applied on cells containing a speck (Figure 2). It confirms the previously observed depletion of free ASC in the cytoplasm but still reveals the presence of ASC clusters outside the speck. In comparison to free ASC in the cytoplasm of unstimulated cells, a small fraction of these clusters are larger and contain more localizations. In the following, the investigations focus on the supramolecular structure of the speck using dSTORM and DNA-PAINT (Figures 3 + 4). The measurements resolved filaments protruding from the dense core of the endogenous, unperturbed ASC specks. The number of observed filaments per speck rarely exceeds 10, which is lower than what is shown in literature for the speck formed after overexpression. Measuring the diameter of the filaments shows an average value of 37.1 nm, which agrees with the value for a single filament reported in the literature [147]. Applying antibody as well as nanobody staining of the structure, revealed an overall spherical shape and a subset of the ASC specks exhibited a hollow center after antibody but not after nanobody staining. Furthermore, nanobody staining shows significantly smaller specks. The combination of both labeling approaches on the same sample (Figure 4) showed that the nanobody stains the dense center of the speck but is washed away from the lower density periphery due to its low binding affinity. The larger antibody complexes, on the other hand, are predominately detected in the periphery of the speck since they have limited access to the dense center. These results provide an explanation for the ring-like structure of the ASC speck reported in the literature. Single cell analysis of ASC redistribution into the speck allows sorting of the specks into a quasi-temporal order, which strikingly suggested that the endogenous speck mainly becomes denser but only marginally larger during its formation (Figure 5). The obtained data is summarized in a model for the endogenous ASC speck (Figure 6). It describes the speck as an overall spherical complex with a dense core from which a limited number of filaments reach out into the periphery. The results suggest that the speck forms from a loose scaffold that becomes denser over time. Interestingly the presented data shows that larger cells form larger specks and a recent study revealed that the NLRP3 inflammasome forms at the MTOC [190] whose size also depends on the cell size at least in *C. elegans* [191] which might suggest that the MTOC serves as a scaffold for NLRP3 inflammasome formation.

4.5   Outlook

The experiments underlying this chapter investigated the ASC component of the NLRP3 inflammasome. In the future, the location of other inflammasome forming proteins within the NLRP3 inflammasome complex should be investigated. Here, the NLRP3 protein and caspase-1 are the next targets to address. It is known that caspase-1 associates with the ASC speck. However, there are currently two models for the association of caspase-1 with the ASC speck reported in the literature. One model suggests that the caspase-1 molecules binding to the ASC speck associate into filaments, which reach out into the periphery [149, 192]. The other model suggests that not filamentious multimers of caspase-1 associate with the speck, but rather single caspase-1 molecules that transiently associate with the speck [193]. Future experiments should address if one of those models is true or whether caspase-1 is associated in a different way with the endogenous inflammasome. On the long term, it would be interesting to investigate the intracellular organization of other inflammasomes such as the AIM2 and the NLRC4 inflammasome. These experiments will potentially reveal the similarities and differences in the supramolecular organization of the inflammasome-forming molecules between different inflammasomes. Several studies suggest that different inflammasome sensor proteins can be recruited into the same macromolecular inflammasome complex [194, 157, 195, 196, 197]. It will be important to address how these inflammasome sensors are associated with the macromolecular inflammasome

complex.

As the method to address these questions, I suggest high throughput Exchange-(DNA-)PAINT [198] with nanobodies directed against the inflammasome proteins. Such an approach would be well suited to address the listed biological questions because of several reasons. It would enable the high spatial resolution needed to get insight into the molecular interactions within the inflammasome complex, provide the multiplexing capabilities needed to image multiple targets within the inflammasome complex and enable imaging of a large number of inflammasome complexes, which is needed to address the heterogenity present in these complexes. To implement this approach, one would need to combine an automated super-resolution localization microscopy setup as previously described by the group of Dr. Jonas Ries at the Molecular Biology Laboratory in Heidelberg, Germany [199] with a microfluidic device for automated exchange DNA-PAINT as previously demonstrated [200]. As the labeling strategy I would suggest nanobodies directed against the inflammasome proteins which are conjugated with a DNA-PAINT docking strand. Nanobodies against inflammasome proteins are currently developed by the group of Dr. Florian Ingo Schmidt at the Institute for Innate Immunity at the Universität Bonn, Germany [201]. To enable DNA-PAINT measurements at widefield illumination the recently developed fluorogenic imager strand probes should be applied [89]. Although those experiments are very likely to greatly extend the current knowledge about the endogenous inflammasome complex, they might not provide sufficient resolution to resolve the finest details of the complexes. This will, for example, most likely be the case in the dense core of the structure. This limitation could be addressed by several means. One could imagine investigating the structure using the recently developed method MINFLUX [202], which is capable of very high resolution. Furthermore, one could combine the high spatial resolution of electron microscopy with the molecular specificity of fluorescence super-resolution microscopy as it has been done for other multiprotein complexes [203]. Another approch would be to perform expansion microscopy in combination with super-resolution localisation microscopy [180]. In addition to the listed questions regarding the association of the proteins within the inflammasome complex, there are open questions regarding the intracellular localization of the inflammasome complexes. A recent publication showed that the NLRP3 and the pyrin inflammasome form at the microtubule-organizing center [190]. However, for the AIM2 and NLRC4 inflammasome, it was shown that they do not form at this location and it remains to be elucidated where in the cell these complexes form. Even though resolution below the diffraction limit is not required to address these questions, they certainly profit from super-resolution imaging.

These experiments addressing the structure and intracellular location of the inflammasome complexes should be complemented with live cells experiments. Due to the previously mentioned tendency of the inflammasome proteins to multimerize, it would be decisive to perform these experiments in cells expressing the inflammasome proteins at the endogenous level. Here, the BlaER1 cell lines, utilized in this chapter of the thesis as well as in the experiments described in appendix A, could be used for initial experiments. These cells express either ASC-TagRFP or ASC-TagRFP and NLRP3-eGFP under the control of the endogenous ASC and NLRP3 promotor. Such experiments could, for example, address the kintetics of inflammasome assembly. Furhermore, they would allow one to study what happens to individual inflammasome complexes over time. Since these experiments would image inflammasome formation on the level of single cells, they would profit from an experimental system that allows to activate inflammasome formation in a specific cell. To this end, an inflammasome activating ligand that can be specifically activated by a flash of laser light would be very valuable. Such a system potentially could be established for the AIM2 inflammasome, which is activated by double-stranded DNA [204]. The probe 6-nitropiperonyloxymethyl (NPOM) can be attached to DNA and can be specifically cleaved off by a flash of laser light [205] as has also been demonstrated in single cells [206]. One could imagine to design a DNA molecule that is rendered unable to activate AIM2 inflammasome through caging with NPOM. When this molecule is transfected into cells the NPOM group could be cleaved off in single cells by a

short exposure with laser light that would turn the DNA molecule back into an AIM2 activator. Hence, AIM2 inflammasome activation could be specifically activated and followed in single cells.

5 Measurements on the interaction of DPPA3 + UHRF1 in the nucleus of mouse embryonic stem cells

Publication underlying this chapter (to be found in appendix B, section B.3):

Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive DNA demethylation in mammals
by Christopher B. Mulholland, Atsuya Nishiyama, Joel Ryan, Ryohei Nakamura, Merve Yiğit, Ivo M. Glück, Carina Trummer, Weihua Qin, Michael D. Bartoschek, Franziska R. Traube, Edris Parsa, Enes Ugur, Miha Modic, Aishwarya Acharya, Paul Stolz, Christoph Ziegenhain, Michael Wierer, Wolfgang Enard, Thomas Carell, Don C. Lamb, Hiroyuki Takeda, Makoto Nakanishi, Sebastian Bultmann, Heinrich Leonhardt. *Nature Communications*, 11(1), 5972 (2020).

5.1 Biological background

After fertilization, the zygote undergoes differentiation. The ability of a cell to differentiate into other cell types is called cell potency and the larger the number of cell types a cell can differentiate into, the higher its potency. Sorted from higher to lower potency, there are totipotent, pluripotent, multipotent, oligopotent and unipotent cells. A hallmark of differentiation is the establishment of epigenetic marks in the genome, one of which is methylation of cytosines into 5-methylcytosine. A prominent effect of this alteration is stabilization of the differentiation status of the cell and restriction of its potency [207]. Oocyte and sperm are two highly differentiated cell types and, as such, their genome contains epigenetic marks including 5-methylcytosine. However, in mammals, these marks need to be erased after gamete fusion to enable pluripotency of the developing cells [208, 209]. The ten-eleven translocation (TET) methylcytosine dioxygenase enzymes are key in this process by either directly demethylating cytosines [210, 211, 212, 213] or driving hypomethylation through the repression of methylation maintenance during replication of the DNA [214, 215]. The study containing the data recorded for and discussed in this chapter further elucidates the role of the proteins TET1 and TET2 in global demethylation and investigates their role in maintaining naïve pluripotency in mouse embryonic stem cells (ESCs) [216]. The authors discovered that the gene encoding Developmental pluripotency-associated protein 3 (DPPA3/PGC7/STELLA) is activated by Tet1-driven demethylation. Furthermore, they unreavel that DPPA3 binds to the protein UHRF1 and thereby prevents it from binding to chromatin. If not associated with DPPA3, UHRF1 contributes to methylation maintenance [217]. The described signalling cascade establishes a novel mechanism underlying global hypomethylation in mammalian embryonic stem cells (ESCs).

The experiments to which I contributed as part of this study investigated the interaction of UHRF1 and DPPA3 in the nucleus of ESCs by RICS and ccRICS. The experiments and the obtained results will be summarized in the following section.

5.2   Summary of the results that I contributed to the paper

Paper to be found in appendix B, section B.3

5.2.1   Analysis in the nucleus

The experiments to which I contributed as part of this study were carried out together with Dr. Joel Ryan and Christopher B. Mulholland from the research group of Professor Heinrich Leonhardt in the Biology Department at the Ludwig-Maximilians-Universität München, Germany. They investigated the mechanism by which DPPA3 inhibits UHRF1 binding to chromatin. Based on the knowledge that the two proteins can interact, two possible scenarios for the repression mechanism were addressed: 1) DPPA3 competes with UHRF1 for binding to chromatin and 2) DPPA3 interacts either directly or indirectly with UHRF1 thereby repressing its chromatin binding function. Both scenarios were addressed by measuring the diffusion and potential interaction of the proteins in the nucleus of mouse ESCs using arbitrary-region RICS (section 2.6.3). The experiments were carried out in cells expressing UHRF1-eGFP instead of wildtype UHRF1 (UHRF1-eGFP was expressed from the endogenous UHRF1 promotor). This expression is denoted as "UHRF1-eGFP" in the following. Additionally wildtype DPPA3 expression was knocked out in these cells and a doxycycline-inducible Sleeping Beauty construct encoding for DPPA3-mScarlet was stably integrated. Expression of DPPA3-mScarlet was induced by incubating the cells with doxycycline prior to the experiments. This expression is denoted as "DPPA3-mScarlet" in the following. We recorded confocal image series of the nucleus of these cells using PIE of two the fluorescent proteins (Figure 5.1A). In the first step we investigated the distribution of UHRF1-eGFP and DPPA3-mScarlet in the obtained images. Nuclear DPPA3-mScarlet was homogenously distributed whereas UHRF1-eGFP was partly enriched in foci indicative for chromatin binding (Figure 5.2A). To get insight into the mobility of UHRF1 and DPPA3 we next calculated the RICS spatial autocorrelation function in the nucleus of the cells for both proteins. They both showed a bound/immobile component (on the RICS timescale) and a predominant mobile population. For DPPA3-mScarlet, $88.4 \pm 5.2\%$ was mobile (Figure 5.1F) diffusing at a speed of 7.18 ± 1.87 $\mu m^2/s$ (Figure 5.2F). UHRF1-eGFP showed a mobile population of $60.6 \pm 13.7\%$ (Figure 5.1G) diffusing at the roughly the same speed as DPPA3-mScarlet (Figure 5.2H).

To check whether the proteins interact we repeated the measurements in cells expressing a DPPA3 mutant unable to interact with UHRF1 (DPPA3-KRR-mScarlet) instead of the DPPA3-mScarlet fusion. Again we first investigated the distribution of the proteins in the recorded images. The images showed no obvious difference in the nuclear DPPA3 distribution whereas UHRF1-eGFP appeared to a much greater extent in the foci (Figure 5.2B). The extracted diffusion behaviour stayed largely unaffected for DPPA3 (Figure 5.1F, Figure 5.2B, F) while the population of mobile UHRF1-eGFP molecules strongly decreased to 32% (Figure 5.1G) which corresponds to the observations made in the images. This suggested that DPPA3 wildtype binds to UHRF1 thereby mobilizing it and suppressing it from binding to chromatin. This was further supported by the fact that the fraction of mobile UHRF1-eGFP molecules increased with increasing DPPA3-mScarlet-to-UHRF1-eGFP ratio in the nucleus (Figure 5.2I).

A competitive binding process between UHRF1 and DPPA3 could be completely ruled out by the observation that the diffusion and distribution of DPPA3-mScarlet stayed unaffected in cells that did not express UHRF1 (Figure 5.1F, Figure 5.2C, F). These cells expressed free eGFP under the endogenous UHRF1 promotor. Free eGFP showed a homogenous distribution in the nucleus, as expected (Figure 5.2C) and the obtained diffusion coefficient of slightly below 50 $\mu m^2/s$ (Figure 5.2H) corresponded to the value reported in the literature [218].

Next, we set out to confirm the interaction of UHRF1 and DPPA3 by ccRICS. To demonstrate the sensitivity of our method, we measured the cross-correlation in cells expressing free eGFP and free mScarlet and in cells expressing an eGFP-mScarlet tandem construct. The measurements showed no cross-correlation (Figure 5.1D, H) and a strong cross-correlation (Figure 5.1E, H), respectively. As expected, the fluorescent species were entirely mobile and homogenously distributed in these cells (Figure 5.1G, Figure 5.2D, E). Based on these results, a nuclear interaction between UHRF1 and DPPA3 was confirmed by calculating the RICS spatial cross-correlation in cells expressing DPPA3-mScarlet and UHRF1-eGFP (Figure 5.1B) whereas cells expressing the DPPA3-KRR-mScarlet mutant did not show an interaction of the proteins (Figure 5.1C). As we expected DPPA3 to interact with UHRF1 thereby mobilizing it and preventing its binding to chromatin, we calculated the fraction of cross-correlating mobile molecules. Consistent with our assumption, we observed a clear population of cross-correlating mobile molecules in cells expressing DPPA3-mScarlet and UHRF1-eGFP while this population disappeared in cells expressing the DPPA3-KRR-mScarlet mutant (Figure 5.1H).

The ccRICS measurements strongly indicated that DPPA3 and UHRF1 interact. Since DPPA3-mScarlet is much smaller than UHRF1-eGFP (45 kDa vs. 125 kDa) one would expect a change in the diffusion coefficient of DPPA3-mScarlet in a situation in which it cannot interact with UHRF1-eGFP. However the diffusion coefficient of DPPA3-mScarlet only showed a very minor increase in cells expressing the DPPA3-KRR-Scarlet mutant (which is unable to interact with UHRF1) and in cells expressing no UHRF1 compared to cells expressing UHRF1-eGFP (Figure 5.2F). Since generally a halving of the molecular weight of the diffusing species only leads to a 26% increase in the diffusion coefficient this suggests that DPPA3-mScarlet interacts with another molecule or molecule complex which stays bound in a situation in which DPPA3-mScarlet and UHRF1-eGFP do not interact. The observation furthermore suggests that DPPA3-mScarlet in complex with the unknown interation partner has roughly the same molecular weight as UHRF1-eGFP. The diffusion coefficient of $7.18 \pm 1.87 \ \mu m^2$/s we measured for DPPA3-mScarlet and UHRF1-eGFP in cells expressing both proteins roughly fits to the value expected in the above described scenario in the nucleus at 37°C [218].
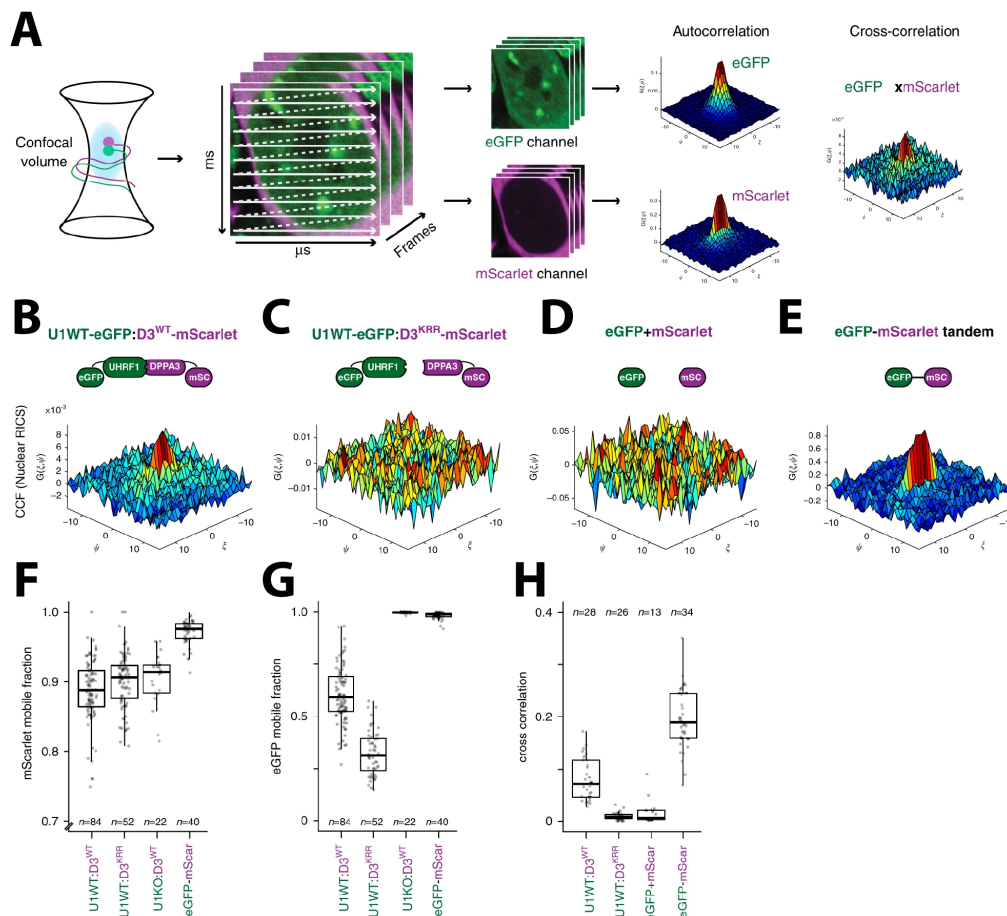
Figure 5.1: **Nuclear diffusion of UHRF1 and DPPA3 in mouse embryonic stem cells - Part I**
A) Schematic illustrating the RICS data acquisition and analysis workflow used for the described experiments. The data was recorded on a confocal laser scanning setup in which a confocal volume (left) is raster-scanned across individual living cells expressing eGFP and mScarlet constructs. The obtained images allowed extraction of spatiotemporal information about the expressed proteins on the microsecond and millisecond timescale. The spatial autocorrelation of each fluorescent protein as well as the cross-correlation between the proteins was calculated.
B-E) Representative spatial cross-correlation functions obtained on pixels in the nucleus of the different cells lines investigated. For cells expressing UHRF1-eGFP and DPPA3-mScarlet, a clear cross-correlation between the proteins was observed (B). This correlation was abrogated when DPPA3-mScarlet was replaced by the DPPA3-KRR-Scarlet mutant, which cannot interact with UHRF1 (C). Negative and positive controls were performed in cells expressing free eGFP and free mScarlet, which showed no cross-correlation (D) and cells expressing an eGFP-mScarlet tandem construct, which showed a clear cross-correlation (E). F-H) Box plots showing the mobile fraction of the mScarlet species (F) and the eGFP species (G) and the fraction of cross-correlating mobile molecules obtained on the different cell lines. The data was obtained by fitting the spatial cross-correlation function with a fit function containing a mobile and an immobile component. In the box plots each data point represents a value obtained on a single cell. The limits of the boxes represent the upper and lower quartiles and the darker horizontal line within the boxes represents the median value. The whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge.
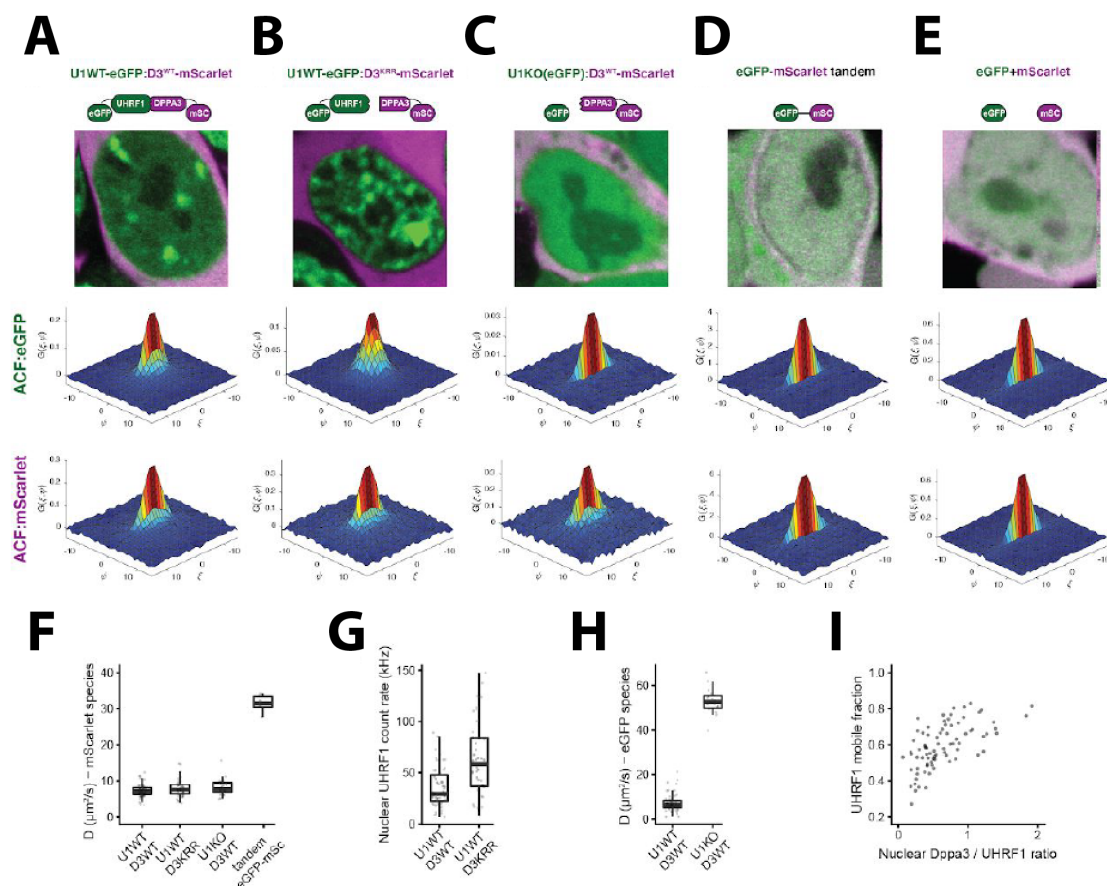Adapted from [216]. Reprinted with permission from Nature Publishing Group

Figure 5.2: **Nuclear diffusion of UHRF1 and DPPA3 in mouse embryonic stem cells - Part II**
A-E) Representative confocal laser scanning images of the cell lines indicated at the top and the corresponding spatial auto-correlation functions (ACF) obtained on the individual channels. The images show a merge of the images recorded in the two channels (for each channel a sum projection of 250 images is shown). In the presence of DPPA3-mScarlet, UHRF1-eGFP is mostly homogenously distributed within the nucleus and only a minor fraction appears enriched in the foci indicative for UHRF1 chromatin binding. The shape of the autocorrelation functions indicates mostly freely diffusing proteins. For the eGFP channel, a slight broadening in the direction of the slow time scale axis is observed, which is indicative of the fraction of bound/immobile UHRF1 molecules (A). Exchange of DPPA3-mScarlet with the DPPA3-KRR-mScarlet mutant leads to an enrichment of UHRF1-eGFP in the foci. A corresponding decrease in UHRF1-eGFP mobility is also observed in the eGFP autocorrelation function, which shows a strong broadening in the direction of the slow time axis. The shape of the DPPA3-mScarlet ACF remains mostly unchanged (B). When free eGFP is expressed instead of UHRF1-eGFP (C) as well as for the control cell lines expressing the eGFP-mScarlet tandem construct (D) or eGFP and mScarlet as free fluorescent proteins (E), a homogenous distribution of the eGFP signal is observed. The shapes of the corresponding ACFs indicate freely diffusing proteins in all cases. The slighlty broadened shape of the mScarlet ACF in c indicates that there is a minor population of immobile DPPA3-mScarlet molecules in these cells.

Continued Figure 5.2: F - H) Box plots showing the diffusion coefficient of the mobile mScarlet species (F), the nuclear UHRF1 count rate (G) and the diffusion coefficient of the mobile eGFP species (H) for the indicated cell lines. The diffusion coefficient of the mScarlet species stays largely unchanged for all cell lines except for the cells expressing the eGFP-mScarlet fusion construct, which shows a higher diffusion coefficient (F). The total amount of UHRF1-eGFP within the nucleus stays largely constant between cells additionally expressing DPPA3-mScarlet or DPPA3-KRR-mScarlet (G). UHRF1-eGFP diffuses roughly at the same speed as the DPPA3-mScarlet constructs while free eGFP diffuses much faster (H). In the box plots, each data point represents a value obtained from a single cell. The limits of the boxes represent the upper and lower quartiles and the darker horizontal line within the boxes represents the median value. The whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge.

I) Scatter plot of the fraction of mobile UHRF1 molecules against the ratio of the mean pixel intensity in the DPPA3-mScarlet to the mean pixel intensity in the UHRF1-eGFP channel. A positive correlation is observable.

Adapted from [216]. Reprinted with permission from Nature Publishing Group.

5.2.2   Analysis in the cytoplasm

An obvious question arising from the obtained results from the measurements in the nucleus is whether DPPA3 and UHRF1 are already interacting in the cytoplasm. In order to obtain insight into this question, the cytoplasmic fraction of the obtained images was analyzed. These data are not part of the above mentioned publication and are preliminary since data acquisition focused on the nucleus and the number of pixels recorded in the cytoplasm was very limited. In addition only a small number of cells were analyzed leading to low statistics underlying the analysis. These measurements would need to be repeated before definite conclusions can be drawn. We started off by calculating the spatial autocorrelation function of the two proteins. As for the measurements in the nucleus the data contained an immobile and a mobile fraction for DPPA3-mScarlet and UHRF1-eGFP. The analysis suggested that DPPA3-mScarlet is also predominantly mobile in the cytoplasm (~88% mobile population) but is diffusing faster (~ 11 μm^2/s) than in the nucleus. This diffusion characteristics were largely unchanged in cells expressing the DPPA3-KRR-mScarlet mutant and did also not change in the absence of UHRF1 (Figure 5.3A, C).

Next diffusion of UHRF1 was investigated. In the presence of DPPA3-mScarlet, about 90% of the UHRF1-eGFP proteins are mobile, diffusing with at a speed of about 8 μm^2/s (Figure 5.3B, D) which is an increase in the fraction of mobile molecules compared to the nucleus while the diffusion coefficient stays approximately the same. Disrupting the interaction of DPPA3 and UHRF1 reduces the fraction of mobile UHRF1-eGFP slightly to around 80% while the diffusion coefficient of the protein showed a larger spread and a tendency towards higher values (Figure 5.3B, D). For the free eGFP control, almost all proteins were mobile and diffused at a speed of slightly below 50 μm^2/s, which was consistent with the value obtained in the nucleus (Figure 5.2H) and reported in the literature [218]. We did not observe a cross-corrlation signal between the proteins in the cytoplasm. Taken together, these data suggest that there is no interaction of the proteins in the cytoplasm and that the interaction starts in the nucleus.
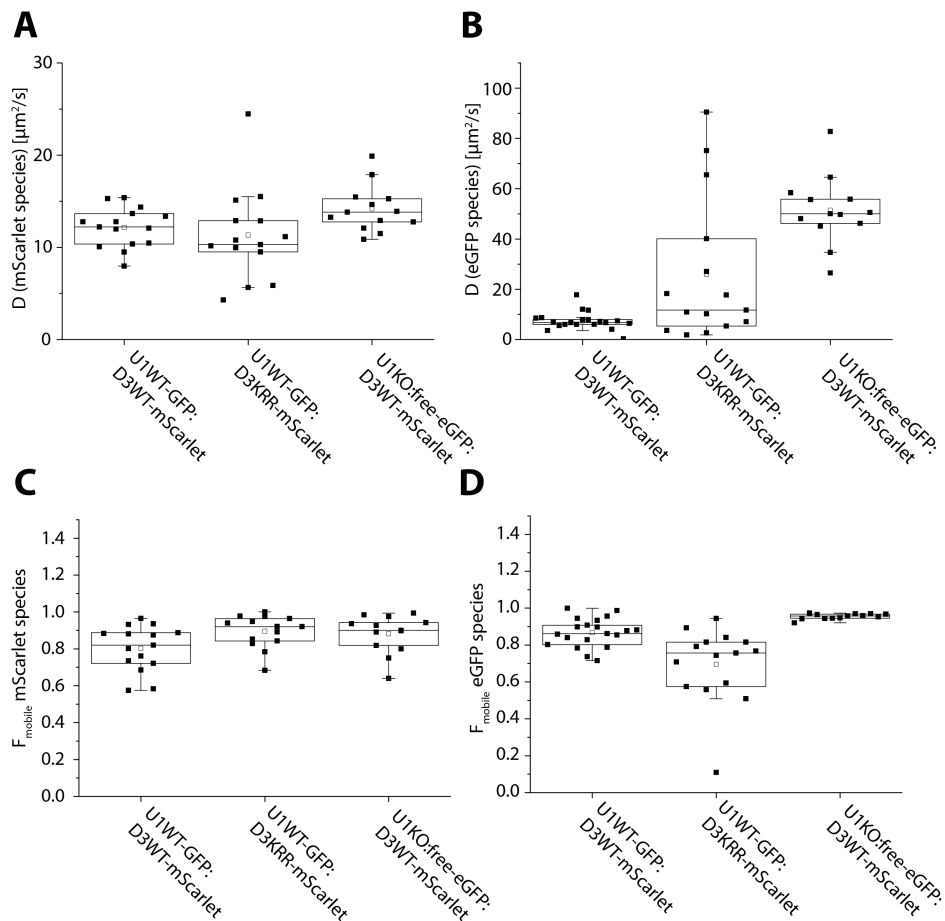
Figure 5.3: **Cytoplasmic diffusion of UHRF1 and DPPA3 in mouse embryonic stem cells**
Box plots showing the following parameters for the cytoplasmic fraction of the indicated cell lines: The diffusion coefficient of the mScarlet- (A) and eGFP-labeled (B) proteins as well as the mobile fraction of the mScarlet- (C) and eGFP-labeled (D) proteins.

The diffusion properties of the mScarlet-labeled protein stays largely unchanged between the investigated cell lines (A + C). Disruption of the interaction between UHRF1-eGFP and DPPA3-mScarlet by expression of the DPPA3-KRR-mScarlet mutant leads to a larger spread in the observed UHRF1-eGFP diffusion coefficients with a tendency towards higher values and a decrease in the fraction of mobile molecules. Free eGFP expressed instead of UHRF1 shows a diffusion coefficient of ~ 50 μm^2/s with almost all of the protein being mobile (B + D). Each black square represents a value obtained from a single cell. The upper and lower limits of the box represent the upper and lower quartiles, the central line the median and the unfilled square the mean of the distribution. The whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge.

5.2.3   Outlook

In the future the interaction of UHRF1 and DPPA3 in the cytoplasm needs to be addressed in more detail. Furthermore future experiments will investigate how the proteins are transported across the nuclear membrane. The latter question could for example be addressed using the pair correlation approach [219].

5.3   Brief description of the data analysis

Raster-scanned images were recorded in two spectrally separated channels. Using TCSPC with PIE (section 2.6.4) cross-talk free data could be collected when only green photons after green excitation and red photons after red excitation were used to construct the images. Using ARICS, the nucleus excluding nuceoli and bright aggregates, or the cytoplasm were selected for analysis. A moving average of three frames corresponding three seconds was subtracted to filter out components that are immobile on this time scale prior to calculation of the spatial correlation functions. The spatial autocorrelation functions as well as cross-correlation functions were calculated according to equation 20 and the obtained correlation functions were fitted with a 2-component model including a mobile and an immobile fraction according to the following equation:

$$G_{fit}\left(\xi, \psi\right) = A_{mob}G_{fit,mob}\left(\xi, \psi\right) + A_{imm}\exp\left(-\delta r^2 \omega_{imm}^{-2}\left(\xi^2 + \psi^2\right)\right) + y_0 \tag{27}$$

in which $A_{mob}G_{fit,mob}\left(\xi, \psi\right)$ represents the equation for free 3D diffusion of a RICS correlation function as given by equation 23 and A_{mob} and A_{imm} are the amplitudes of the mobile and immobile fractions, respectively. The second summand represents the immobile component (on the RICS timescale), which was fitted with a symmetrical 2D Gaussian. Here, ω_{imm} represents the half-width of the Gaussian at $\frac{1}{e^2}$ of the maximum and, for objects smaller than the diffraction limit, it represents ω_r of the PSF. y_0 denotes a baseline offset.

The average number of mobile molecules (on the RICS time scale) in the confocal volume can be derived from the fit by calculating

$$N_{mob} = \left(\frac{\gamma}{A_{mob}}\right)\left(\frac{2\Delta F}{2\Delta F + 1}\right) \tag{28}$$

where γ is the shape factor of $2^{-\frac{2}{3}}$ for a 3D Gaussian volume and the factor $\frac{2\Delta F}{2\Delta F+1}$ corrects the amplitude for the above mentioned moving average subtraction. ΔF thereby denotes the number of frames that were averaged before and after each frame for moving average substraction and was taken as 1 for the data presented here.

The fraction of mobile and immobile molecules were calculated by relating the obtained amplitudes, e.g. $F_{mob} = \frac{A_{mob}}{A_{mob}+A_{imm}}$. However, it has to be noted that these are not absolute estimates of the fractions since the amplitudes are not only depending on the average number of fluorescent molecules but also on their molecular brightness [220]. As such equation 28 only holds true if the components have the same molecular brightness.

The fraction of cross-correlating mobile molecules (denoted as "cross-correlation" in Figure 5.1) was calculated by dividing the amplitude of the mobile component obtained from the cross-correlation fit by the amplitude of the mobile component obtained from the autocorrelation fit of the DPPA3-mScarlet species.

The ratio of UHRF1-to-DPPA3 protein present in the nucleus was calculated by relating the average count rates per pixel obtained for both proteins in the nucleus.

6 Summary and Outlook

This thesis is based on methods that exploit fluorescence emitted by single molecules. It presents experiments in which single-molecule fluorescence is used to study structures and dynamics in single cells. The applications thereby range from the tracking of single virus particles as they enter the host cell to the study of the supramolecular structure of an intracellular protein complex central in the innate immune system, up to the investigation of the interaction of molecules diffusing in the nucleus and cytoplasm of single living cells. In these experiments, the unique advantages of fluorescence microscopy and spectroscopy become apparent. The experiments make use of the specificity, the sensitivity, the live-cell compatibility and the multiplexing capabilities of fluorescence. Furthermore, the experiments illustrate some of the rapid developments in the field of fluorescence microscopy in recent years: The investigations of the supramolecular structure of the endogenous ASC speck applied the recently developed imaging modalities of DNA-PAINT and flat-field-optimized dSTORM, in combination with the recently developed labeling approach of nanobodies. The experiments investigating the diffusion of proteins inside cells exploited recent developments in the area of gene editing as they were performed on cells generated by the CRISPR/Cas9 technology.

The investigations also point out that, in order to fully understand the processes inside cells, both the dynamics and the structures need to be understood. Gaining the highest spatial resolution is equally important to high temporal resolution combined with live-cell compatibility. In both areas, fluorescence microscopy has rapidly developed in the past years. Next to DNA-PAINT, also the recently developed method of MINFLUX [202] enables spatial resolution that was previously unreached in fluorescence microscopy. Lattice light-sheet microscopy [221] enables live-cell measurements with minimized photobleaching and phototoxicity. In the project that investigated the endogenous ASC speck, quasi-dynamic information were extracted from fixed samples. Similar approaches have been taken before [199] and they are likely to be fruitful in the future especially in light of the current efforts to turn super-resolution microscopy into a high throughput method [29, 199, 222, 66]. However, the true value of such measurements becomes apparent when they can be related to data obtained on living cells.

Next to the manifold advantages of fluorescence imaging, the method in its current form also has limitations. One limitation concerns the achievable spatial resolution. Despite recent developments, the spatial resolution achievable by electron microscopy still reaches beyond what is possible by super-resolution fluorescence microscopy. A second limitation of fluorescence microscopy is the necessity to fluorescently tag the molecules to be investigated. This bears the risk of altering the fuctionality of the biological system and needs to be carefully evaluated for every experiment. An inherent disadvantage of almost all imaging modalities is the lack of throughput and the fact that the user selects the field of view to be recorded. This introduces the risk of non-representative results as rare phenotypes in the sample might be missed or overrepresented.

All of the limitations listed above have been addressed through the development of new technologies in recent years: To keep the benefit of molecular specificity offered by fluorescence microscopy but additionally gain higher resolution information on the same sample, fluorescence microscopy has been combined with electron microscopy. Important methods to name here are correlative light and electron microscopy (CLEM) [203, 130, 223] and fluorescence/electron array tomography (FM/EM-AT) [224]. Another approach that could potentially offer insights into details that are currently not observable by fluorescence microscopy is expansion microscopy in combination with super-resolution microscopy [225, 180]. The limitations put onto fluorescence microscopy by the necessity to specifically label the target structure have been addressed by new labeling strategies. Next to the constant development of new fluorescent proteins and dyes, smaller protein binders, such as nanobodies ([182]), aptamers [226, 227, 228] and affimers [229] have been introduced. Furthermore, new protein tags, labeling strategies

based on the introduction of unnatural amino acids into proteins [230] and the CRISPR/Cas9 gene editing technology [231] are important advances in the field.

The lack of throughput and inherent user bias in biological imaging has been met by the development of automated imaging and image analysis. Important contributions in this area were made, by the laboratory of Professor Jan Ellenberg at the European Molecular Biology Laboratory in Heidelberg [232] and by several laboratories in the field of super-resolution microscopy [199, 233, 222, 66]. Also, the idea of the smart and gentle microscope has been proposed [234].

In all mentioned areas, further developments are to be expected in the future. From these developements, also the research questions addressed during this thesis would benefit. To decipher what causes the observed intermediate step in foamy virus fusion, it would be highly interesting to fix the sample once the intermediate step is observed in living cells and subsequently perform CLEM on the sample. Such an experiment would potentially reveal what component constrains the virus capsid and envelope from moving appart during the intermediate step.

Concerning the structure of the endogenous inflammasome complexes, future studies will hopefully address the association of other inflammasome proteins within the endogenous complex. For those experiments, high throughput DNA-PAINT seems to be the method of choice. Such an approach would benefit from the high resolution and multiplexing capabilities offered by DNA-PAINT and combine it with the possibility to image many inflammasome complexes. Since the inflammasome complex is very heterogenous, high statistics are especially important for a comprehensive understanding of its supramolecular organization. Furthermore, automatization would enable very slow data collection (achieved by a low imager strand concentration) over extended timeperiods, which is central for achieving sufficient resolution and sampling of dense complexes. When coupled with automatic exchange of imager strands by a microfluidic device, multiple targets within the inflamamsome could be imaged sequentially. One could envision a setup that first identifies cells in a sample where an inflammasome has been formed using low magnification. Recognition of a target cell could, for example, be based on a fluorescent marker that is linked to the inflammasome complex such as a fluorophore-conjugated DNA-PAINT docking strand. Subsequently, the microscope would, in a second step, automatically focus on the inflammasome complex and perform automated Exchange-PAINT measurements by employing a microfluidic device. Similar approaches have been implemented [235], e.g. for genome imaging [200], and they would very likely give novel insights into the composition of the endogenous inflammasome complex. Such a workflow could be combined well with the latest labeling strategies, such as nanobodies directed against the inflammasome components [201]. In order to also get insights into the very dense regions of the structure, a combination with electron or expansion microscopy would be highly interesting. A combination with electron microscopy would furthermore offer insights into the cellular ultrastructure surrounding the inflammasome complex. Beyond inflammasome complexes the suggested strategies would be well suited to investigate other multiprotein complexes inside the cell for which the term "Supramolecular organizing centers" (SMOCS) has been coined [236].

The experiments investigating the diffusion and interaction of molecules within the nucleus and cytoplasm of single living cells by RICS would also benefit to a great extent from an automated imaging approach. For the investigations on the interaction of the proteins UHRF1 and DPPA3 in the nucleus of embryonic stem cells, such an approach could increase the statistics by automatically detecting and measuring many cells while keeping the benefits of single cell resolution. To enable the microscope to automatically distinguish the cytoplasma and the nucleus of the cells, markers specifically staining the cytoplasm and the nucleus could be expressed in the cells. The microscope, in a first step, could record a snap shot of the cells at low magnification. Here, the two markers could be detected in two widefield channels. Subsequently, the microscope would zoom into single cells and record RICS data either in the nucleus or in the cytoplasm. The expressed markers could, in

this case, also be used to automatically define masks for cell compartments that would enable automatization of the analysis. In a similar fashion, data recording could be performed for the experiments on the alternative inflammasome. Here, the microscope would need to identify cells in which the alternative pathway has been activated. This could be implemented based on a fluorescent marker for mature IL-1β, as this cytokine is processed into its mature form once the alternative pathway has been activated in a cell. RICS measurement would then be performed on the identified cells.

The importance of imaging for biological research was underpinned by the Nobel Prize for the discovery and development of the green fluorescent protein in 2008, super-resolved fluorescence microscopy in 2014 and cryo-electron microscopy in 2017. Given the above described developments, I think that imaging will continue to provide many novel and unexpected insights in the near and far future.

A Measurements on ASC in the alternative NLRP3 inflammasome pathway

A.1 Biological background

Formation of a macrometer-sized inflammasome complex leading to IL-1β release and cell death, described in chapter 4, is generally regarded as a process that is triggered in two steps: First, an NFκB-activating stimulus upregulates the expression of NLRP3 and IL-1β and second, an inflammasome-activating molecule binds to NLRP3 thereby initiating inflammasome assembly [174]. A classical example for the first stimulus is activation of Toll-like receptor 4 (TLR4) by Lipopolysaccharide (LPS) [237]. This two-step activation mechanism for the "classical" NLRP3 inflammasome was mainly established in the murine system. However, even before the inflammasome was discovered, it was observed that, in human monocytes, LPS alone leads to the maturation and secretion of IL-1β [238, 239, 240, 241]. Investigations of the underlying mechanism were, however, hindered by the lack of knockout cell lines. The discovery and establishment of the CRISPR/Cas9 gene editing system enabled the generation of these cell lines. Based on these technical advances, our collaboration partners from the research group of Professor Veit Hornung in the Biochemistry Department at the Ludwig-Maximilians-Universität München, Germany dissected an "alternative" NLRP3 inflammasome activation pathway that drives the observed IL-1β response triggered by LPS stimulation [181]. Hornung and colleagues discovered the pathway in a human B cell cell line which can be transdifferentiated into monocytes [242]. Knockout of the proteins NLRP3, ASC and caspase-1 in these cells revealed that the LPS-dependent IL-1β response depends on these proteins as also the case for the "classical" inflammasome pathway. However, in contrast to the "classical" pathway, the observed response is independent of K$^+$efflux, does not show ASC speck formation and does not result in pyroptotic cell death. Furthermore, Gaidt et al. showed that the response depends on TLR4, TRIF, RIPK1, FADD and activated caspase-8 upstream of NLRP3.

Although NLRP3, ASC and caspase-1 are clearly indispensable for this pathway, a physical interaction of these proteins has not yet been shown. Tagging endogenous ASC with a fluorescent protein showed a homogenous distribution of the protein in unstimulated and LPS-stimulated cells when imaged by diffraction-limited fluorescence microscopy. Based on their well-established interaction in the "classical" inflammasome pathway, it is tempting to speculate that theses proteins interact in oligomers of a few molecules in the "alternative" pathway. Diffraction-limited light microscopy imaging techniques lack resolution and sensitivity to detect such an oligomerization process and, even for super-resolution light microscopy techniques offering the highest resolution and sensitivity, it remains extremly challenging. An alternative approach to investigate molecular intractions inside cells is to correlate fluctuations in the detected signal of fluorescently labeled molecules. Especially the cross-correlation between two differently labeled molecules very sensitively detects molecular interactions. Since these methods are mostly applied to living cells, they avoid potential artifacts due to fixation and also reveal information of the underlying dynamics of the molecules. For these reasons, I applied RICS (section 2.6.3) to addressed the hypothetical interaction of inflammasome proteins in the "alternative" pathway and measured the diffusion properties of ASC and NLRP3 in transdifferentiated BlaER1 monocytes expressing only ASC-TagRFP or ASC-TagRFP and NLRP3-eGFP on an endogenous level.

A.2 Experimental procedure

A.2.1 Cell culture

If not stated differently all reagents where purchased from Thermo Fisher Scientific, Massachusetts, USA. BlaER1 B cells were cultivated at 37°C and 5% CO_2 in culture medium consisting of Roswell Park Memorial Institute (RPMI) 1640 medium (21875034) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS) (10500064), 1 mM sodium pyruvate (11360039) and 100 U/ml Penicillin/Streptomycin (15140122), and maintained at a density between $1x10^5$ and $1x10^6$ cells per milliliter. Cells were transdifferentiated into BlaER1 monocytes/macrophages for three nights by seeding them at 60,000 cells/cm^2 in culture medium supplemented with 50 ng/ml IL-3 (200-03 B, Peprotech, New Jersey, USA), 50 ng/ml M-CSF (300-25 B, Peprotech) and 500 nM β-Estradiol (E8875, Sigma-Aldrich) on cell culture-treated plastic bottom slides (80826, ibidi, Gräfelfing, Germany). On the fourth day, cells were trypsinized and transferred to glass bottom slides coated with 100 μg/ml poly-l-ornithine (A-004-C, Merck-Millipore, Massachusetts, USA) at the same density in culture medium without growth factors and incubated overnight at 37°C and 5% CO_2. Differentiation of the cells into monocytes was confirmed by a change in morphology from completely round cells into cells with cellular extrusions (Figure A.2A, E). The alternative inflammasome pathway was activated by incubating the cells with 200 ng/ml LPS-EB (tlrl-3pelps, Invivogen) for 14 hours. RICS measurements were conducted within 4 hours after removing the cells from the incubator on the home-built confocal microscope setup described in section 2.4.3 at 37°C. For cells in which the alternative inflammasome was activated, the LPS stimulus remained on the cells during the measurements or was exchanged by fresh medium (Figure A.2H, left vs. middle box plot).

A.2.2 Raster image correlation spectroscopy measurements

Data were recorded on the home-built confocal setup described in section 2.4.3. Microscope alignment was verified prior to RICS measurements by performing a single-point FCS measurement on an Atto565 solution in H_2O, at 10 μW laser power of the 561 nm laser (measured in the collimated space before entering the galvo-scanning mirror system), 1 μm above the surface of the coverslip. Prior to FCS recording the glass surface the coverslip was coated for 20 min by incubation with 2 mg/ml Bovine Serum Albumin (BSA) (B9000S, New England Biolabs, Frankfurt am Main, Germany) in H_2O to prevent sticking of the dye.

The obtained autocorrelation function was fitted with the model for free 3D diffusion given in equation 19 in which the diffusion coefficient was fixed to 535 μm^2/s for measurements at 37°C (Figure A.1). The obtained molecular brightness served as an indication for setup alignment and was compared to previously measured values obtained under conditions when the setup was optimally aligned. The focus dimensions, ω_r and w_z, retrieved from the autocorrelation function were taken as inputs for the analysis of the following RICS experiments.
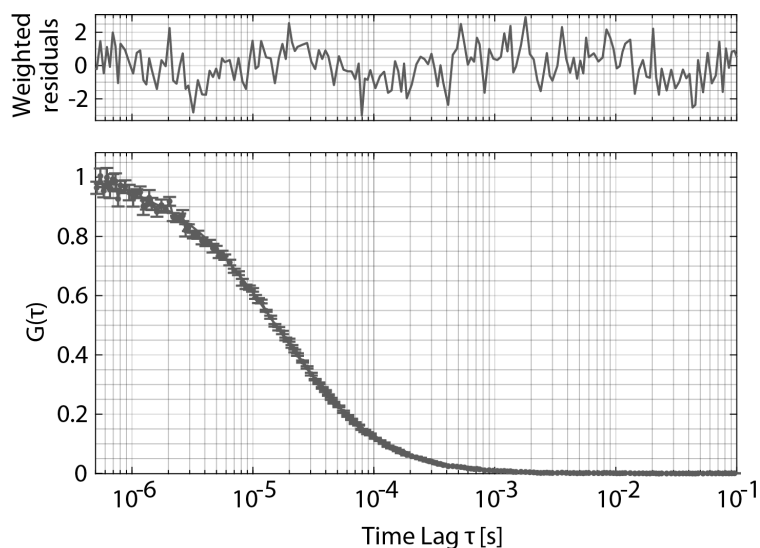
Figure A.1: **Exemplary single-point FCS calibration curve of Atto565 in water at 37°C to verify proper setup alignment prior to raster image correlation spectroscopy measurements**
The dots represent the measured data and the solid line is a fit with the function for free 3D diffusion (equation 19). The upper graph shows the weighted residuals of the fit. Error bars indicate the standard deviation of the measured data points derived from FCS curves calculated for the individual time bins of the underlying photon stream

For RICS measurements, ASC-TagRFP was excited using a 561 nm laser at 3 µW (measured in the collimated space before entering the galvo-scanning mirror system). Single cells were imaged in a field of view of 14 µm x 14 µm consisting of 300 x 300 pixels, corresponding to a pixel size of 46.7 nm. In total, 200 frames were recorded with a pixel dwell time of 11 µs resulting in a line and frame time of 3.3 ms and 1 s, respectively. Excitation was separated from emission using a 405/488/561/635 polychroic mirror and a 660 longpass filter reflected the fluorescence onto the avalanche photodiode detector in front of which a 630/75 emission filter was applied.

Images were reconstructed from the raw photon stream and analysed using the Microtime Image Analysis software developed in our group [243]. Prior to image correlation static features on the timescale of three seconds were filtered out by a moving average subtraction and an arbitrary region selection was applied to select the cell outline and used to exclude intracellular vesicles, aggregates and the very edge of the cell to avoid artefacts in the correlation (Figure A.2A right, E right) [101]. The spatial autocorrelation function (SACF) was calculated for each frame according to equation 20 and averaged over all frames.

A.3   Results

I started off by measuring the diffusion of ASC-TagRFP molecules in the cytosol of unstimulated cells. The obtained average SACF could be fit well with the model function for free 3D diffusion of a single species given by equation 23 as indicated by the low weighted residuals of the fit (Figure A.2B, C). This suggests that ASC-TagRFP is freely diffusing in the cytosol of unstimulated cells and the measurements showed a diffusion coefficient of 10.9 µm^2/s +/- 0.2 SEM for mobile ASC-TagRFP molecules (Figure A.2D). This agrees with the expected value for a ~ 48 kDa protein (21.6 kDa ASC + 26.1 kDa TagRFP) [41] and indicates that ASC-TagRFP is mainly present as monomers or dimers in unstimulated cells. To verify

that the contribution of autofluorescence to the measurements can be omitted, wildtype cells were measured under the same conditions which showed a neglectible count rate (data not shown).

Repetition of the measurements in cells stimulated with LPS to activate the "alternative" inflammasome pathway showed largely the same diffusion behaviour of ASC-TagRFP data suggesting that there is no change in the oligomerization status of ASC between both conditions (Figure A.2F - H). For measurements in which the LPS stimulus remained on the cells, exactly the same average diffusion coefficient of 10.9 μm^2/s +/- 0.3 SEM was obtained (Figure A.2H, left box plot). The measurements for which the stimulus was removed and exchanged by fresh medium showed a slightly higher average diffusion coefficient of 12.2 μm^2/s +/- 0.4 SEM (Figure A.2H, middle box plot). Although this result might be statistically, significant its biological relevance is highly questionable and therefore the two data sets were pooled, which resulted in an average diffusion coefficient of 11.3 μm^2/s +/- 0.3 SEM, which is insignificantly different to the value obtained for in unstimulated cells. However, one has to keep in mind the limited sensitivity of the diffusion coefficient as a read out for a change in oligomerization when judging these results: a doubling of the molecular mass of the diffusing species only leads to a 26% increase in the diffusion coefficient. Especially in situations in which multiple species with different oligomerization status are present in the same sample a quantitative analysis becomes very difficult based on the autocorrelation. For detecting molecular interactions the cross-correlation function is much more sensitive. Thus, we next set out to address a possible interaction of ASC and NLRP3 after activation of the "alternative" pathway by measuring the cross-correlation in cells expressing ASC-TagRFP and NLRP3-eGFP. However, our attempts to retrieve a correlation function for NLRP3-eGFP from the cells so far proved unsucessful due to the fact that endogenous NLRP3 expression is very low. Although RICS is generally sensitive enough to extract information from the diffusion of single eGFP molecules this does not seem to be the case for the intracellular measurements described here. This is most likely due to the increased background and autofluorescence and the resulting low signal-to-noise inside the cells.
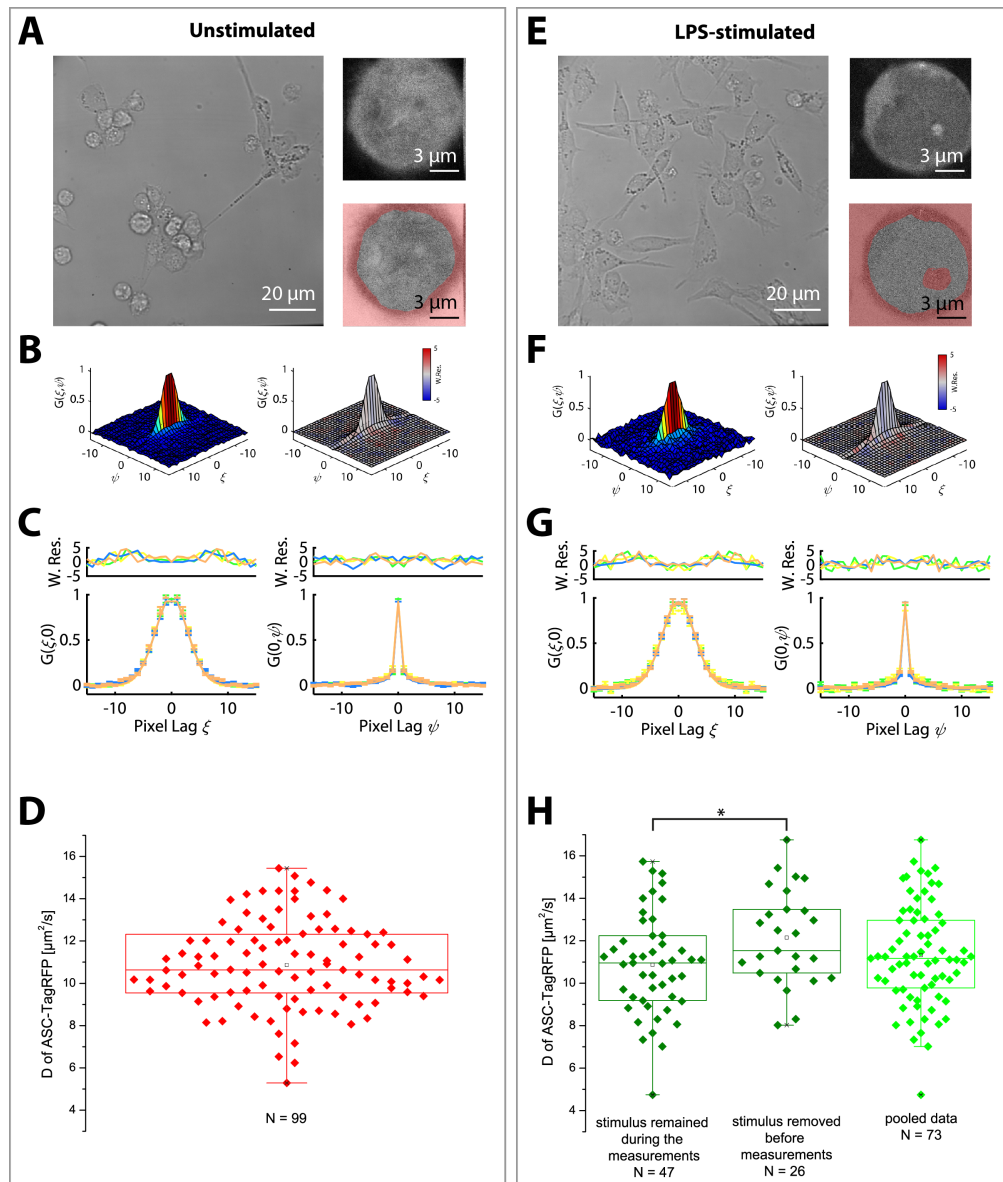
Figure A.2: Determination of the diffusion coefficient of ASC-TagRFP in BlaER1 cells by raster image correlation spectroscopy

A) A brightfield image of BlaER1 transdifferentiated monocytes is shown; Note the characteristic macrophage-like phenotype indicated by cellular extrusions from the cells; right top: An exemplary confocal image of a BlaER1 monocyte as recorded for ASC-TagRFP RICS analysis (top) and the corresponding analysis region selection (red areas are excluded from the analysis) (bottom)

B) An exemplary frame-averaged RICS spatial autocorrelation function (SACF) obtained from ASC-TagRFP in a single cell (left) and a surface representation of the weighted residuals of the applied fit model (right). The SACF is color-coded according to the correlation value with blue representing low correlation and red representing high correlation. In the surface representation of the weighted residuals red and blue encode for positive and negative weighted residuals of the applied fit model.

Continued Figure A.2:

C) Exemplary projections of the fit along the *x*- (left) and *y*-axis (right) of the SACF; individual cells are color-coded; error bars indicate the standard deviation of the fit calculated from the individual frames of the recorded image stack

D) Box plot of the ASC-TagRFP diffusion coefficient obtained by the RICS measurements; The upper and lower limits of the box represent the 25% and the 75% percentile, respectively. The central line represents the median and the square the mean. The crosses represent the 1% and the 99% percentile, respectively. The whiskers extend to the minimum and the maximum, respectively.

E)-H) Panels show data as mentioned for A)-D) for data recorded on LPS-stimulated cells; The difference between the left and the middle boxplot was confirmed using the two sample T test resulting in a P value of 0.027;

Graphs in D) and H) shows pooled data from 10 independent experiments

A.4   Discussion and outlook

A very promising experiment would be to address the potential interaction of ASC in the alternative pathway based on cross-correlation measurements in cells expressing ASC-TagRFP and ASC-eGFP on both allels, respectively. In such an experiment, one could take advantage of the above mentioned high sensitivity of the cross-correlation function as well as the sufficiently high expression level of ASC. There are several reasons why the results presented here have to be juged with care. The morphology of the cells compared with data reported in literature [178] suggest that the degree of differentiation of the investigated cells was variable between preparations and in some cases might be questionable. Possible reasons for that are manifold ranging from the activity of the applied growth factors to the integrity of the cells before the start of the experiment. It worth pointing out the complexity of the underlying experimental protocol here. Although the obtained results on the diffusion properties of ASC are very convincing and most likely hold true in stronger differentiated cells, the data presented here are preliminary and experiments need to be repeated.

Furthermore, one cannot say for sure whether the alternative inflammasome pathway was successfully activated in the measured cells. Future experiments with an established readout that monitors alternative inflammasome activation on a single-cell level would enable targeted measurements of cells in which the alternative inflammasome is activated. Such a readout could be, for example, based on IL-1β maturation. IL-1β is expressed as pro-IL-1β and is cleaved into its active form by caspase-1 and subsequently released from the cell. As caspase-1 becomes activated by the alternative inflammasome pro-IL-1β maturation could serve as a readout for alternative inflammasome activation in single cells. Such an assay could exploit Förster-Resonance-Energy-Transfer (FRET) and utilize an IL-1β construct in which immature IL-1β is labeled with a donor-acceptor pair at the pro-domain and at a suitable position within the sequence of the active interleukin, respectively and maturation would be readout as a decrease in the FRET signal. Similarly, a readout based on cross-correlation spectroscopy is imaginable in which IL-1β maturation would separate two fluorescent proteins attached to the immature protein at the pro-domain and at a suitable position within the sequence of the active interleukin thereby leading to a decrease in cross-correlation. Alternatively, a dequenching mechanism invoing a quencher attached to the pro-domain of IL-1β and a fluorescent probe at a suitable position within the sequence of the active IL-1β could be exploited. One could also imagine a readout based on the release of IL-1β from the cells. How suitable these ideas are in practice would need to be carefully evaluated before their implementation.

Another approach to tackle the problem of single-cell activation would be to increase statistics and thereby reveal even small differences in the underlying cell populations. It was with this idea in mind that the data sets presented here comprise a comparably large number of cells. However, automated data acquisition could increase statistics even further while keeping the benefits of single-cell resolution. Such an approach would also benefit from the above mentioned single-cell activation readout. One could imagine a microscope setup which, in a first step, automatically identifies cells producing mature IL-1β (thereby confirming the activation of the alternative inflammasome), for example based on fluorescence in an additional lower magnification widefield channel, and, in a second step, records RICS data on the identified cells. Similar ideas of automated microscopy have been suggested [232, 234]. It is likely that many biological questions would benefit from such automated approaches since it removes the user bias inherent to the majority of current imaging approaches in general and single-cell imaging in particular.

References

[1] Nirmalya Bag and Thorsten Wohland. Imaging fluorescence fluctuation spectroscopy: New tools for quantitative bioimaging. *Annual Review of Physical Chemistry*, 65(1):225–248, 2014.

[2] Sergey Y. Tetin. *Fluorescence fluctuation spectroscopy (FFS)*, volume 519. Academic Press, 2013.

[3] George G. Stokes. Xxx. on the change of refrangibility of light. *Philosophical Transactions of the Royal Society*, 142:463–562, 1852.

[4] A. Jablonski. Efficiency of anti-stokes fluorescence in dyes. *Nature*, 131(3319):839–840, 1933.

[5] M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln. *Annalen der Physik*, 389(20):457–484, 1927.

[6] Michael Kasha. Characterization of electronic transitions in complex molecules. *Discussions of the Faraday Society*, 9(0):14–19, 1950.

[7] O. Shimomura, F. H. Johnson, and Y. Saiga. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, aequorea. *Journal of Cellular and Comparative Physiology*, 59:223–39, 1962.

[8] M Chalfie, Y Tu, G Euskirchen, WW Ward, and DC Prasher. Green fluorescent protein as a marker for gene expression. *Science*, 263(5148):802–805, 1994.

[9] Roger Heim, Andrew B. Cubitt, and Roger Y. Tsien. Improved green fluorescence. *Nature*, 373(6516):663–664, 1995.

[10] Roger Y. Tsien. The green fluorescent protein. *Annual Review of Biochemistry*, 67(1):509–544, 1998.

[11] Dmitriy M. Chudakov, Mikhail V. Matz, Sergey Lukyanov, and Konstantin A. Lukyanov. Fluorescent proteins and their applications in imaging living cells and tissues. *Physiological Reviews*, 90(3):1103–1163, 2010.

[12] Brendan P. Cormack, Raphael H. Valdivia, and Stanley Falkow. Facs-optimized mutants of the green fluorescent protein (GFP). *Gene*, 173(1):33–38, 1996.

[13] E. M. Merzlyak, J. Goedhart, D. Shcherbo, M. E. Bulina, A. S. Shcheglov, A. F. Fradkov, A. Gaintzeva, K. A. Lukyanov, S. Lukyanov, T. W. Gadella, and D. M. Chudakov. Bright monomeric red fluorescent protein with an extended fluorescence lifetime. *Nature Methods*, 4(7):555–7, 2007.

[14] Daphne S. Bindels, Lindsay Haarbosch, Laura van Weeren, Marten Postma, Katrin E. Wiese, Marieke Mastop, Sylvain Aumonier, Guillaume Gotthard, Antoine Royant, Mark A. Hink, and Theodorus W. J. Gadella. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nature Methods*, 14(1):53–56, 2017.

[15] Luke D. Lavis and Ronald T. Raines. Bright building blocks for chemical biology. *ACS Chemical Biology*, 9(4):855–866, 2014.

[16] M. Ormö, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington. Crystal structure of the aequorea victoria green fluorescent protein. *Science*, 273(5280):1392–5, 1996.

[17] John Frederick William Herschel. Iv. amórfwga, no. i. - on a case of superficial colour presented by a homogeneous liquid internally colourless. *Philosophical Transactions of the Royal Society*, 135:143–145, 1845.

[18] Hermann Helmholtz. Die theoretischen Grenzen für die Leistungsfähigkeit der Mikroskope. *Annalen der Physik*, pages 557–584, 1874.

[19] August Köhler. Mikrophotographische Untersuchungen mit ultraviolettem Licht. *Zeitschrift für Wissenschaftliche Mikroskopie*, 21:128–165 and 273–304, 1904.

[20] Karl Reichert. Das Fluoreszenzmikroskop. *Phys. Z.*, 12:1010–1011, 1911.

[21] H. Lehmann. Das Lumineszenz-Mikroskop seine Grundlagen und seine Anwendungen. *Zeitschrift für wissenschaftliche Mikroskopie*, 30:417–470, 1913.

[22] Philipp Ellinger and August Hirt. Mikroskopische Untersuchungen an lebenden Organen. *Naunyn-Schmiedebergs Archiv für experimentelle Pathologie und Pharmakologie*, 145(4):193–210, 1929.

[23] Albert H. Coons, Hugh J. Creech, R. Norman Jones, and Ernst Berliner. The demonstration of pneumococcal antigen in tissues by the use of fluorescent antibody. *The Journal of Immunology*, 45(3):159, 1942.

[24] E. Abbe. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9(1):413–468, 1873.

[25] G. B. Airy. On the diffraction of an object-glass with circular aperture. *Transactions of the Cambridge Philosophical Society*, 5:283, 1835.

[26] John William Strutt. Investigations in optics, with special reference to the spectroscope. *Philosophical Magazine*, 8(49):261–274, 1879.

[27] H. Kirshner, F. Aguet, D. Sage, and M. Unser. 3D PSF fitting for fluorescence microscopy: implementation and localization application. *Journal of Microscopy*, 249(1):13–25, 2013.

[28] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, 1949.

[29] Kyle M. Douglass, Christian Sieben, Anna Archetti, Ambroise Lambert, and Suliana Manley. Super-resolution imaging of multiple cells by optimised flat-field epi-illumination. *Nature Photonics*, 10(11):705–708, 2016.

[30] E. J. Ambrose. The movements of fibrocytes. *Experimental Cell Research*, Suppl 8:54–73, 1961.

[31] D Axelrod. Cell-substrate contacts illuminated by total internal reflection fluorescence. *Journal of Cell Biology*, 89(1):141–145, 1981.

[32] N. L. Thompson, T. P. Burghardt, and D. Axelrod. Measuring surface dynamics of biomolecules by total internal reflection fluorescence with photobleaching recovery or correlation spectroscopy. *Biophysical Journal*, 33(3):435–454, 1981.

[33] Makio Tokunaga, Naoko Imamoto, and Kumiko Sakata-Sogawa. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nature Methods*, 5(2):159–161, 2008.

[34] Marvin Minsky. "Microscopy apparatus", 1961.

[35] H. Naora. Microspectrophotometry and cytochemical analysis of nucleic acids. *Science*, 114(2959):279–80, 1951.

[36] J. G. White, W. B. Amos, and M. Fordham. An evaluation of confocal versus conventional imaging of biological structures by fluorescence light microscopy. *Journal of Cell Biology*, 105(1):41–48, 1987.

[37] Ulrich Kubitschek. *Fluorescence Microscopy: From Principles to Biological Applications*, volume 2nd Edition. Wiley-VCH, 2017.

[38] K. A. Zachariasse. Einzelphotonenzählung: Time-correlated single photon counting. *Nachrichten aus Chemie, Technik und Laboratorium*, 33(10):896–896, 1985.

[39] Barbara K. Müller, Evgeny Zaychikov, Christoph Bräuchle, and Don C. Lamb. Pulsed interleaved excitation. *Biophysical Journal*, 89(5):3508–3522, 2005.

[40] Jelle Hendrix, Waldemar Schrimpf, Matthias Höller, and Don C. Lamb. Pulsed interleaved excitation fluctuation imaging. *Biophysical Journal*, 105(4):848–861, 2013.

[41] J. Hendrix, V. Baumgärtel, W. Schrimpf, S. Ivanchenko, M. A. Digman, E. Gratton, H. G. Kräusslich, B. Müller, and D. C. Lamb. Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-interacting gag oligomers. *Journal of Cell Biology*, 210(4):629–46, 2015.

[42] M. D. Egger and M. Petráň. New reflected-light microscope for viewing unstained brain and ganglion cells. *Science*, 157(3786):305–7, 1967.

[43] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682, 2012.

[44] T. W. Ridler and S. Calvard. Picture thresholding using an iterative selection method. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(8):630–632, 1978.

[45] Liang-Kai Huang and Mao-Jiun J. Wang. Image thresholding by minimizing the measures of fuzziness. *Pattern Recognition*, 28(1):41–51, 1995.

[46] G. Seisenberger, M. U. Ried, T. Endress, H. Büning, M. Hallek, and C. Bräuchle. Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. *Science*, 294(5548):1929–32, 2001.

[47] Shu-Lin Liu, Zhi-Gang Wang, Hai-Yan Xie, An-An Liu, Don C. Lamb, and Dai-Wen Pang. Single-virus tracking: From imaging methodologies to virological applications. *Chemical Reviews*, 120(3):1936–1979, 2020.

[48] Daniel L. Floyd, Justin R. Ragains, John J. Skehel, Stephen C. Harrison, and Antoine M. van Oijen. Single-particle kinetics of influenza virus membrane fusion. *Proceedings of the National Academy of Sciences*, 105(40):15382–15387, 2008.

[49] Michael J. Rust, Melike Lakadamyali, Feng Zhang, and Xiaowei Zhuang. Assembly of endocytic machinery around individual influenza viruses during viral entry. *Nature Structural & Molecular Biology*, 11(6):567–573, 2004.

[50] Melike Lakadamyali, Michael J. Rust, Hazen P. Babcock, and Xiaowei Zhuang. Visualizing infection of individual influenza viruses. *Proceedings of the National Academy of Sciences*, 100(16):9280–9285, 2003.

[51] Lucas Pelkmans, Jürgen Kartenbeck, and Ari Helenius. Caveolar endocytosis of simian virus 40 reveals a new two-step vesicular-transport pathway to the er. *Nature Cell Biology*, 3(5):473–483, 2001.

[52] Lucas Pelkmans, Daniel Püntener, and Ari Helenius. Local actin polymerization and dynamin recruitment in sv40-induced internalization of caveolae. *Science*, 296(5567):535–539, 2002.

[53] Lucas Pelkmans, Thomas Bürli, Marino Zerial, and Ari Helenius. Caveolin-stabilized membrane domains as multi-functional transport and sorting devices in endocytic membrane traffic. *Cell*, 118(6):767–780, 2004.

[54] G. F. Elphick, W. Querbes, J. A. Jordan, G. V. Gee, S. Eash, K. Manley, A. Dugan, M. Stanifer, A. Bhatnagar, W. K. Kroeze, B. L. Roth, and W. J. Atwood. The human polyomavirus, JCV, uses serotonin receptors to infect cells. *Science*, 306(5700):1380–3, 2004.

[55] Vilja Pietiäinen, Varpu Marjomäki, Paula Upla, Lucas Pelkmans, Ari Helenius, and Timo Hyypiä. Echovirus 1 endocytosis into caveosomes requires lipid rafts, dynamin ii, and signaling events. *Molecular Biology of the Cell*, 15(11):4911–4925, 2004.

[56] A. Dupont, K. Stirnnagel, D. Lindemann, and D. C. Lamb. Tracking image correlation: combining single-particle tracking and image correlation. *Biophysical Journal*, 104(11):2373–82, 2013.

[57] Norman Bobroff. Position measurement with a resolution and noise-limited instrument. *Review of Scientific Instruments*, 57(6):1152–1157, 1986.

[58] Russell E. Thompson, Daniel R. Larson, and Watt W. Webb. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal*, 82(5):2775–2783, 2002.

[59] A. Yildiz, J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman, and P. R. Selvin. Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science*, 300(5628):2061–5, 2003.

[60] Stefan W. Hell and Jan Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19(11):780–782, 1994.

[61] Thomas A. Klar and Stefan W. Hell. Subdiffraction resolution in far-field fluorescence microscopy. *Optics Letters*, 24(14):954–956, 1999.

[62] R. Heintzmann and C. Cremer. Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating. In *European Conference on Biomedical Optics*.

[63] M. G. L. Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy*, 198(2):82–87, 2000.

[64] E. Betzig. Proposed method for molecular optical imaging. *Optics Letters*, 20(3):237–239, 1995.

[65] W. E. Moerner and L. Kador. Optical detection and spectroscopy of single molecules in a solid. *Physical Review Letters*, 62(21):2535–2538, 1989.

[66] Dora Mahecic, Ilaria Testa, Juliette Griffié, and Suliana Manley. Strategies for increasing the throughput of super-resolution microscopies. *Current Opinion in Chemical Biology*, 51:84–91, 2019.

[67] Christian Sieben, Niccolo Banterle, Kyle M. Douglass, Pierre Gönczy, and Suliana Manley. Multicolor single-particle reconstruction of protein complexes. *Nature Methods*, 15(10):777–780, 2018.

[68] Ralf Jungmann, Christian Steinhauer, Max Scheible, Anton Kuzyk, Philip Tinnefeld, and Friedrich C. Simmel. Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on dna origami. *Nano Letters*, 10(11):4756–4761, 2010.

[69] George H. Patterson and Jennifer Lippincott-Schwartz. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science*, 297(5588):1873, 2002.

[70] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz, and Harald F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, 2006.

[71] Daniel Sage, Hagai Kirshner, Thomas Pengo, Nico Stuurman, Junhong Min, Suliana Manley, and Michael Unser. Quantitative evaluation of software packages for single-molecule localization microscopy. *Nature Methods*, 12(8):717–724, 2015.

[72] Mike Heilemann, Sebastian van de Linde, Mark Schüttpelz, Robert Kasper, Britta Seefeldt, Anindita Mukherjee, Philip Tinnefeld, and Markus Sauer. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie International Edition*, 47(33):6172–6176, 2008.

[73] Yiming Li, Markus Mund, Philipp Hoess, Joran Deschamps, Ulf Matti, Bianca Nijmeijer, Vilma Jimenez Sabinina, Jan Ellenberg, Ingmar Schoen, and Jonas Ries. Real-time 3D single-molecule localization using experimental point spread functions. *Nature Methods*, 15(5):367–369, 2018.

[74] Fang Huang, Tobias M. P. Hartwich, Felix E. Rivera-Molina, Yu Lin, Whitney C. Duim, Jane J. Long, Pradeep D. Uchil, Jordan R. Myers, Michelle A. Baird, Walther Mothes, Michael W. Davidson, Derek Toomre, and Joerg Bewersdorf. Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. *Nature Methods*, 10(7):653–658, 2013.

[75] Yina Wang, Joerg Schnitzbauer, Zhe Hu, Xueming Li, Yifan Cheng, Zhen-Li Huang, and Bo Huang. Localization events-based sample drift correction for localization microscopy with redundant cross-correlation algorithm. *Optics Express*, 22(13):15982–15991, 2014.

[76] H. P. Kao and A. S. Verkman. Tracking of single fluorescent particles in three dimensions: use of cylindrical optics to encode particle position. *Biophysical Journal*, 67(3):1291–1300, 1994.

[77] Bo Huang, Wenqin Wang, Mark Bates, and Xiaowei Zhuang. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, 319(5864):810–813, 2008.

[78] Gleb Shtengel, James A. Galbraith, Catherine G. Galbraith, Jennifer Lippincott-Schwartz, Jennifer M. Gillette, Suliana Manley, Rachid Sougrat, Clare M. Waterman, Pakorn Kanchanawong, Michael W. Davidson, Richard D. Fetter, and Harald F. Hess. Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. *Proceedings of the National Academy of Sciences*, 106(9):3125–3130, 2009.

[79] Sri Rama Prasanna Pavani, Michael A. Thompson, Julie S. Biteen, Samuel J. Lord, Na Liu, Robert J. Twieg, Rafael Piestun, and W. E. Moerner. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proceedings of the National Academy of Sciences*, 106(9):2995–2999, 2009.

[80] Ulrike Endesfelder, Sebastian Malkusch, Franziska Fricke, and Mike Heilemann. A simple method to estimate the average localization precision of a single-molecule localization microscopy experiment. *Histochemistry and Cell Biology*, 141(6):629–638, 2014.

[81] Robert P. J. Nieuwenhuizen, Keith A. Lidke, Mark Bates, Daniela Leyton Puig, David Grünwald, Sjoerd Stallinga, and Bernd Rieger. Measuring image resolution in optical nanoscopy. *Nature Methods*, 10(6):557–562, 2013.

[82] A. Descloux, K. S. Grußmayer, and A. Radenovic. Parameter-free image resolution estimation based on decorrelation analysis. *Nature Methods*, 16(9):918–924, 2019.

[83] Michael J. Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–796, 2006.

[84] J. Fölling, M. Bossi, H. Bock, R. Medda, C. A. Wurm, B. Hein, S. Jakobs, C. Eggeling, and S. W. Hell. Fluorescence nanoscopy by ground-state depletion and single-molecule return. *Nature Methods*, 5(11):943–5, 2008.

[85] Alexey Sharonov and Robin M. Hochstrasser. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proceedings of the National Academy of Sciences*, 103(50):18911–18916, 2006.

[86] Joerg Schnitzbauer, Maximilian T. Strauss, Thomas Schlichthaerle, Florian Schueder, and Ralf Jungmann. Super-resolution microscopy with DNA-PAINT. *Nature Protocols*, 12(6):1198–1228, 2017.

[87] Ralf Jungmann, Maier S. Avendano, Mingjie Dai, Johannes B. Woehrstein, Sarit S. Agasti, Zachary Feiger, Avital Rodal, and Peng Yin. Quantitative super-resolution imaging with qPAINT. *Nature Methods*, 13(5):439–442, 2016.

[88] Alexander Auer, Maximilian T. Strauss, Thomas Schlichthaerle, and Ralf Jungmann. Fast, background-free DNA-PAINT imaging using fret-based probes. *Nano Letters*, 17(10):6428–6434, 2017.

[89] Kenny KH Chung, Zhao Zhang, Phylicia Kidd, Yongdeng Zhang, Nathan D Williams, Bennett Rollins, Yang Yang, Chenxiang Lin, David Baddeley, and Joerg Bewersdorf. Fluorogenic probe for fast 3D whole-cell DNA-PAINT. *bioRxiv*, page 2020.04.29., 2020.

[90] F. Schueder, J. Stein, F. Stehr, A. Auer, B. Sperl, M. T. Strauss, P. Schwille, and R. Jungmann. An order of magnitude faster DNA-PAINT imaging by optimized sequence design and buffer conditions. *Nature Methods*, 16(11):1101–1104, 2019.

[91] Sebastian Strauss and Ralf Jungmann. Up to 100-fold speed-up and multiplexing in optimized DNA-PAINT. *Nature Methods*, 17(8):789–791, 2020.

[92] P. Blumhardt, J. Stein, J. Mücksch, F. Stehr, J. Bauer, R. Jungmann, and P. Schwille. Photo-induced depletion of binding sites in DNA-PAINT microscopy. *Molecules*, 23(12), 2018.

[93] Douglas Magde, Elliot Elson, and W. W. Webb. Thermodynamic fluctuations in a reacting system—measurement by fluorescence correlation spectroscopy. *Physical Review Letters*, 29(11):705–708, 1972.

[94] Elliot L. Elson and Douglas Magde. Fluorescence correlation spectroscopy. I. conceptual basis and theory. *Biopolymers*, 13(1):1–27, 1974.

[95] Douglas Magde, Elliot L. Elson, and Watt W. Webb. Fluorescence correlation spectroscopy. ii. an experimental realization. *Biopolymers*, 13(1):29–61, 1974.

[96] R. Rigler, Ü. Mets, J. Widengren, and P. Kask. Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. *European Biophysics Journal*, 22(3):169–175, 1993.

[97] N. O. Petersen, P. L. Höddelius, P. W. Wiseman, O. Seger, and K. E. Magnusson. Quantitation of membrane receptor distributions by image correlation spectroscopy: concept and application. *Biophysical Journal*, 65(3):1135–46, 1993.

[98] Mamta Srivastava and Nils O. Petersen. Image cross-correlation spectroscopy: A new experimental biophysical approach to measurement of slow diffusion of fluorescent molecules. *Methods in Cell Science*, 18(1):47–54, 1996.

[99] M. A. Digman, P. Sengupta, P. W. Wiseman, C. M. Brown, A. R. Horwitz, and E. Gratton. Fluctuation correlation spectroscopy with a laser-scanning microscope: exploiting the hidden time structure. *Biophysical Journal*, 88(5):L33–6, 2005.

[100] M. A. Digman, C. M. Brown, P. Sengupta, P. W. Wiseman, A. R. Horwitz, and E. Gratton. Measuring fast dynamics in solutions and cells with a laser scanning microscope. *Biophysical Journal*, 89(2):1317–27, 2005.

[101] J. Hendrix, T. Dekens, W. Schrimpf, and D. C. Lamb. Arbitrary-region raster image correlation spectroscopy. *Biophysical Journal*, 111(8):1785–1796, 2016.

[102] Michelle A. Digman, Paul W. Wiseman, Alan R. Horwitz, and Enrico Gratton. Detecting protein complexes in living cells from laser scanning confocal image sequences by the cross correlation raster image spectroscopy method. *Biophysical Journal*, 96(2):707–716, 2009.

[103] Thomas Weidemann, Malte Wachsmuth, Michael Tewes, Karsten Rippe, and Jörg Langowski. Analysis of ligand binding by two-colour fluorescence cross-correlation spectroscopy. *Single Molecules*, 3(1):49–61, 2002.

[104] A. S. Khan, J. Bodem, F. Buseyne, A. Gessain, W. Johnson, J. H. Kuhn, J. Kuzmak, D. Lindemann, M. L. Linial, M. Löchelt, M. Materniak-Kornas, M. A. Soares, and W. M. Switzer. Spumaretroviruses: Updated taxonomy and nomenclature. *Virology*, 516:158–164, 2018.

[105] Kathrin Plochmann, Anne Horn, Eva Gschmack, Nicole Armbruster, Jennifer Krieg, Tatiana Wiktorowicz, Conrad Weber, Kristin Stirnnagel, Dirk Lindemann, Axel Rethwilm, and Carsten Scheller. Heparan sulfate is an attachment factor for foamy virus entry. *Journal of Virology*, 86(18):10028–10035, 2012.

[106] Marcus Picard-Maureau, Gergely Jarmy, Angelika Berg, Axel Rethwilm, and Dirk Lindemann. Foamy virus envelope glycoprotein-mediated entry involves a ph-dependent fusion process. *Journal of Virology*, 77(11):6587–6587, 2003.

[107] C. Petit, M. L. Giron, J. Tobaly-Tapiero, P. Bittoun, E. Real, Y. Jacob, N. Tordo, H. De Thé, and A. Saïb. Targeting of incoming retroviral gag to the centrosome involves a direct interaction with the dynein light chain 8. *Journal of Cell Science*, 116(Pt 16):3433–42, 2003.

[108] G. S. Patton, O. Erlwein, and M. O. McClure. Cell-cycle dependence of foamy virus vectors. *Journal of General Virology*, 85(Pt 10):2925–2930, 2004.

[109] G. Trobridge and D. W. Russell. Cell cycle requirements for transduction by foamy virus vectors compared to those of oncovirus and lentivirus vectors. *Journal of Virology*, 78(5):2327–35, 2004.

[110] J. Lehmann-Che, M. L. Giron, O. Delelis, M. Löchelt, P. Bittoun, J. Tobaly-Tapiero, H. de Thé, and A. Saïb. Protease-dependent uncoating of a complex retrovirus. *Journal of Virology*, 79(14):9244–53, 2005.

[111] P. A. Goepfert, K. L. Shaw, Jr. Ritter, G. D., and M. J. Mulligan. A sorting motif localizes the foamy virus glycoprotein to the endoplasmic reticulum. *Journal of Virology*, 71(1):778–84, 1997.

[112] A. Moebes, J. Enssle, P. D. Bieniasz, M. Heinkelein, D. Lindemann, M. Bock, M. O. McClure, and A. Rethwilm. Human foamy virus reverse transcription that occurs late in the viral replication cycle. *Journal of Virology*, 71(10):7305–11, 1997.

[113] S. F. Yu, M. D. Sullivan, and M. L. Linial. Evidence that the human foamy virus genome is dna. *Journal of Virology*, 73(2):1565–1572, 1999.

[114] J. Roy, W. Rudolph, T. Juretzek, K. Gärtner, M. Bock, O. Herchenröder, D. Lindemann, M. Heinkelein, and A. Rethwilm. Feline foamy virus genome and replication strategy. *Journal of Virology*, 77(21):11324–31, 2003.

[115] D. N. Baldwin and M. L. Linial. The roles of pol and env in the assembly pathway of human foamy virus. *Journal of Virology*, 72(5):3658–65, 1998.

[116] N. Fischer, M. Heinkelein, D. Lindemann, J. Enssle, C. Baum, E. Werder, H. Zentgraf, J. G. Müller, and A. Rethwilm. Foamy virus particle formation. *Journal of Virology*, 72(2):1610–5, 1998.

[117] G. S. Patton, S. A. Morris, W. Chung, P. D. Bieniasz, and M. O. McClure. Identification of domains in gag important for prototypic foamy virus egress. *Journal of Virology*, 79(10):6392–9, 2005.

[118] A. Stange, I. Mannigel, K. Peters, M. Heinkelein, N. Stanke, M. Cartellieri, H. Göttlinger, A. Rethwilm, H. Zentgraf, and D. Lindemann. Characterization of prototype foamy virus gag late assembly domain motifs and their role in particle egress and infectivity. *Journal of Virology*, 79(9):5466–76, 2005.

[119] B. G. Achong, P. W. Mansell, M. A. Epstein, and P. Clifford. An unusual virus in cultures from a human nasopharyngeal carcinoma. *Journal of the National Cancer Institute*, 46(2):299–307, 1971.

[120] O. Herchenröder, R. Renne, D. Loncar, E. K. Cobb, K. K. Murthy, J. Schneider, A. Mergia, and P. A. Luciw. Isolation, cloning, and sequencing of simian foamy viruses from chimpanzees (SFVcpz): high homology to human foamy virus (HFV). *Virology*, 201(2):187–99, 1994.

[121] M. Schweizer, R. Turek, H. Hahn, A. Schliephake, K. O. Netzer, G. Eder, M. Reinhardt, A. Rethwilm, and D. Neumann-Haefelin. Markers of foamy virus infections in monkeys, apes, and accidentally infected humans: appropriate testing fails to confirm suspected foamy virus prevalence in humans. *AIDS Research and Human Retroviruses*, 11(1):161–70, 1995.

[122] W. Heneine, M. Schweizer, P. Sandstrom, and T. Folks. Human infection with foamy viruses. *Current Topics in Microbiology and Immunology*, 277:181–96, 2003.

[123] W. M. Switzer, V. Bhullar, V. Shanmugam, M. E. Cong, B. Parekh, N. W. Lerche, J. L. Yee, J. J. Ely, R. Boneva, L. E. Chapman, T. M. Folks, and W. Heneine. Frequent simian foamy virus infection in persons occupationally exposed to nonhuman primates. *Journal of Virology*, 78(6):2780–9, 2004.

[124] Jean-Yves Tinevez, Nick Perry, Johannes Schindelin, Genevieve M Hoopes, Gregory D Reynolds, Emmanuel Laplantine, Sebastian Y Bednarek, Spencer L Shorte, and Kevin W Eliceiri. Trackmate: An open and extensible platform for single-particle tracking. *Methods*, 115:80–90, 2017.

[125] A. Dupont, I. M. Glück, D. Ponti, K. Stirnnagel, S. Hütter, F. Perrotton, N. Stanke, S. Richter, D. Lindemann, and D. C. Lamb. Identification of an intermediate step in foamy virus fusion. *Viruses*, 12(12), 2020.

[126] K. Stirnnagel, D. Schupp, A. Dupont, V. Kudryavtsev, J. Reh, E. Müllers, D. C. Lamb, and D. Lindemann. Differential ph-dependent cellular uptake pathways among foamy viruses elucidated using dual-colored fluorescent particles. *Retrovirology*, 9:71, 2012.

[127] G. H. Patterson, S. M. Knobel, W. D. Sharif, S. R. Kain, and D. W. Piston. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophysical Journal*, 73(5):2782–90, 1997.

[128] T. Pietschmann, M. Heinkelein, M. Heldmann, H. Zentgraf, A. Rethwilm, and D. Lindemann. Foamy virus capsids require the cognate envelope protein for particle export. *Journal of Virology*, 73(4):2613–21, 1999.

[129] Nicole Stanke, Annett Stange, Daniel Lüftenegger, Hanswalter Zentgraf, and Dirk Lindemann. Ubiquitination of the prototype foamy virus envelope glycoprotein leader peptide regulates subviral particle release. *Journal of Virology*, 79(24):15074–15083, 2005.

[130] Pascal de Boer, Jacob P. Hoogenboom, and Ben N. G. Giepmans. Correlated light and electron microscopy: ultrastructure lights up! *Nature Methods*, 12(6):503–513, 2015.

[131] F. Martinon, K. Burns, and J. Tschopp. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proil-β. *Molecular Cell*, 10(2):417–26, 2002.

[132] Kerry E. Conway, Beth B. McConnell, Claire E. Bowring, Carlton D. Donald, Stephen T. Warren, and Paula M. Vertino. TMS1, a novel proapoptotic caspase recruitment domain protein, is a target of methylation-induced gene silencing in human breast cancers. *Cancer Research*, 60(22):6236–6242, 2000.

[133] J. Masumoto, S. Taniguchi, K. Ayukawa, H. Sarvotham, T. Kishino, N. Niikawa, E. Hidaka, T. Katsuyama, T. Higuchi, and J. Sagara. ASC, a novel 22-kda protein, aggregates during apoptosis of human promyelocytic leukemia HL-60 cells. *Journal of Biological Chemistry*, 274(48):33835–8, 1999.

[134] A D Howard, M J Kostura, N Thornberry, G J Ding, G Limjuco, J Weidner, J P Salley, K A Hogquist, D D Chaplin, and R A Mumford. IL-1-converting enzyme requires aspartic acid residues for processing of the IL-1 beta precursor at two distinct sites and does not cleave 31-kDa IL-1 alpha. *The Journal of Immunology*, 147(9):2964–2969, 1991.

[135] N. A. Thornberry, H. G. Bull, J. R. Calaycay, K. T. Chapman, A. D. Howard, M. J. Kostura, D. K. Miller, S. M. Molineaux, J. R. Weidner, J. Aunins, K. O. Elliston, J. M. Ayala, F. J. Casano, J. Chin, G. J.-F. Ding, L. A. Egger, E. P. Gaffney, G. Limjuco, O. C. Palyha, S. M. Raju, A. M. Rolando, J. P. Salley, T.-T. Yamin, T. D. Lee, J. E. Shively, M. MacCross, R. A. Mumford, J. A. Schmidt, and M. J. Tocci. A novel heterodimeric cysteine protease is required for interleukin-1 beta processing in monocytes. *Nature*, 356(6372):768–74, 1992.

[136] L. Wang, G. A. Manji, J. M. Grenier, A. Al-Garawi, S. Merriam, J. M. Lora, B. J. Geddes, M. Briskin, P. S. DiStefano, and J. Bertin. Pypaf7, a novel pyrin-containing apaf1-like protein that regulates activation of nf-kappa b and caspase-1-dependent cytokine processing. *Journal of Biological Chemistry*, 277(33):29874–80, 2002.

[137] T. Fernandes-Alnemri, J. Wu, J. W. Yu, P. Datta, B. Miller, W. Jankowski, S. Rosenberg, J. Zhang, and E. S. Alnemri. The pyroptosome: a supramolecular assembly of asc dimers mediating inflammatory cell death via caspase-1 activation. *Cell Death and Differentiation*, 14(9):1590–1604, 2007.

[138] Kay Hofmann, Philipp Bucher, and Jurg Tschopp. The CARD domain: a new apoptotic signalling motif. *Trends in Biochemical Sciences*, 22(5):155–156, 1997.

[139] Edvards Liepinsh, Raitis Barbals, Edgar Dahl, Anatoly Sharipo, Eike Staub, and Gottfried Otting. The death-domain fold of the ASC pyrin domain, presenting a basis for PYRIN/PYRIN recognition. *Journal of Molecular Biology*, 332(5):1155–1163, 2003.

[140] Ali C. Sahillioglu, Fidan Sumbul, Nesrin Ozoren, and Turkan Haliloglu. Structural and dynamics aspects of ASC speck assembly. *Structure*, 22(12):1722–1734, 2014.

[141] P. R. Vajjhala, R. E. Mirams, and J. M. Hill. Multiple binding sites on the pyrin domain of ASC protein allow self-association and interaction with NLRP3 protein. *Journal of Biological Chemistry*, 287(50):41732–43, 2012.

[142] Parimala R. Vajjhala, Sebastian Kaiser, Sarah J. Smith, Qi-Rui Ong, Stephanie L. Soh, Katryn J. Stacey, and Justine M. Hill. Identification of multifaceted binding modes for pyrin and ASC pyrin domains gives insights into pyrin inflammasome assembly. *The Journal of Biological Chemistry*, 289(34):23504–23519, 2014.

[143] James J. Chou, Hiroshi Matsuo, Hanjun Duan, and Gerhard Wagner. Solution structure of the raidd card and model for CARD/CARD interaction in caspase-2 and caspase-9 recruitment. *Cell*, 94(2):171–180, 1998.

[144] E. de Alba. Structure and interdomain dynamics of apoptosis-associated speck-like protein containing a card (ASC). *Journal of Biological Chemistry*, 284(47):32932–41, 2009.

[145] Kristof Kersse, Jelle Verspurten, Tom Vanden Berghe, and Peter Vandenabeele. The death-fold superfamily of homotypic interaction motifs. *Trends in Biochemical Sciences*, 36(10), 2011.

[146] Lorenzo Sborgi, Francesco Ravotti, Venkata P. Dandey, Mathias S. Dick, Adam Mazur, Sina Reckel, Mohamed Chami, Sebastian Scherer, Matthias Huber, Anja Böckmann, Edward H. Egelman, Henning Stahlberg, Petr Broz, Beat H. Meier, and Sebastian Hiller. Structure and assembly of the mouse asc inflammasome by combined NMR spectroscopy and cryo-electron microscopy. *Proceedings of the National Academy of Sciences*, 112(43):13237–13242, 2015.

[147] R. J. T. Nambayan, S. I. Sandin, D. A. Quint, D. M. Satyadi, and E. de Alba. The inflammasome adapter ASC assembles into filaments with integral participation of its two death domains, PYD and CARD. *Journal of Biological Chemistry*, 294(2):439–452, 2019.

[148] Xin Cai, Jueqi Chen, Hui Xu, Siqi Liu, Qiu-Xing Jiang, Randal Halfmann, and ZhijianJ. Chen. Prion-like polymerization underlies signal transduction in antiviral immune defense and inflammasome activation. *Cell*, 156(6):1207–1222, 2014.

[149] Alvin Lu, Venkat G. Magupalli, Jianbin Ruan, Qian Yin, ManinjayÂ K. Atianand, Matthijn R. Vos, Gunnar F. Schröder, Katherine A. Fitzgerald, Hao Wu, and Edward H. Egelman. Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes. *Cell*, 156(6):1193–1206, 2014.

[150] Mathias S. Dick, Lorenzo Sborgi, Sebastian Rühl, Sebastian Hiller, and Petr Broz. ASC filament formation serves as a signal amplification mechanism for inflammasomes. *Nature Communications*, 7(1):11929, 2016.

[151] Yang Li, Tian-Min Fu, Alvin Lu, Kristen Witt, Jianbin Ruan, Chen Shen, and Hao Wu. Cryo-em structures of ASC and NLRC4 CARD filaments reveal a unified mechanism of nucleation and activation of caspase-1. *Proceedings of the National Academy of Sciences*, 115(43):10845–10852, 2018.

[152] Bernardo S. Franklin, Lukas Bossaller, Dominic De Nardo, Jacqueline M. Ratter, Andrea Stutz, Gudrun Engels, Christoph Brenker, Mark Nordhoff, Sandra R. Mirandola, Ashraf Al-Amoudi, Matthew S. Mangan, Sebastian Zimmer, Brian G. Monks, Martin Fricke, Reinhold E. Schmidt, Terje Espevik, Bernadette Jones, Andrew G. Jarnicki,

Philip M. Hansbro, Patricia Busto, Ann Marshak-Rothstein, Simone Hornemann, Adriano Aguzzi, Wolfgang Kasten-müller, and Eicke Latz. The adaptor ASC has extracellular and 'prionoid' activities that propagate inflammation. *Nature Immunology*, 15(8):727–737, 2014.

[153] N. Richards, P. Schaner, A. Diaz, J. Stuckey, E. Shelden, A. Wadhwa, and D. L. Gumucio. Interaction between pyrin and the apoptotic speck protein (asc) modulates asc-induced apoptosis. *Journal of Biological Chemistry*, 276(42):39320–9, 2001.

[154] Fátima Martin-Sánchez, Vincent Compan, Ana Tapia-Abellán, Ana I. Gómez-Sanchez, Maria C. Baños, Florian I. Schmidt, and Pablo Pelegrin. Asc oligomer favor caspase-1 card domain recruitment after intracellular potassium efflux. *bioRxiv*, page 2020.01.27., 2020.

[155] Mie Moriya, Shunichiro Taniguchi, Peter Wu, Edvards Liepinsh, Gottfried Otting, and Junji Sagara. Role of charged and hydrophobic residues in the oligomerization of the PYRIN domain of ASC. *Biochemistry*, 44(2):575–583, 2005.

[156] Michael Beilharz, Dominic De Nardo, Eicke Latz, and Bernardo S. Franklin. *Measuring NLR Oligomerization II: Detection of ASC Speck Formation by Confocal Microscopy and Immunofluorescence*, pages 145–158. Springer New York, New York, NY, 2016.

[157] Si Ming Man, Lee J. Hopkins, Eileen Nugent, Susan Cox, Ivo M. Glück, Panagiotis Tourlomousis, John A. Wright, Pietro Cicuta, Tom P. Monie, and Clare E. Bryant. Inflammasome activation causes dual recruitment of NLRC4 and NLRP3 to the same macromolecular complex. *Proceedings of the National Academy of Sciences*, 111(20):7403–7408, 2014.

[158] F. I. Schmidt, A. Lu, J. W. Chen, J. Ruan, C. Tang, H. Wu, and H. L. Ploegh. A single domain antibody fragment that recognizes the adaptor asc defines the role of asc domains in inflammasome assembly. *Journal of Experimental Medicine*, 213(5):771–90, 2016.

[159] Alberto Baroja-Mazo, Fatima Martin-Sanchez, Ana I. Gomez, Carlos M. Martinez, Joaquin Amores-Iniesta, Vincent Compan, Maria Barbera-Cremades, Jordi Yagüe, Estibaliz Ruiz-Ortiz, Jordi Anton, Segundo Bujan, Isabelle Couillin, David Brough, Juan I. Arostegui, and Pablo Pelegrin. The NLRP3 inflammasome is released as a particulate danger signal that amplifies the inflammatory response. *Nature Immunology*, 15(8):738–748, 2014.

[160] Katharina S. Schneider, Christina J. Gross, Roland F. Dreier, Benedikt S. Saller, Ritu Mishra, Oliver Gorka, Rosalie Heilig, Etienne Meunier, Mathias S. Dick, Tamara Cikovic, Jan Sodenkamp, Guillaume Médard, Ronald Naumann, Jürgen Ruland, Bernhard Kuster, Petr Broz, and Olaf Gross. The inflammasome drives GSDMD-independent secondary pyroptosis and IL-1 release in the absence of caspase-1 protease activity. *Cell Reports*, 21(13):3846–3859, 2017.

[161] F. Hoss, J. F. Rodriguez-Alcazar, and E. Latz. Assembly and regulation of ASC specks. *Cellular and Molecular Life Sciences*, 74(7):1211–1229, 2017.

[162] Nicole B. Bryan, Andrea Dorfleutner, Yon Rojanasakul, and Christian Stehlik. Activation of inflammasomes requires intracellular redistribution of the apoptotic speck-like protein containing a caspase recruitment domain. *The Journal of Immunology*, 182(5):3173–3182, 2009.

[163] Nicole B. Bryan, Andrea Dorfleutner, Sara J. Kramer, Chawon Yun, Yon Rojanasakul, and Christian Stehlik. Differential splicing of the apoptosis-associated speck like protein containing a caspase recruitment domain (ASC) regulates inflammasomes. *Journal of Inflammation*, 7(1):23, 2010.

[164] Sonal Khare, Andrea Dorfleutner, Nicole B. Bryan, Chawon Yun, Alexander D. Radian, Lucia de Almeida, Yon Rojanasakul, and Christian Stehlik. An NLRP7-containing inflammasome mediates recognition of microbial lipopeptides in human macrophages. *Immunity*, 36(3):464–476, 2012.

[165] Silvia Lucena Lage, Venina Marcela Dominical, Chun-Shu Wong, and Irini Sereti. Evaluation of canonical inflammasome activation in human monocytes by imaging flow cytometry. *Frontiers in Immunology*, 10(1284), 2019.

[166] Jiang-yuan Li, Ke Gao, Tong Shao, Dong-dong Fan, Chong-bin Hu, Cen-cen Sun, Wei-ren Dong, Ai-fu Lin, Lixin Xiang, and Jian-zhong Shao. Characterization of an NLRP1 inflammasome from zebrafish reveals a unique sequential activation mechanism underlying inflammatory caspases in ancient vertebrates. *The Journal of Immunology*, 201(7):1946–1966, 2018.

[167] J. Masumoto, S. Taniguchi, and J. Sagara. Pyrin N-terminal homology domain- and caspase recruitment domain-dependent oligomerization of ASC. *Biochemical and Biophysical Research Communication*, 280(3):652–5, 2001.

[168] Beth B. McConnell and Paula M. Vertino. Activation of a caspase-9-mediated apoptotic pathway by subcellular redistribution of the novel caspase recruitment domain protein TMS1. *Cancer Research*, 60(22):6243–6247, 2000.

[169] B. B. McConnell and P. M. Vertino. TMS1/ASC: the cancer connection. *Apoptosis*, 9(1):5–18, 2004.

[170] Tom Monie and Joseph Boyle. Insights into assembly of the macromolecular inflammasome complex. *Inflammasome*, 1, 2015.

[171] Lan Hoang Chu, Anu Gangopadhyay, Andrea Dorfleutner, and Christian Stehlik. An updated view on the structure and function of pyrin domains. *Apoptosis*, 20(2):157–173, 2015.

[172] E. de Alba. Structure, interactions and self-assembly of ASC-dependent inflammasomes. *Archives of Biochemistry Biophysics*, 670:15–31, 2019.

[173] E. I. Elliott and F. S. Sutterwala. Initiation and perpetuation of NLRP3 inflammasome activation and assembly. *Immunological Reviews*, 265(1):35–52, 2015.

[174] Franz G. Bauernfeind, Gabor Horvath, Andrea Stutz, Emad S. Alnemri, Kelly MacDonald, David Speert, Teresa Fernandes-Alnemri, Jianghong Wu, Brian G. Monks, Katherine A. Fitzgerald, Veit Hornung, and Eicke Latz. Cutting edge: NF-kB activating pattern recognition and cytokine receptors license NLRP3 inflammasome activation by regulating NLRP3 expression. *The Journal of Immunology*, 183(2):787–791, 2009.

[175] Sanjeev Mariathasan, David S. Weiss, Kim Newton, Jacqueline McBride, Karen O'Rourke, Meron Roose-Girma, Wyne P. Lee, Yvette Weinrauch, Denise M. Monack, and Vishva M. Dixit. Cryopyrin activates the inflammasome in response to toxins and ATP. *Nature*, 440(7081):228–232, 2006.

[176] Shigeru Tsuchiya, Michiko Yamabe, Yoshiko Yamaguchi, Yasuko Kobayashi, Tasuke Konno, and Keiya Tada. Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *International Journal of Cancer*, 26(2):171–176, 1980.

[177] Heike Schwende, Edith Fitzke, Petra Ambs, and Peter Dieter. Differences in the state of differentiation of THP-1 cells induced by phorbol ester and 1,25-dihydroxyvitamin D3. *Journal of Leukocyte Biology*, 59(4):555–561, 1996.

[178] M. M. Gaidt, F. Rapino, T. Graf, and V. Hornung. Modeling primary human monocytes with the trans-differentiation cell line BLaER1. *Methods in Molecular Biology*, 1714:57–66, 2018.

[179] Anne E. Carpenter, Thouis R. Jones, Michael R. Lamprecht, Colin Clarke, In Han Kang, Ola Friman, David A. Guertin, Joo Han Chang, Robert A. Lindquist, Jason Moffat, Polina Golland, and David M. Sabatini. Cellprofiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, 7(10):R100, 2006.

[180] Fabian U. Zwettler, Sebastian Reinhard, Davide Gambarotto, Toby D. M. Bell, Virginie Hamel, Paul Guichard, and Markus Sauer. Molecular resolution imaging by post-labeling expansion single-molecule localization microscopy (Ex-SMLM). *Nature Communications*, 11(1):3388, 2020.

[181] M. M. Gaidt, T. S. Ebert, D. Chauhan, T. Schmidt, J. L. Schmid-Burgk, F. Rapino, A. A. Robertson, M. A. Cooper, T. Graf, and V. Hornung. Human monocytes engage an alternative inflammasome pathway. *Immunity*, 44(4):833–46, 2016.

[182] Jonas Ries, Charlotte Kaplan, Evgenia Platonova, Hadi Eghlidi, and Helge Ewers. A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. *Nature Methods*, 9(6):582–584, 2012.

[183] Anna Szymborska, Alex de Marco, Nathalie Daigle, Volker C. Cordes, John A. G. Briggs, and Jan Ellenberg. Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging. *Science*, 341(6146):655–658, 2013.

[184] Shotaro Otsuka, Anna Szymborska, and Jan Ellenberg. *Chapter 10 - Imaging the Assembly, Structure, and Function of the Nuclear Pore Inside Cells*, volume 122, pages 219–238. Academic Press, 2014.

[185] Charlotte Kaplan and Helge Ewers. Optimized sample preparation for single-molecule localization-based superresolution microscopy in yeast. *Nature Protocols*, 10(7):1007–1021, 2015.

[186] L. J. Fulcher, L. D. Hutchinson, T. J. Macartney, C. Turnbull, and G. P. Sapkota. Targeting endogenous proteins for degradation through the affinity-directed protein missile system. *Open Biology*, 7(5), 2017.

[187] Judith Scholz, Hüseyin Besir, Claudia Strasser, and Sabine Suppmann. A new method to customize protein expression vectors for fast, efficient and background free parallel cloning. *BMC Biotechnology*, 13(1):12, 2013.

[188] Tino Pleiner, Mark Bates, Sergei Trakhanov, Chung-Tien Lee, Jan Erik Schliep, Hema Chug, Marc Böhning, Holger Stark, Henning Urlaub, and Dirk Görlich. Nanobodies: site-specific labeling for super-resolution imaging, rapid epitope-mapping and native protein complex isolation. *eLife*, 4:e11349, 2015.

[189] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231.

[190] Venkat G. Magupalli, Roberto Negro, Yuzi Tian, Arthur V. Hauenstein, Giuseppe Di Caprio, Wesley Skillern, Qiufang Deng, Pontus Orning, Hasan B. Alam, Zoltan Maliga, Humayun Sharif, Jun Jacob Hu, Charles L. Evavold, Jonathan C. Kagan, Florian I. Schmidt, Katherine A. Fitzgerald, Tom Kirchhausen, Yongqing Li, and Hao Wu. HDAC6 mediates an aggresome-like mechanism for NLRP3 and pyrin inflammasome activation. *Science*, 369(6510):eaas8995, 2020.

[191] M. Decker, S. Jaensch, A. Pozniakovsky, A. Zinke, K. F. O'Connell, W. Zachariae, E. Myers, and A. A. Hyman. Limiting amounts of centrosome material set centrosome size in c. elegans embryos. *Current Biology*, 21(15):1259–67, 2011.

[192] Alvin Lu, Yang Li, Florian I. Schmidt, Qian Yin, Shuobing Chen, Tian-Min Fu, Alexander B. Tong, Hidde L. Ploegh, Youdong Mao, and Hao Wu. Molecular basis of caspase-1 polymerization and its inhibition by a new capping mechanism. *Nature Structural & Molecular Biology*, 23(5):416–425, 2016.

[193] Dave Boucher, Mercedes Monteleone, Rebecca C. Coll, Kaiwen W. Chen, Connie M. Ross, Jessica L. Teo, Guillermo A. Gomez, Caroline L. Holley, Damien Bierschenk, Katryn J. Stacey, Alpha S. Yap, Jelena S. Bezbradica, and Kate Schroder. Caspase-1 self-cleavage is an intrinsic mechanism to terminate inflammasome activity. *Journal of Experimental Medicine*, 215(3):827–840, 2018.

[194] Petr Broz, Kim Newton, Mohamed Lamkanfi, Sanjeev Mariathasan, Vishva M. Dixit, and Denise M. Monack. Redundant roles for inflammasome receptors NLRP3 and NLRC4 in host defense against salmonella. *Journal of Experimental Medicine*, 207(8):1745–1755, 2010.

[195] Yan Qu, Shahram Misaghi, Kim Newton, Allie Maltzman, Anita Izrael-Tomasevic, David Arnott, and Vishva M. Dixit. NLRP3 recruitment by NLRC4 during salmonella infection. *Journal of Experimental Medicine*, 213(6):877–885, 2016.

[196] Caroline Junqueira, Angela Crespo, Shahin Ranjbar, Jacob Ingber, Blair Parry, Sagi Ravid, Luna B. de Lacerda, Mercedes Lewandrowski, Sarah Clark, Felicia Ho, Setu M. Vora, Valerie Leger, Caroline Beakes, Justin Margolin, Nicole Russell, Lee Gehrke, Upasana Das Adhikari, Lauren Henderson, Erin Janssen, Douglas Kwon, Chris Sander, Jonathan Abraham, Michael Filbin, Marcia B. Goldberg, Hao Wu, Gautam Mehta, Steven Bell, Anne E. Goldfeld, and Judy Lieberman. SARS-CoV-2 infects blood monocytes to activate NLRP3 and AIM2 inflammasomes, pyroptosis and cytokine release. *medRxiv*, page 2021.03.06., 2021.

[197] SangJoon Lee, Rajendra Karki, Yaqiu Wang, Lam Nhat Nguyen, Ravi C. Kalathur, and Thirumala-Devi Kanneganti. AIM2 forms a complex with pyrin and ZBP1 to drive panoptosis and host defence. *Nature*, 2021.

[198] Ralf Jungmann, Maier S. Avendano, Johannes B. Woehrstein, Mingjie Dai, William M. Shih, and Peng Yin. Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and exchange-PAINT. *Nature Methods*, 11(3):313–318, 2014.

[199] Markus Mund, Johannes Albertus van der Beek, Joran Deschamps, Serge Dmitrieff, Philipp Hoess, Jooske Louise Monster, Andrea Picco, Francois Nédélec, Marko Kaksonen, and Jonas Ries. Systematic nanoscale analysis of endocytosis links efficient vesicle formation to patterned actin nucleation. *Cell*, 174(4):884–896.e17, 2018.

[200] Guy Nir, Irene Farabella, Cynthia Pérez Estrada, Carl G. Ebeling, Brian J. Beliveau, Hiroshi M. Sasaki, S. Dean Lee, Son C. Nguyen, Ruth B. McCole, Shyamtanu Chattoraj, Jelena Erceg, Jumana AlHaj Abed, Nuno M. C. Martins, Huy Q. Nguyen, Mohammed A. Hannan, Sheikh Russell, Neva C. Durand, Suhas S. P. Rao, Jocelyn Y. Kishi, Paula Soler-Vila, Michele Di Pierro, José N. Onuchic, Steven P. Callahan, John M. Schreiner, Jeff A. Stuckey, Peng Yin, Erez Lieberman Aiden, Marc A. Marti-Renom, and C. Ting Wu. Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. *PLOS Genetics*, 14(12):e1007872–e1007872, 2018.

[201] Florian I. Schmidt. From atoms to physiology: what it takes to really understand inflammasomes. *The Journal of Physiology*, 597(22):5335–5348, 2019.

[202] Francisco Balzarotti, Yvan Eilers, Klaus C. Gwosch, Arvid H. Gynna, Volker Westphal, Fernando D. Stefani, Johan Elf, and Stefan W. Hell. Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science*, pages 606–612, 2016.

[203] Anna Löschberger, Sebastian van de Linde, Marie-Christine Dabauvalle, Bernd Rieger, Mike Heilemann, Georg Krohne, and Markus Sauer. Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution. *Journal of Cell Science*, 125(3):570–575, 2012.

[204] Veit Hornung, Andrea Ablasser, Marie Charrel-Dennis, Franz Bauernfeind, Gabor Horvath, Daniel R. Caffrey, Eicke Latz, and Katherine A. Fitzgerald. AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature*, 458(7237):514–518, 2009.

[205] Hrvoje Lusic, Douglas D. Young, Mark O. Lively, and Alexander Deiters. Photochemical DNA activation. *Organic Letters*, 9(10):1903–1906, 2007.

[206] Y. Liu, R. S. Zou, S. He, Y. Nihongaki, X. Li, S. Razavi, B. Wu, and T. Ha. Very fast CRISPR on demand. *Science*, 368(6496):1265–1269, 2020.

[207] Robert J. Schmitz, Zachary A. Lewis, and Mary G. Goll. DNA methylation: Shared and divergent features across eukaryotes. *Trends in Genetics*, 35(11):818–827, 2019.

[208] M. Monk, M. Boubelik, and S. Lehnert. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development*, 99(3):371–82, 1987.

[209] J. P. Sanford, H. J. Clark, V. M. Chapman, and J. Rossant. Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes & Developement*, 1(10):1039–46, 1987.

[210] M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, and A. Rao. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 324(5929):930–5, 2009.

[211] Y. F. He, B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He, and G. L. Xu. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, 333(6047):1303–7, 2011.

[212] S. Cortellino, J. Xu, M. Sannai, R. Moore, E. Caretti, A. Cigliano, M. Le Coz, K. Devarajan, A. Wessels, D. Soprano, L. K. Abramowitz, M. S. Bartolomei, F. Rambow, M. R. Bassi, T. Bruno, M. Fanciulli, C. Renner, A. J. Klein-Szanto, Y. Matsumoto, D. Kobi, I. Davidson, C. Alberti, L. Larue, and A. Bellacosa. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell*, 146(1):67–79, 2011.

[213] S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, and Y. Zhang. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, 333(6047):1300–3, 2011.

[214] S. K. Howlett and W. Reik. Methylation levels of maternal and paternal genomes during preimplantation development. *Development*, 113(1):119–27, 1991.

[215] N. Rougier, D. Bourchis, D. M. Gomes, A. Niveleau, M. Plachot, A. Paldi, and E. Viegas-Péquignot. Chromosome methylation patterns during mammalian preimplantation development. *Genes & Developement*, 12(14):2108–13, 1998.

[216] C. B. Mulholland, A. Nishiyama, J. Ryan, R. Nakamura, M. Yigit, I. M. Glück, C. Trummer, W. Qin, M. D. Bartoschek, F. R. Traube, E. Parsa, E. Ugur, M. Modic, A. Acharya, P. Stolz, C. Ziegenhain, M. Wierer, W. Enard, T. Carell, D. C. Lamb, H. Takeda, M. Nakanishi, S. Bultmann, and H. Leonhardt. Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive DNA demethylation in mammals. *Nature Communiications*, 11(1):5972, 2020.

[217] Y. Li, Z. Zhang, J. Chen, W. Liu, W. Lai, B. Liu, X. Li, L. Liu, S. Xu, Q. Dong, M. Wang, X. Duan, J. Tan, Y. Zheng, P. Zhang, G. Fan, J. Wong, G. L. Xu, Z. Wang, H. Wang, S. Gao, and B. Zhu. Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. *Nature*, 564(7734):136–140, 2018.

[218] Nicolas Dross, Corentin Spriet, Monika Zwerger, Gabriele Müller, Waldemar Waldeck, and Jörg Langowski. Mapping eGFP oligomer mobility in living cell nuclei. *PLOS ONE*, 4(4):e5041–e5041, 2009.

[219] Michelle A. Digman and Enrico Gratton. Imaging barriers to diffusion by pair correlation functions. *Biophysical Journal*, 97(2):665–673, 2009.

[220] Don C. Lamb, Andreas Schenk, Carlheinz Röcker, C. Scalfi-Happ, and G. Ulrich Nienhaus. Sensitivity enhancement in fluorescence correlation spectroscopy of multiple species using time-gated detection. *Biophysical Journal*, 79(2):1129–1138, 2000.

[221] B. C. Chen, W. R. Legant, K. Wang, L. Shao, D. E. Milkie, M. W. Davidson, C. Janetopoulos, X. S. Wu, J. A. Hammer, Z. Liu, B. P. English, Y. Mimori-Kiyosue, D. P. Romero, A. T. Ritter, J. Lippincott-Schwartz, L. Fritz-Laylin, R. D. Mullins, D. M. Mitchell, J. N. Bembenek, A. C. Reymann, R. Böhme, S. W. Grill, J. T. Wang, G. Seydoux, U. S. Tulu, D. P. Kiehart, and E. Betzig. Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution. *Science*, 346(6208):1257998, 2014.

[222] Katharina Brandstetter, Tilo Zülske, Tobias Ragoczy, David Hörl, Eric Haugen, Eric Rynes, John A. Stamatoyannopoulos, Heinrich Leonhardt, Gero Wedemann, and Hartmann Harz. Differences in nanoscale organization of DNase I hypersensitive and insensitive chromatin in single human cells. *bioRxiv*, page 2021.05.27., 2021.

[223] David P. Hoffman, Gleb Shtengel, C. Shan Xu, Kirby R. Campbell, Melanie Freeman, Lei Wang, Daniel E. Milkie, H. Amalia Pasolli, Nirmala Iyer, John A. Bogovic, Daniel R. Stabley, Abbas Shirinifard, Song Pang, David Peale, Kathy Schaefer, Wim Pomp, Chi-Lun Chang, Jennifer Lippincott-Schwartz, Tom Kirchhausen, David J. Solecki, Eric Betzig, and Harald F. Hess. Correlative three-dimensional super-resolution and block-face electron microscopy of whole vitreously frozen cells. *Science*, 367(6475):eaaz5357, 2020.

[224] Kristina D. Micheva and Stephen J. Smith. Array tomography: A new tool for imaging the molecular architecture and ultrastructure of neural circuits. *Neuron*, 55(1):25–36, 2007.

[225] F. Chen, P. W. Tillberg, and E. S. Boyden. Optical imaging. expansion microscopy. *Science*, 347(6221):543–8, 2015.

[226] M. A. de Castro, B. Rammner, and F. Opazo. Aptamer stainings for super-resolution microscopy. *Methods in Molecular Biology*, 1380:197–210, 2016.

[227] Felipe Opazo, Matthew Levy, Michelle Byrom, Christina Schäfer, Claudia Geisler, Teja W. Groemer, Andrew D. Ellington, and Silvio O. Rizzoli. Aptamers as potential tools for super-resolution microscopy. *Nature Methods*, 9(10):938–939, 2012.

[228] Sebastian Strauss, Philipp C. Nickels, Maximilian T. Strauss, Vilma Jimenez Sabinina, Jan Ellenberg, Jeffrey D. Carter, Shashi Gupta, Nebojsa Janjic, and Ralf Jungmann. Modified aptamers enable quantitative sub-10-nm cellular DNA-PAINT imaging. *Nature Methods*, 15(9):685–688, 2018.

[229] Glenn Carrington, Darren Tomlinson, and Michelle Peckham. Exploiting nanobodies and affimers for superresolution imaging in light microscopy. *Molecular Biology of the Cell*, 30(22):2737–2740, 2019.

[230] Pooja Laxman, Shirin Ansari, Katharina Gaus, and Jesse Goyette. The benefits of unnatural amino acid incorporation as protein labels for single molecule localization microscopy. *Frontiers in Chemistry*, 9(161), 2021.

[231] Adrian Pickar-Oliver and Charles A. Gersbach. The next generation of crispr-cas technologies and applications. *Nature Reviews Molecular Cell Biology*, 20(8):490–507, 2019.

[232] Christian Conrad, Annelie Wünsche, Tze Heng Tan, Jutta Bulkescher, Frank Sieckmann, Fatima Verissimo, Arthur Edelstein, Thomas Walter, Urban Liebel, Rainer Pepperkok, and Jan Ellenberg. Micropilot: automation of fluorescence microscopy-based imaging for systems biology. *Nature Methods*, 8(3):246–249, 2011.

[233] Marcel Stefko, Baptiste Ottino, Kyle M. Douglass, and Suliana Manley. Autonomous illumination control for localization microscopy. *Optics Express*, 26(23):30882–30900, 2018.

[234] Nico Scherf and Jan Huisken. The smart and gentle microscope. *Nature Biotechnology*, 33(8):815–818, 2015.

[235] Pedro Almada, Pedro M. Pereira, Sian Culley, Ghislaine Caillol, Fanny Boroni-Rueda, Christina L. Dix, Guillaume Charras, Buzz Baum, Romain F. Laine, Christophe Leterrier, and Ricardo Henriques. Automating multimodal microscopy with nanoj-fluidics. *Nature Communications*, 10(1):1223, 2019.

[236] Q. Qiao and H. Wu. Supramolecular organizing centers (SMOCs) as signaling machines in innate immune activation. *Science China Life Sciences*, 58(11):1067–72, 2015.

[237] Alexander Poltorak, Xiaolong He, Irina Smirnova, Mu-Ya Liu, Christophe Van Huffel, Xin Du, Dale Birdwell, Erica Alejos, Maria Silva, Chris Galanos, Marina Freudenberg, Paola Ricciardi-Castagnoli, Betsy Layton, and Bruce Beutler. Defective lps signaling in C3H/HeJ and C57BL/10ScCr mice: Mutations in tlr4 gene. *Science*, 282(5396):2085–2088, 1998.

[238] D. J. Hazuda, J. C. Lee, and P. R. Young. The kinetics of interleukin 1 secretion from activated monocytes. differences between interleukin 1 alpha and interleukin 1 beta. *Journal of Biological Chemistry*, 263(17):8473–8479, 1988.

[239] Gerhard Lonnemann, Stefan Endres, Jos W. M. Van Der Meer, Joseph G. Cannon, Karl M. Koch, and Charles A. Dinarello. Differences in the synthesis and kinetics of release of interleukin 1α, interleukin 1$\beta$and tumor necrosis factor from human mononuclear cells. *European Journal of Immunology*, 19(9):1531–1536, 1989.

[240] Ralf R. Schumann, Claus Belka, Dirk Reuter, Norbert Lamping, Carsten J. Kirschning, Joerg R. Weber, and Dagmar Pfeil. Lipopolysaccharide activates caspase-1 (interleukin-1-converting enzyme) in cultured monocytic and endothelial cells. *Blood*, 91(2):577–584, 1998.

[241] V. A. Danis, A. J. Kulesz, D. S. Nelson, and P. M. Brooks. Cytokine regulation of human monocyte interleukin-1 (IL-1) production in vitro. enhancement of IL-1 production by interferon (IFN) gamma, tumour necrosis factor-α, IL-2 and IL-1, and inhibition by IFN-α. *Clinical and Exprimental Immunology*, 80(3):435–43, 1990.

[242] Francesca Rapino, Eloy Â F. Robles, Jose Â A. Richter-Larrea, Eric Â M. Kallin, Jose Â A. Martinez-Climent, and Thomas Graf. Cebpαinduces highly efficient macrophage transdifferentiation of b lymphoma and leukemia cell lines and impairs their tumorigenicity. *Cell Reports*, 3(4):1153–1163, 2013.

[243] W. Schrimpf, A. Barth, J. Hendrix, and D. C. Lamb. PAM: A framework for integrated analysis of imaging, single-molecule, and ensemble fluorescence data. *Biophysical Journal*, 114(7):1518–1528, 2018.

List of Figures

List of Tables

List of Abbreviations

ACF	autocorrelation function
AdPROM	affinity-directed protein missile
AIM2	absent in melanoma 2
ALC	andor laser combiner
AOTF	acousto-optic tunable filter
APD	avalanche photodiodes
ARICS	arbitrary-region raster image correlation spectroscopy
AU	airy units
CARD	caspase activation and recruitment domain
cc	cross-correlation
CCD	charge-coupled device
CLEM	correlative light and electron microscopy
CRISPR	clustered regularly interspaced short palindromic repeats
CSU	confocal spinning-disk unit
DBSCAN	density-based spatial clustering of applications with noise
DIC	differential interference contrast
DMEM	Dulbecco's modified eagle medium
DNA-PAINT	DNA point accumulation for imaging in nanoscale topography
DPPA3	developmental pluripotency-associated protein 3
dSTORM	direct stochastic optical reconstruction microscopy
eGFP	enhanced green fluorescent protein
EM	electron microscopy
EMCCD	electron-multiplying charge-coupled device
ESC	embryonic stem cells
FADD	fas-associated protein with death domain

FBS	fetal bovine serum
FCS	fluorescence correlation spectroscopy
FFS	fluorescence fluctuation spectroscopy
FIFI	flat illumination for field-independent imaging
FP	fluorescent proteins
FPGA	field programmable gate array
FRET	förster-resonance-energy-transfer
FWHM	full width at half maximum
GFP	green fluorescent protein
GSDIM	ground-state depletion with individual molecule return
GUI	graphical user interface
HILO	highly inclined and laminated optical sheet microscopy
ICS	image correlation spectroscopy
IL	interleukin
ISC	intersystem crossing
KO	knockout
LED	light emitting diode
LPS	lipopolysaccharide
LSCM	laser scanning confocal microscope
M-CSF	macrophage colony-stimulating factor
MLA	microlens arrays
MTOC	microtubule-organizing center
NA	numerical aperture
NGS	normal goat serum
NLRC4	NLR family CARD domain-containing protein 4
NLRP3	NOD-, LRR- and pyrin domain-containing protein 3

NMR	nuclear magnetic resonance
NPOM	6-nitropiperonyloxymethyl
paGFP	photoactivatable green fluorescent protein
PALM	photoactivated localization microscopy
PAMP	pathogen-associated molecular pattern
PBS	phosphate-buffered saline
PCA	protocatechuic acid
PCR	polymerase chain reaction
PCU	Andor Precision Control Unit
PFA	paraformaldehyde
PFS	perfect focus system
PFV	prototype foamy virus
PIE	pulsed-interleaved excitation
PMT	photomultiplier tube
PRR	pattern recognition receptor
PSF	point spread function
PYD	pyrin domain
RFP	red fluorescent protein
RICS	raster image correlation spectroscopy
RIPK1	receptor-interacting serine/threonine-protein kinase 1
RPMI	Roswell Park Memorial Institute
SACF	spatial autocorrelation function
SACF	spatial autocorrelation function
sCMOS	scientific complementary metal–oxide–semiconductor
SD	standard deviation
SDCM	spinning-disk confocal microscope

SEM	standard error of the mean
SFVmac	macaque simian foamy virus
SIM	structured illumination microscopy
SMOCS	supramolecular organizing centers
STED	stimulated emission depletion
SVT	single virus tracking/tracing
TAC	time-to-amplitude converter
TCEP	tris(2-carboxyethyl)phosphin
TCSPC	time-correlated single photon counting
TET	ten-eleven translocation
TIRF	internal reflection fluorescence
TLR4	toll-like receptor 4
TrIC	tracking image correlation
TRIF	TIR-domain-containing adapter-inducing interferon-b
UV	ultraviolet

Acknowledgements

B Appended papers

B.1   Paper 1:Identification of an Intermediate Step in Foamy Virus Fusion

Reproduced with permission from *Viruses* 2020, 12, 1472. Copyright 2020 MDPI

viruses

MDPI

Article

Identification of an Intermediate Step in Foamy Virus Fusion

Aurélie Dupont [1,2,3,†], Ivo M. Glück [1,2,†], Dorothee Ponti [1,2], Kristin Stirnnagel [4,5], Sylvia Hütter [4,5], Florian Perrotton [1], Nicole Stanke [4,5], Stefanie Richter [4,5], Dirk Lindemann [4,5,*] and Don C. Lamb [1,2,6,7,*]

[1]   Department of Chemistry, Ludwig Maximilians-Universität München, Butenandtstraße 5-13, 81377 München, Germany; aurelie.dupont@univ-grenoble-alpes.fr (A.D.); Ivo.Glueck@cup.lmu.de (I.M.G.); dorothee_ponti@gmx.de (D.P.); florian.perrotton@ens-lyon.fr (F.P.)

[2]   Center for Nano Science (CENS), Ludwig Maximilians-Universität München, Butenandtstraße 5-13, 81377 München, Germany

[3]   LIPhy, University Grenoble Alpes, CNRS, F-38000 Grenoble, France

[4]   Medical Faculty "Carl Gustav Carus", Institute of Virology, Technische Universität Dresden, Fetscherstr. 74, 01307 Dresden, Germany; kstirnnagel@web.de (K.S.); sylvia.huetter86@gmail.com (S.H.); nicole.stanke@tu-dresden.de (N.S.); stefanie.richter4@tu-dresden.de (S.R.)

[5]   CRTD/DFG-Center for Regenerative Therapies, Technische Universität Dresden, Fetscherstr. 105, 01307 Dresden, Germany

[6]   Nanosystems Initiative München (NIM), Ludwig Maximilians-Universität München, Butenandtstraße 5-13, 81377 München, Germany

[7]   Center for Integrated Protein Science (CIPSM), Ludwig Maximilians-Universität München, Butenandtstraße 5-13, 81377 München, Germany

*   Correspondence: dirk.lindemann@tu-dresden.de (D.L.); d.lamb@lmu.de (D.C.L.)

†   These authors contributed equally to this study.

check for updates

Abstract: Viral glycoprotein-mediated membrane fusion is an essential step for productive infection of host cells by enveloped viruses; however, due to its rarity and challenges in detection, little is known about the details of fusion events at the single particle level. Here, we have developed dual-color foamy viruses (FVs) composed of eGFP-tagged prototype FV (PFV) Gag and mCherry-tagged Env of either PFV or macaque simian FV (SFVmac) origin that have been optimized for detection of the fusion process. Using our recently developed tracking imaging correlation (TrIC) analysis, we were able to detect the fusion process for both PFV and SFVmac Env containing virions. PFV Env-mediated fusion was observed both at the plasma membrane as well as from endosomes, whereas SFVmac Env-mediated fusion was only observed from endosomes. PFV Env-mediated fusion was observed to happen more often and more rapidly than as for SFVmac Env. Strikingly, using the TrIC method, we detected a novel intermediate state where the envelope and capsids are still tethered but separated by up to 400 nm before final separation of Env and Gag occurred.

1. Introduction

Several genera of spumaviruses constitute the Spumaretrovirinae subfamily of retroviruses. Spumaviruses, also known as foamy viruses (FVs), are endemic to a wide range of vertebrates [1,2]. They are thought to be the oldest retroviruses, having emerged >450 million years ago coinciding with the origin of jawed vertebrates [3] and display a remarkably stable co-speciation with their hosts.

Their largest genus, *Simiispumavirus*, combines all known FV species identified in different non-human primates of new- and old-world monkey, and ape origin. Prototype FV (PFV) is the best studied primate FV isolate. It was the first retrovirus discovered in humans, originally described as human FV, but was later recognized as being derived from a zoonotic transmission of a chimpanzee FV to man [4,5]. A hallmark of FVs, setting them apart from all other retroviruses, is their apparent apathogenicity, not only in their natural hosts but also in zoonotically infected humans. Although the FV genome structure is typical for a complex orthoretrovirus, FVs were grouped into a separate subfamily because research showed that their replication strategy deviates in several aspects from that of all other retroviruses [2]. Examples of special features of FVs in comparison to orthoretroviruses are a strictly viral glycoprotein (Env)-dependent particle egress that involves budding predominantly at intracellular membranes; a release of capsid-less, Env-containing subviral particles, and an extremely broad tropism that employs heparin sulfate attachment and currently unknown specific entry receptor(s).

FV structural protein synthesis and virion morphogenesis is also characterized by several special features that make these viruses an interesting tool for single particle tracing analysis using fluorescently labeled virions. First, FV Env biosynthesis is unique amongst retroviruses as the glycoprotein precursor is not cotranslationally processed by cellular signal peptidase, removing the N-terminal signal peptide, during translation into the secretory pathway [6–9]. Instead, a full-length Env precursor is translated and initially adopts a membrane topology with both N- and C-terminus located in the cytoplasm (Figure 1a) [7,8]. Only during cell surface transport is the Env precursor postranslationally modified and proteolytically processed by furin or furin-like proteases into the mature leader peptide (gp18^{LP}), surface (gp80^{SU}) and transmembrane (gp48^{TM}) subunits [6,9]. All three subunits are integral components of a heterotrimeric glycoprotein complex in released PFV virions with leader peptide (LP) and transmembrane (TM) subunits adopting a type II and type I membrane topology, respectively (Figure 1a,b). The extracellular surface (SU) subunit appears to be associated with LP and TM subunit through non-covalent interactions. On released FV virions the mature Env glycoprotein complex forms prominent spike structures that are organized in an elaborate network of interlocked hexagons [10,11].

FV virion release, unlike orthoretroviruses, is characterized by a strict dependence of Gag and Env co-expression as the FV Gag proteins lack membrane targeting or membrane association signals [12,13]. This Gag feature results in an accumulation of naked preassembled capsid at the centrosome and prevents an orthoretroviral-like release of virus-like particle release in the absence of Env coexpression. The Env-dependence of FV budding and release is the consequence of a direct interaction of the N-terminus of Gag in capsids preassembled at the centrosome and the N-terminal cytoplasmic domain of the Env LP subunit [7,8,14,15]. The active participation of FV glycoproteins in virion morphogenesis is also emphasized by their ability to induce the release of capsid-less subviral particles, similar to what is observed for the hepatitis B virus S protein [16,17].

Target cell entry of most FVs is thought to require endocytosis and involve a viral glycoprotein-mediated fusion of viral and cellular lipid membranes that is predominantly dependent on low pH [18,19]. Only PFV Env was previously reported to possess significant fusion activity at neutral pH, which is in line with the observation that PFV Env containing retroviral particles can fuse with target cells at the plasma membrane or from within endocytic vesicles.

Most of what has been learned regarding virus fusion has been obtained using bulk experiments. However, direct information regarding the kinetics and details of the fusion process are missing. With the development of single virus tracing (SVT) techniques [20,21], it is now possible to follow the viral–host cell interactions of a single virion and thereby elucidate new details regarding the fusion processes. Pioneering experiments revealed that the adeno-associated virus runs through several stages of motion, each characterized by distinct diffusion characteristics during the infection pathway into living cells [21]. Groundbreaking work on influenza viruses determined the kinetics of hemifusion and content mixing using an in vitro system [22] and characterized the infection process in living cells [23,24]. Several studies applied SVT to elucidate details of the fusion process of Simian virus 40 [25–27] as well as the one of the human polyoma virus [28] and the echovirus I [29].

Figure 1. Characterization of the virus particles used in this study. (**a**) Schematic illustration of the prototype foamy viruses (PFV) Env membrane topology at the state of the precursor (upper graphic) and oligomeric, mature glycoprotein complex (lower graphic). Color coding of the individual domains or subunits is indicated in panels (**b**–**e**). (**b**) Schematic representation of the PFV particle structure. pr: precursor protein; p: protein; gp: glycoprotein; gp18^{LP}: Env leader peptide subunit; gp80^{SU}: Env surface subunit; gp48^{TM}: Env transmembrane subunit; p85$^{PR-RT-RH}$: Pol protease-reverse-transcriptase-RNaseH subunit; p40^{IN}: Pol integrase subunit; pr71^{Gag}: Gag precursor; p68^{Gag}: Gag p68 subunit. (**c**–**e**) Schematic outline of the PFV Gag, PFV, or macaque simian foamy virus (SFVmac) Env, and PFV vector genome variants. Graphical illustration of the PFV Gag and PFV or SFVmac Env protein structure as well as the PFV viral genome structure of the different variants employed in this study. Numbers indicate the amino acid positions and the subunit processing sites within the translated precursor proteins are indicated by dashed lines. The viral RNA genome sizes are given in base pairs. NH$_2$: N-terminus; COOH: C-terminus; LP: Env leader peptide domain; SU: Env surface domain; TM: Env transmembrane domain; Ch: mCherry open reading frame; L: glycine-serine linker peptide; p68: Gag p68 domain; p3: Gag p3 domain; GFP: eGFP open reading frame; ***: Gag glycine-arginine box deletions; R: long terminal repeat (LTR) repeat regions; U5: LTR unique 5′ region; ΔU3: enhancer–promoter deleted LTR unique 3′ region; ©: Cap structure; A$_n$: poly A tail. (**f**) Virus particles spotted on a glass slide and recorded under widefield illumination using alternating laser excitation. Scale bar 10 μm. (**g**) Quantification of the particles used in this study according to their colocalization in both channels. Colocalizing green particles: Green particles colocalizing with a red particle signal; Colocalizing red particles: Red particles colocalizing with a green particle signal. Error bars show standard error of the mean determined from three fields of view.

Here, we use dual-color FV constructs and our recently developed tracking image correlation (TrIC) approach [30] to visualize the fusion event in both PFV (SFVpsc) and SFVmac (SFVmcy). The fusion of PFV was observed both at the plasma membrane as well as from endosomes and was observed to happen on the timescale of 10 to 20 min. In contrast, fusion of SFVmac was only observed from endosomes and occurred on timescales longer than was measured for the fusion of PFV in endosomes. This is consistent with the higher fusogenicity of PFV Env at neutral pH as observed previously [18,19]. Both types of viruses exhibited an intermediate state where the fluorescence signal from the envelope and capsid are separated by 100 to 400 nm but are still tethered together. This intermediate persists on the timescale of 6 to 11 min and is independent of the properties of the packaged RNA.

2. Materials and Methods

2.1. Expression Constructs and Virus Preparation

The 4-component PFV vector system consisting of the packaging plasmids containing either authentic or expression-optimized ORF encoding wildtype or mutant PFV (SFVpsc) Gag, PFV Pol, PFV Env or SFVmac (SFVmcy) Env and PFV transfer vectors encoding a lacZ reporter gene was described previously [18,31,32]. In addition to fluorescent protein tagged wild-type constructs, the following mutant variants were used for this study: PFV Env: pcoPE iCS1 (iFuse) encoding a surface–transmembrane subunit furin cleavage site variant by R571T mutation [18]; PFV Gag: pcoPG4 iNAB1 (iNAB) encoding a nucleic acid binding deficient variant by deletion of glycine-arginine rich (GR) boxes I to III [33]; and PFV Pol: pcoPP2 or pcziPol iRT encoding a variant with catalytically inactive reverse transcriptase by YVDD312–315GAAA mutation. A schematic outline of FV Gag, Env packaging and PFV transfer vector variants are shown in Figure 1c–e.

The dual-colored FV particles (Gag-GFP and mCherry-Env) were prepared and concentrated as explained previously [18]. Briefly, most PFV Env (PE) containing supernatants were generated by cotransfection of 293T cells with the transfer vector pMD11 (wt vgRNA) or puc2MD11 MS2Bas (long vgRNA) and expression-optimized packaging plasmids encoding PFV Pol (pcoPP2), PFV Env (pcoPE Ch, pcoPE Ch iCS1) and PFV Gag (pcoPG4, pcoPG4 CeGFP, pcoPG4 CeGFP iNAB1) at a ratio of 28:2:1:4. SFVmac Env (SE) and some PE containing supernatants were produced by cotransfection of the transfer vector pMD11 and packaging plasmids encoding PFV Gag (pcziGag-CeGFP), PFV Pol (pcziPol iRT), and PFV Env (pczPE iCS2) or SFVmcy (SFVmac) Env (pciSE Ch) at a ratio of 1:1:1:1. Cell-free viral supernatants were harvested 48 h post transfection and viral particles concentrated by ultracentrifugation or ultrafiltration and aliquots snap-frozen on dry ice and stored at −80 °C until further use.

2.2. Spinning Disk Confocal Microscope

Experiments were performed on a modified Andor Revolution system spinning disk confocal microscope (Andor Technology, Belfast, UK) (Figure A1). The system is built using a Nikon TE2000E (Nikon, Tokyo, Japan) microscopy body, a spinning-disk unit (CSU10; Yokogawa Electric Corporation, Musashino, Japan), an OptoSplit II (Cairn Research Ltd., Faversham, UK) for separating the eGFP and mCherry emissions and an EMCCD camera (DU-897 Ixon, Andor Technology, Belfast, UK) for detection. The excitation was controlled using an acousto-optic tunable filter (Gooch and Housego, Ilminster, UK) and the excitation and fluorescence emission were separated using a quadruple-band dichroic beam splitter (Di01-T405/488/568/647; Semrock, Rochester, NY, USA). The eGFP and mCherry signals were separated using a dichroic beamsplitter (BS562) and the respective emission filters (HC525/50, and ET605/70), all purchased from AHF Analysentechnik AG (Tübingen, Germany). Z-stacks were recorded over 20 min with an exposure time of 130 ms/frame/plane and 15–25 z positions spaced by 300 nm were acquired per z-stack. This resulted in a complete three-dimensional (3D) image every ~3 to 5 s.

2.3. Live-Cell Imaging Experiments

An overview of the live-cell imaging experiments is given in Figure A2. HeLa cells were cultivated in Dulbecco's Modified Eagle Medium (DMEM) + 10% fetal bovine serum (FBS) at 37 °C in a 5% CO_2 atmosphere and split every 2 to 3 days. Cells were seeded at 2×10^4 cells per well in an 8-well Nunc LabTek II chambered coverglass slide coated with collagen A-solution (Sigma-Aldrich, St. Louis, MO, USA) according to the manufacturer's protocol one day prior to experiments. On the day of the experiment, the cells were washed once with phosphate-buffered saline (PBS) solution and the virus particles were added at a density of 40 to 80 particles per cell in Leibovitz's L15 medium containing 10% FBS. To allow attachment of the particles to the cell surface while avoiding virus uptake into the cells, cells were incubated with the virus at 4 °C for 10 min. Subsequently, the cells were rinsed with

cold L15 medium and the imaging was started immediately after mounting the sample holder on the microscope stage and warming the cells to 37 °C to synchronize the uptake of the viruses.

Data was recorded over 20 min on single cells by acquiring multiple *z*-stacks spanning the entire cell volume with z-planes spaced 300 nm apart. EGFP and mCherry-labeled virus particles were excited in parallel with 488 nm and 561 nm continuous wave lasers. Figure A1 depicts the microscope setup used All cell culture reagents were purchased from Thermo Fisher Scientific, Waltham, Massachusetts, USA, if not stated differently.

2.4. Data Analysis

The analysis was done using the TrIC software (Figure A3) previously developed in our lab [30]. Briefly, the virus is tracked in 3 dimensions in the eGFP channel. Single viruses are tracked in 2D either manually, automatically using TrackMate [34] or using a home-written wavelet tracking method [35]. The subpixel accurate 3D-trajectory is obtained in a second step by fitting the particle image with a 2D-gaussian function for the *x-y* position and a 1D-gaussian function for the *z* position. A box about the particle is taken (2.94 μm × 2.94 μm × ~5 μm or 21 pixels × 21 pixels × entire *z*-stack) and a 3-dimensional image cross-correlation is performed between the eGFP and mCherry channels. When a particle is detected in both channels, a positive correlation signal is observed. To determine a threshold for the amplitude of the correlation function, we also randomize the pixels in the voxel/box and perform the same analysis. Using the average value and standard deviation of the randomized signal from a sliding window of 10 3D-images, we define a threshold of the mean plus 3 standard deviations for the randomized image as the threshold for a positive cross-correlation signal. The software provides the background-corrected intensity of both fluorescence labels, subpixel accurate coordinates of the tracked virus in 3D for both detection channels, the instantaneous velocity of the particle, the 3D-colocalization status and the relative distance of the fluorescence labels along the track (Figure A3).

To test the precision of our tracking method, we tracked dual-color fluorescent beads (TetraSpeck microspheres, 0.1 μm, Thermo Fisher Scientific, Waltham, MA, USA) in a glycerol solution (Figure A4). This trajectory shows the accuracy limit of our method. The mean relative distance was 25 nm over the ten minutes tracked with a standard deviation of 10 nm. For testing the resolution in the case of the dual-color viruses, we performed experiments on a fusion-deficient mutant, iFuse, and the mean relative distance was found to be 46 nm with a standard deviation of 20 nm, giving a 95% confidence boundary at 86 nm (data not shown). The threshold for separation of the two colors was then set to 100 nm.

2.5. Estimation of the Cell Surface

As the cells were pre-incubated with the virus on ice in the refrigerator (~4 °C), most of the viruses are located at the cell membrane during the first frames of the movies. We took advantage of this situation to estimate the cell surface. The visible viruses were automatically detected in the first *z*-stack and their position was automatically determined in 3D via Gaussian fitting (see Figure A5a). The cell surface was then reconstructed from the single virus positions by a nearest neighbor interpolation and plotted as a 3D-surface together with a virus track to help determine the location of viral fusion with respect to the cell membrane (see Figure A5b).

3. Results

3.1. Characterization of Foamy Virus Particles

With the aim of visualizing the fusion process of individual viruses, we further characterized the fluorescence-labeled FV particles that we designed previously [18]. Here, PFV and SFVmac Env-containing particles were labeled with eGFP and mCherry. It has been shown that endocytosis plays a significant role in the uptake of FV [18,19]. For endosomal uptake, the viruses encounter a decrease in pH from early endosomes (pH ~ 6.5), to late endosomes (pH ~ 5.5), to lysosomes (pH ~ 4.5).

As it is known that eGFP fluorescence is quenched at acidic pH values [36], we tested the stability of the fluorescent proteins incorporated into the virus particles under these conditions (Figure A6). Particles labeled with eGFP attached to the Gag protein and mCherry attached to the Env protein were sedimented on a glass slide. The initial pH of 7.0 was then dropped to pH 5.5. A small reduction in the Gag-eGFP fluorescence upon lowering of the pH was observed, but the eGFP signal was still easily detectable (Figure A6a, left graph). mCherry-Env showed no change in fluorescence intensity (Figure A6a, right graph). For comparison, we first permeabilized the particles with Triton X-100 to allow access of the protons to the eGFP-labeled capsid. As expected, the eGFP fluorescence showed strong quenching at pH 5.5, a value typically found in endosomes (Figure A6b, left graph) whereas the mCherry fluorescence stayed unaffected (Figure A6b, right graph). Hence, the choice of label position with eGFP attached to the capsid works well even when the viruses have been taken up in endosomes. The eGFP is well shielded within the virus lipid envelope.

Secondly, a high fusogenicity is needed for these studies. Hence, as we have shown previously, we generated the viruses by mixing unlabeled Gag proteins with eGFP-tagged Gag at a ratio of 3:1 [18]. This mixture keeps the infectivity of the virus at near wild-type levels whereas labeling 100% of the Gag proteins with eGFP reduces the infectivity of the particles by approximately 90%. Labeling of the Env did not significantly alter the infectivity of the virus. Although a drop in infectivity, as determined by reporter gene expression, may not necessarily indicate a decrease in fusogenicity, a high infectivity does guarantee that the fusogenicity is also high. Thus, we used constructs using a mixture of 3:1 Gag:Gag-eGFP. From bulk infectivity assay and time-lapsed spinning disk confocal microscopy, we verified the fusogenicity of the viruses and could show that a significant fraction of PFV particles (~15%) underwent fusion during the first 30 min [18].

Lastly, to simplify detection of the fusion process, it is important that a high fraction of the prepared viral particles contain both labels. Fortunately, FVs cannot bud without the envelope protein [12,13]. Hence, 93 ± 1% of viral particles released into the cell culture supernatant were fluorescently labeled with Env-mCherry (Figure 1e,f). Conversely, about half of the viral envelopes were missing a detectible capsid signal [18]. This high number of presumably empty virus-like particles could be due to the optimized gene expression used to generate the virus [31,32], the known capability of FV to form sub-viral particles [17] and/or due to the attachment of a fluorescent protein to the N-terminus of a FV glycoprotein. Although a higher capsid incorporation would simplify the fusion measurements, 50% is the limitation of this viral system. However, with the high envelope labeling efficiency, the virus preparation is still well suited for investigating fusion using single virus tracing. A detailed characterization of the virus preparations is given in Table 1.

Table 1. Characteristics of dual-color labeled samples.

Label of Virus Prep	vgRNA	Viral Gag	Viral Env	Gag Ratio Tagged:Untagged	∑ Green Particles	% Coloc. Green ± SEM	∑ Red Particles	% Coloc. Red ± SEM
PFV wt [1,2]	wt	PG EG	PE Ch	1:0	6666	97 ± 0.8	12,649	55 ± 0.5
PFV wt [2]	wt	PG EG	PE Ch	1:3	2365	98 ± 0.2	7683	33 ± 0.7
PFV wt [1,2]	wt	PG EG	PE Ch	1:3	1476	93 ± 1	4022	33 ± 1
PFV iFuse [1]	wt	PG EG	PE Ch iCS1	1:3	542	95 ± 1	1159	48 ± 2
PFV no vRNA	-	PG EG	PE Ch	1:3	168	100 ± 0	351	48 ± 6
PFV long vRNA	long	PG EG	PE Ch	1:3	152	94 ± 6	235	64 ± 5
SFV wt [1]	wt	PG EG	SE Ch	1:3	944	93 ± 0.8	3290	30 ± 1

[1] Characterization of these particle lots was already described in [18]. [2] Different lots of dual-colored, wildtype PFV particles.

3.2. Investigation of Individual Prototype Foamy Virus Env-Mediated Fusion Events

To gain detailed insights into the fusion process, we performed live cell measurements using the fluorescently labeled FV particles described above. Cells were incubated with viral particles at ~4 °C.

Subsequently, the cells were rinsed with cold L15 medium and the imaging was started immediately after warming the cells to 37 °C to synchronize the uptake of the viruses. Confocal *z*-stacks were collected over approximately 20 min and the data analyzed as outlined in the Section 2.

Figure 2 and Movie S1 show the results of typical fusion of a PFV. Panel a of Figure 2 depicts the differential interference contrast (DIC) image of an infected cell overlaid with the trajectory of the infecting particle in three dimensions (3D). At the beginning of the movie, the virus is located at the plasma membrane. Both PFV Gag-eGFP and PFV mCherry-Env signals are observable, represented by the yellow color-coding in the trajectory. During the initial stage, the virus undergoes limited movement and slow photobleaching of the eGFP signal can be observed (Figure 2(bi)). Subsequently, the fluorescence signals begin to separate. At 9.7 min, the capsid is transported into the cytoplasm of the cell (green part of the trace) whereas the envelope remains on the plasma membrane (Figure 2a). To provide additional support that fusion actually occurred at the plasma membrane, we developed another method to help define the location of the plasma membrane. Before starting the acquisition of a movie, cells were incubated at low temperature to synchronize virus entry. As a consequence during the first *z*-stack, most viruses were still located on the cell surface. By automatically determining the 3D-position of all the fluorescent particles present in the first *z*-stack of the movie, the 3D-cell shape was inferred (see Materials and Methods, Figure A5). Comparison of the 3D-position of the tracked virus to the reconstructed 3D cell shape was then used to determine the location of the virus during the fusion. This method was helpful in identifying plasma fusion events but could also indicate fusion from a particle in the actin cortex just underneath the plasma membrane as the accuracy of the method is not sufficient to resolve these two cases.

The results of the TrIC analysis are shown in Figure 2b. The cross-correlation amplitude is a measure of the similarity of the image data in the two channels and a positive correlation above the control indicates a clear viral signal in both channels. At 9.7 min, the cross-correlation amplitude (Figure 2(biii)) between the two channels shows a clear drop indicating that the mCherry- Env signal is no longer in the box around the viral Gag-eGFP being tracked. This is a clear marker of color separation and indicates complete fusion of the virus. The drop in cross-correlation amplitude coincides with the loss of fluorescence intensity in the mCherry channel and with a sudden increase in the instantaneous velocity to values above 1 μm/s (Figure 2(bii)). The high velocity and directionality of motion indicates that the capsid is being actively transported towards the cell center. The transport velocity of 0.5 to 1 μm/s is consistent with values observed for direct transport of internalized viruses and endosomes along microtubules [23,24,37–39]. Thus, we assign the post-fusion active transport processes to the capsid hijacking cellular motor proteins and being transported along the microtubule network (summarized in [40]).

The image cross-correlation analysis also provides information about the relative distance between the signals in the two channels. As discussed in the materials and methods, we consider color separation to be significant when the relative distance increases over 100 nm. Figure 2(biv) depicts this information over the course of the observed fusion event. It can be subdivided into three stages. Initially, the separation between the eGFP and mCherry signals is approximately 100 nm or less (blue), which is within the detection limit. In stage 2, which starts about four minutes after the beginning of the track, the distance increases to values between 100 nm and 400 nm (cyan). This stage lasts for about 6 min. In stage 3 (green), separation rises above 400 nm until the mCherry-Env signal disappears from the tracking box around the Gag-eGFP signal, at which point the fusion is complete. Figure 2c shows single frames representative for the three stages from the recorded movie data. Plotting the relative distance of the envelope from the capsid signal in three dimensions (Figure 2d) reveals in stage 2 motion of the envelope around the capsid until the separation is completed.
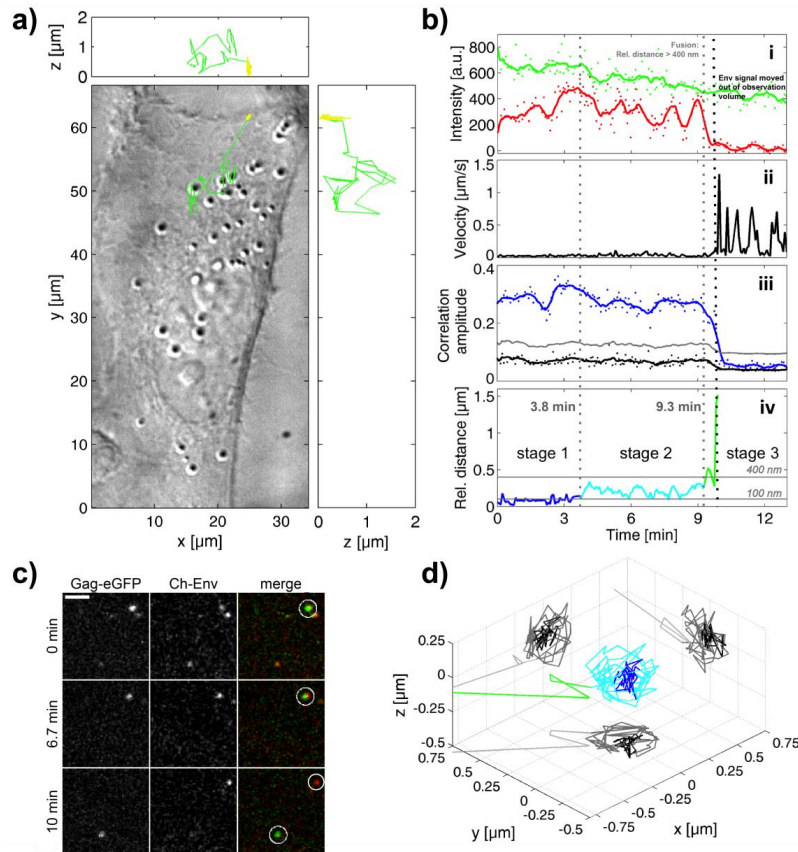
Figure 2. PFV Env-mediated fusion at the plasma membrane. (**a**) A differential interference contrast (DIC) image of a cell where the trajectory of the tracked PFV has been overlaid. The proportion of the track where the particle is double-labeled is plotted in yellow whereas movement of the capsid towards the nucleus/microtubule organizing center after color separation is plotted in green. *x-z* and *y-z* projects of the trajectory as shown in the upper and right panels respectively. (**b**) Results of the tracking imaging correlation (TrIC) analysis along the trajectory are shown: (**i**) The background-corrected fluorescence intensity of the Gag-eGFP channel (green) and the mCherry-Env channel (red) are plotted as a function of time. (**ii**) The instantaneous velocity of the Gag-eGFP signal is plotted as a function of time. (**iii**) The amplitude of the cross-correlation of the TrIC analysis for the data (blue), the randomized control (grey) and threshold (black) are plotted as a function of time. (**iv**) The TrIC analysis also provides the relative distance between the fluorescence signals in the two channels. The distance of the Gag-eGFP to the mCherry-Env is plotted over time, with distances <100 nm marked in blue (stage 1), between 100 and 400 nm marked in cyan (stage 2), and >400 nm marked in green (stage 3). Solid lines were generated using running average of ~30 s. (**c**) A close up of three frames from the movie showing the tracked virus (circled in white) at different stages of the fusion process: bound to the plasma membrane (0 min), during stage 2 (6.7 min), and after fusion (10 min) where the capsid has been transported within the cell. Left: eGFP channel; Middle: mCherry channel; Right: merged image. Scale bar: 4 μm. (**d**) A 3D-representation of the relative position of the mCherry-Env signal with respect to the Gag-eGFP signal distance color-coded according to the three stages shown in panel biv. The circular movement of the envelope signal around the capsid in stage 2 (cyan) is clearly visible.

3.3. Fusion Mediated by Simian Foamy Virus Env

As a comparison, we also investigated the fusion mediated by SFVmac Env (Figure 3 and Movie S2). For SFVmac Env, this viral particle was internalized, complete with envelope, within the first minute. The overlay of the 3D trajectory with the DIC image of the cell shows that the particle was actively transported towards the nucleus and that capsid release occurred internally (Figure 3a). Instantaneous velocities of up to 600 nm/s where measured (Figure 3(bii)). After 21 min, the particle fused with the surrounding cellular membrane (probably with an endosome) as indicated by a sudden drop of the mCherry-Env signal (Figure 3(bi)), a drop of the correlation amplitude (Figure 3(biii)) and an increase of the relative distance between the eGFP and mCherry signal to values above 400 nm (Figure 3(biv)).

Again, for this event, the relative distance between the two signals can be divided into three stages (Figure 3(biv)): a stage in which the two signals colocalize within the detection limit of 100 nm (blue), an intermediate stage in which the relative distance varies between 100 and 400 nm (cyan) and finally full separation (green). Figure 3c shows three time points of the underlying image data highlighting the three stages. In the first two time points, there is at least some colocalization of the two signals. The lowest section shows the phase of full separation after ~22.3 min. We can also plot the 3D-distance between the eGFP labeled capsid and the mCherry-Env signal (Figure 3d). Here again, during the intermediate stage, the envelope and capsid signal move with respect to each other.

3.4. Analysis of All Observed Events

We have demonstrated that it is possible to visualize fusion events using the fluorescently labeled FV particles harboring PFV Gag and PFV or SFVmac Env at different steps of viral uptake as schematically depicted in Figure 4a. Figure 4 gives a summary of all the observed single virus particles and fusion events observed for both types of FV particles. The number of fusion events detected was limited and thus the statistics are limited. From 520 detected PFV Env-containing particles, we tracked 88 individual virions and observed 13 fusion events (Figure 4b). Four fusion events were detected at the plasma membrane and nine events occurred from endosomes. The plasma membrane events were categorized as such when the particles were located on the plasma membrane (within the resolution of our microscope) and they have not undergone motion with velocities above 0.05 μm/s prior to fusion, consistent with transport velocities of particles transported on the plasma membrane [41,42]. The possibility of PFV Env being able to fuse at the plasma membrane is consistent with its ability to fuse to a significant extent already at neutral pH [18,19]. For fusion events categorized as occurring from endosomes, the particles demonstrated clear active transport before the fusion event was detected. In contrast to the plasma membrane fusion event shown in Figure 2 and Movie S1, the mCh-Env signal was typically lost within 5 to 15 s after the endosomal fusion event was completed, which is consistent with what has been reported for other viruses [24,43]. Disappearance of the mCh-Env signal is attributed to dilution of viral glycoproteins in host cell membranes after the fusion process and was observed both for fusion at the plasma membrane and in endosomes.

From the measurements, we can also gain insights into the kinetics of the fusion event. For fusion from the plasma membrane, the average time until fusion was 19 min. For fusion events from endosomes, it was necessary to separate the events into cases where the entire uptake was observed (Figure 4(ai)) and cases where the viruses had already undergone endocytosis before being detected (Figure 4(aii)). For the latter, it was not always possible to calculate the entire time until fusion. However, a minimum time could be estimated and, in either case, the time it took fusion to occur from the plasma membrane was significantly longer. The short lag time between entry until fusion in the case of endocytosed particles relative to fusion at the plasma membrane is in agreement with the reported pH-dependency of PFV fusion and the fact that early endosomes already have a slightly acidic pH around 6.5 to 6.0 [44–46]. In addition, early endosome fusions can be induced within 1 to 5 min after virus uptake (e.g., vesicular stomatitis virus, VSV), whereas late endosome fusion events (e.g., influenza A virus, INF) take longer, ranging from 10 to 20 min [44]. The kinetics of the fusion process are given in Figure 4c.
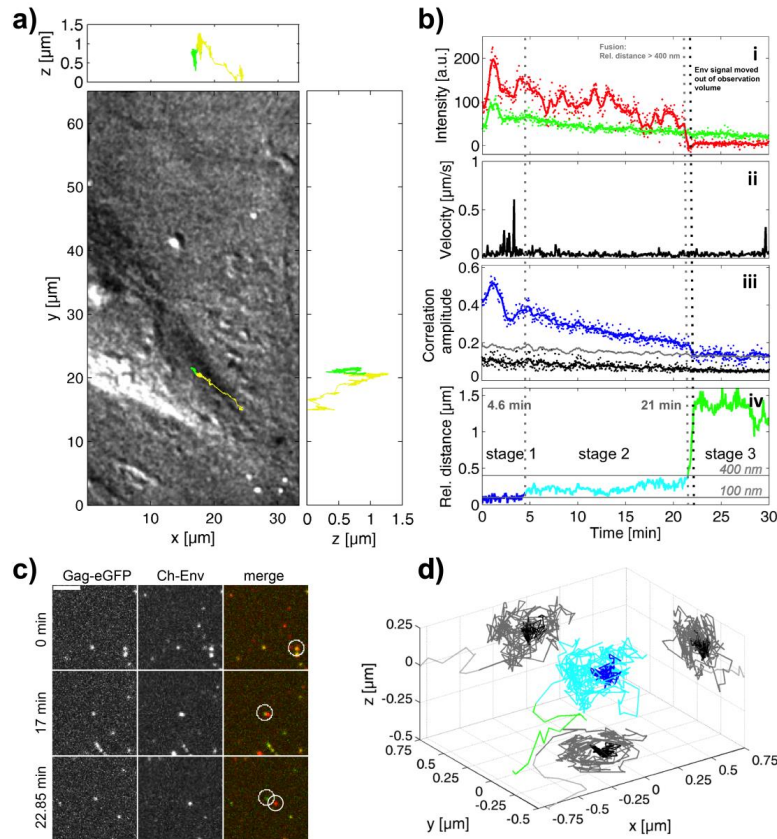
Figure 3. SFVmac Env-mediated fusion from the endosome. (**a**) A DIC image of a cell where the trajectory of the tracked SFVmac has been overlaid. The proportion of the track where the particle is double-labeled is plotted in yellow and motion of the capsid after color separation is plotted in green. *x-z* and *y-z* projects of the trajectory as shown in the upper and right panels respectively. (**b**) Results of the TrIC analysis along the trajectory are shown: (**i**) The background-corrected fluorescence intensity of the Gag-eGFP channel (green) and the mCherry-Env channel (red) are plotted as a function of time. (**ii**) The instantaneous velocity of the Gag-eGFP signal is plotted as a function of time. (**iii**) The amplitude of the cross-correlation of the TrIC analysis for the data (blue), the randomized control (grey) and threshold are plotted as a function of time. (**iv**) The relative distance between the eGFP-capsid and mCherry-Env signal is plotted over time. The distance of the Gag-eGFP to the mCherry-Env is plotted over time, with distances <100 nm marked in blue (stage 1), between 100 and 400 nm marked in cyan (stage 2), and >400 nm marked in green (stage 3). Solid lines were generated using running average of ~30 s. (**c**) A close up of three frames from the movie showing the tracked virus (circled in white) at different stages of the fusion process: bound to the plasma membrane (0 min), just after the virus has been internalized (17 min), and after the fusion is completed (22.85 min). Left: eGFP channel; Middle: mCherry channel; Right: merged image. Scale bar: 4 μm. (**d**) A 3D-representation of the relative position of the mCherry-Env signal with respect to the Gag-eGFP signal distance color-coded according to the three stages shown in panel (biv). The circular movement of the envelope signal around the capsid in stage 2 (cyan) is clearly visible.

b) Statistics

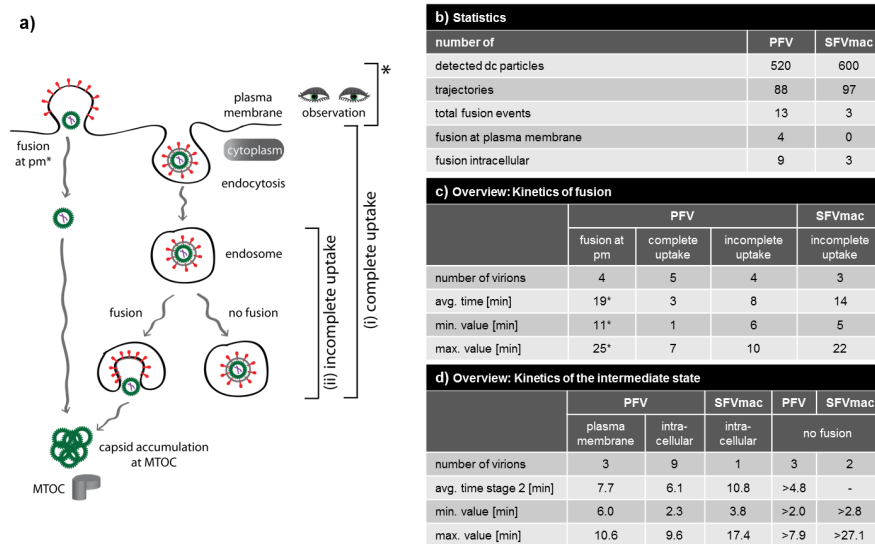| number of	PFV	SFVmac
detected dc particles	520	600
trajectories	88	97
total fusion events	13	3
fusion at plasma membrane	4	0
fusion intracellular	9	3

c) Overview: Kinetics of fusion

	PFV			SFVmac
	fusion at pm	complete uptake	incomplete uptake	incomplete uptake
number of virions	4	5	4	3
avg. time [min]	19*	3	8	14
min. value [min]	11*	1	6	5
max. value [min]	25*	7	10	22

d) Overview: Kinetics of the intermediate state

	PFV		SFVmac	PFV	SFVmac
	plasma membrane	intra-cellular	intra-cellular	no fusion	
number of virions	3	9	1	3	2
avg. time stage 2 [min]	7.7	6.1	10.8	>4.8	-
min. value [min]	6.0	2.3	3.8	>2.0	>2.8
max. value [min]	10.6	9.6	17.4	>7.9	>27.1

Figure 4. Overview of the detected fusion events. (**a**) Schematic illustration of the different types of fusions events detected during cellular uptake of dual-labeled FV particles. Events were classified into fusion at the plasma membrane and fusion from the endosome. Fusion events from an endosome were further divided into complete events (where the particle could be tracked from attachment at the plasma membrane until completion of the fusion from the endosome) and incomplete events for which the particle was only detected when it had already been taken up into an endosome. (**b**) Statistics of detected and tracked particles and detected fusion events. (**c**) Overview of the kinetics of fusion. * time at 37 °C until fusion. (**d**) Overview of the kinetics of the intermediate state.

For SFVmac Env containing particles, we detected 600 viruses, tracked 97 particles, and detected three fusion events (Figure 4b). With these low statistics, it is not possible to make any definitive statements. What we did observe is that all three fusion events occurred from endosomes, consistent with the strong pH dependence of SFVmac Env-mediated fusion [18]. Although we can only estimate a minimum time until endosome fusion for SFVmac, it was significantly slower than that for endosomal fusion of PFV (Figure 4c). This is also consistent with the difference in uptake observed for PFV and SFVmac Env containing particles and may suggest that fusion of SFVmac occurs from late endosomes [18].

The PFV Env mediated uptake pathway has similarities with many other viruses. The uptake pathway resembles that of human immunodeficiency virus 1 (HIV-1) in that PFV infects target cells by fusion at the plasma membrane or by endocytosis [47]. PFV uptake also shows similarities with VSV, which fuses early after uptake in early endosomes [48]. In addition, the PFV fusion process can also be activated in maturing or late endosomes on the way towards the cell center as indicated by the fusion events observed in between 7 and 10 min after virus entry (Figure 4c). In contrast, we suggest that SFVmac Env containing particles fuse similar to INF viruses predominantly in late endosomal compartments [49]. The average minimum duration from uptake till fusion of 13.6 min supports this model.

3.5. A Puzzling Delay in Fusion

One of the surprising results of this investigation is the discovery of an intermediate step (stage 2) in the fusion process where the envelope and capsid signals are in close proximity, but are no-longer fully overlapping. Using the TrIC analysis, we detect a physical separation between the envelope

and capsid signals of 100 to 400 nm. This is clearly greater than the accuracy of our tracking method. Both PFV and SFVmac Env-containing particles exhibited this intermediate step that was observed in all fusion events with the exception of one fusion event at the plasma membrane (Figure 4d). Occasionally, we observed viruses in the intermediate state that did not undergo complete fusion within the measurement time, but we never observed a virus return from the intermediate state into stage 1. Hence, the intermediate stage appears to be an important step in completing the fusion process in FV Env mediated entry, so we investigated it in more detail.

The first question we asked is what causes the two components to stay in close proximity. This could be caused by either spatial confinement, such as both components being trapped in an endosome or a pocket in the actin cortex, or the two components could be tethered together by an unknown component. To distinguish between these two alternatives, we compared the mean square displacement (MSD) analysis on the relative trajectory of envelope with respect to capsid and on the absolute trajectories of the Gag-eGFP and the mCh-Env during stage 2 (Figure 5a). If the two components are moving independently, the relative motion between them will be the sum of two random motions and will have a higher diffusion coefficient than the absolute diffusion coefficients. On the other hand, if the two components are tethered together, the relative diffusion coefficient will be less than that of the absolute diffusion coefficients. The early times need to be compared, as confinement in an endosome will limit the sensitivity of the MSD analysis to the differences in motion. Figure 5a shows a MSD analysis of the absolute diffusion of the Gag-eGFP and mCh-Env in comparison to the relative motion for an SFVmac Env fusion event from an endosome. The relative diffusion coefficient is clearly smaller than the absolute diffusion coefficients, suggesting that the two components are tethered together. This is true for all fusion events measured.

Next, we investigated the average separation and duration of the intermediate state (stage 2) for the different categories of fusion events. Image-wise histograms of the envelope-capsid separation is plotted in Figure 5b and shows a peak with lower average separation for PFV Env constructs (blue, green, red) in comparison to SFVmac Env (black). A small difference is also observable when investigating the average separation per event (Figure 5d). When separating the PFV Env events into fusion from the plasma membrane and from endosomes, the image-wise histograms show a narrow peak in the separation for endosomal fusion events compared to a broader histogram for fusion from the plasma membrane (Figure 5c). Although the statistics are too low to make a significant statement, this could be due to confinement in endosomes limiting the maximum possible separation. The fusion events from endosomes for SFVmac Env (Figure 5b, black) exhibit an average larger separation then the endosomal fusion events measured for PFV Env (Figure 5b–d, black versus light blue), which would suggest that fusion occurs from different types of endosomes. When investigating the separation and duration of the intermediate state (Figure 5d–f), no significant trend is observable. The average duration of stage 2 in SFVmac Env mediated fusion events is longer, but this can be attributed to one very long event out of three. Hence, the statistics are not sufficient to make any significant statements here. Moreover, the scatterplot of duration versus average separation shows no correlation (Figure 5f).

Furthermore, we considered what could be physically tethering the envelope and capsid together over hundreds of nanometers. One possible candidate would be the viral genome, although one would expect the genome to be packaged inside the capsid during fusion and not having physical contact with the Env protein. We prepared several variants of PFV Env containing particles with differences in the packed genome (Figure 1e,g and Table 1). The first variant (no vRNA) were particles containing wild type PFV Gag-eGFP but no viral genome was co-expressed during virus productions. In this case, the distance distribution and duration of the intermediate state was not significantly different for SFVmac Env (SFV) or for endosomal fusion from PFV Env containing particles (PFV endo) with regular PFV genomes (Figure 5d,e). However, this PFV variant, though it does not contain a viral genome, still packages nonspecific cellular mRNA [33]. As a next step, we generated a PFV variant (iNAB) that was incapable of binding and incorporating any kind of RNA due to deletion of the PFV Gag nucleic acid binding domain [33]. However, we were unable to detect any fusion events for this

construct (data not shown). For the final variant (long vRNA), we replaced for virus production the regular transfer vector expressing the encapsidated viral RNA genome by a transfer vector expressing a viral RNA genome that had been extended by 1319 bases, or an increase of 20% (Figure 1e). For this PFV variant, we saw no significant difference in the average separation or the duration of the fusion event (Figure 5e). Details of the intermediate state are summarized in Figure 4d.
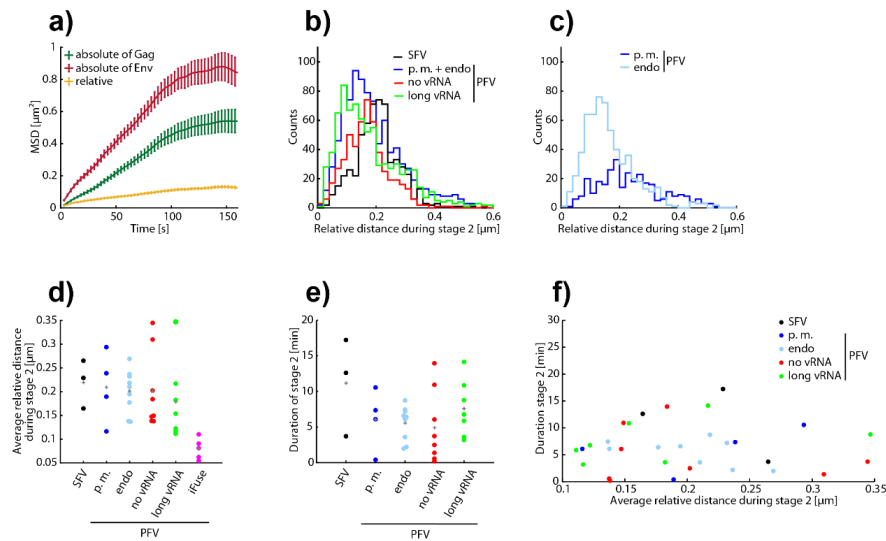

Figure 5. The intermediate stage during the fusion process. (**a**) A mean-square displacement (MSD) analysis of the absolute trajectory of the envelope signal (red), the absolute trajectory of the capsid (green) and of the relative trajectory of the envelope with respect to the capsid (yellow) during stage 2 of a SFVmac fusion event. The relative diffusion coefficient is clearly smaller than the absolute diffusion coefficients indicating that the two signals are tethered together. The Env protein diffuses faster than the Gag protein. Error bars show the standard error of the mean of the squared displacement. (**b**) Image-wise histograms of the relative distance of the capsid and envelope component within the intermediate step of the fusion process. PFV particles (blue) show a shift towards lower distances compared to SFVmac (black) or particles without a viral genome (red). (**c**) Image-wise histograms of the relative separation of envelope and capsid depending on whether fusion occurred from the plasma membrane (dark blue) or from endosomes (light blue). (**d**) A comparison of the average separation of the envelope from the capsid plotted per event for different types of FV particles. (**e**) A comparison of the average duration of the intermediate state per event for different types of FV particles. (**f**) A scatter plot of the average envelope-capsid separation versus duration of the intermediate fusion state for different types of FV particles. FV particle types: SFVmac virions at endosomes (SFV); PFV virions at the plasma membrane (PFV p.m.); PFV virions at endosomes (PFV endo); PFV virions that do not contain viral RNA (PFV no vRNA); PFV virions with longer viral RNA (PFV long vRNA); fusion incompetent PFV virions (PFV iFuse).

4. Discussion

Using the recently developed dual-colored PFV virions containing Gag-eGFP and mCherry-Env labels, we were capable of detecting individual fusion events. Strikingly, in 185 viral particle tracks, we were able to observe a total of 16 fusion events, characterized by a loss of colocalization and separation of Gag-eGFP and mCherry Env signals. This is in contrast to a previous study for HIV-1 by Koch and colleagues who used an ecotropic murine leukemia virus (eMLV) Env-YFP combined with HIV-1 MA-mCherry to study particle fusion [50]. From more than 20,000 2D trajectories,

they were able to detect 28 fusion events of rapid color separation. In addition, the authors identified 45 events of simultaneous loss of MA and Env signal, which is most likely due to endocytosis of double-labeled particles. Simultaneous disappearance of MA and Env signal was also observed for particles pseudotyped with fusion-deficient Env proteins, although to a lower extent. Hence, FV is much more fusogenic than eMLV pseudotyped HIV-1.

In our study, we detected for PFV Env containing particles fusion both at the plasma membrane as well as from endosomes whereas fusion of SFVmac Env containing particles was only detected from endosomes. This is in line with previous results demonstrating that Env protein of PFV is the only FV Env examined so far that already has a significant fusion activity at neutral pH, thereby enabling fusion at the plasma membrane of the host cells [18,19].

However, even PFV Env mediated fusion is known to be enhanced by low pH [19]. In agreement with this, we also observed PFV Env-mediated fusion events after endocytic uptake of viral particles. By comparing the kinetics of viral fusion after endocytic uptake between PFV Env- and SFVmac Env containing particles, we found that the time from uptake to fusion was also faster for PFV Env in comparison to SFVmac Env containing particles. This may suggest that pH-triggered PFV Env mediated fusion takes place in a different endosomal compartment with higher pH, perhaps in early endosomes, whereas SFVmac Env mediated fusion may only be triggered by the lower pH found in late endosomes. However, further single particle tracking studies including fluorescent markers for endosomal subpopulations are required to verify this hypothesis.

Using the TrIC approach, we detected a novel intermediate state during the fusion process where the envelope and capsid separate on the range of 200 to 400 nm that occurs for 6 to 10 min before complete fusion occurs. During this intermediate state, the envelope and capsid are still tethered. We showed that the fusion intermediate is not just due to co-confinement and could rule out that it is the viral genome that tethers the envelope and capsid together during this intermediate state. Fusion of membrane enveloped virus particles with host cell membranes is a multi-step process involving membrane merging via a universal "cast-and-fold" mechanism. Fusion-protein-mediated membrane merging is characterized by different intermediate structures including stalk formation, hemifusion, pore formation, pore growth, and, finally, capsid delivery [51,52]. It may be possible that the ability to detect the intermediate state during fusion of FV Env is dependent on the location of the fluorescent tag on the FV glycoprotein. All glycoproteins used in this study had the fluorescent protein tag fused to the cytoplasmic N-terminus of the Env LP-subunit, which adopts a type II membrane topology (Figure 1a). LP is thought to be an integral component of the heterotrimeric FV glycoprotein complex, also containing SU- and TM subunits [10]. Structural glycoprotein complex rearrangements in the multi-step fusion process may lead to release of LP from the other subunits during early phases and a higher mobility of the free, tagged LP subunit within the surrounding membrane(s) resulting in a greater distance to the tagged FV capsid. Alternatively, or in addition, the unique organization of trimeric PFV glycoprotein complexes in intertwined hexagonal networks on the surface of virions [8,10] may be responsible for the occurrence of the intermediate state. Perhaps expansion of the FV Env-mediated fusion pore, which is required for release of the capsid into the cytoplasm, is progressing slower than for other retroviral glycoproteins because the transmembrane helices of $gp18^{LP}$ and $gp48^{TM}$ may move as a block during fusion as suggested by Effantin and colleagues [8,10]. This may enable detection of an intermediate state with an increased distance between Gag and LP labels and their tethering for a certain time period.

In future studies, it would therefore be interesting to identify functional FV Env variants with fluorescent tags in the SU- or TM subunit and determine whether fusion events of such dual-colored FV virions also allow identification of an intermediate stage 2 fusion state. Furthermore, placing this fusion intermediate state in context with the other steps in the fusion process and to determine the actual biomolecules responsible for tethering the capsid and Env proteins during the fusion process, is of great interest for follow-up studies.

Abbreviations

DIC	differential interference contrast
DMEM	Dulbecco's Modified Eagle Medium
eGFP	enhanced green fluorescent protein
Env	envelope
eMLV	ecotropic murine leukemia virus
FBS	fetal bovine serum
FV	foamy virus
Gag	group specific antigen
HIV-1	human immunodeficiency virus 1
INF	influenza A virus
LP	leader peptide
mRNA	messenger RNA
MSD	mean square displacement
PBS	phosphate-buffered saline
PFV	prototype foamy virus
Pol	polymerase
RNA	ribonucleic acid
SFVmac	macaque simian foamy virus
SFVmcy	simian foamy virus Macaca cyclopis
SFVpsc	simian foamy virus Pan troglodytes schweinfurthii
SU	surface
SVT	single virus tracing
TM	transmembrane
TrIC	tracking image correlation
vRNA	viral RNA
VSV	vesicular stomatitis virus
YFP	yellow fluorescent protein

Appendix A

Figure A1. Spinning disk confocal microscope setup. A modified confocal spinning disk setup based on an Andor Revolution system was used. Continues wave laser lines of 405 nm, 488 nm, 561 nm and 640 nm wavelength built into the Andor Laser Combiner were directed into a single mode fiber. Continues wave laser lines of 445 nm, 514 nm, 594 nm and 685 nm were combined with the appropriate dichroic mirrors and directed into a single mode fiber together with the other laser lines and all laser lines were coupled into a Yokogawa CSU-10 unit and focused onto the sample by a 100×, 1.49 NA objective. The fluorescence signal was split with the appropriate dichroics and focused onto 3 Andor iXon cameras equipped with the appropriate emission filters. A piezo element inserted into the stage enabled the recording of *z*-stacks and a heatable sample chamber allowed live cell measurements at 37 °C. A perfect focus system was used to compensate for *z*-drift during data acquisition.

Figure A2. Sample preparation and data acquisition. Precooled HeLa cells were incubated double-labeled FV particles at 4 °C to allow membrane attachment and synchronize uptake. Immediately after transferring the sample onto the 37 °C warm microscope stage virus uptake was monitored by acquiring *z*-stacks of single cells decorated with not more than approx. 40 particles for a total time of about 20 min.

Figure A3. Data analysis via the TrIC software. Data was analyzed using the TrIC algorithm [30] in a home-written software in Matlab. The general workflow of the analysis and a screen shot of the data analysis program is shown. Starting in the top left, the program window shows: the instantaneous velocity, *x-y* projection of the trajectory, the z position as a function of time, frame overview, the image correlation analysis along the trajectory, a background-corrected fluorescence intensity of the two channels and images of the tracked and untracked channels and their overlay at the position of the cursor. Control parameters and functionalities in the program are given on the right.

Figure A4. TrIC analysis on a dual-color fluorescent bead. As a positive control, we measured the diffusion of a dual-color fluorescent bead in a glycerol solution. (**a**) The 3D-trajectory and 2D-projections of a tracked bead. (**b**) The corresponding TrIC analysis: (**i**) The background-corrected fluorescence intens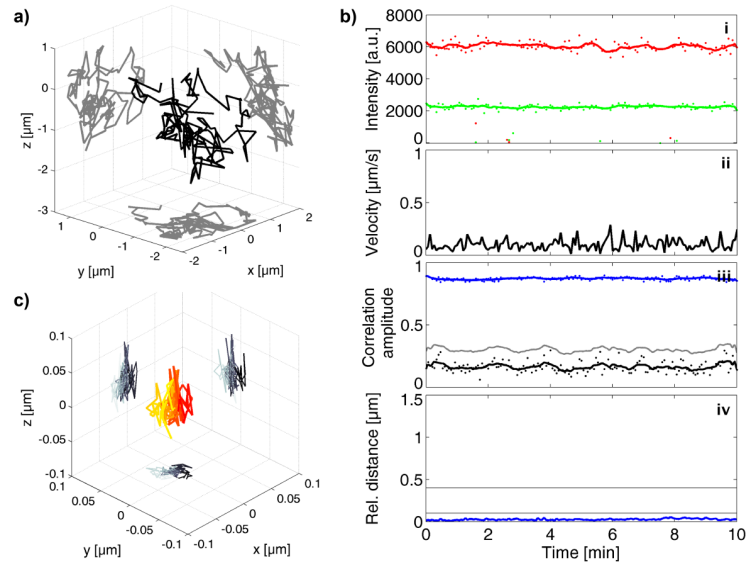ity of the green and the red signal are plotted as a function of time. (**ii**) The instantaneous velocity is plotted as a function of time. (**iii**) The amplitude of the cross-correlation of the TrIC analysis for the data (blue), the randomized control (black) and threshold (grey) are plotted as a function of time. (**iv**) The relative distance between the eGFP-capsid and mCherry-Env signal is plotted over time. (**c**). A 3D-representation of the relative position of the red signal with respect to the green signal color-coded as a function of time. The results show a high correlation amplitude and a relative distance of less than 100 nm over the whole trajectory.
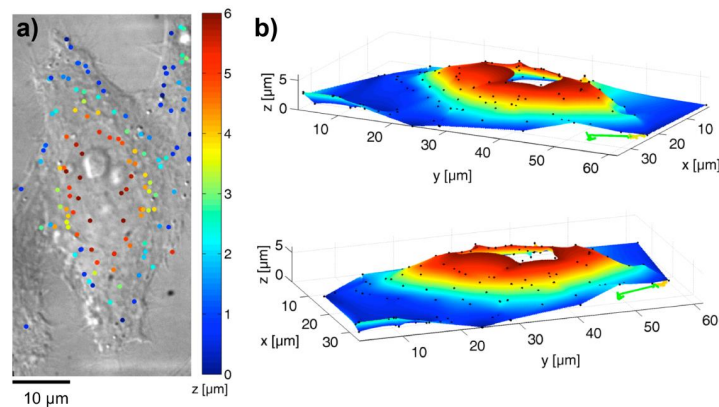
Figure A5. Estimation of the cell surface. The 3D-position of the virus particles was determined by 3D-Gaussian fitting on the first *z*-stack of the movie when the viruses are still bound to the plasma membrane of the cell. (**a**) An overlay of the virus particles' positions with the DIC image of the cell. The *z*-position is color-coded. (**b**) The cell surface was reconstructed by interpolation of the virus's positions and plotted as a 3D-surface together with the virus track (yellow and green after fusion).
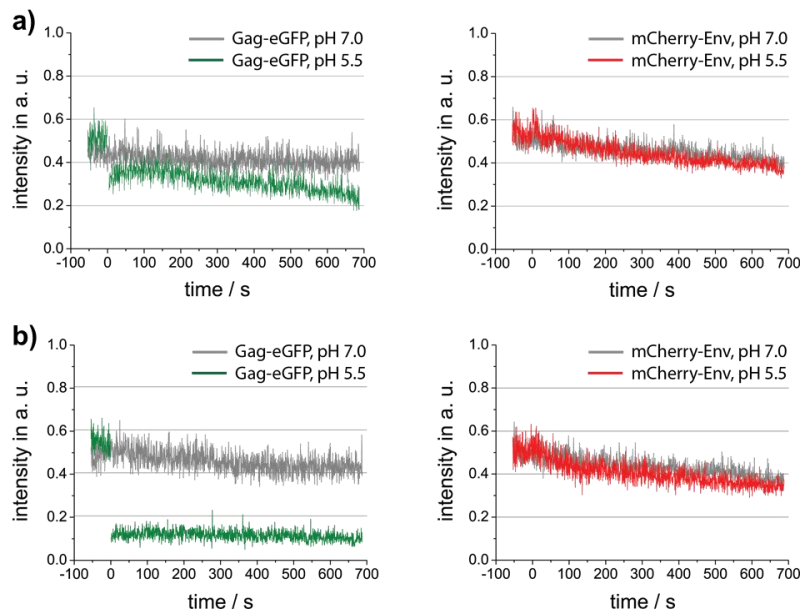
Figure A6. Stability of the fluorescent proteins on dual-color foamy virus particles at different pH values. Intensity traces of double-tagged PFV Env containing particles at pH 7.0 (gray) and pH 5.5 (green). At time point 0, citrate buffer was added to decrease the pH. The background-corrected intensities of ten particles were determined and the average is shown. The intensity of (**a**) intact particles and (**b**) control measurements on fixed and permeabilized particles measured on a coverslip are shwon. Note that the eGFP fluorescence showed a significant drop at pH 5.5.

References

1. Khan, A.S.; Bodem, J.; Buseyne, F.; Gessain, A.; Johnson, W.; Kuhn, J.H.; Kuzmak, J.; Lindemann, D.; Linial, M.L.; Lochelt, M.; et al. Spumaretroviruses: Updated taxonomy and nomenclature. *Virology* **2018**, *516*, 158–164. [CrossRef] [PubMed]

2. Rethwilm, A.; Lindemann, D. Foamy Viruses. In *Fields Virology*, 6th ed.; Knipe, D.M., Howley, P.M., Eds.; +20]Lippincott Williams & Wilkins, a Wolters Kluwer Business: Philadelphia, PA, USA, 2013; Volume 2, pp. 1613–1632.

3. Aiewsakun, P.; Katzourakis, A. Marine origin of retroviruses in the early Palaeozoic Era. *Nat. Commun.* **2017**, *8*, 13954. [CrossRef] [PubMed]

4. Herchenröder, O.; Renne, R.; Loncar, D.; Cobb, E.K.; Murthy, K.K.; Schneider, J.; Mergia, A.; Luciw, P.A. Isolation, cloning, and sequencing of simian foamy viruses from chimpanzees (SFVcpz): High homology to human foamy virus (HFV). *Virology* **1994**, *201*, 187–199. [CrossRef] [PubMed]

5. Achong, B.G.; Mansell, P.W.; Epstein, M.A.; Clifford, P. An unusual virus in cultures from a human nasopharyngeal carcinoma. *J. Natl. Cancer Inst.* **1971**, *46*, 299–307. [PubMed]

6. Duda, A.; Stange, A.; Lüftenegger, D.; Stanke, N.; Westphal, D.; Pietschmann, T.; Eastman, S.W.; Linial, M.L.; Rethwilm, A.; Lindemann, D. Prototype foamy virus envelope glycoprotein leader peptide processing is mediated by a furin-like cellular protease, but cleavage is not essential for viral infectivity. *J. Virol.* **2004**, *78*, 13865–13870. [CrossRef] [PubMed]

7. Lindemann, D.; Pietschmann, T.; Picard-Maureau, M.; Berg, A.; Heinkelein, M.; Thurow, J.; Knaus, P.; Zentgraf, H.; Rethwilm, A. A particle-associated glycoprotein signal peptide essential for virus maturation and infectivity. *J. Virol.* **2001**, *75*, 5762–5771. [CrossRef]

8.   Wilk, T.; Geiselhart, V.; Frech, M.; Fuller, S.D.; Flügel, R.M.; Löchelt, M. Specific interaction of a novel foamy virus env leader protein with the n-terminal gag domain. *J. Virol.* **2001**, *75*, 7995–8007. [CrossRef]

9.   Geiselhart, V.; Bastone, P.; Kempf, T.; Schnolzer, M.; Löchelt, M. Furin-mediated cleavage of the feline foamy virus Env leader protein. *J. Virol.* **2004**, *78*, 13573–13581. [CrossRef]

10. Effantin, G.; Estrozi, L.F.; Aschman, N.; Renesto, P.; Stanke, N.; Lindemann, D.; Schoehn, G.; Weissenhorn, W. Cryo-electron Microscopy Structure of the Native Prototype Foamy Virus Glycoprotein and Virus Architecture. *PLoS Pathog.* **2016**, *12*, e1005721. [CrossRef]

11. Wilk, T.; de Haas, F.; Wagner, A.; Rutten, T.; Fuller, S.; Flügel, R.M.; Löchelt, M. The intact retroviral Env glycoprotein of human foamy virus is a trimer. *J. Virol.* **2000**, *74*, 2885–2887. [CrossRef]

12. Pietschmann, T.; Heinkelein, M.; Heldmann, M.; Zentgraf, H.; Rethwilm, A.; Lindemann, D. Foamy virus capsids require the cognate envelope protein for particle export. *J. Virol.* **1999**, *73*, 2613–2621. [CrossRef] [PubMed]

13. Fischer, N.; Heinkelein, M.; Lindemann, D.; Enssle, J.; Baum, C.; Werder, E.; Zentgraf, H.; Müller, J.G.; Rethwilm, A. Foamy virus particle formation. *J. Virol.* **1998**, *72*, 1610–1615. [CrossRef] [PubMed]

14. Goldstone, D.C.; Flower, T.G.; Ball, N.J.; Sanz-Ramos, M.; Yap, M.W.; Ogrodowicz, R.W.; Stanke, N.; Reh, J.; Lindemann, D.; Stoye, J.P.; et al. A Unique Spumavirus Gag N-terminal Domain with Functional Properties of Orthoretroviral Matrix and Capsid. *PLoS Pathog.* **2013**, *9*, e1003376. [CrossRef] [PubMed]

15. Reh, J.; Stange, A.; Götz, A.; Rönitz, M.; Grosse, A.; Lindemann, D. An N-terminal domain helical motif of Prototype Foamy Virus Gag with dual functions essential for particle egress and viral infectivity. *Retrovirology* **2013**, *10*, 45. [CrossRef] [PubMed]

16. Shaw, K.L.; Lindemann, D.; Mulligan, M.J.; Goepfert, P.A. Foamy virus envelope glycoprotein is sufficient for particle budding and release. *J. Virol.* **2003**, *77*, 2338–2348. [CrossRef] [PubMed]

17. Stanke, N.; Stange, A.; Lüftenegger, D.; Zentgraf, H.; Lindemann, D. Ubiquitination of the Prototype Foamy Virus Envelope Glycoprotein Leader Peptide Regulates Subviral Particle Release. *J. Virol.* **2005**, *79*, 15074–15083. [CrossRef]

18. Stirnnagel, K.; Schupp, D.; Dupont, A.; Kudryavtsev, V.; Reh, J.; Müllers, E.; Lamb, D.C.; Lindemann, D. Differential pH-dependent cellular uptake pathways among foamy viruses elucidated using dual-colored fluorescent particles. *Retrovirology* **2012**, *9*, 71. [CrossRef]

19. Picard-Maureau, M.; Jarmy, G.; Berg, A.; Rethwilm, A.; Lindemann, D. Foamy virus envelope glycoprotein-mediated entry involves a pH-dependent fusion process. *J. Virol.* **2003**, *77*, 4722–4730. [CrossRef]

20. Liu, S.L.; Wang, Z.G.; Xie, H.Y.; Liu, A.A.; Lamb, D.C.; Pang, D.W. Single-Virus Tracking: From Imaging Methodologies to Virological Applications. *Chem. Rev.* **2020**, *120*, 1936–1979. [CrossRef]

21. Seisenberger, G.; Ried, M.U.; Endress, T.; Büning, H.; Hallek, M.; Bräuchle, C. Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. *Science* **2001**, *294*, 1929–1932. [CrossRef]

22. Floyd, D.L.; Ragains, J.R.; Skehel, J.J.; Harrison, S.C.; van Oijen, A.M. Single-particle kinetics of influenza virus membrane fusion. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 15382–15387. [CrossRef] [PubMed]

23. Rust, M.J.; Lakadamyali, M.; Zhang, F.; Zhuang, X. Assembly of endocytic machinery around individual influenza viruses during viral entry. *Nat. Struct. Mol. Biol.* **2004**, *11*, 567–573. [CrossRef] [PubMed]

24. Lakadamyali, M.; Rust, M.J.; Babcock, H.P.; Zhuang, X. Visualizing infection of individual influenza viruses. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9280–9285. [CrossRef] [PubMed]

25. Pelkmans, L.; Burli, T.; Zerial, M.; Helenius, A. Caveolin-stabilized membrane domains as multifunctional transport and sorting devices in endocytic membrane traffic. *Cell* **2004**, *118*, 767–780. [CrossRef]

26. Pelkmans, L.; Puntener, D.; Helenius, A. Local actin polymerization and dynamin recruitment in SV40-induced internalization of caveolae. *Science* **2002**, *296*, 535–539. [CrossRef]

27. Pelkmans, L.; Kartenbeck, J.; Helenius, A. Caveolar endocytosis of simian virus 40 reveals a new two-step vesicular-transport pathway to the ER. *Nat. Cell Biol.* **2001**, *3*, 473–483. [CrossRef]

28. Elphick, G.F.; Querbes, W.; Jordan, J.A.; Gee, G.V.; Eash, S.; Manley, K.; Dugan, A.; Stanifer, M.; Bhatnagar, A.; Kroeze, W.K.; et al. The human polyomavirus, JCV, uses serotonin receptors to infect cells. *Science* **2004**, *306*, 1380–1383. [CrossRef]

29. Pietiainen, V.; Marjomaki, V.; Upla, P.; Pelkmans, L.; Helenius, A.; Hyypia, T. Echovirus 1 endocytosis into caveosomes requires lipid rafts, dynamin II, and signaling events. *Mol. Biol. Cell* **2004**, *15*, 4911–4925. [CrossRef]

30. Dupont, A.; Stirnnagel, K.; Lindemann, D.; Lamb, D.C. Tracking image correlation: Combining single-particle tracking and image correlation. *Biophys. J.* **2013**, *104*, 2373–2382. [CrossRef]

31. Stirnnagel, K.; Lüftenegger, D.; Stange, A.; Swiersy, A.; Müllers, E.; Reh, J.; Stanke, N.; Grosse, A.; Chiantia, S.; Keller, H.; et al. Analysis of prototype foamy virus particle-host cell interaction with autofluorescent retroviral particles. *Retrovirology* **2010**, *7*, 45. [CrossRef]

32. Müllers, E.; Uhlig, T.; Stirnnagel, K.; Fiebig, U.; Zentgraf, H.; Lindemann, D. Novel functions of prototype foamy virus Gag glycine-arginine-rich boxes in reverse transcription and particle morphogenesis. *J. Virol.* **2011**, *85*, 1452–1463. [CrossRef] [PubMed]

33. Hamann, M.V.; Müllers, E.; Reh, J.; Stanke, N.; Effantin, G.; Weissenhorn, W.; Lindemann, D. The cooperative function of arginine residues in the Prototype Foamy Virus Gag C-terminus mediates viral and cellular RNA encapsidation. *Retrovirology* **2014**, *11*, 87. [CrossRef] [PubMed]

34. Tinevez, J.Y.; Perry, N.; Schindelin, J.; Hoopes, G.M.; Reynolds, G.D.; Laplantine, E.; Bednarek, S.Y.; Shorte, S.L.; Eliceiri, K.W. TrackMate: An open and extensible platform for single-particle tracking. *Methods* **2017**, *115*, 80–90. [CrossRef] [PubMed]

35. Henss, A.K.; Sakong, S.; Messer, P.K.; Wiechers, J.; Schuster, R.; Lamb, D.C.; Gross, A.; Wintterlin, J. Density fluctuations as door-opener for diffusion on crowded surfaces. *Science* **2019**, *363*, 715–718. [CrossRef]

36. Patterson, G.H.; Knobel, S.M.; Sharif, W.D.; Kain, S.R.; Piston, D.W. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys. J.* **1997**, *73*, 2782–2790. [CrossRef]

37. Verdeny-Vilanova, I.; Wehnekamp, F.; Mohan, N.; Sandoval Alvarez, A.; Borbely, J.S.; Otterstrom, J.J.; Lamb, D.C.; Lakadamyali, M. 3D motion of vesicles along microtubules helps them to circumvent obstacles in cells. *J. Cell Sci.* **2017**, *130*, 1904–1916. [CrossRef]

38. Lakadamyali, M.; Rust, M.J.; Zhuang, X. Ligands for clathrin-mediated endocytosis are differentially sorted into distinct populations of early endosomes. *Cell* **2006**, *124*, 997–1009. [CrossRef]

39. Suomalainen, M.; Nakano, M.Y.; Keller, S.; Boucke, K.; Stidwill, R.P.; Greber, U.F. Microtubule-dependent plus- and minus end-directed motilities are competing processes for nuclear targeting of adenovirus. *J. Cell Biol.* **1999**, *144*, 657–672. [CrossRef]

40. Hsieh, M.J.; White, P.J.; Pouton, C.W. Interaction of viruses with host cell molecular motors. *Curr. Opin. Biotechnol.* **2010**, *21*, 633–639. [CrossRef]

41. Schelhaas, M.; Ewers, H.; Rajamaki, M.L.; Day, P.M.; Schiller, J.T.; Helenius, A. Human papillomavirus type 16 entry: Retrograde cell surface transport along actin-rich protrusions. *PLoS Pathog.* **2008**, *4*, e1000148. [CrossRef]

42. Bausinger, R.; von Gersdorff, K.; Braeckmans, K.; Ogris, M.; Wagner, E.; Brauchle, C.; Zumbusch, A. The transport of nanosized gene carriers unraveled by live-cell imaging. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 1568–1572. [CrossRef] [PubMed]

43. Lehmann, M.J.; Sherer, N.M.; Marks, C.B.; Pypaert, M.; Mothes, W. Actin- and myosin-driven movement of viruses along filopodia precedes their entry into cells. *J. Cell Biol.* **2005**, *170*, 317–325. [CrossRef] [PubMed]

44. Mercer, J.; Schelhaas, M.; Helenius, A. Virus entry by endocytosis. *Annu. Rev. Biochem.* **2010**, *79*, 803–833. [CrossRef] [PubMed]

45. Hubbard, A.L. Endocytosis. *Curr. Opin. Cell Biol.* **1989**, *1*, 675–683. [CrossRef]

46. Mellman, I.; Fuchs, R.; Helenius, A. Acidification of the endocytic and exocytic pathways. *Annu. Rev. Biochem.* **1986**, *55*, 663–700. [CrossRef]

47. Daecke, J.; Fackler, O.T.; Dittmar, M.T.; Kräusslich, H.G. Involvement of clathrin-mediated endocytosis in human immunodeficiency virus type 1 entry. *J. Virol.* **2005**, *79*, 1581–1594. [CrossRef]

48. Johannsdottir, H.K.; Mancini, R.; Kartenbeck, J.; Amato, L.; Helenius, A. Host cell factors and functions involved in vesicular stomatitis virus entry. *J. Virol.* **2009**, *83*, 440–453. [CrossRef]

49. Ruigrok, R.W.; Hewat, E.A.; Wade, R.H. Low pH deforms the influenza virus envelope. *J. Gen. Virol.* **1992**, *73*, 995–998. [CrossRef]

50. Koch, P.; Lampe, M.; Godinez, W.J.; Müller, B.; Rohr, K.; Kräusslich, H.G.; Lehmann, M.J. Visualizing fusion of pseudotyped HIV-1 particles in real time by live cell microscopy. *Retrovirology* **2009**, *6*, 84. [CrossRef]

51. Barrett, C.T.; Dutch, R.E. Viral Membrane Fusion and the Transmembrane Domain. *Viruses* **2020**, *12*, 693. [CrossRef]

52. Harrison, S.C. Viral membrane fusion. *Virology* **2015**, *479–480*, 498–507. [CrossRef] [PubMed]

B.2   Paper 2: Nanoscale organization of the endogenous ASC speck

Nanoscale organization of the endogenous ASC speck

Ivo M. Glück[1,2], Grusha Primal Mathias[1,3], Sebastian Strauss[4,5], Thomas S. Ebert[6], Che Stafford[6], Ganesh Agam[1,2], Suliana Manley[8], Veit Hornung[5,6], Ralf Jungmann[4,5], Christian Sieben[7,§,*], Don C. Lamb[1,2,*]

[1] Department of Chemistry, Ludwig Maximilians-Universität München, Butenandtstraße 5-13, 81377 München, Germany

[2] Center for Nano Science (CENS), Ludwig Maximilians-Universität München, Butenandtstraße 5-13, 81377 München, Germany

[3] Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Bonn, Germany und Universität Köln, Köln, Germany

[4] Faculty of Physics and Center for Nanoscience, Ludwig Maximilian University, Munich, Germany

[5] Max Planck Institute of Biochemistry, Martinsried, Germany

[6] Gene Center and Department of Biochemistry, Ludwig-Maximilians-Universität, Munich, Germany

[7] Laboratory of Experimental Biophysics, École Polytechnique Fédérale de Lausanne, BSP 427 (Cubotron UNIL), Rte de la Sorge, CH-1015 Lausanne, Switzerland

[§] Current address: Nanoscale Infection Biology, Helmholtz Centre for Infection Research, Inhoffenstr. 7, 38124 Braunschweig, Germany

* Correspondence should be addressed to

Christian Sieben, christian.sieben@helmholtz-hzi.de

or

Don C. Lamb, d.lamb@lmu.de

Abstract

The NLRP3 inflammasome is a central component of the innate immune system. Its activation leads to the formation of a supramolecular assembly of the inflammasome adaptor protein ASC, commonly referred to as the "ASC speck". Different models of the overall structure of the ASC speck, as well as the entire NLRP3 inflammasome, have been reported in the literature. While many experiments involve overexpression or *in vitro* reconstitution of recombinant ASC, the cytoplasmic endogenous ASC speck remains difficult to study due to its relatively small size and structural variability.

Here, we use a combination of fluorescence imaging techniques including dual-color 3D super-resolution imaging (dSTORM and DNA-PAINT) to visualize the endogenous ASC speck following NLRP3 inflammasome activation. We observe that the complex varies in diameter between ~800 and 1000 nm and is composed of a dense core from which filaments extend into the periphery. We used a combination of anti-ASC antibodies as well as much smaller nanobodies for labeling, and find that the larger complexes reliably label the lower-density periphery whereas the nanobody, which has a lower binding affinity, is more efficient in labeling the dense core. Imaging whole cells using dSTORM, furthermore, allowed us to sort the imaged structures into a pseudo-time sequence suggesting that the endogenous ASC speck becomes denser but not much larger during its formation. The reported results add an important piece of information towards a comprehensive understanding of the supramolecular structure of the endogenous inflammasome complex.

Introduction

Inflammasomes are a class of large, multiprotein complexes that assemble upon activation of cellular pattern recognition receptors (PRRs) [1]. Being part of the innate immune system, inflammasomes can sense the presence of non-self biomolecules or the perturbation of cellular homeostasis. The key components include a sensor protein, the adaptor protein Apoptosis-associated speck-like protein containing a Caspase activation and recruitment domain (ASC) [2, 3] and the inflammatory caspase 1 (CASP1) [4-6]. The largest group of sensor proteins belongs to the family of nucleotide-binding oligomerization domain-like receptors (NLRs). Among these, NLRP3 (NOD-, LRR- and pyrin domain-containing protein 3) has been shown to play a critical role in many infectious, as well as sterile inflammatory conditions. Although the molecular mode of action remains to be elucidated, potassium efflux, as it occurs in the context of membrane damage, appears to be a key signal upstream of NLRP3 activation [7]. NLRP3 activation triggers recruitment of ASC and CASP1 leading to a single micrometer-sized assembly. For the structure of ASC within the inflammasome, Masumoto *et al.* coined the term "ASC speck" [3, 8]. ASC is composed of two interaction domains connected by a semi-flexible linker: a Pyrin domain (PYD) [9] and a Caspase activation and recruitment domain (CARD) [10, 11]. The individual domains have a high tendency for homotypic interactions due to their complementarity in structure and charge, leading to the assembly of the large multiprotein inflammasome complex [12-15].

The details of how the ASC speck is organized is the subject of intensive research. ASC (22 kDa) is soluble at low pH and in a chaotropic solution, but at physiological pH, the protein can assemble into filaments as observed *in vitro* by solid-state NMR spectroscopy and electron microscopy (EM) [16-21]. The ASC speck can also be purified from inflammasome-activated cells expressing ASC endogenously, where it was found to take on diverse morphologies, including a star-shaped assembly [8], isolated filaments, or amorphous aggregates potentially composed of intertwined filaments [19, 22]. In cells, ASC specks were visualized by microscopy using immunofluorescence or expression of ASC tagged with a fluorescent protein. Upon overexpression, the resulting speck was typically much larger than 1 μm and occasionally showed filaments protruding from the edge of the structure [20, 22-25]. In contrast, diffraction-limited immunofluorescence imaging of the endogenous ASC speck revealed a spot of about 1 μm in diameter. Due to the propensity of ASC to self-assemble, it is unclear whether the endogenous structure resembles the complexes observed *in vitro* or upon ASC overexpression. Interestingly, one approach to labeling ASC using a EGFP-labeled nanobody directed against the CARD domain revealed an intermediate, filamentous structure during speck formation, but no ASC specks were observed [26]. Higher resolution *in situ* studies, either by EM [27, 28] or super-resolution fluorescence STED (stimulated emission depletion) microscopy [29] resolved the endogenous speck as an amorphous aggregate potentially made up of intertwined filaments. In contrast, other studies describe the ASC speck as a hollow, ring-shaped complex [3, 6, 23, 30-38], and based on this observation, different models for ASC speck and inflammasome formation have been proposed [34, 35, 39-42]. Hence, despite its relevance for understanding inflammasome formation, the nanoscale organization of the endogenous ASC speck remains controversial.

Here, we performed a systematic study of fluorescence labelling strategies, and used quantitative widefield microscopy, confocal microscopy and single-molecule localization microscopy (i.e. 3D dual-color dSTORM and 3D DNA-PAINT) to investigate the organization of the endogenous ASC speck. Our data resolved filaments protruding from a dense core of the endogenous ASC speck, supporting the amorphous nature of the complex. By using two complementary labelling approaches comparing antibody- with nanobody-labelled ASC, we could probe the organization of and density differences within the ASC speck. We found that nanobodies labelled the center of the speck while the antibody was predominantly detected in the periphery of the complex, occasionally exhibiting a hollow center. Thus, our results reconcile the disparate structures reported in the literature.

Furthermore, we analyzed the redistribution of ASC into the speck using single-cell analysis, allowing us to sort specks with respect to the degree of ASC recruitment. Our results indicate that endogenous specks mainly become denser but only slightly larger during inflammasome assembly.

Results

Endogenous specks vary strongly in ASC content

We investigated the organization of the endogenous ASC speck in THP-1 cells. The cells were primed using lipopolysaccharide (LPS) [43] followed by stimulation with the bacterial, potassium-efflux-inducing ionophore nigericin [44]. After 90 minutes, the cells were fixed using paraformaldehyde and stained with a primary monoclonal antibody against ASC in combination with a secondary Alexa Fluor 647-tagged F(ab')$_2$ fragment. To gain an initial insight into the distribution of ASC in the cell, we imaged the cells at low magnification (10x) using confocal microscopy. In unstimulated cells, ASC was distributed throughout the cell, including the nucleus (**Figure 1A**, left panel, **Supplementary Figure S1A**). Upon NLRP3 inflammasome activation, ASC relocated into the characteristic single perinuclear speck in about 30% of the cells (**Figure 1A**, right panel) [45]. We could also observe extracellular specks as previously reported [22, 27] (**Supplementary Figure S1B**). Next, we used high magnification (60x) widefield microscopy to image 59 individual ASC specks, which appear as spherical complexes ranging in size between approximately 0.5 - 1 μm diameter (**Figure 1B**). Analysis of the integrated intensity revealed that the amount of incorporated ASC as reflected by antibody binding can vary by almost one order of magnitude (**Figure 1C**). The widefield microscopy images also confirmed that ASC is distributed throughout unstimulated cells (**Figure 2A**, left panel), and that, in cells showing a speck, the cytoplasmic ASC was almost completely redistributed into a single, bright ASC spot (**Figure 2A**, right panel). These observations are consistent with previous reports of the ASC speck upon activation of the NLRP3 inflammasome [46].
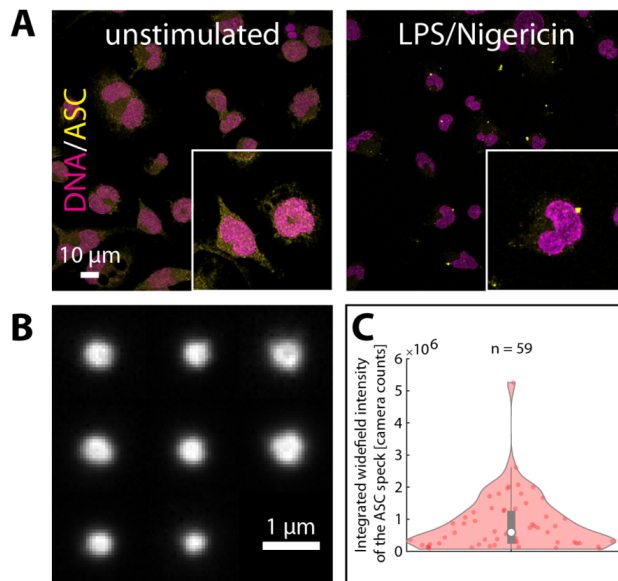
Figure 1 - LPS + Nigericin stimulation of THP-1 cells induces ASC speck formation with variable size and ASC content. A) Confocal images (maximum projections) of fluorescently labeled ASC (shown in yellow) in unstimulated (left panel) and in LPS + Nigericin-stimulated THP-1 cells (right panel). The formation of ASC specks after stimulation is clearly visible. DNA was stained using DAPI (shown in magenta). B) A montage of ASC specks immunostained using a primary antibody and a secondary Alexa Fluor 647-conjugated F(ab')$_2$ fragment and imaged by diffraction-limited widefield illumination with 60x magnification. The speck size varies and the appearance of a darker center is observed for some of the specks. C) A violin plot of the integrated widefield intensity of ASC specks imaged as described for panel B. There is a large variation in the intensity, which serves as an indicator for the amount of protein incorporated into the speck. Pooled data from three independent cell preparations is shown.

<u>Non-speck incorporated ASC clusters maintain a similar size distribution during ASC speck formation</u>

To study the distribution of ASC at the nanoscale, we used direct stochastic optical reconstruction microscopy (dSTORM) [47]. Reconstructed dSTORM images further confirmed the cytoplasmic ASC distribution in unstimulated cells and its redistribution into a single speck after inflammasome activation. Interestingly, the increased detection sensitivity of dSTORM allowed us to visualize a previously undetected non-speck-bound ASC population in the cytoplasm of activated cells (**Figure 2B**). We performed a clustering analysis based on the local density of molecules, which distinguishes cytoplasmic ASC clusters from background localizations (**Supplementary Figure S2 and Supplementary Figure S3**). We used a Density-Based Spatial Clustering of Applications with Noise (DBSCAN) analysis [48] (**Figure 2C**), which revealed a decrease in the cluster density from an average of 0.40 clusters per μm^2 in unstimulated cells to 0.18 clusters per μm^2 in cells showing a speck (**Figure 2D**). This is consistent with the observed ASC recruitment (**Figure 2A**). Next, from the detected clusters, we investigated the number of localizations per cluster and the cluster size by calculating the radius of gyration (R_g). We found the size of non-speck incorporated clusters to be very similar between unstimulated and stimulated cells ($R_g < 20$ nm and localizations/cluster < 60) (**Figure 2E, F**). However, in speck-containing cells, an additional cytoplasmic cluster population appears with a larger size and more localizations (R_g 20-80 nm and localizations/cluster > 60). We also applied a Ripley's K clustering analysis [49] as an alternative approach for characterizing the spatial distribution of non-speck incorporated ASC. This analysis (**Figure 2G**) confirmed the decrease in cytoplasmic cluster density in cells containing a speck as shown by the shift of the amplitude of the obtained L(r)-r curve towards higher values and an increase of the L(r)-r maximum value consistent with a small population of larger clusters.
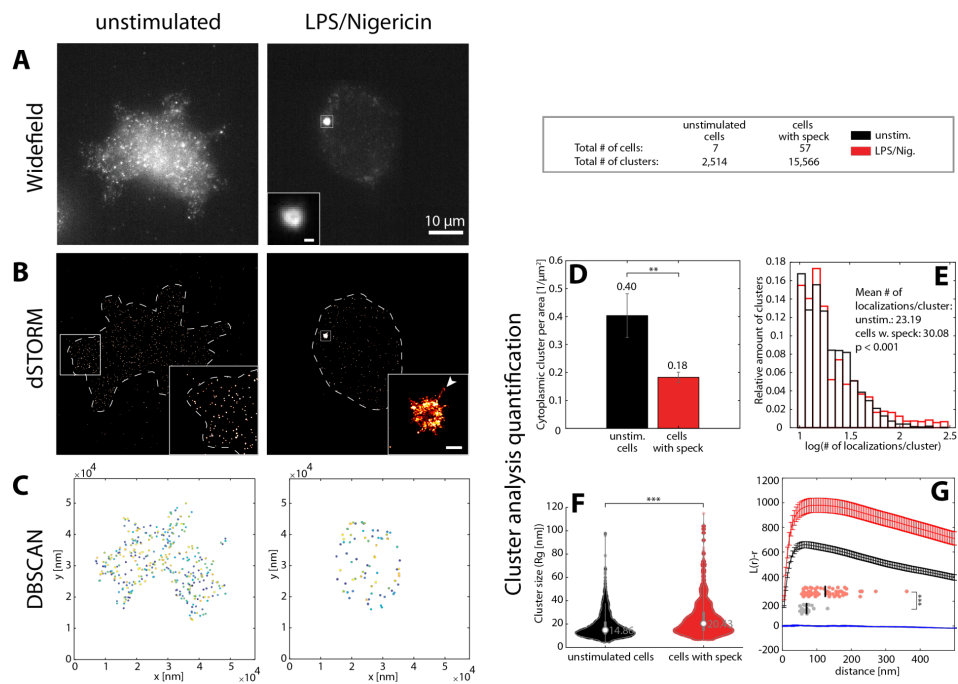
Figure 2 – Whole-cell dSTORM super-resolution microscopy of endogenous ASC in THP-1 cells. A, B) The distribution of ASC in non-stimulated cells (left panels) and the redistribution of ASC into the speck after stimulation with LPS and Nigericin (right panels) observed using diffraction-limited widefield imaging (A) and dSTORM (B). dSTORM additionally resolves filaments protruding from the speck core (one example is highlighted by the arrow). Scale bars in insets correspond to 500 nm. C) DBSCAN clustering of (non-speck bound) cytosolic ASC localizations confirms depletion of the protein from the cytosol after speck formation. Color coding is used to distinguish individual clusters. D – G) The results of the DBSCAN analysis are shown. D) The non-speck incorporated ASC cluster density decreases in cells exhibiting a speck due to recruitment of the protein into the speck. The error bars indicate the standard error of the mean calculated from the analysis results of individual cells. E), F) The distribution of the number of localizations (E) and violin plots of the radius of gyration (R_g) (F) for unstimulated cells and cells forming a speck. The distributions are similar with an additional population observable in speck-producing cells with more localizations and a larger size compared to non-stimulated cells. White circles in the violin plots indicate the median of the distribution. The statistical significance in D-F was assessed by a Two-Sample Kolmogorov-Smirnov Test (*) $p<0.05$, (**) $p<0.01$, (***) $p<0.001$. G) A Ripley's K analysis of the ASC clusters. The Ripley's K function confirms the decrease in cluster density after speck formation as well as the increase in cluster size. The curves show the mean $\pm$ SD of 100 ROI from 40 cells imaged from three independent experiments. For the estimation of the curve's maximum (G, lower part), we measured from the individual ROIs). The same data was taken and the positions of the individual localizations were randomized at the same density and analyzed for comparison. The statistical significance was assessed by one-sided t-test (**) $p<0.005$, (***) $p<0.001$. Data on unstimulated cells was obtained on a single experiment. Data on cells with speck was obtained on three independent cell preparations.

Table 1 - Summary of measurement parameters on unstimulated cells and LPS + Nigericin-treated cells containing an ASC speck *All values occurred just once; MAD: Median absolute deviation of the median; AF 647: Alexa Fluor 647

		unstimulated cells	LPS + Nigericin-treated, Speck-containing cells				
		Primary antibody + secondary AF 647-conjugated F(ab')2 fragment	Primary antibody + secondary AF 647-conjugated F(ab')2 fragment w + w/o additional DL755-conjugated nanobody	Primary antibody + secondary AF 647-conjugated F(ab')2 fragment only	Primary antibody + secondary AF 647-conjugated F(ab')2 fragment w additional DL755-conjugated nanobody	AF 647-conjugated nanobody only	DL 755-conjugated nanobody w additional primary antibody + secondary AF 647-conjugated F(ab')2 fragment staining
# of cells or specks		7	57 (cluster analysis) 60 (speck analysis)	21	39	92	42
# of clusters		2,514	15,566	-	-	-	-
Mean cytoplasmic clusters/area [1/μm^2]		0.18	0.40	-	-	-	-
Mean # of localizations per cluster		23.19	30.08	-	-	-	-
Median cluster size (2D-R_g) [nm]		14.86	20.43	-	-	-	-
Mode cluster size (2D-R_g) [nm]		6.04	7.05	-	-	-	-
xy Eccentricity of speck	Mean	-	1.23	1.14	1.29	1.33	1.5
	Median	-	1.19	1.13	1.24	1.23	1.4
	Mode	-	1.02	1.02	1.04	1.07	1.03
	MAD	-	0.08	0.04	0.07	0.11	0.19
2D-R_g of speck [nm]	Mean	-	325.2	332.6	321.2	298.1	158.3
	Median	-	321.3	324.0	319.7	281.9	133.9
	Mode	-	80.4	80.4	242.1	310.6	19.9
	MAD	-	30.9	35.9	29.0	36.5	49.9
Number of localizations per speck	Mean	-	5010	7209	3826	5338	3848
	Median	-	4113	7180	3307	2749	2787
	Mode	-	N/A*	1062	N/A*	1297	345
	MAD	-	1430	2691	954	1872	2061

<u>Super-resolution microscopy reveals distinct morphologies of the endogenous ASC speck</u>

Next, we investigated the nanoscale organization of the ASC speck itself [50, 51]. We labelled endogenous ASC using a primary antibody in combination with a secondary F(ab')$_2$ fragment. A large proportion of the specks appeared as round, amorphous structures with a diameter of about 1 μm exhibiting a rough surface with short protrusions (**Supplementary Figure S4A**). Interestingly, for a large number of the specks, the higher resolution obtained by dSTORM imaging resolved ASC filaments reaching out from the dense core of the speck (**Figure 3A**). The number of clearly resolved filaments per structure varied but we rarely observed more than ~10. We measured the diameter of the filaments using a Gaussian fitting of the intensity profile of multiple cross sections along each filament. The filament diameter was derived from the full width half maximum of the intensity profile, where we found a median value of 37.1 nm (**Figure 3B**). Considering the size of a primary/secondary F(ab')$_2$ antibody complex (~11 nm) [52] used for labeling and the experimental localization precision of ~ 10 nm (see Materials and Methods), the measured value corresponds to an actual thickness of the filament of ~15 nm. When compared to values obtained for filaments formed by ASC *in vitro* and studied by EM (16 nm) [17], this value indicates that the majority of filaments are isolated single filaments.

To validate these results, we also imaged the endogenous ASC specks using DNA-PAINT super-resolution microscopy [53, 54]. Here, we used BlaER1 cells, transdifferentiated into monocytes/macrophages in which ASC was endogenously tagged with TagRFP [55-57]. This alternative system again revealed the morphology of the endogenous ASC speck including filaments and a dense core. Strikingly, the filaments were much longer than the ones observed by dSTORM in THP-1 cells (**Figure 3C** and **Supplementary Figure S4B**).

Next, we performed 3D dSTORM imaging of the endogenous ASC specks. Consistent with the 2D dSTORM images, the specks had an overall spherical shape with filaments occasionally reaching out from a dense core (**Figure 3D**, **Supplementary Figure S5**, **Supplementary Figure S6, Supplementary Movie S1** and **Supplementary Movie S2**). The two-dimensional eccentricity for the ASC specks was determined to be 1.19, corresponding to more or less round specks as observed by eye, with a median radius of gyration of 321 ± 30.9 nm Median absolute deviation of the median (MAD) (**Figure 3E**, **F**, and Material and Methods). For some ASC specks, we observed a ring-like appearance (**Figure 3H - L** and **Supplementary Figure S4A + Supplementary Figure S6**). Some of the structures appeared less extended along the z-axis resulting in an overall disk-like shape (**Figure 3I + Supplementary Movie S3 + Supplementary Movie S4**) while others had a spherical, hollow shape (**Figure 3K + L, Supplementary Movie S5**). Although some specks appeared ring-like, the signal was not entirely excluded from the center. We hypothesized that local density differences may be responsible for the ring-like signal, perhaps due to the steric exclusion of primary and secondary antibody complexes from the dense center of the structure.

To test this hypothesis, we performed 3D super-resolution imaging using a 3-fold smaller ASC nanobody [26] to stain endogenous specks. Following nanobody labelling and dSTORM imaging, the specks appeared as amorphous structures with no obvious organization (**Figure 3M – P, Supplementary Figure S7** and **Supplementary Figure S8**). The structure remained spherical (**Figure 3E**) while the overall size of the specks, determined from the radius of gyration, was smaller (**Figure 3F**) and the total number of localizations within the speck decreased compared to specks stained with primary antibody and secondary F(ab')$_2$ fragment (**Figure 3G**). None of the 134 structures observed exhibited a hollow center. DNA-PAINT confirmed our observation of a smaller speck size and also resolved short filamentous extensions at the edge of the structure (**Figure 3O + P**, **Supplementary Figure S7B**, and **Supplementary Movie S6**). The absence of ring-like structures after nanobody labelling is consistent with our hypothesis that the dense regions within the speck limit the

accessibility of the labeling probe. The smaller radius of gyration cannot be entirely explained by the high labeling density in the center, and we attribute it to the comparably low binding-affinity of the nanobody (apparent binding constant: 159.5 ± 1.5 nM [26]). The nanobody has a single binding site, compared to two for a normal primary antibody, potentially explaining the higher labeling efficiency of the outer, lower-density regions of the speck by the antibody.
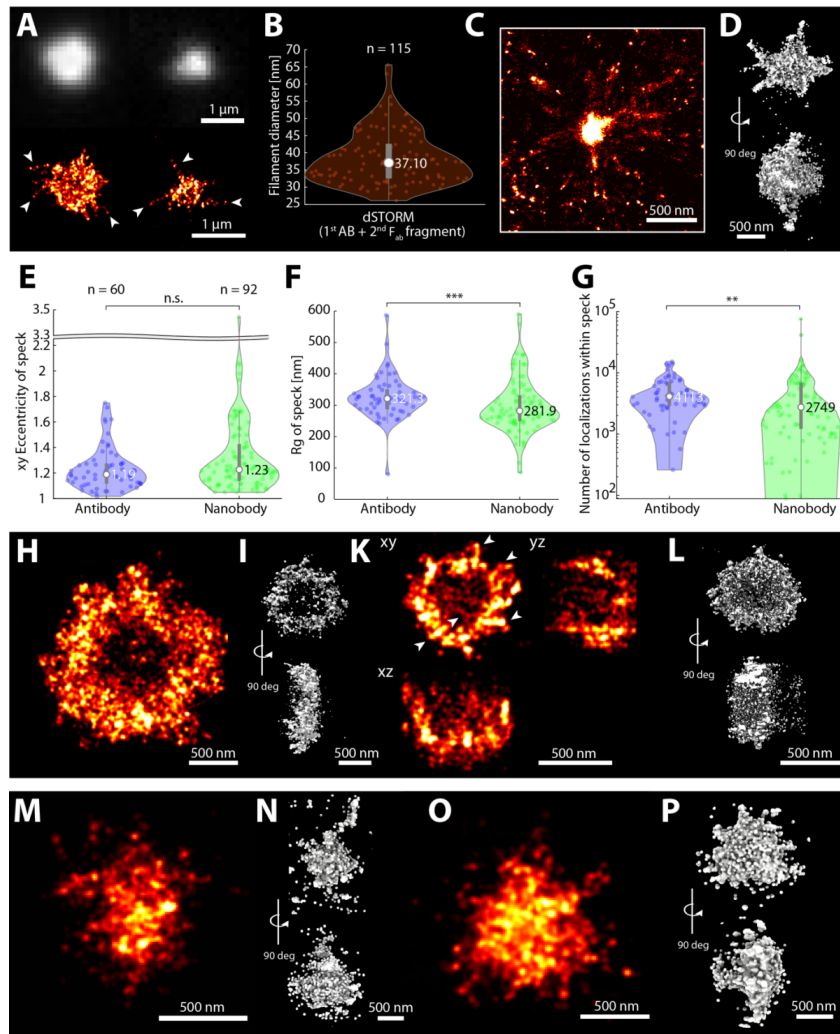
Figure 3 – Super-resolution imaging of the endogenous ASC speck. A) Two ASC specks stained with primary antibody and secondary F(ab')₂ fragment in THP-1 cells exhibiting different sizes are shown after diffraction-limited widefield imaging (upper images) together with their 2D dSTORM reconstructions (lower images). B) The distribution of ASC filament diameters as measured in 2D dSTORM reconstructions on multiple cross-sections per filament. The circle in the violin plot indicates the median of the distribution. Data was obtained on three independent cell preparations and 18 individual specks. C) An ASC speck in BlaER1 cells imaged using DNA-PAINT super-resolution microscopy. Filaments are clearly observed reaching out from a dense core. D) A 3D dSTORM reconstruction of a single ASC speck stained with primary antibody and secondary F(ab')₂ fragment in THP-1 cells. E–G) Comparison of the endogenous ASC speck parameters determined using primary antibody and secondary F(ab')₂ fragment labeling (blue) and nanobody labeling (green). From the 2D dSTORM images, violin plots of the eccentricity (E), the radius of gyration (F) and the number of localizations (G) were determined. White circles in the violin plots indicate the median of the distribution. Statistical significance was assessed by a two-sided Two-Sample Kolmogorov-Smirnov Test (*) p<0.05, (**) p<0.01, (***) p<0.001. Data on antibody- and nanobody-stained specks were obtained from three and two independent cell preparations, respectively. H-L). A characteristic 2D reconstruction (H) and a 3D reconstruction of an ASC speck (I) as well as a z-stack projection of a 3D dSTORM reconstruction (K) and a 3D reconstruction of another ASC speck (L). Both exhibit a ring-like structure. In panel K, the arrows highlight filament-like structures. M-P) ASC specks stained with anti-ASC nanobodies. 2D (M) and 3D (N) reconstructions of two individual ASC specks imaged by dSTORM. O-P) 2D (O) and 3D (P) representations of an ASC speck imaged using DNA-PAINT. Short filamentous structures are resolvable at the edge of the dense speck core.

<u>Two-color super-resolution imaging confirms accessibility differences within endogenous ASC specks</u>

To examine whether both the high-density core structure and low-density filaments are present on the same ASC speck, we performed two-color super-resolution microscopy using both nanobody and antibody labeling. Diffraction-limited widefield imaging showed that both labels specifically stained the ASC speck (**Figure 4A**). Consistent with our previous observation, the signal resulting from nanobody staining was smaller compared to the one obtained from antibody staining. Dual-color dSTORM reconstructions revealed that the antibody-labeled specks have a diameter of about one micrometer in widefield and an R_g of 319.7 ± 29 nm MAD (compared to 324 ± 36 nm MAD under single-labeling conditions, **Table 1**) with the nanobody staining localized in the center of the speck (**Figure 4B** and **Supplementary Figure S9**). Antibody labeling also resolved filaments and, as observed in the single-color antibody staining, a subset of the ASC specks appeared hollow (**Figure 4C**). Strikingly, in these particles, the nanobody staining was more compact, with a significantly smaller R_g (134 ± 50 nm MAD, compared to 282 ± 36.5 nm MAD under single labeling conditions, **Table 1**). We speculate that the higher-affinity antibody outcompetes the nanobody in the lower density regions of the speck, while only the nanobody is small enough to penetrate the dense core of the inflammasome. Aligning the dual-color structures along the center of mass of the antibody signals confirmed the observation that nanobody staining is confined to the center while the antibody complex was found more towards the speck periphery (**Figure 4D**). Hence, different labeling approaches bring out different features of the ASC speck. While it is typically beneficial to use the smallest available labels [58, 59], other factors such as binding affinity and density of the target structure can also play a role.
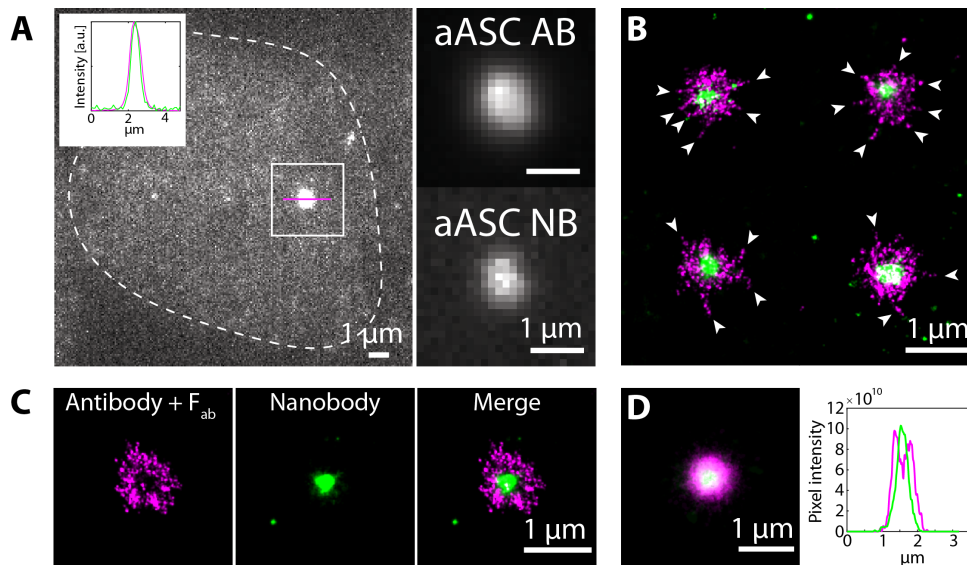
Figure 4 - Dual-color dSTORM imaging of ASC specks simultaneously stained with primary antibodies/secondary F(ab')₂ fragments and nanobodies against ASC. A) Diffraction-limited widefield imaging of an ASC speck stained with primary antibody and Alexa Fluor 647-conjugated secondary F(ab')₂ fragment (the signal of which is shown here) and DyLight 755-conjugated nanobody. The dashed line shows the cell outline. The intensity profile of the ASC speck normalized to the peak of the intensity distribution along a cross-section of the speck (magenta line) is shown in the top left of the image (antibody signal: magenta; nanobody signal: green) showing that both staining approaches stain the same structure. The right part of the panel shows the boxed area split into the signal after antibody (top) and nanobody (bottom) staining. B) Four representative dual-color dSTORM reconstructions of ASC specks stained simultaneously with both antibodies and nanobodies (magenta: Alexa Fluor 647; green: DyLight 755). Arrows point towards filaments reaching out from the dense speck core. C) Dual-color dSTORM reconstruction of a speck showing a ring-like appearance after antibody + F(ab')₂-staining (left) labeling, staining of the dense core by the nanobody (center) and the merge of both labeling strategies (right). D) Alignment of dual-color-labeled specks (n = 35) along the center of mass of the antibody + F(ab')₂ fragment staining (left) and the intensity profile along a cross-section through the aligned structure (right). Data was obtained on a single cell preparation.

ASC forms a scaffold that increases in density with time

From the wealth of information that we gathered using widefield and super-resolution microscopy, we further quantitatively analyzed our data. Since the labeling of ASC with nanobodies was less robust, we limited our analysis to the data collected using primary antibody and secondary F(ab')₂ fragment labeling. We first manually segmented the cell and the speck so that we could calculate parameters dependent on the characteristics of the cells and the specks they contained. Widefield images were collected before super-resolution microscopy was performed. Hence, the information from both imaging modalities was available from the same cells. By plotting the total widefield intensity as a function of cell area, a clear correlation was observed (**Supplementary Figure S10A**) showing that larger cells express more ASC protein. Similarly, there is a positive correlation between the total number of localizations and cell size (**Supplementary Figure S10B**). In fact, the total number of localizations measured using dSTORM correlates well with the total widefield intensity, as one would expect (**Supplementary Figure S10C + D**). Small corrections for day-to-day variations were performed as discussed in the material and methods (**Materials and Methods** and **Supplementary Figure S11**). In contrast, the intensity normalized by the cell area is relatively constant

(**Supplementary Figure S10E**), suggesting that the concentration of ASC is constant across different cells.

Next, we investigated how the size of the speck varies with cell properties. The size of the speck, determined either by manual segmentation of the speck or via the calculation of the 2D radius of gyration, was found to increase with cell size (**Supplementary Figure S10F, Figure 5A**). To investigate how the radius of gyration depends on other parameters, we normalized out the cell-area dependence (**Supplementary Figure S10K** see **Materials and Methods**). Interestingly, the radius of gyration only weakly depends on the amount of ASC within the speck (**Supplementary Figure S10L, M**). The same trend was found when we quantified the speck size manually via the occupied area **(Supplementary Figure S10G - I)**.

Speck assembly is a dynamic process, and recruitment of ASC to the perinuclear speck upon NLRP3 inflammasome activation occurs stochastically in the different cells we measured [18, 46]. Thus, when fixed, different cells represent different stages of the assembly process. We used the ratio of ASC signal in the speck with respect to the total amount of ASC within the entire cell as a metric for the progression of speck assembly, i.e. pseudo time. For the ASC signal, we utilized the number of localizations as this is more reliable. We observed that the widefield intensity of the speck, as well as the number of localizations within the speck increased with ASC recruitment along the pseudo time axis, as expected (**Supplementary Figure S10N - O**). In addition, the cytosolic, non-speck ASC signal consistently decreased with the pseudo time regardless of whether we quantified it using the number of localizations, the cluster density or the localization density (**Supplementary Figure S10R - T**). A clear decrease in the integrated widefield intensity with the pseudo time is not observed (**Supplementary Figure S10Q**), we attribute this to the low level of signal remaining in the cytosol upon ASC recruitment to the speck, which suggests that the autofluorescence signal a significant fraction of the entire signal. A plot of the radius of gyration versus pseudo time showed little change in the size of the speck with pseudo time (**Supplementary Figure 10U**). This also holds true for the speck area (**Supplementary Figure 10V**), but the observation that the speck size depends on cell area could potentially confound the trend. Hence, we examined the normalized speck size as a function of pseudo time. The normalized radius of gyration of the speck (corrected for the correlation with cell area) increases only slightly during the course of ASC recruitment (**Figure 5B**). Similarly, when we normalized out the increase in speck area with cell size, only a small increase with pseudo time is observed **(Supplementary Figure S10P)**. We also plotted cell area with pseudo time (**Supplementary Figure S10W**), and observed no correlation. Although this was expected, it also verifies that the endogenous ASC speck formation is complete before pyroptosis and cell shrinkage is triggered. If the radius of gyration does not depend on pseudo time but the ASC content does, then we would expect the ASC density in the speck to increase with time. We investigated the speck density (localizations per area) and found it to be largely independent of the cell area (**Supplementary Figure S10X**) but clearly increasing with ASC recruitment (**Figure 5C**). In line with this observation, there is a clear correlation between the speck density and the amount of ASC within the speck, calculated either via the integrated widefield intensity within the speck (**Supplementary Figure S10Y**) or the number of localizations within the speck (**Supplementary Figure S10Z**). Hence, we conclude that the density and, to a much lesser extent, the size of the speck increases with ASC recruitment.

To visualize potential structural rearrangements of the speck, we plotted the endogenous ASC specks as a function of pseudo time (**Figure 5D**). No clear structural progression was visible. To look for more subtle changes during ASC recruitment, we divided specks into three groups according to the fraction of recruited ASC. We aligned the complexes in each group to their center of mass. The resulting sum projection showed an increased tendency to form a ring-like structure at later stages of

recruitment, which is consistent with the observation of the specks become denser with time and thereby exclude the antibody from the center of the speck (**Figure 5D** lower panel).
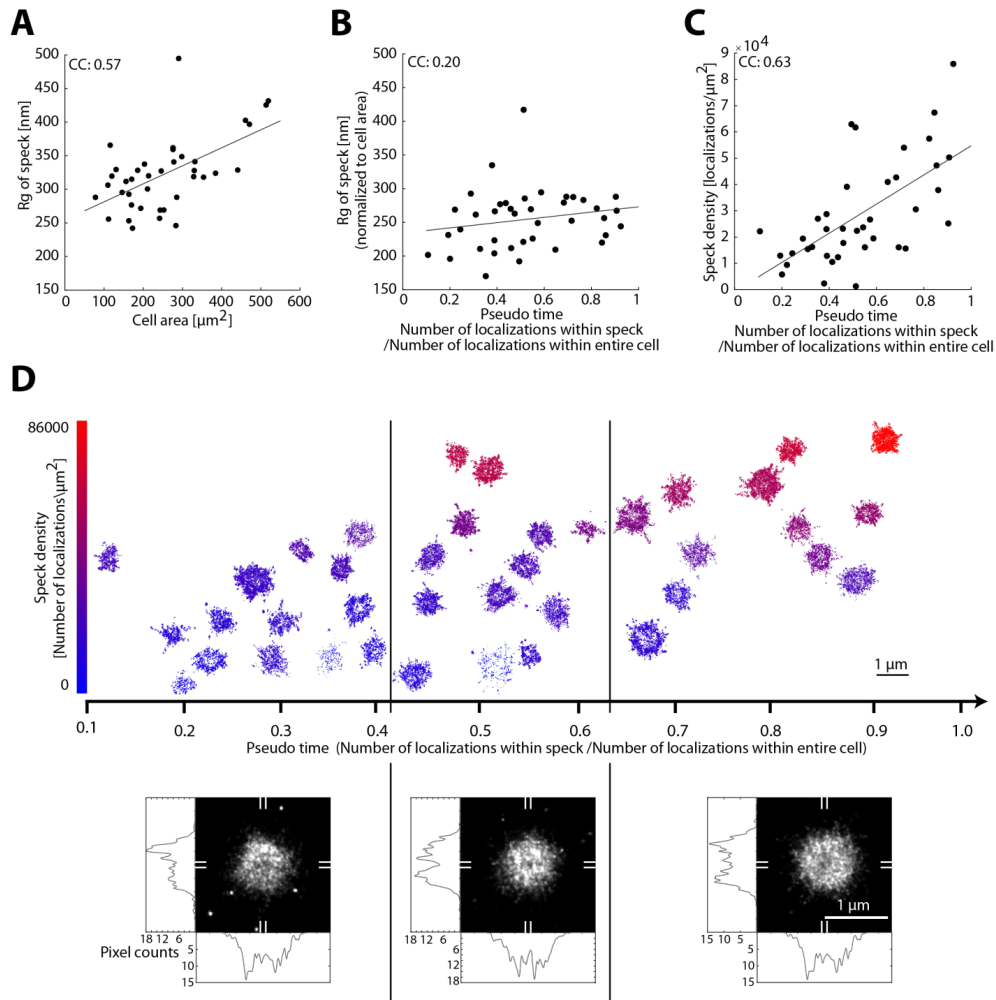

Figure 5 – Dynamics of endogenous ASC speck formation. A) – C) Scatter plots of speck size (measured as the radius of gyration) as a function of the cell area (A), the speck size, normalized to cell area versus the fraction of the total localizations located within the speck (pseudo time) (B) and the speck density (measured as number of localizations per μm^2) as a function of pseudo time (C). Larger cells form larger specks and the normalized speck size stays almost constant with increasing pseudo time whereas the speck density increases with pseudo time. CC: Pearson correlation coefficient. D) Plot of individual speck structures as a function of the pseudo time. The speck density is color-coded (scale bar is shown on the left) and specks are positioned in the vertical direction approximately according to their density. Specks were separated into three time bins (separated by the vertical black lines) and the super-resolution structures aligned and summed (lower panels). Horizontal and vertical cross-sections are shown (determined by averaging along the marked 9 pixel wide regions indicated on the periphery with white lines). The averaged speck structure appears more ring-like at late time points compared to early time points. Data was obtained on three independent cell preparations.

Taken together, our data suggest that the speck forms initially as a loose scaffold of intertwined filaments whose size depends on the cell area and which becomes denser, but not much larger over time by recruiting and incorporating more ASC (schematically shown in **Figure 6**).
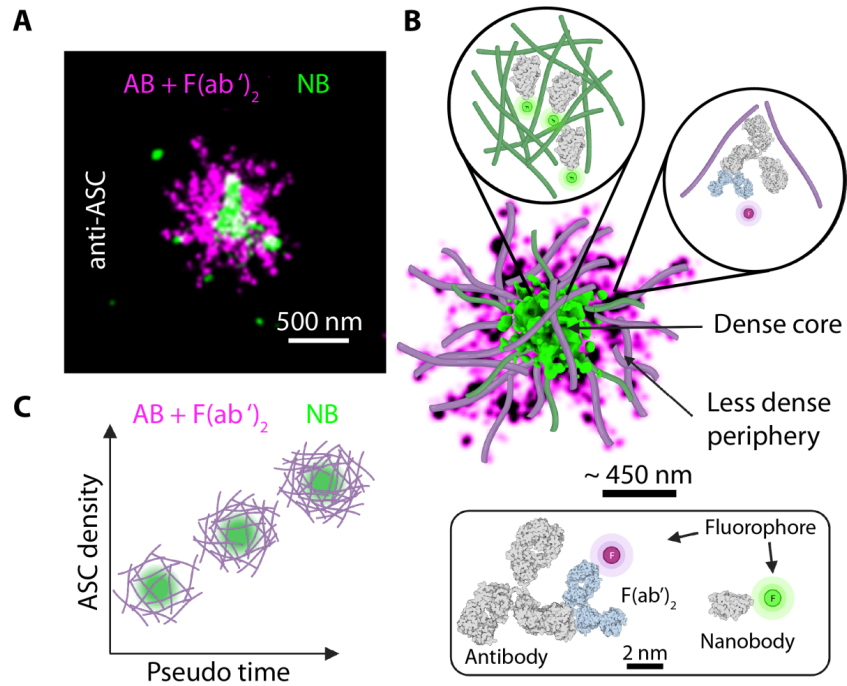

Figure 6 – Model of endogenous ASC speck formation. A) 2-color dSTORM reconstruction of an ASC speck stained with primary anti-ASC antibody and secondary Alexa Fluor 647-conjugated F(ab')$_2$ fragment (magenta) and a DyLight 755-conjugated anti-ASC nanobody (green). The nanobody signal is concentrated in the center of the structure whereas the primary antibody / secondary F(ab')$_2$ fragment staining is also observed in the periphery of the structure. B) Schematic model for the supramolecular structure of the ASC speck based on measured data overlaid with modeled filaments. The structure is characterized by a dense core and a less dense periphery. The zoom ins illustrate that the nanobody is able to penetrate into the dense core of the speck whereas the antibody preferably labels the less dense periphery of the structure. The insert at the bottom of panel B illustrates the size difference between the labeling probes. C) Illustration of speck formation over time. The ASC speck forms from intertwined filaments assembling into a scaffold whose size scales with the cell size. Early stages of speck formation are characterized by a loose assembly into which the antibody (magenta) and nanobody (green) can penetrate. Further ASC recruitment into the speck leads to denser structures but only a marginal increase in its size. Antibody and F(ab')$_2$ fragment staining is sterically excluded from the dense center of the structure resulting in an overall ring-like appearance. The smaller nanobody stains the dense core of the speck but is washed away from the less dense regions of the structure due to its lower binding affinity. AB: Antibody; F(ab')$_2$: F(ab')$_2$ fragment; NB; Nanobody

Discussion

The supramolecular NLRP3 inflammasome complex is a central component of the innate immune system, driving the maturation of the proinflammatory cytokines IL-1β and IL-18 as well as pyroptosis, a proinflammatory form of cell death. The protein ASC is critical during inflammasome formation and, upon cell stimulation, is recruited into a single, condensed structure, the ASC speck. ASC speck formation and organization are difficult to study due to the small size and heterogeneity of the complex. We used a combination treatment of LPS and nigericin, which led to robust induction of ASC specks with a broad distribution in ASC content (**Figure 1**). In diffraction limited wide-field microscopy, the complexes appear as spherical structures with very few morphological features. We then used dSTORM and DNA-PAINT, super-resolution techniques, to investigate the organization of endogenous ASC in the speck as well as in the cytoplasm. dSTORM measurements were performed on a total of 358 cells and 251 specks were analyzed in detail. This allowed us to reconcile several controversies regarding the reported structures of the ASC speck.

Our cluster analysis of the non-speck bound cytosolic ASC fraction revealed that the oligomeric size of the vast majority of ASC signals remained unchanged between unstimulated and speck-containing cells (**Figure 2E + F**). Considering the size of the labels, we measured an average R_g in unstimulated cells of ~15 nm (**Figure 2F**). This is in good agreement with an atomic force microscopy (AFM) study that measured the dimensions of full-length human ASC assemblies, which organized into disc-like oligomers of 1 nm in height and ~ 12 nm in diameter *in vitro* [11]. Upon stimulation, we observe a small increase in the number of larger clusters in the cytosol. This could be due to preassembly of a portion of ASC into higher-order ASC oligomers that later associate into the speck as has been suggested in the literature [8, 60]. Western blot analysis of inflammasome activated cells found different levels of ASC multimerization in addition to the ASC speck with a large proportion of the protein being dimeric [35].

In our systematic measurements of ASC speck formation, the cells either showed a single speck or the ASC protein appeared homogenously distributed throughout the cell. We did not observe any concentration dependence of ASC as a function of distance from the speck (e.g. **Figure 2C**), consistent with previous observations [46] made in cells overexpressing ASC. This is understandable since the entire pool of ASC is recruited into the ASC speck within a few minutes during complex formation [8, 18, 46, 61]. We were still able to detect cytosolic ASC in speck containing cells. The percentage of ASC recruitment we observed was typically 20% or above (with one exception) suggesting that the initial recruitment of ASC is faster than the fixation processes. This is consistent with live-cell imaging measurements in where the increase in speck size occurred over a time interval of ~100 s [46]. Our pseudo time observations indicate that it is mostly the speck density but not the speck size that increases with the percentage of ASC being recruited during the later stages of assembly.

The specks we recorded were smaller in size than those measured previously upon overexpression of ASC [22, 25]. This is in line with our observation that the radius of gyration increases with total ASC content (**Figure 5A**). Moreover, NLRP3 inflammasome formation has recently been shown to occur at the microtubule organizing center (MTOC) [62] and a correlation has been observed between the size of the MTOC and the cell size (at least in *C. elegans*)[63]. This could provide an additional explanation for the size dependence of the ASC speck, in the case that the MTOC provides a scaffold for assembly of the speck. To reliably quantify the attributes of the amorphous, heterogeneous speck, we combined the results from 251 specks. The endogenous ASC speck has a size variation, determined from the 2D radius of gyration, ranging from 250 to 500 nm with an average radius of gyration of 321 ± 31 nm MAD (**Figure 3F**). The radius of gyration of the ASC speck corresponds the previously published value of ~ 600 nm diameter for the endogenous structure (measured from cross-sections through the

speck) [35]. Typically, one would expect the size of the speck to increase during assembly and hence the distribution of sizes to depend, in part, on the stage of assembly at which they are measured. However, when calculating the amount of ASC located within the speck relative to the total amount of ASC within the cell (i.e. our pseudo time), we found that the speck size depends only to a minor degree on the amount of recruited ASC (**Figure 5B**).

We found the majority of specks appeared as amorphous objects with an overall spherical structure, as determined from the calculated eccentricity. These results are consistent with previous observations [8, 29]. In addition, we observed that some specks exhibit filamentous extensions protruding from the dense core. Although several studies suggested that the ASC speck is made up of intertwined filaments [20, 22, 26, 27], this has not yet been confirmed for the endogenous, unperturbed structure inside cells. We observe a dense core with filaments protruding from the endogenous ASC speck, which would be consistent with this hypothesis.

Although we could not analyze the observed filaments in detail due to the limited labeling density, we did observed filaments of varying length and thickness. Some specks exhibit many short and faint fibrils (**Figure 3 M-P**) while others show a small number of longer filaments protruding from the edge of the speck center (**Figure 3C**). The thin fibrils are reminiscent of the fibrils resolved by EM at the edge of the *in vitro* formed ASC assembly [22]. A similar variation in thickness of the filaments protruding from the dense speck core has been observed in immortalized ASC[-/-] bone marrow-derived macrophages (BMDMs) [20]. Interestingly, electron microscopy data of *in vitro* assembled structure from full-length human ASC protein suggested that individual ASC filaments can laterally stack via the exposed CARD domains and the authors raise the question of whether this could also happen in the endogenous structure [17]. The average filament thickness we obtained on the above-mentioned filaments agrees with the thickness reported for a single filament ($\sim$16 nm) [17] suggesting that, for the majority of filaments, lateral stacking does not take place.

It is important to mention that the majority of our experiments were conducted in caspase-1 knock-out cells. Hence, we cannot exclude that the lack of caspase-1 binding to ASC has an influence on the appearance of the speck. Experiments performed using DNA-PAINT on BlaER1 cells that express caspase-1 showed the same dense, amorphous core with protruding filaments. This strongly suggests that the measured structures are indicative of the morphology of the ASC speck and is consistent with what has been observed in experiments in cells with ASC overexpression [20, 22, 25]. However, in comparison to the previous observations of the speck formed after overexpression [22, 25], we found fewer and shorter filaments protruding from the endogenous ASC speck core with the majority of filaments being 500 nm long or shorter. It is conceivable that overexpression leads a larger number of filaments emanating from the formed speck. It is also possible that caspase-1 influences the filament length as our measurements in BlaER1 cells that express caspase-1 showed longer filaments than the ones observed in THP-1 caspase-1 knock-out cells (**Figure 3C** and **Supplementary Figure S4B**). The longer filament lengths are in line with a previous report where *in vitro* experiments measured filament lengths of 500-2000 nm for full length mouse ASC [16]. However, we cannot entirely rule out that the differences in filament length are due the details of the super-resolution approach used.

An overall ring-like appearance of the ASC speck has been proposed since its discovery [3]. However, the question of how the ring-like assembly and the irregular structure made up of intertwined filaments relate to each other remains unanswered [39-42]. Here, we provide new data that offers an explanation for the observed ring-like structure. Complementing previous studies, which showed a ring-like assembly of ASC following antibody labelling, we used labels of different sizes, primary antibodies labeled with secondary F(ab')$_2$ fragments and nanobodies. While primary antibody plus secondary F(ab')$_2$ fragment

labeling occasionally showed ring-like structures, simultaneous labeling with nanobodies revealed a dense core in the speck (**Figure 4**). Hence, we conclude that the endogenous speck can be divided into two different regions: a dense core and a less dense periphery (**Figure 6**). This would imply that the ring shape originates from the primary and/or secondary antibody being less likely to penetrate into the dense center of the structure and thus leading to a decrease in labeling efficiency at the center of the structure. Conversely, the less dense structures at the periphery are less efficiently labeled by the nanobody as it has only one binding site per protein with, in this case, a relatively low binding affinity (~160 nM). This also demonstrates the importance of verifying super-resolution structures using different labeling strategies and approaches.

Conclusion

Taken together, we investigated the nanoscale organization of the endogenous ASC speck using super-resolution microscopy. We found that the speck size was heterogenous and correlated with cell size. The speck contained a dense core with filaments protruding from the center of the speck. This was observed for dSTORM and DNA-PAINT measurements in THP-1 cells as well as DNA-PAINT measurements in BlaER1 cells. These results are consistent with the model of the ASC speck being an assembly of intertwined filaments. By using differently sized labels, we found that the ring-like appearance of the speck is a result of the labeling approach and limited access of the primary antibody and/or F(ab')$_2$ fragment to the dense core of the speck. Conversely, the small nanobody was able to label the center of the speck but was less efficient in labeling the low density periphery regions due to the lower affinity of this nanobody (**Figure 4**). Hence, it is important to verify super-resolution structures using different labeling approaches. Finally, by measuring the fraction of recruited ASC, we hypothesize that speck formation starts with a loose scaffold that becomes denser but only marginally larger during ASC speck formation.

Materials and Methods

If not stated differently all reagents were purchased from Thermo Fisher Scientific, Massachusetts, USA.

Cultivation of cells and activation of the NLRP3 inflammasome

For the experiments performed with THP-1 caspase-1 knock-out cells, the THP-1 cells were cultivated at 37°C, 5% CO_2 in Roswell Park Memorial Institute 1640 medium (21875034) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS) (10500064), 1 mM sodium pyruvate (11360039) and 100 U/ml Penicillin/Streptomycin (15140122), and maintained at a density between 1×10^5 and 1×10^6 cells per milliliter. One day prior to seeding the cells, coverslips (1.5, Menzel Gläser, 18 mm) were coated with 0.01% poly-l-ornithine solution (A-004-C, Merck-Millipore, Massachusetts, USA) in the dark. Cells were seeded at a density of ~ 75×10^3/cm^2 in culture medium supplemented with 100 ng/ml Phorbol 12-myristate 13-acetate (PMA) (BML-PE160-0001, Enzo Life Sciences, Lörrach, Germany) and differentiated into macrophage-like cells for three days. To increase NLRP3 protein expression, cells were primed for three hours with 1 µg/ml Lipopolysaccharide (LPS) from *E. coli* K12 cells (Ultrapure tlrl-peklps, Invivogen, San Diego, California, USA). To suppress cell death due to caspase activity, cells were subsequently treated with 20 µM of the pan-caspase inhibitor Z-VAD (tlrl-vad, Invivogen) for 60 mins. The NLRP3 inflammasome response was activated by incubating the cells with 10 µM Nigericin (N7143-5MG, Sigma Aldrich, Missouri, USA) for 90 mins.

ASC immunofluorescence with primary antibody and secondary F(ab')₂ fragment and for double-labeling with an ASC nanobody

All steps were performed at room temperature if not stated differently. Cells were washed once with phosphate-buffered saline (PBS) solution to remove remaining serum proteins, fixed for 15 mins in the dark with 4% paraformaldehyde (PFA) (E15710-S, Electron Microscopy Sciences, Pennsylvania, USA) in PBS. Subsequently, PFA was rinsed once with PBS and then quenched by rinsing once with 0.1 M NH_4Cl (254134, Sigma-Aldrich) in PBS followed by a 15 min incubation with 0.1 M NH_4Cl in PBS. Permeabilization and blocking was done for 30 mins in 10% normal goat serum (NGS) (16201), 0.5% Triton X-100 (T8787, Sigma-Aldrich) in PBS followed by a washing step with PBS and 30 mins incubation in Image-iTFX Signal enhancer (I36933). After two additional PBS washing steps, cells were incubated overnight at 4°C with purified monoclonal mouse anti-human ASC antibodies (TMS-1) (clone HASC-71) (Biolegend, California, USA) at a final concentration of 10 µg/ml diluted in 10% NGS + 0.5% Triton X-100 in PBS. 25 µl of this solution was applied to the fixed cells on a coverslip. The plasmid encoding for the ASC nanobody was provided by Prof. Dario Alessi, University of Dundee via the MRC – Protein Phosphorylation and Ubiquitylation Unit [DU 54832]. The nanobody encoding sequence with a C-terminal cysteine was amplified for labeling and expressed at the protein production core facility of the Max-Planck-Institute for Biochemistry in Martinsried, Germany. Parts of the nanobody were conjugated to Alexa Fluor 647 in house as described in [64] and at Nanotag Biotechnologies GmbH, Göttingen, Germany. In the case of dual-color labeling, DyLight755-conjugated anti-ASC nanobody was purchased from Nanotag and mixed into the antibody solution at 1 µg/ml final concentration. Non-specific sticking of antibody and nanobody was minimized by washing three times with 0.1% Triton X-100 in PBS. Labeling with secondary F(ab')₂-goat anti-mouse IgG (H+L) cross-adsorbed Alexa Fluor 647-conjugated F(ab')₂ fragment (A-21237) was performed at 200 ng/ml final concentration in 10% NGS + 0.5% Triton X-100 in PBS for 1 hr followed by three washing steps with 0.1% Triton X-100 in PBS and postfixation with 3% PFA in PBS for 10 mins. The specificity of the staining procedure was confirmed by staining THP-1 caspase-1 knock-out cells only with the F(ab')₂ fragment without previous administration of the primary antibody as well as by applying the staining protocol to THP-1 ASC knock-out cells (**Supplementary Figure S10**).

ASC immunofluorescence with ASC nanobody

THP-1 caspase-1 knock-out cells were activated for the NLRP3 inflammasome, fixed and quenched as described above followed by permeabilization with 0.05% Saponin (47036, Sigma-Aldrich), 1% BSA (A7030, Sigma-Aldrich), and 0.05% NaN_3 (S2002, Sigma-Aldrich) in PBS for 20 mins. Afterwards, the sample was wash for 2 mins with PBS and then the sample was blocked for 30 mins with Image-iTFX Signal Enhancer (R37107). The sample was then washed twice for 2 mins with PBS and stained against ASC with 1 ug/ml Alexa Fluor 647-conjugated nanobody and postfixed for 10 mins in 3% PFA in PBS. The specificity of the staining was confirmed using THP-1 ASC knock-out cells (**Supplementary Figure S13**).

Confocal imaging

For the confocal imaging shown in Figure 1, the samples were embedded in ProLong Gold (P10144) on standard glass slides for at least one day before imaging. Microscopy was then performed on a Leica SP8 STED 3x equipped with a 470-670nm white light laser and a 100 x PlanApo /NA 1.4 objective. The spinning disk confocal microscopy images shown in **Supplementary Figure S12A** were measured on a Zeiss Cell Observer SD. Alexa Fluor 647 and DAPI were excited using a 639 nm and 405 nm laser, respectively. Fluorescence was separated using a 660 nm longpass filter and

recorded on the two Evolve 512 electron-multiplying charge-coupled device cameras (Photometrics) of the system equipped with a 525/50 and a 690/50 bandpass filter, respectively.

dSTORM imaging and analysis

Samples were imaged on a flat field-optimized widefield setup as described before [51]. Briefly, lasers of 405 nm, 642 nm and 750 nm were expanded and reflected into a 60 x objective (CFI60 PlanApo Lambda 60x / NA 1.4, Nikon) using a custom appropriate dichroic mirror (ZT405/561/642/750/850rpc, Chroma). Fluorescence emission was imaged onto a sCMOS camera (Prime, Photometrics) using one of two emission filters (ET700/75M and ET810/90m, Chroma) combined with a short-pass filter (FF01-842/SP, Semrock) and a tube lens of 200 mm focal length. The microscope was controlled using Micromanager [65, 66]. Widefield images were collected prior to dSTORM recordings. We typically recorded between 20-80k frames at 10 ms exposure time. For 3D dSTORM, we introduced a cylindrical lens (f=1000 mm, Thorlabs LJ1516RM-A) into the emission path. Single- and dual-color single-molecule localization microscopy (SMLM) imaging was carried out with an optimized SMLM buffer, as described previously [67]. Single molecules were localized using an sCMOS-specific localizer routine introduced by Huang *et al.* [68] and included in a custom MatLab program used for data analysis. To exclude non-specific localizations, filter parameters were adjusted using datasets acquired using ASC knock-out THP-1 cells. Drift correction was performed using the Redundancy Cross-Correlation (RCC) algorithm introduced by Wang *et al.* [69]. The cells and specks were manually segmented. Cytoplasmic localizations were clustered using DBSCAN where a cluster was defined as a group of 10 – 300 localizations within a search radius of 70 nm [48] or Ripley's K function [49]. Super-resolution images were reconstructed using ThunderStorm [70] with 10x magnification and applying a Gaussian blur of one pixel (10.6 nm). For dual-color data sets, the localizations in the DyLight 755 channel were registered to the ones in the Alexa Fluor 647 channel by an affine transformation calculated from widefield images of immobilized TetraSpeck Microspheres (T7279) on a plasma cleaned coverslip recorded in both channels. 3D representations were rendered using Chimera X [71].

The eccentricity, the radius of gyration and the number of localizations of the specks were calculated from the manually segmented localizations within the structure. The radius of gyration was calculated as the square root of the sum of the variances of the x and the y coordinates of the individual localizations within each speck. As, in most cases, the specks are amorphous, spherical structures and we have significantly better resolution in the radial dimension, we limited the calculated radius of gyration to two-dimensions. The eccentricity of the individual specks was determined from the localizations by first calculating the covariance matrix of the individual localizations. From the covariance matrix, the eigenvectors were calculated, which provide a measure for how elliptical the distribution of localizations is and calculates the direction of the major and minor axes of the ellipse. We then take the square root of eigenvectors, which give a measure of length for the different axes. Finally, the eccentricity is calculated by taking the ratio of the major axis (square root of the maximum eigenvector) to the minor axis (square root of the minimum eigenvector). Circular objects have an eccentricity near 1. For our microscope we determined an experimental localization precision of σ_{xy} = 12 nm for Alexa Fluor 647 and σ_{xy} = 21 nm for DyLight755 [51].

For the quantitative image analysis, measurements of the integrated widefield intensity were corrected by subtracting background counts determined individually for each field of view by averaging pixel counts in an area absent of cellular signal. To calibrate for the experimental differences observed between measurement days, the widefield intensity per cell area and the number of detected localizations per widefield intensity were corrected separately for each measurement day. For this, we used the correlation of fluorescence intensity with cell size and the number of localizations versus

widefield intensity. The correlation was determined for each day by fitting the data with a line and the slope was scaled to correspond to the maximum of all experiment days. A correction factor was determined for each measurement day and used to rescale the measurements to the measurement day with the largest slope. The scaled widefield intensity and number of localizations were then used for the ensuing analyses. **Supplementary Figure S11** illustrates the applied correction procedure.

The cell area dependence of the radius of gyration (R_g) was normalized out by first plotting the R_g against the cell area and fitting the data with a line. Subsequently for each data point, the product of the slope of the line and the value of the cell area was calculated and subtracted from the R_g value. The speck area was normalized analogously.

<u>BlaER1 cell preparation and DNA-PAINT imaging</u>

BlaER1 ASC-TagRFP cells were cultivated as described for the THP-1 cells. Cells were transdifferentiation for 7 days at 60,000 cells/cm^2 in medium containing 50 ng/ml IL-3 (200-03 B, Peprotech, New Jersey, USA), 50 ng/ml M-CSF (300-25 B, Peprotech) and 500 nM β-Estradiol (E8875, Sigma-Aldrich) in cell culture-treated plastic bottom slides (80826, ibidi, Gräfelfing, Germany). Cells were then trypsinized, transferred to glass bottom slides coated with poly-l-ornithine as described for THP-1 cells in cultivation medium without growth factors at the same density and incubated overnight at 37°C, 5% CO$_2$. The next day, the NLRP3 inflammasome was activated by priming the cells for 14 hrs with 200 ng/ml Lipopolysaccharide (LPS) from *E. coli* K12 cells in medium followed by caspase inhibition with 20 μM Z-VAD-FMK incubation for 60 mins as described for THP-1 cells and 3 hrs incubation with 6.5 μM Nigericin. Cells were fixed and stained as described for THP-1 cells with the exception that a secondary donkey anti-mouse antibody (715-005-151, Jackson ImmunoResearch) conjugated to a P3 docking strand (TTCTTCATTA) was used instead of the F(ab')$_2$ fragment. Specks were identified by excitation of TagRFP with a 561 nm laser, followed by bleaching the signal and DNA-PAINT data were recorded using a P3-Cy3B imager strand (AATGAAGA-Cy3B) at 1 nM concentration and HiLo illumination at 1.2 kW/cm² excitation with the same laser at 100 ms camera exposure time for 8000 frames.

THP-1 cells imaged by DNA-PAINT were first labeled with a low amount of Alexa Fluor 647-conjugated anti-ASC nanobody to facilitate identification of the speck. Subsequently, the nanobody conjugated to a P3 docking strand was applied and the structure was imaged using the complementary imager strand labeled with Cy3B at 0.5 nM concentration. Images were recorded for 20,000 – 30,000 frames with an exposure time of 100 ms and a laser power density of ~200 kW/cm².

Super-resolution images were reconstructed using the Picasso [51] and ThunderStorm [64] softwares and rendered with a Gaussian blur of one pixel corresponding to 13 nm.

Authors contributions

I. M. G. and D. C. L. designed experiments, I. M. G., C. S. and D. C. L. analyzed the data, I. M. G., G. A. and G. M. P. conjugated the nanobody, I. M. G. and G. M. P. prepared immunofluorescence samples, C. S. and I. M. G. recorded dSTORM data, S. S. and I. M. G. recorded DNA-PAINT data, C. S. prepared THP-1 knock-out cells, T. S. E. and C. S. helped with establishing the immunofluorescence protocols, D. C. L., V. H., S. M., R. J. and I. M. G. acquired funding, I. M. G., C. S. and D. C. L. wrote the manuscript with input from all other authors.

Conflicts of Interest

The authors declare no conflict of interest.

References

1.   Martinon, F., Burns, K. & Tschopp, J. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. *Molecular cell* **10**, 417-426 (2002).

2.   Conway, K.E. *et al.* TMS1, a Novel Proapoptotic Caspase Recruitment Domain Protein, Is a Target of Methylation-induced Gene Silencing in Human Breast Cancers. *Cancer Research* **60**, 6236-6242 (2000).

3.   Masumoto, J. *et al.* ASC, a novel 22-kDa protein, aggregates during apoptosis of human promyelocytic leukemia HL-60 cells. *The Journal of biological chemistry* **274**, 33835-33838 (1999).

4.   Howard, A.D. *et al.* IL-1-converting enzyme requires aspartic acid residues for processing of the IL-1 beta precursor at two distinct sites and does not cleave 31-kDa IL-1 alpha. *The Journal of Immunology* **147**, 2964-2969 (1991).

5.   Thornberry, N.A. *et al.* A novel heterodimeric cysteine protease is required for interleukin-1 beta processing in monocytes. *Nature* **356**, 768-774 (1992).

6.   Wang, L. *et al.* PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates activation of NF-kappa B and caspase-1-dependent cytokine processing. *The Journal of biological chemistry* **277**, 29874-29880 (2002).

7.   Gaidt, M.M. & Hornung, V. The NLRP3 Inflammasome Renders Cell Death Pro-inflammatory. *Journal of Molecular Biology* **430**, 133-141 (2018).

8.   Fernandes-Alnemri, T. *et al.* The pyroptosome: a supramolecular assembly of ASC dimers mediating inflammatory cell death via caspase-1 activation. *Cell Death & Differentiation* **14**, 1590-1604 (2007).

9.   Liepinsh, E. *et al.* The Death-domain Fold of the ASC PYRIN Domain, Presenting a Basis for PYRIN/PYRIN Recognition. *Journal of Molecular Biology* **332**, 1155-1163 (2003).

10.   Hofmann, K., Bucher, P. & Tschopp, J. The CARD domain: a new apoptotic signalling motif. *Trends in Biochemical Sciences* **22**, 155-156 (1997).

11.   de Alba, E. Structure and interdomain dynamics of apoptosis-associated speck-like protein containing a CARD (ASC). *The Journal of biological chemistry* **284**, 32932-32941 (2009).

12.   Sahillioglu, Ali C., Sumbul, F., Ozoren, N. & Haliloglu, T. Structural and Dynamics Aspects of ASC Speck Assembly. *Structure* **22**, 1722-1734 (2014).

13.   Vajjhala, P.R., Mirams, R.E. & Hill, J.M. Multiple binding sites on the pyrin domain of ASC protein allow self-association and interaction with NLRP3 protein. *The Journal of biological chemistry* **287**, 41732-41743 (2012).

14.   Vajjhala, P.R. *et al.* Identification of multifaceted binding modes for pyrin and ASC pyrin domains gives insights into pyrin inflammasome assembly. *The Journal of biological chemistry* **289**, 23504-23519 (2014).

15.   Chou, J.J., Matsuo, H., Duan, H. & Wagner, G. Solution Structure of the RAIDD CARD and Model for CARD/CARD Interaction in Caspase-2 and Caspase-9 Recruitment. *Cell* **94**, 171-180 (1998).

16.   Sborgi, L. *et al.* Structure and assembly of the mouse ASC inflammasome by combined NMR spectroscopy and cryo-electron microscopy. *Proceedings of the National Academy of Sciences* **112**, 13237-13242 (2015).

17.   Nambayan, R.J.T., Sandin, S.I., Quint, D.A., Satyadi, D.M. & de Alba, E. The inflammasome adapter ASC assembles into filaments with integral participation of its two Death Domains, PYD and CARD. *The Journal of biological chemistry* **294**, 439-452 (2019).

18.   Cai, X. *et al.* Prion-like Polymerization Underlies Signal Transduction in Antiviral Immune Defense and Inflammasome Activation. *Cell* **156**, 1207-1222 (2014).

19.   Lu, A. *et al.* Unified Polymerization Mechanism for the Assembly of ASC-Dependent Inflammasomes. *Cell* **156**, 1193-1206 (2014).

20.   Dick, M.S., Sborgi, L., Rühl, S., Hiller, S. & Broz, P. ASC filament formation serves as a signal amplification mechanism for inflammasomes. *Nature Communications* **7**, 11929 (2016).

21.   Li, Y. *et al.* Cryo-EM structures of ASC and NLRC4 CARD filaments reveal a unified mechanism of nucleation and activation of caspase-1. *Proceedings of the National Academy of Sciences* **115**, 10845-10852 (2018).

22.   Franklin, B.S. *et al.* The adaptor ASC has extracellular and 'prionoid' activities that propagate inflammation. *Nature Immunology* **15**, 727-737 (2014).

23.   Richards, N. *et al.* Interaction between pyrin and the apoptotic speck protein (ASC) modulates ASC-induced apoptosis. *The Journal of biological chemistry* **276**, 39320-39329 (2001).

24.   Martín-Sánchez, F. *et al.* ASC oligomer favor caspase-1[CARD] domain recruitment after intracellular potassium efflux. *bioRxiv*, 2020.2001.2027.921239 (2020).

25.   Hoss, F., Rodriguez-Alcazar, J.F. & Latz, E. Assembly and regulation of ASC specks. *Cellular and molecular life sciences : CMLS* **74**, 1211-1229 (2017).

26.   Schmidt, F.I. *et al.* A single domain antibody fragment that recognizes the adaptor ASC defines the role of ASC domains in inflammasome assembly. *J Exp Med* **213**, 771-790 (2016).

27.   Baroja-Mazo, A. *et al.* The NLRP3 inflammasome is released as a particulate danger signal that amplifies the inflammatory response. *Nature Immunology* **15**, 738-748 (2014).

28.   Qiao, Q. & Wu, H. Supramolecular organizing centers (SMOCs) as signaling machines in innate immune activation. *Science China. Life sciences* **58**, 1067-1072 (2015).

29.   Schneider, K.S. *et al.* The Inflammasome Drives GSDMD-Independent Secondary Pyroptosis and IL-1 Release in the Absence of Caspase-1 Protease Activity. *Cell Reports* **21**, 3846-3859 (2017).

30.   Bryan, N.B., Dorfleutner, A., Rojanasakul, Y. & Stehlik, C. Activation of Inflammasomes Requires Intracellular Redistribution of the Apoptotic Speck-Like Protein Containing a Caspase Recruitment Domain. *The Journal of Immunology* **182**, 3173-3182 (2009).

31.   Bryan, N.B. *et al.* Differential splicing of the apoptosis-associated speck like protein containing a caspase recruitment domain (ASC) regulates inflammasomes. *Journal of Inflammation* **7**, 23 (2010).

32.   Khare, S. *et al.* An NLRP7-Containing Inflammasome Mediates Recognition of Microbial Lipopeptides in Human Macrophages. *Immunity* **36**, 464-476 (2012).

33.   Lage, S.L., Dominical, V.M., Wong, C.-S. & Sereti, I. Evaluation of Canonical Inflammasome Activation in Human Monocytes by Imaging Flow Cytometry. *Frontiers in Immunology* **10** (2019).

34.   Li, J.-y. *et al.* Characterization of an NLRP1 Inflammasome from Zebrafish Reveals a Unique Sequential Activation Mechanism Underlying Inflammatory Caspases in Ancient Vertebrates. *The Journal of Immunology* **201**, 1946-1966 (2018).

35.   Man, S.M. *et al.* Inflammasome activation causes dual recruitment of NLRC4 and NLRP3 to the same macromolecular complex. *Proceedings of the National Academy of Sciences* **111**, 7403-7408 (2014).

36.   Masumoto, J., Taniguchi, S. & Sagara, J. Pyrin N-terminal homology domain- and caspase recruitment domain-dependent oligomerization of ASC. *Biochem Biophys Res Commun* **280**, 652-655 (2001).

37.   McConnell, B.B. & Vertino, P.M. Activation of a Caspase-9-mediated Apoptotic Pathway by Subcellular Redistribution of the Novel Caspase Recruitment Domain Protein TMS1. *Cancer Research* **60**, 6243-6247 (2000).

38.   McConnell, B.B. & Vertino, P.M. TMS1/ASC: the cancer connection. *Apoptosis* **9**, 5-18 (2004).

39.   Monie, T. & Boyle, J. Insights into assembly of the macromolecular inflammasome complex. *Inflammasome* **1** (2015).

40.   Chu, L.H., Gangopadhyay, A., Dorfleutner, A. & Stehlik, C. An updated view on the structure and function of PYRIN domains. *Apoptosis* **20**, 157-173 (2015).

41.   de Alba, E. Structure, interactions and self-assembly of ASC-dependent inflammasomes. *Archives of biochemistry and biophysics* **670**, 15-31 (2019).

42.   Elliott, E.I. & Sutterwala, F.S. Initiation and perpetuation of NLRP3 inflammasome activation and assembly. *Immunological reviews* **265**, 35-52 (2015).

43.   Bauernfeind, F.G. *et al.* Cutting Edge: NF-κB Activating Pattern Recognition and Cytokine Receptors License NLRP3 Inflammasome Activation by Regulating NLRP3 Expression. *The Journal of Immunology* **183**, 787-791 (2009).

44.   Mariathasan, S. *et al.* Cryopyrin activates the inflammasome in response to toxins and ATP. *Nature* **440**, 228-232 (2006).

45.   Gammon, K. Inflammation: A complex problem. *Nature* **502**, S86-S87 (2013).

46.   Cheng, J. *et al.* Kinetic properties of ASC protein aggregation in epithelial cells. *Journal of Cellular Physiology* **222**, 738-747 (2010).

47.   Heilemann, M. *et al.* Subdiffraction-Resolution Fluorescence Imaging with Conventional Fluorescent Probes. *Angewandte Chemie International Edition* **47**, 6172-6176 (2008).

48.   Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. in Kdd, Vol. 96 226-2311996).

49.   Ripley, B.D. The second-order analysis of stationary point processes. *Journal of Applied Probability* **13**, 255-266 (2016).

50.   Sieben, C., Banterle, N., Douglass, K.M., Gönczy, P. & Manley, S. Multicolor single-particle reconstruction of protein complexes. *Nature Methods* **15**, 777-780 (2018).

51.   Douglass, K.M., Sieben, C., Archetti, A., Lambert, A. & Manley, S. Super-resolution imaging of multiple cells by optimized flat-field epi-illumination. *Nature Photonics* **10**, 705-708 (2016).

52.   Zwettler, F.U. *et al.* Molecular resolution imaging by post-labeling expansion single-molecule localization microscopy (Ex-SMLM). *Nature Communications* **11**, 3388 (2020).

53.   Jungmann, R. *et al.* Single-Molecule Kinetics and Super-Resolution Microscopy by Fluorescence Imaging of Transient Binding on DNA Origami. *Nano Letters* **10**, 4756-4761 (2010).

54.   Schnitzbauer, J., Strauss, M.T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-resolution microscopy with DNA-PAINT. *Nature Protocols* **12**, 1198-1228 (2017).

55.   Rapino, F. *et al.* C/EBPα Induces Highly Efficient Macrophage Transdifferentiation of B Lymphoma and Leukemia Cell Lines and Impairs Their Tumorigenicity. *Cell Rep* **19**, 1281 (2017).

56.   Gaidt, M.M. *et al.* Human Monocytes Engage an Alternative Inflammasome Pathway. *Immunity* **44**, 833-846 (2016).

57.   Gaidt, M.M., Rapino, F., Graf, T. & Hornung, V. Modeling Primary Human Monocytes with the Trans-Differentiation Cell Line BLaER1. *Methods in molecular biology (Clifton, N.J.)* **1714**, 57-66 (2018).

58.   Moore, R.P. & Legant, W.R. Improving probes for super-resolution. *Nature Methods* **15**, 659-660 (2018).

59.   Ganji, M., Schlichthaerle, T., Eklund, A.S., Strauss, S. & Jungmann, R. Quantitative Assessment of Labeling Probes for Super-Resolution Microscopy Using Designer DNA Nanostructures. *ChemPhysChem* **22**, 911-914 (2021).

60.   Santiveri, C.M., Oroz, J. & de Alba, E. A ring-like model for ASC self-association via the CARD domain. *Inflammasome* **1**, 44-54 (2014).

61.   Sagoo, P. *et al.* In vivo imaging of inflammasome activation reveals a subcapsular macrophage burst response that mobilizes innate and adaptive immunity. *Nature medicine* **22**, 64-71 (2016).

62.   Magupalli, V.G. *et al.* HDAC6 mediates an aggresome-like mechanism for NLRP3 and pyrin inflammasome activation. *Science* **369**, eaas8995 (2020).

63.   Decker, M. *et al.* Limiting amounts of centrosome material set centrosome size in C. elegans embryos. *Current biology : CB* **21**, 1259-1267 (2011).

64.   Pleiner, T. *et al.* Nanobodies: site-specific labeling for super-resolution imaging, rapid epitope-mapping and native protein complex isolation. *eLife* **4**, e11349 (2015).

65.   Edelstein, A., Amodaj, N., Hoover, K., Vale, R. & Stuurman, N. Computer Control of Microscopes Using μManager. *Current Protocols in Molecular Biology* **92**, 14.20.11-14.20.17 (2010).

66.   Edelstein, A.D. *et al.* Advanced methods of microscope control using μManager software. *Journal of Biological Methods; Vol 1, No 2 (2014)* (2014).

67. Olivier, N., Keller, D., Gönczy, P. & Manley, S. Resolution Doubling in 3D-STORM Imaging through Improved Buffers. *PLOS ONE* **8**, e69004 (2013).

68. Huang, F. *et al.* Video-rate nanoscopy using sCMOS camera–specific single-molecule localization algorithms. *Nature Methods* **10**, 653-658 (2013).

69. Wang, Y. *et al.* Localization events-based sample drift correction for localization microscopy with redundant cross-correlation algorithm. *Opt. Express* **22**, 15982-15991 (2014).

70. Ovesný, M., Křížek, P., Borkovec, J., Svindrych, Z. & Hagen, G.M. ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics (Oxford, England)* **30**, 2389-2390 (2014).

71. Pettersen, E.F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein science : a publication of the Protein Society* **30**, 70-82 (2021).

Supplementary Information

Nanoscale organization of the endogenous ASC speck

Ivo M. Glück[1,2], Grusha Primal Mathias[1,3], Sebastian Strauss[4,5], Thomas S. Ebert[6], Che Stafford[6], Ganesh Agam[1,2], Suliana Manley[8], Veit Hornung[5,6], Ralf Jungmann[4,5], Christian Sieben[7,§,*], Don C. Lamb[1,2,*]

[1] Department of Chemistry, Ludwig Maximilians-Universität München, Butenandtstraße 5-13, 81377 München, Germany

[2] Center for Nano Science (CENS), Ludwig Maximilians-Universität München, Butenandtstraße 5-13, 81377 München, Germany

[3] Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Bonn, Germany und Universität Köln, Köln, Germany

[4] Faculty of Physics and Center for Nanosciency, Ludwig Maximilian University, Munich, Germany

[5] Max Planck Institute of Biochemistry, Martinsried, Germany

[6] Gene Center and Department of Biochemistry, Ludwig-Maximilians-Universität, Munich, Germany

[7] Laboratory of Experimental Biophysics, École Polytechnique Fédérale de Lausanne, BSP 427 (Cubotron UNIL), Rte de la Sorge, CH-1015 Lausanne, Switzerland

[§] Current address: Nanoscale Infection Biology, Helmholtz Centre for Infection Research, Inhoffenstr. 7, 38124 Braunschweig, Germany

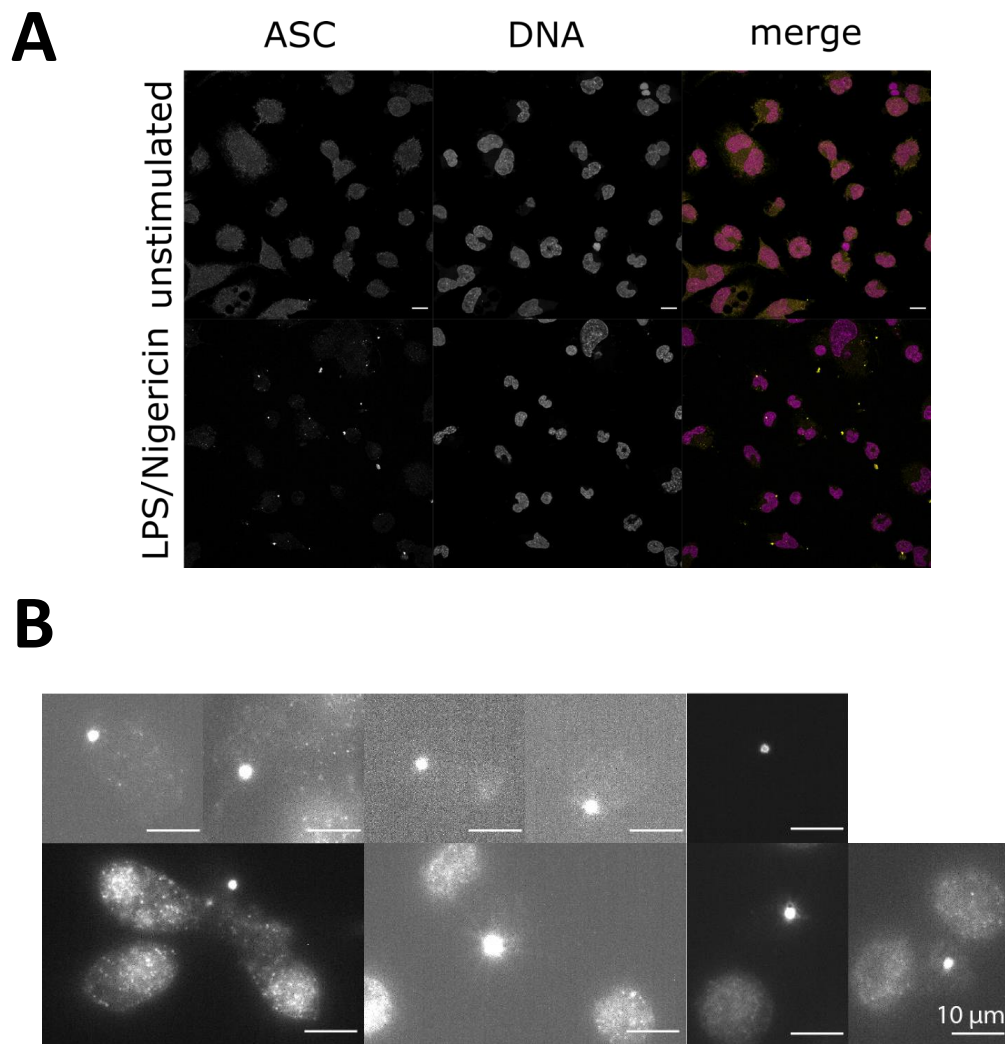* Correspondence should be addressed to

Christian Sieben, christian.sieben@helmholtz-hzi.de

or

Don C. Lamb, don.lamb@cup.uni-muenchen.de

Table of Contents

Supplementary Figure S1: Diffraction-limited images of fluorescently labeled ASC in unstimulated and stimulated THP-1 cells and of extracellular ASC specks. A) Confocal images (maximum projections) of fluorescently labeled ASC in unstimulated (upper panel) and in LPS + Nigericin-stimulated THP-1 cells (lower panel). Images as shown in Figure 1. The scale bar corresponds to 10 µm. B) Widefield images of extracellular ASC specks stained with primary ASC antibody and secondary Alexa Fluor 647-conjugated F(ab')$_2$ fragment.
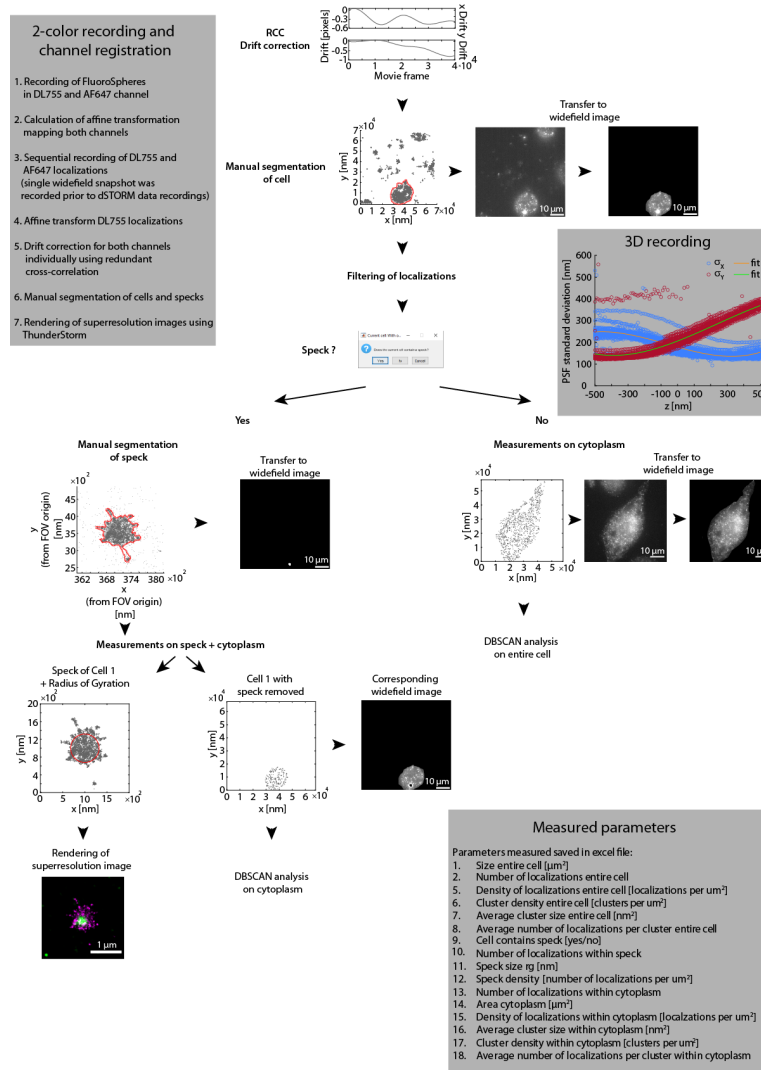
Supplementary Figure S2: Applied image data analysis workflow. Widefield and dSTORM Microscopy experiments were performed on fixed cells. After localizing individual molecules and performing drift correction, the cells were manually segmented from the surroundings using the localizations and the obtained mask was applied to the widefield image. The localizations were filtered to remove unspecific and low quality localizations. In cells containing an ASC speck, the speck was manually segmented and the obtained mask was applied to the widefield image. A super-resolution image was rendered from the localizations. Finally, various parameters for the entire cell, the cytosol and the ASC speck were calculated (listed in the box in lower, right corner). Cytosolic localizations were analyzed using the DBSCAN analysis (Supplementary Figure S3). The axial positions for 3D images were obtained by recording and fitting a calibration curve measured on TetraSpeck Microspheres (see gray box to the upper right). The 2-channel registration was performed as described in the box in the upper left. RCC: Redundancy cross-correlation.
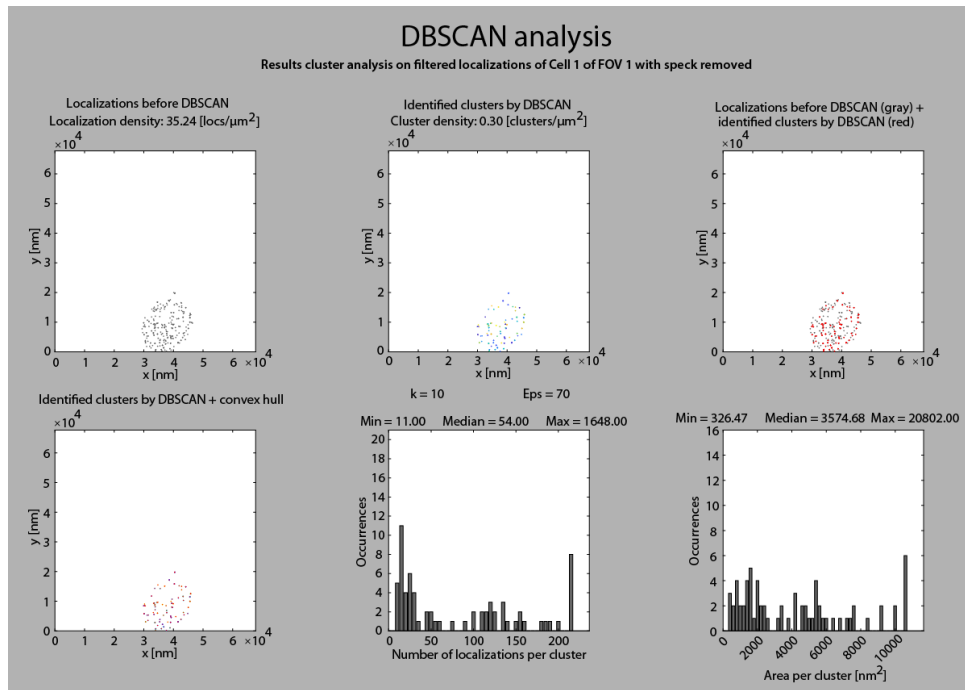
Supplementary Figure S3: The DBSCAN analysis. A screen shot of the DBSCAN analysis of the cytosolic non-speck bound ASC signal of the cell shown in Supplementary Figure S2 is depicted. The localization density before DBSCAN is shown in the top left panel. The top middle panel depicts the identified clusters after an analysis using a search radius of 70 nm and 10-300 localizations per cluster. Individual clusters are color-coded. The top right panel depicts an overlay of the identified clusters (red) with the unclustered localizations (gray). The bottom left panel depicts the clusters together with their convex hull. The bottom middle and right panels show histograms of the number of localizations per cluster and the area per cluster derived by calculation of the convex hull, respectively.
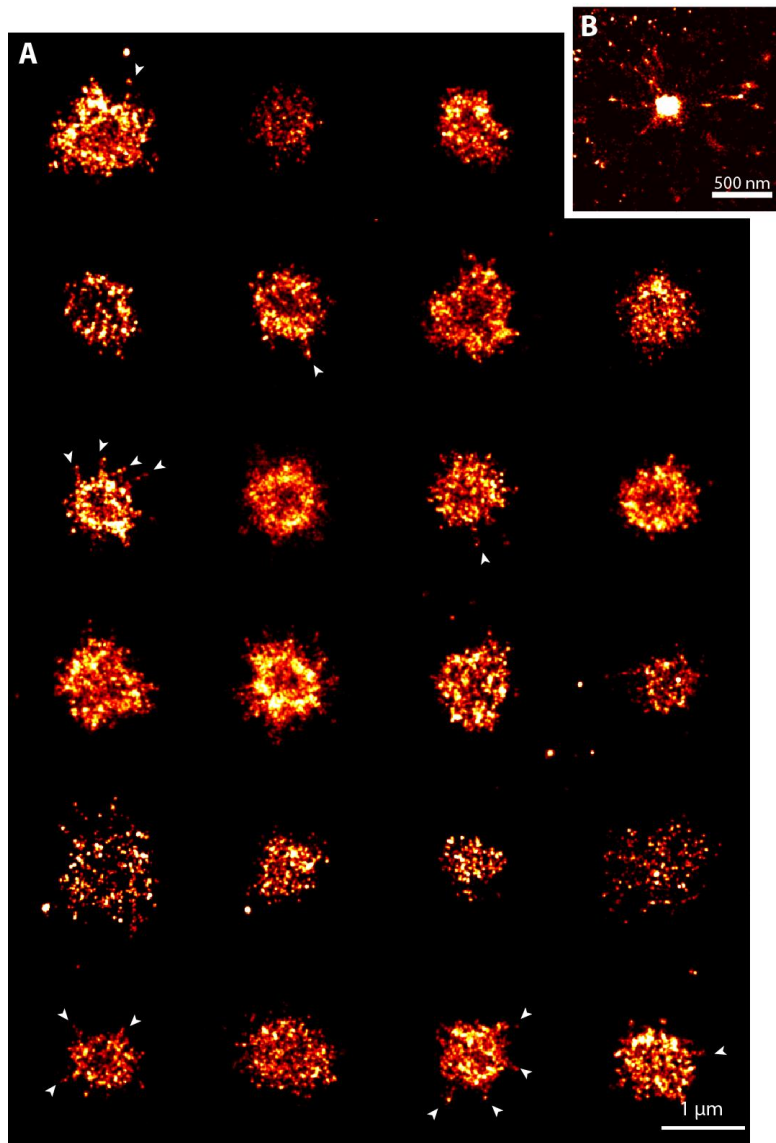
Supplementary Figure S4: Two-dimensional super-resolution images of endogenous ASC specks.
A) A gallery of endogenous ASC specks labeled with an ASC antibody and an Alexa Fluor 647-conjugated F(ab')$_2$ fragment in THP-1 cells imaged using dSTORM. Some structures appear as ring-like structures and other as dense amorphous complexes. In several structures, filamentous extensions are observed, which are indicated by the arrows. B) An ASC speck observed using DNA-PAINT labeled with an ASC antibody and a secondary antibody conjugated with a P3 docking strand for DNA-PAINT imaging. Data was obtained on three independent cell preparations.
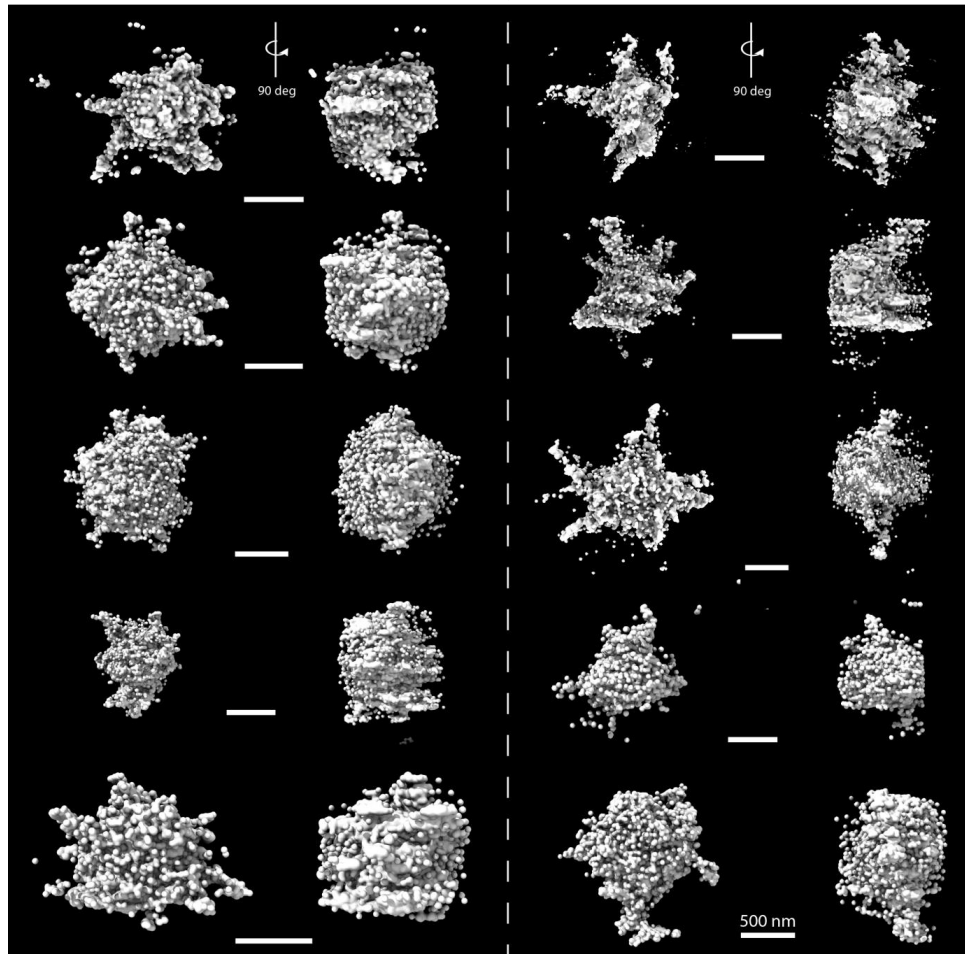
Supplementary Figure S5: Three-dimensional super-resolution images of endogenous ASC specks. 3D renderings of ASC specks that were stained using an ASC antibody and a secondary Alexa Fluor 647-conjugated F(ab')$_2$ fragment and measured using dSTORM. Data was obtained on two independent cell preparations.
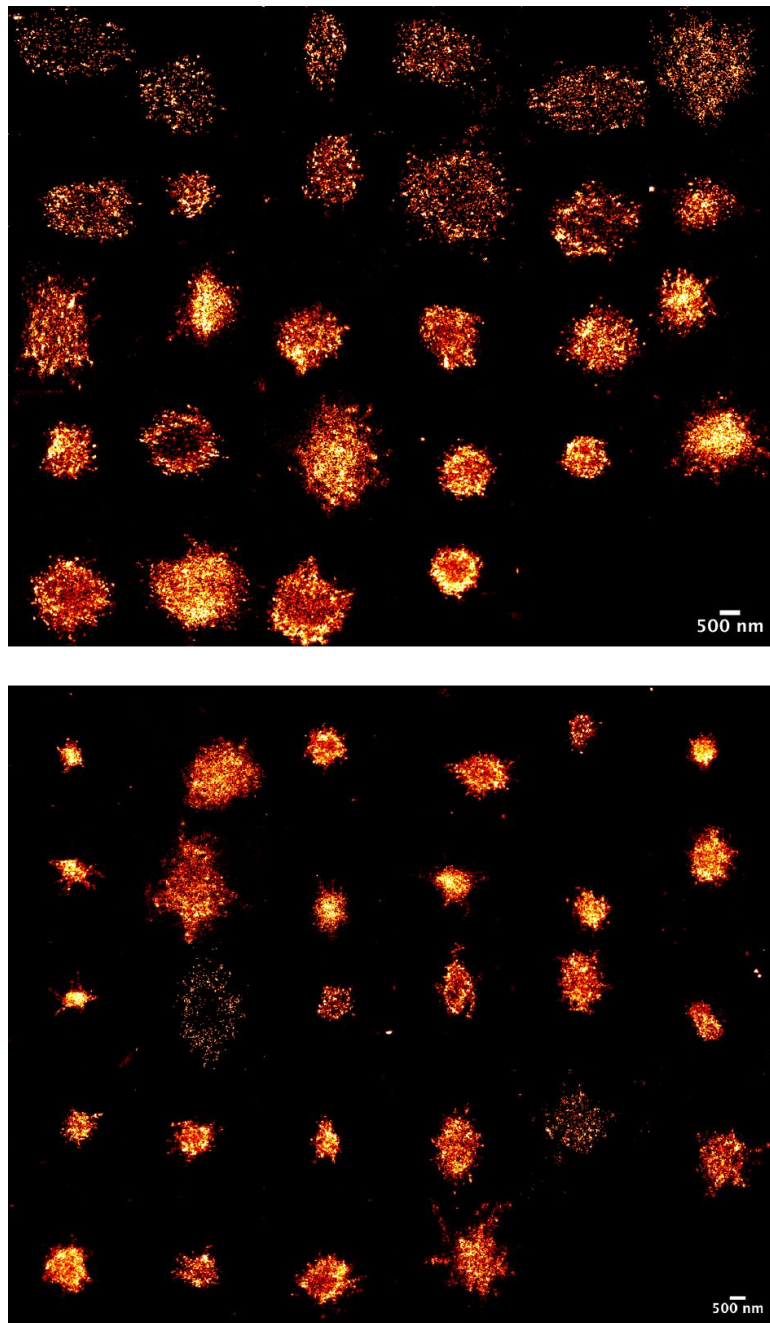
Supplementary Figure S6: Two-dimensional renderings of 3D super-resolution images of endogenous ASC specks. The ASC specks were measured in 3D using dSTORM and stained with an ASC antibody and an Alexa Fluor 647-conjugated secondary F(ab')$_2$ fragment. Data was obtained on two independent cell preparations.
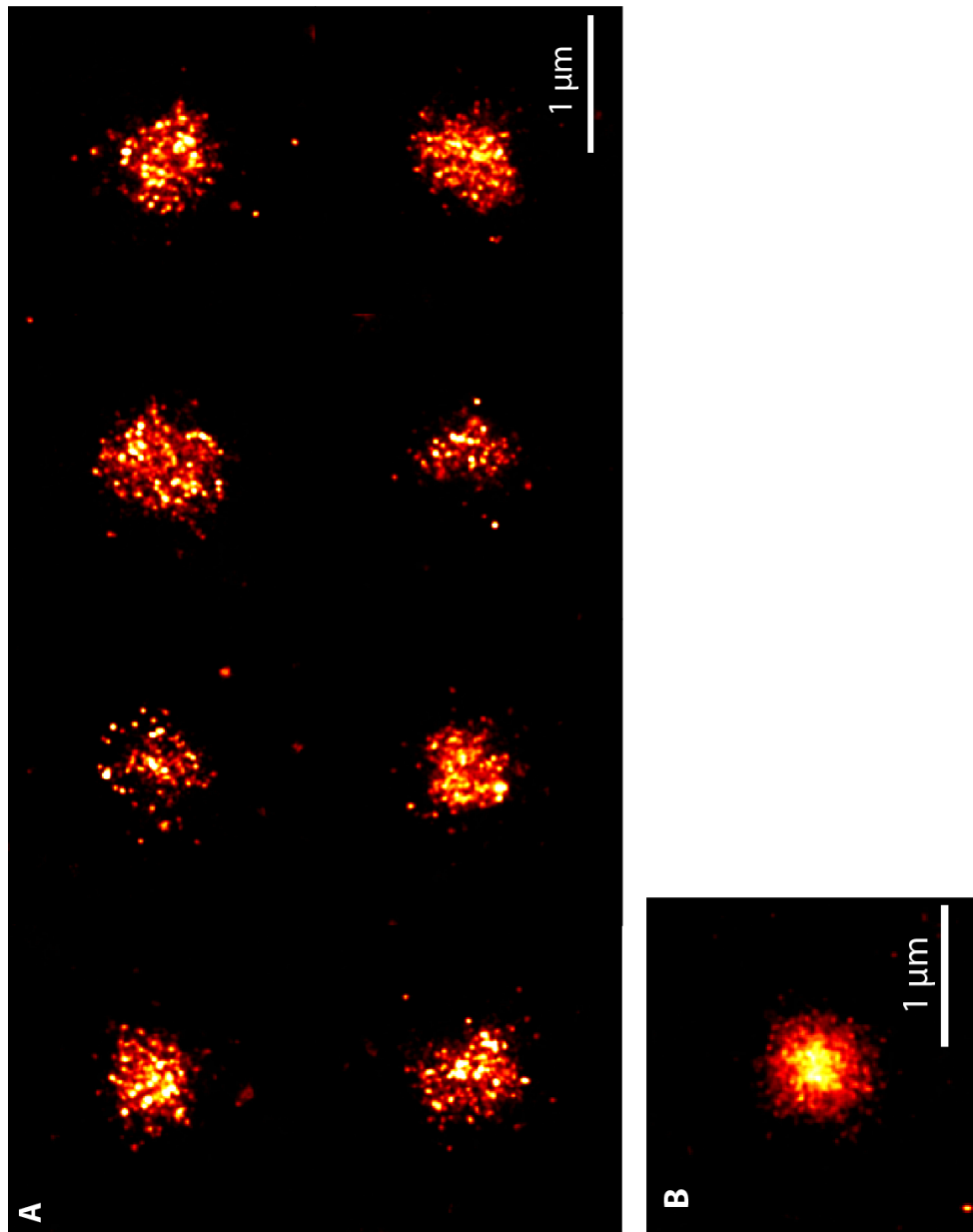
Supplementary Figure S7: Nanobody labeling of endogenous ASC specks. A) ASC specks were stained with Alexa Fluor 647-conjugated nanobody and imaged using 2D STORM. Data was obtained on a single cell preparation. B) An ASC speck stained with a P3 DNA-PAINT docking strand-conjugated nanobody and imaged by 2D DNA-PAINT.
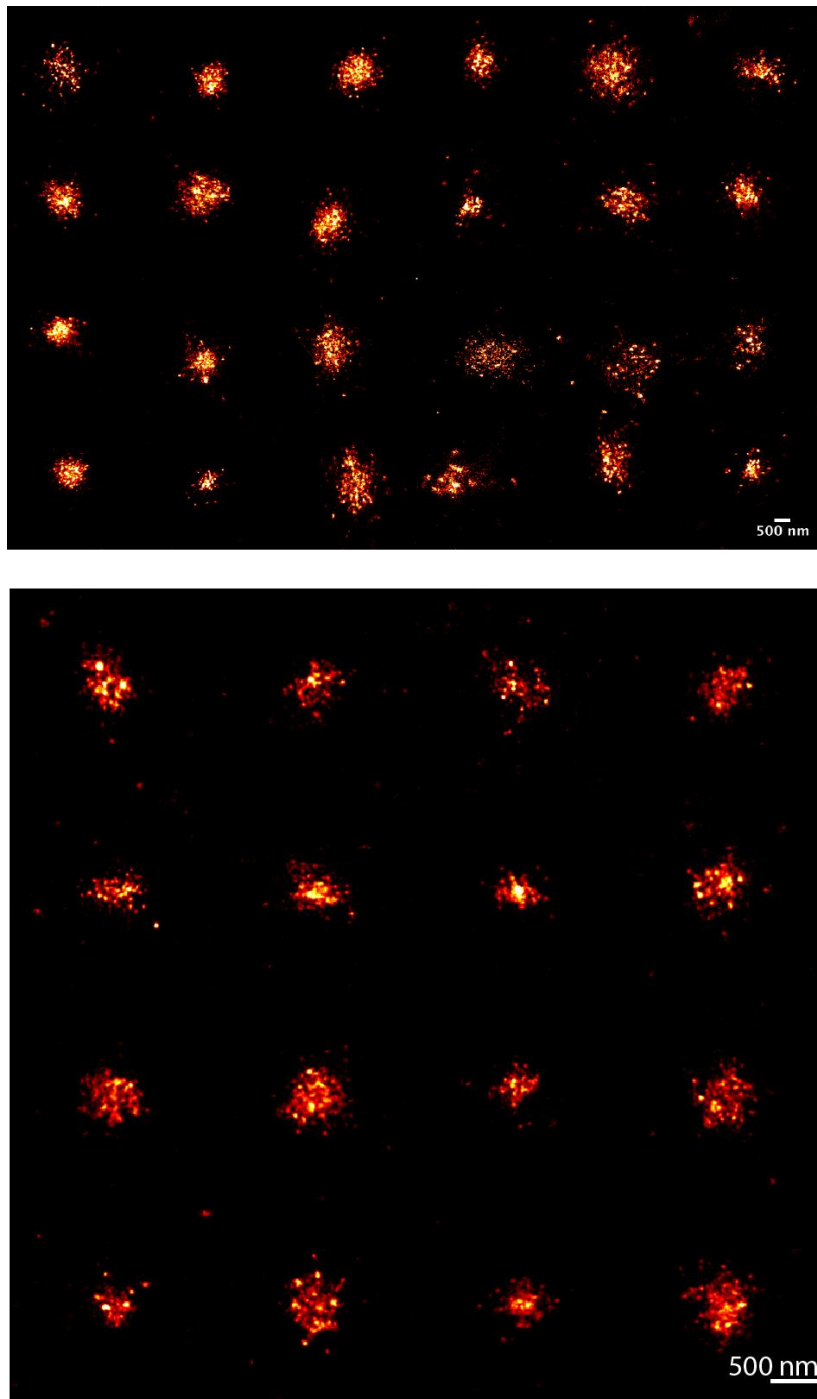
Supplementary Figure S8: Two-dimensional renderings of nanobody-labeled endogenous ASC specks imaged in three-dimensions using dSTROM. The ASC nanobody was conjugated with Alexa Fluor 647.
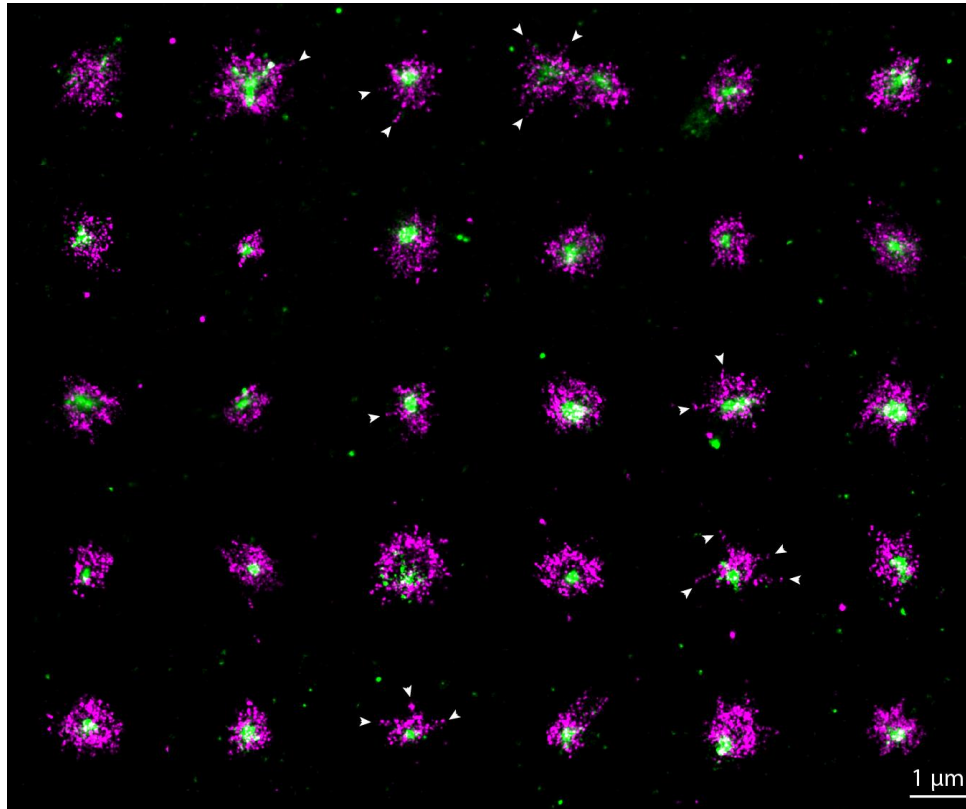
Supplementary Figure S9: Two-color super-resolution images of endogenous ASC specks. 2D STORM images were collected of ASC specks labeled with primary ASC antibody and secondary Alexa Fluor 647-conjugated F(ab')$_2$ fragment (magenta), and with DyLight 755-conjugated ASC nanobody (green). Arrows indicate filamentous extensions reaching out from the speck center. Data was obtained on a single cell preparation.

12

219

N

O

P

Q

R

S

T

13

220

U

Rg of speck

CC: 0.12

Pseudo time
Number of localizations within speck
/Number of localizations within entire cell

V

Speck area [μm^2]

CC: 0.12

Pseudo time
Number of localizations within speck
/Number of localizations within entire cell

W

Cell area [μm^2]

CC: -0.07

Pseudo time
Number of localizations within speck
/Number of localizations within entire cell

X

Speck density [localizations/μm^2]

CC: 0.09

Cell area [μm^2]

Y

Speck density [localizations/μm^2]

CC: 0.49

Integrated WF intensity of speck [counts]

Z

Speck density [localizations/μm^2]

CC: 0.86

Number of localizations within speck

14

221

Supplementary Figure S10: Correlation plots for various parameters obtained from the analysis of single cells. Panels A – E relate the number of localizations, the widefield intensity and the cell area with each other. Panels F – I depicts how the size of the speck depends on cell area and ASC concentration in the speck. Panels J - M depicts how the radius of gyration Rg, depends on cell area and ASC concentration in the speck. Panels N – X depict how various parameters depend on the pseudo time. Specifically, the amount of ASC in the speck (N, O), the speck area (P), the amount of ASC in the cytosol (Q, R), the cytosolic cluster density (S) and localization density (T), the radius of gyration, Rg (U), the speck area (V), the cell area (W) and the speck density (X). Panels Y – Z depict how the speck density depends on the amount of ASC in the speck. Data was obtained on three independent cell preparations. WF: widefield; CC: Pearson correlation coefficient

15
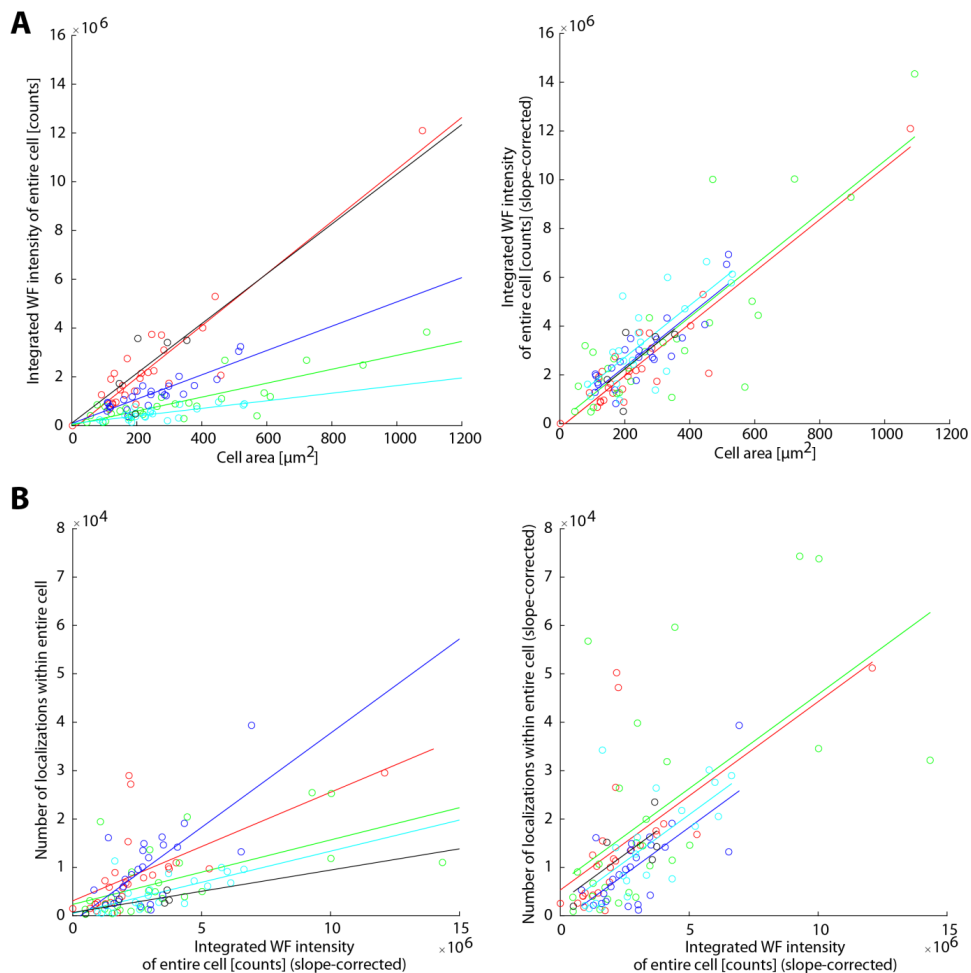
Supplementary Figure S11: Scatter plots corrections. The uncorrected and corrected scatter plots are shown to illustrate the correction to the integrated widefield intensity and localization data for differences between measurement days. Different sample preparations and dSTORM buffer conditions lead to variations in the widefield intensity and number of localizations with respect to ASC content. Hence, we used the correlation between ASC content and cell area to normalize the widefield intensity for different measurement days. For the number of localizations, we used the linear correlation between intensity and number of detected localizations. A) Scatter plot of the integrated widefield intensity of the entire cell against the cell area including the linear fits of individual measurement days (left panel) and the normalized integrated widefield intensity by adjusting the individual slopes to that of the measurement day with the steepest slope against the cell area (right panel). B) Scatter plot of the number of localizations within the entire cell against the normalized integrated widefield intensity of the entire cell including the linear fits of individual measurement days (left panel) and number of localizations within the entire cell normalized with respect to the measurement day with the steepest slope plotted against the normalized integrated widefield intensity of the entire cell (right panel) The individual measurement days are color-coded.

A

THP-1 caspase-1 KO
1st AB + AF647-2nd F_ab only

THP-1 caspase-1 KO
AF647-2nd F_ab only

THP-1 ASC KO
1st AB + AF647-2nd F_ab only

LPS + Nigericin

100 µm

B

THP-1 caspase-1 KO
1st AB + 2nd AF647-F_ab

THP-1 ASC KO
1st AB + 2nd AF647-F_ab

THP-1 caspase-1 KO
unstained

10 µm
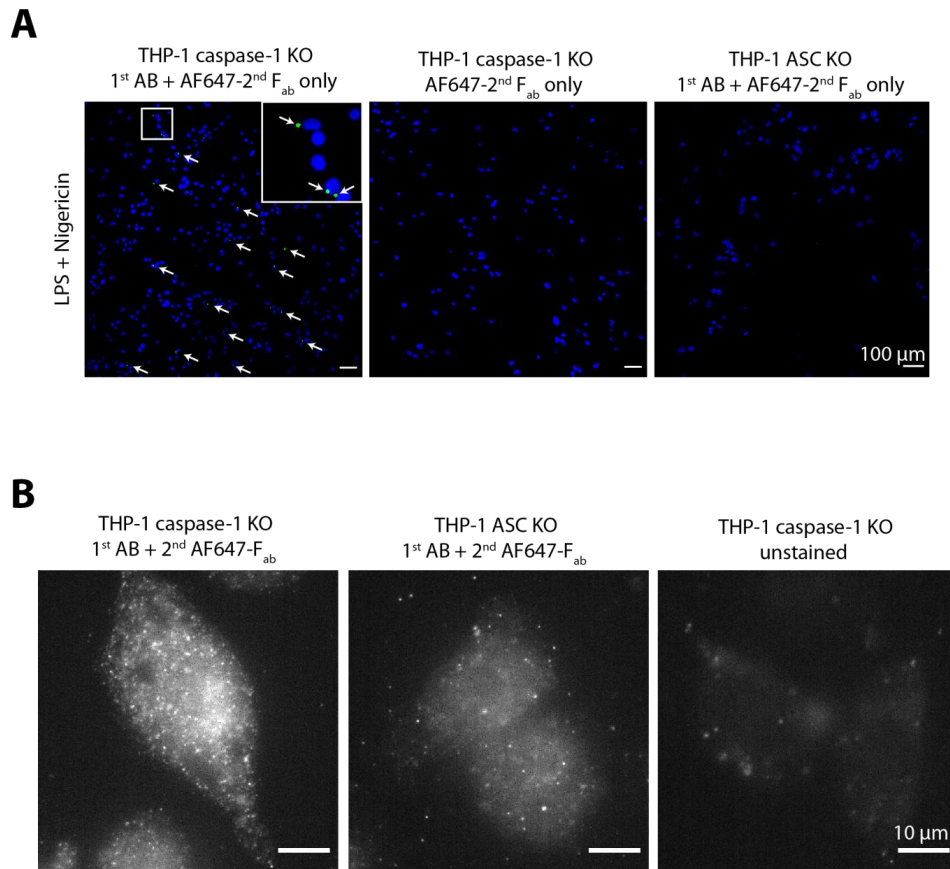
Supplementary Figure S12: Controls for the applied antibody and F(ab')₂ fragment staining in THP-1 knock-out (KO) cells. A) 10x overview images recorded on a confocal spinning disk microscope indicating the specificity of the applied ASC staining; green: Alexa Fluor 647 (summed projections of a recorded Z-stack are shown); blue: DNA-staining with DAPI. No ASC staining was observed in the absence of ASC or primary antibody. B) Higher magnification images (60x) verifying the specificity of ASC staining.
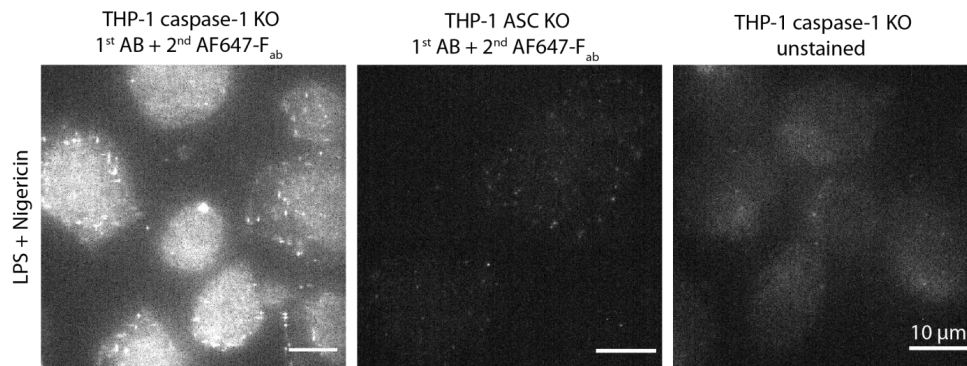
Supplementary Figure S13: Controls for the applied nanobody (AF647-conjugated) staining in THP-1 knock-out (KO) cells. Widefield images recorded at 60x magnification are shown. Stimulated and labeled cells are shown in the left panel. No unspecific signal was observed in the absence of ASC (middle panel) or in unstained cells (right panel) verifying the specificity of ASC staining. All images are displayed using the same intensity scale.

B.3   Paper 3: Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive DNA demethylation in mammals

Reproduced with permission from *Nature Communications* 2020, 11, 5972. Copyright 2020 Nature Publishing Group

nature
COMMUNICATIONS

ARTICLE

Check for updates

Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive DNA demethylation in mammals

Christopher B. Mulholland [1], Atsuya Nishiyama [2,9], Joel Ryan [1,9], Ryohei Nakamura[3], Merve Yiğit[1], Ivo M. Glück[4], Carina Trummer[1], Weihua Qin[1], Michael D. Bartoschek [1], Franziska R. Traube [5], Edris Parsa[5], Enes Ugur[1,6], Miha Modic[7], Aishwarya Acharya [1], Paul Stolz [1], Christoph Ziegenhain[8], Michael Wierer [6], Wolfgang Enard[8], Thomas Carell [5], Don C. Lamb [4], Hiroyuki Takeda [3], Makoto Nakanishi [2], Sebastian Bultmann [1,10 ✉] & Heinrich Leonhardt [1,10 ✉]

Genome-wide DNA demethylation is a unique feature of mammalian development and naïve pluripotent stem cells. Here, we describe a recently evolved pathway in which global hypo-methylation is achieved by the coupling of active and passive demethylation. TET activity is required, albeit indirectly, for global demethylation, which mostly occurs at sites devoid of TET binding. Instead, TET-mediated active demethylation is locus-specific and necessary for activating a subset of genes, including the naïve pluripotency and germline marker *Dppa3* (*Stella, Pgc7*). DPPA3 in turn drives large-scale passive demethylation by directly binding and displacing UHRF1 from chromatin, thereby inhibiting maintenance DNA methylation. Although unique to mammals, we show that DPPA3 alone is capable of inducing global DNA demethylation in non-mammalian species (Xenopus and medaka) despite their evolutionary divergence from mammals more than 300 million years ago. Our findings suggest that the evolution of *Dppa3* facilitated the emergence of global DNA demethylation in mammals.

[1] Department of Biology II and Center for Integrated Protein Science Munich (CIPSM), Human Biology and BioImaging, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany. [2] Division of Cancer Cell Biology, The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. [3] Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. [4] Physical Chemistry, Department of Chemistry, Center for Nanoscience, Nanosystems Initiative Munich and Center for Integrated Protein Science Munich, Ludwig-Maximilians-Universität München, Munich, Germany. [5] Center for Integrated Protein Science (CIPSM) at the Department of Chemistry, Ludwig-Maximilians-Universität München, Munich, Germany. [6] Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany. [7] The Francis Crick Institute and UCL Queen Square Institute of Neurology, London, UK. [8] Department of Biology II, Anthropology and Human Genomics, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany. [9] These authors contributed equally: Atsuya Nishiyama, Joel Ryan. [10] These authors jointly supervised: Sebastian Bultmann, Heinrich Leonhardt. ✉email: bultmann@bio.lmu.de; h.leonhardt@lmu.de

ARTICLE

During early embryonic development the epigenome undergoes massive changes. Upon fertilization, the genomes of highly specialized cell types—sperm and oocyte—need to be reprogrammed in order to obtain totipotency. This process entails decompaction of the highly condensed gametic genomes and global resetting of chromatin states to confer the necessary epigenetic plasticity required for the development of a new organism[1]. At the same time, the genome needs to be protected from the activation of transposable elements (TEs) abundantly present in vertebrate genomes[2]. Activation and subsequent transposition of TEs result in mutations that can have deleterious effects and are passed onto offspring if they occur in the germline during early development[2,3]. The defense against these genomic parasites has shaped genomes substantially[4,5].

Cytosine DNA methylation (5-methylcytosine (5mC)) is a reversible epigenetic mark essential for cellular differentiation, genome stability, and embryonic development in vertebrates[6]. Predominantly associated with transcriptional repression, DNA methylation has important roles in gene silencing, genomic imprinting, and X inactivation[7]. However, the most basic, conserved function of DNA methylation is the stable repression of TEs and other repetitive sequences[8]. Accordingly, the majority of genomic 5mC is located within these highly abundant repetitive elements. Global DNA methylation loss triggers the derepression of transposable and repetitive elements, which leads to genomic instability and cell death, highlighting the crucial function of vertebrate DNA methylation[9–14]. Hence, to ensure continuous protection against TE reactivation, global DNA methylation levels remain constant throughout the lifetime of non-mammalian vertebrates[15–18]. Paradoxically, mammals specifically erase DNA methylation during preimplantation development[19,20], a process that would seemingly expose the developing organism to the risk of genomic instability through the activation of TEs. DNA methylation also acts as an epigenetic barrier to restrict and stabilize cell fate decisions and thus constitutes a form of epigenetic memory. The establishment of pluripotency in mammals requires the erasure of epigenetic memory and as such, global hypomethylation is a defining characteristic of pluripotent cell types including naïve embryonic stem cells (ESCs), primordial germ cells (PGCs), and induced pluripotent stem cells (iPSCs)[21].

In animals, DNA methylation can be reversed to unmodified cytosine by two mechanisms; either actively by Ten-eleven translocation (TET) dioxygenase-mediated oxidation of 5mC in concert with the base excision repair machinery[22–25] or passively by a lack of functional DNA methylation maintenance during the DNA replication cycle[26,27]. Both active and passive demethylation pathways have been implicated in the genome-wide erasure of 5mC accompanying mammalian preimplantation development[28–34]. Despite the extensive conservation of the TET enzymes and DNA methylation machinery throughout metazoa[35], developmental DNA demethylation appears to be unique to placental mammals[19,36–43]. In contrast, 5mC patterns have been found to remain constant throughout early development in all non-mammalian vertebrates examined to date[15,44–48]. This discrepancy implies the existence of yet-to-be-discovered mammalian-specific pathways that orchestrate the establishment and maintenance of global hypomethylation.

Here, we use mouse embryonic stem cells (ESCs) cultured in conditions promoting naïve pluripotency[49–51] as a model to study global DNA demethylation in mammals. By dissecting the contribution of the catalytic activity of TET1 and TET2 to global hypomethylation, we find that TET-mediated active demethylation drives the expression of the Developmental pluripotency-associated protein 3 (DPPA3/PGC7/STELLA). We show that DPPA3 directly binds UHRF1 and triggers its release from chromatin, thereby inhibiting maintenance methylation and

causing global passive demethylation. Although DPPA3 is only found in mammals, we found that DPPA3 can also potently induce global demethylation when introduced into non-mammalian vertebrates. In summary, our study uncovers a novel TET-controlled and DPPA3-driven pathway for passive demethylation in naïve pluripotency in mammals.

Results

TET1 and TET2 indirectly protect the naïve genome from hypermethylation. Mammalian TET proteins, TET1, TET2, and TET3, share a conserved catalytic domain and the ability to oxidize 5mC but exhibit distinct expression profiles during development[52]. Naïve ESCs and the inner cell mass (ICM) of the blastocyst from which they are derived feature high expression of *Tet1* and *Tet2* but not *Tet3*[29,53–55]. To dissect the precise contribution of TET-mediated active DNA demethylation to global DNA hypomethylation in naïve pluripotency we generated isogenic *Tet1* (T1CM) and *Tet2* (T2CM) single as well as *Tet1/Tet2* (T12CM) double catalytic mutant mouse ESC lines using CRISPR/Cas-assisted gene editing (Supplementary Fig. 1). We derived two independent clones for each mutant cell line and confirmed the inactivation of TET1 and TET2 activity by measuring the levels of 5-hydroxymethylcytosine (5hmC), the product of TET-mediated oxidation of 5mC[22] (Supplementary Fig. 1i). While the loss of either *Tet1* or *Tet2* catalytic activity significantly reduced 5hmC levels, inactivation of both TET1 and TET2 resulted in the near total loss of 5hmC in naïve ESCs (Supplementary Fig. 1i) indicating that TET1 and TET2 account for the overwhelming majority of cytosine oxidation in naïve ESCs. We then used reduced representation bisulfite sequencing (RRBS) to determine the DNA methylation state of T1CM, T2CM, and T12CM ESCs as well as wild-type (wt) ESCs. All *Tet* catalytic mutant (T1CM, T2CM, and T12CM) cell lines exhibited severe DNA hypermethylation throughout the genome including promoters, gene bodies, and repetitive elements (Fig. 1a, b and Supplementary Fig. 2a). The increase in DNA methylation was particularly pronounced at LINE-1 (L1) elements of which 97%, 98%, and 99% were significantly hypermethylated in T1CM, T2CM, and T12CM ESCs, respectively (Supplementary Fig. 2b). This widespread DNA hypermethylation was reminiscent of the global increase in DNA methylation accompanying the transition of naïve ESCs to primed epiblast-like cells (EpiLCs)[54,56,57], which prompted us to investigate whether the DNA methylation signature in T1CM, T2CM, and T12CM ESCs resembles that of more differentiated cells. In line with this hypothesis, *Tet* catalytic mutant ESCs displayed DNA methylation levels similar to or higher than those of wt EpiLCs (Supplementary Fig. 2c). Moreover, hierarchical clustering and principal component analyses (PCA) of the RRBS data revealed that ESCs from *Tet* catalytic mutants clustered closer to wt EpiLCs than wt ESCs (Fig. 1c and Supplementary Fig. 2d). In fact, the vast majority of significantly hypermethylated CpGs in *Tet* catalytic mutant ESCs overlapped with those normally gaining DNA methylation during the exit from naïve pluripotency (Fig. 1d). In contrast, T1CM, T2CM, and T12CM transcriptomes are clearly clustered by differentiation stage, indicating that the acquisition of an EpiLC-like methylome was not due to premature differentiation (Supplementary Fig. 2e). When comparing our data to that of TET knockout ESCs[58], we found that the catalytic inactivation of the TET proteins caused a far more severe hypermethylation phenotype than the complete removal of the TET proteins (Supplementary Fig. 2f). Intriguingly, whereas TET1 and TET2 prominently associate with sites of active demethylation (Supplementary Fig. 2g), we found that the majority of sites hypermethylated in *Tet* catalytic mutant ESCs
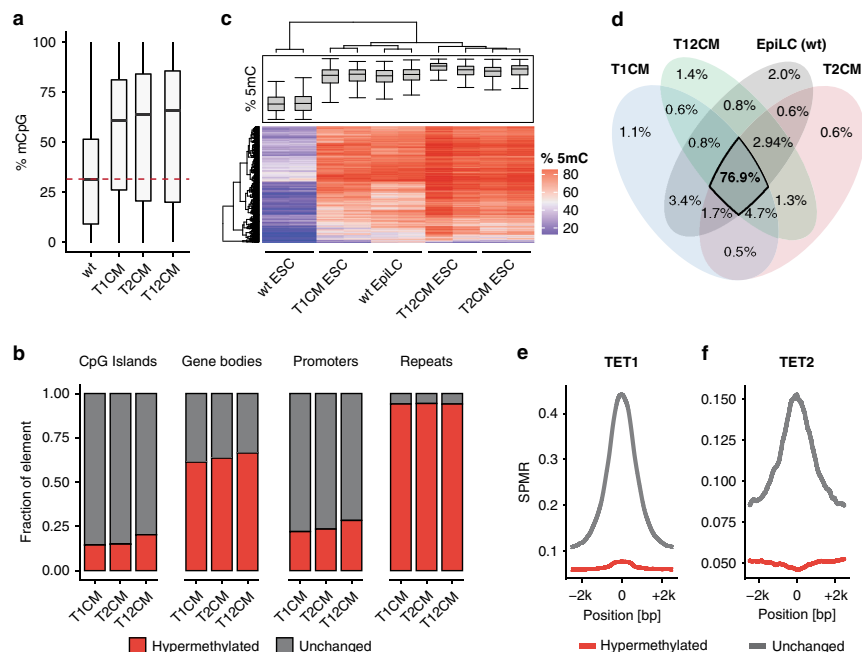
ARTICLE

Fig. 1 TET1 and TET2 prevent hypermethylation of the naïve genome. a Loss of TET catalytic activity leads to global DNA hypermethylation. Percentage of total 5mC as measured by RRBS. For each genotype, $n = 2$ biologically independent samples per condition. **b** Loss of TET catalytic activity leads to widespread DNA hypermethylation especially at repetitive elements. Relative proportion of DNA hypermethylation (q value < 0.05; absolute methylation difference >20%) at each genomic element in T1CM, T2CM, and T12CM ESCs compared to wt ESCs. **c** Heat map of the hierarchical clustering of the RRBS data depicting the top 2000 most variable 1 kb tiles during differentiation of wt ESCs to EpiLCs with $n = 2$ biologically independent samples per genotype and condition. **d** Venn diagram depicting the overlap of hypermethylated (compared to wt ESCs; q value < 0.05; absolute methylation difference >20%) sites among T1CM, T2CM, and T12CM ESCs and wt EpiLCs. **e**, **f** TET binding is not associated with DNA hypermethylation in TET mutant ESCs. Occupancy of (**e**) TET1[66] and (**f**) TET2[67] over 1 kb tiles hypermethylated (dark red) or unchanged (dark gray) in T1CM and T2CM ESCs, respectively (SPMR: Signal per million reads). In the boxplots in (**a**) and (**c**), horizontal black lines within boxes represent median values, boxes indicate the upper and lower quartiles, and whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge. In (**b**) and (**d**), the q-values were calculated with a two-sided Wald test followed by p-value adjustment using SLIM[209].

are not bound by either enzyme (Fig. 1e, f) suggesting that TET1 and TET2 maintain the hypomethylated state of the naïve methylome by indirect means.

TET1 and TET2 control *Dppa3* expression in a catalytically dependent manner. To explore how TET1 and TET2 might indirectly promote demethylation of the naïve genome, we first examined the expression of the enzymes involved in DNA methylation. Loss of TET catalytic activity was not associated with changes in the expression of *Dnmt1*, *Uhrf1*, *Dnmt3a*, and *Dnmt3b* nor differences in UHRF1 protein abundance, indicating the hypermethylation in *Tet* catalytic mutant ESCs is not caused by aberrant upregulation of DNA methylation machinery components (Fig. 2a, Supplementary Fig. 2h). To identify candidate factors involved in promoting global hypomethylation, we compared the transcriptome of hypomethylated wild-type ESCs with those of hypermethylated cells, which included wt EpiLCs as well as T1CM, T2CM, and T12CM ESCs (Fig. 2b). Among the 14 genes differentially expressed in hypermethylated cell lines, the naïve pluripotency factor, *Dppa3* (also known as *Stella* and *Pgc7*), stood out as an interesting candidate due to its reported

involvement in the regulation of global DNA methylation in germ cell development and oocyte maturation[59–62]. In contrast to the core components of the DNA (de)methylation machinery (DNMTs, UHRF1, TETs), which are conserved throughout metazoa, *Dppa3* is only present in mammals, suggesting it might also contribute to the mammal-specific hypomethylation in naïve pluripotency (Fig. 2c).

While normally highly expressed in naïve ESCs and only downregulated upon differentiation[63,64], *Dppa3* was prematurely repressed in T1CM, T2CM, and T12CM ESCs (Fig. 2d). The strongly reduced expression of *Dppa3* in TET mutant ESCs was accompanied by significant hypermethylation of the *Dppa3* promoter (Fig. 2e), consistent with reports demonstrating *Dppa3* to be one of the few pluripotency factors downregulated by promoter methylation upon differentiation in vitro and in vivo[51,63–65]. In contrast to the majority of genomic sites gaining methylation in TET mutant ESCs (Fig. 1e, f), hypermethylation at the *Dppa3* locus occurred at sites bound by both TET1 and TET2 (Fig. 2e)[66,67]. This hypermethylation overlapped with regions at which the TET oxidation product 5-carboxylcytosine (5caC) accumulates in Thymine DNA glycosylase (TDG)-knockdown ESCs (Fig. 2e)[68], indicating that the
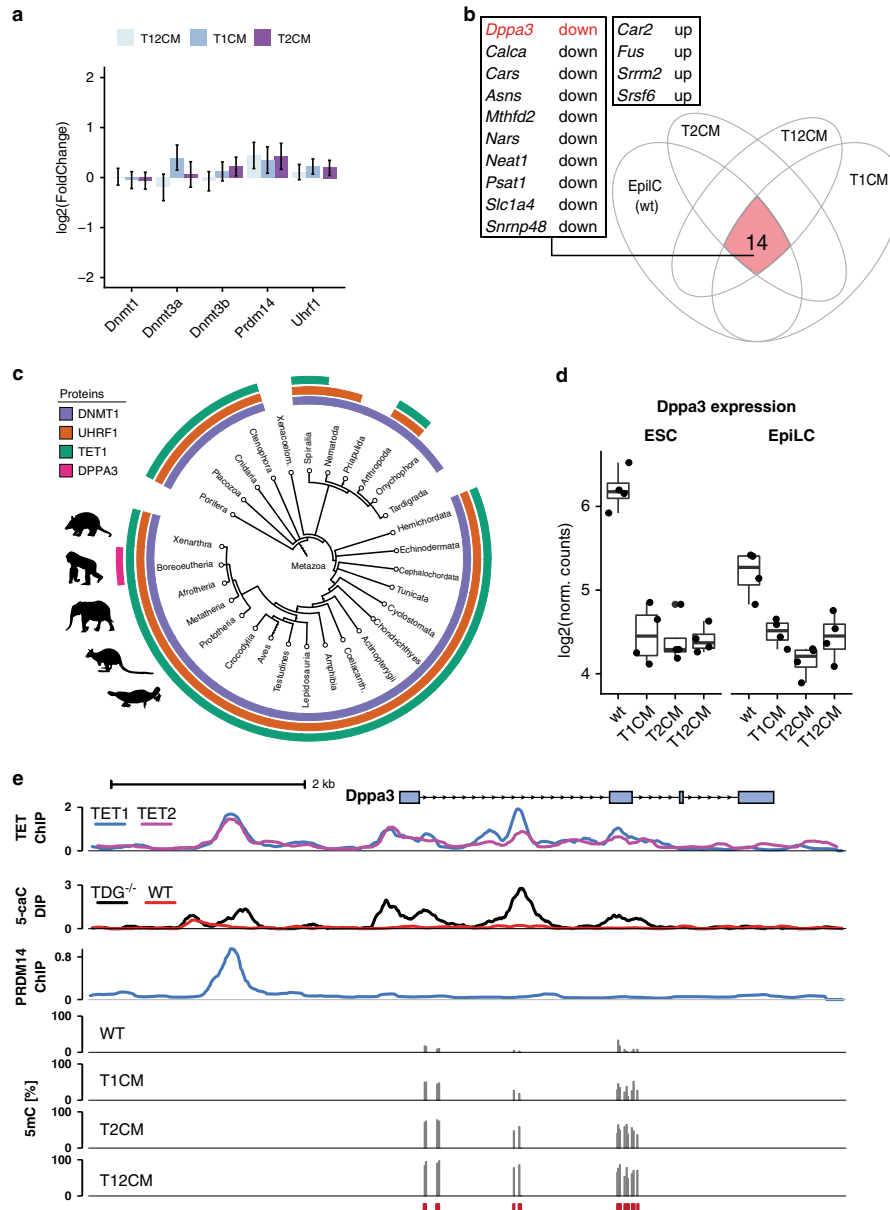
Dppa3 locus is a direct target of TET/TDG-mediated active DNA demethylation in ESCs. To test whether Dppa3 transcription can be induced by DNA demethylation, we analyzed RNA-seq data from conditional Dnmt1 KO ESCs[69]. In the absence of genome-wide DNA methylation, *Dppa3* levels more than doubled, thus confirming our results that the *Dppa3* promoter is sensitive to DNA methylation (Supplementary Fig. 2i).

In addition, *Dppa3* is also a direct target of PRDM14, a PR domain-containing transcriptional regulator known to promote the DNA hypomethylation associated with naïve pluripotency[50,70–72] (Fig. 2e). PRDM14 has been shown to recruit TET1 and TET2 to sites of active demethylation and establish global hypomethylation in naïve pluripotency[50,54,71–73]. As the expression of *Prdm14* was not altered in *Tet* catalytic mutant ESCs (Fig. 2a), we analyzed PRDM14

ARTICLE

Fig. 2 TET1 and TET2 catalytic activity is necessary for *Dppa3* expression. a Expression of genes involved in regulating DNA methylation levels in T1CM, T2CM, and T12CM ESCs as assessed by RNA-seq. Expression is given as the $\log_2$ fold-change compared to wt ESCs. Error bars indicate mean ± SD, $n = 4$ biological replicates. No significant changes observable (Likelihood ratio test). **b** *Dppa3* is downregulated upon loss of TET activity and during differentiation. Venn diagram depicting the overlap (red) of genes differentially expressed (compared to wt ESCs; adjusted p < 0.05) in T1CM, T2CM, T12CM ESCs, and wt EpiLCs. **c** Phylogenetic tree of TET1, DNMT1, UHRF1, and DPPA3 in metazoa. **d** *Dppa3* expression levels as determined by RNA-seq in the indicated ESC and EpiLC lines ($n = 4$ biological replicates). **e** TET proteins bind and actively demethylate the *Dppa3* locus. Genome browser view of the *Dppa3* locus with tracks of the occupancy (Signal pileup per million reads; (SPMR)) of TET1[66], TET2[67], and PRDM14[71] in wt ESCs, 5caC enrichment in wt vs. TDG$^{-/-}$ ESCs[68], and 5mC (%) levels in wt, T1CM, T2CM, and T12CM ESCs (RRBS). Red bars indicate CpGs covered by RRBS. In the boxplots in (**d**), horizontal black lines within boxes represent median values, boxes indicate the upper and lower quartiles, and whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge.

occupancy at the *Dppa3* locus using publicly available ChIP-seq data[71]. This analysis revealed that PRDM14 binds the same upstream region of *Dppa3* occupied by TET1 and TET2 (Fig. 2e). Taken together, these data suggest that TET1 and TET2 are recruited by PRDM14 to maintain the expression of *Dppa3* by active DNA demethylation.

DPPA3 acts downstream of TET1 and TET2 and is required to safeguard the naïve methylome. DPPA3 has been reported to both prevent and promote DNA demethylation depending on the cellular and developmental context[59,61,62,74–78]. However, the function of DPPA3 in naïve pluripotency, for which it is a well-established marker gene[63], remains unclear. To investigate the relationship between *Dppa3* expression and DNA hypomethylation in naïve pluripotency, we established *Dppa3* knockout (Dppa3KO) mouse ESCs (Supplementary Fig. 3a–c) and profiled their methylome by RRBS. Deletion of *Dppa3* led to severe global hypermethylation (Fig. 3a), with substantial increases in DNA methylation observed across all analyzed genomic features, including promoters, repetitive sequences, and imprinting control regions (ICRs) (Supplementary Fig. 3d–f). In particular, transposable elements experienced the most extensive gains in DNA methylation, with >90% of detected LINE and ERVs found hypermethylated in Dppa3KO ESCs (Supplementary Fig. 3e).

A principal component analysis of the RRBS data revealed that Dppa3KO ESCs clustered closer to wt EpiLCs and *Tet* catalytic mutant ESCs rather than wt ESCs (Fig. 3b). Furthermore, we observed a striking overlap of hypermethylated CpGs between *Tet* catalytic mutant and Dppa3KO ESCs (Fig. 3c), suggesting that DPPA3 and TETs promote demethylation at largely the same targets. A closer examination of the genomic distribution of overlapping hypermethylation in *Tet* catalytic mutant and Dppa3KO ESCs revealed that the majority (~90%) of hypermethylated events within repetitive elements are common to both cell lines (Fig. 3d and Supplementary Fig. 3g–j) and are globally correlated with heterochromatic histone modifications (Supplementary Fig. 3k). In contrast, only half of the observed promoter hypermethylation among all cell lines was dependent on DPPA3 (classified as "common", Fig. 3d and Supplementary Fig. 3h–j). This allowed us to identify a set of strictly TET-dependent promoters (N = 1573) (Fig. 3d, Supplementary Fig. 3i and Supplementary Data 1), which were enriched for developmental genes (Fig. 3e and Supplementary Data 2). Intriguingly, these TET-specific promoters contained genes (such as *Pax6*, *Foxa1*, and *Otx2*) that were recently shown to be conserved targets of TET-mediated demethylation during *Xenopus*, zebrafish, and mouse development[79].

DPPA3 appeared to act downstream of TETs as the global increase in DNA methylation in Dppa3KO ESCs was not associated with a reduction in 5hmC levels nor with a downregulation of TET family members (Fig. 3f and Supplementary Fig. 3l). In support of this notion, inducible over-expression of *Dppa3* (Supplementary Fig. 3m–o) completely rescued the observed hypermethylation phenotype at LINE-1

elements in T1CM as well as T2CM ESCs and resulted in a significant reduction of hypermethylation in T12CM cells (Fig. 3g). Strikingly, prolonged induction of *Dppa3* resulted in hypomethylation in wild-type as well as T1CM ESCs (Fig. 3g). Collectively, these results show that TET1 and TET2 activity contributes to genomic hypomethylation in naïve pluripotency by both direct and indirect pathways. Whereas direct and active demethylation protects a limited but key set of promoters, global DNA demethylation occurs as an indirect effect of *Dppa3* activation.

TET-dependent expression of DPPA3 regulates UHRF1 subcellular distribution and controls DNA methylation maintenance in embryonic stem cells. To investigate the mechanism underlying the regulation of global DNA methylation patterns by DPPA3, we first generated an endogenous DPPA3-HALO fusion ESC line to monitor the localization of DPPA3 throughout the cell cycle (Supplementary Fig. 4a, c). Previous studies have shown that DPPA3 binds H3K9me2[77] and that in oocytes its nuclear localization is critical to inhibit the activity of UHRF1[62], a key factor for maintaining methylation. Expecting a related mechanism to be present in ESCs, we were surprised to find that DPPA3 primarily localized to the cytoplasm of ESCs (Fig. 4a). Although present in the nucleus, DPPA3 was far more abundant in the cytoplasmic fraction (Supplementary Fig. 4e). Furthermore, DPPA3 did not bind to mitotic chromosomes indicating a low or absent chromatin association of DPPA3 in ESCs (Fig. 4a). To further understand the mechanistic basis of DPPA3-dependent DNA demethylation in ESCs, we performed FLAG-DPPA3 pulldowns followed by liquid chromatography tandem mass spectrometry (LC-MS/MS) to profile the DPPA3 interactome in naïve ESCs. Strikingly, among the 303 significantly enriched DPPA3 interaction partners identified by mass spectrometry, we found both UHRF1 and DNMT1 (Fig. 4b and Supplementary Data 3), the core components of the DNA maintenance methylation machinery[80,81]. A reciprocal immunoprecipitation of UHRF1 confirmed its interaction with DPPA3 in ESCs (Supplementary Fig. 4g). Moreover, GO analysis of the top 131 interactors of DPPA3 in ESCs showed the two most enriched GO terms to be related to DNA methylation (Supplementary Data 4). These findings are consistent with previous studies implicating DPPA3 in the regulation of maintenance methylation in other cellular contexts[60,62]. We also detected multiple members of the nuclear transport machinery in our DPPA3 interactome (highlighted in purple, Fig. 4b and Supplementary Data 3), which prompted us to investigate whether DPPA3 influences the subcellular localization of UHRF1. Surprisingly, biochemical fractionation experiments revealed UHRF1 to be present in both the nucleus and cytoplasm of naïve wt ESCs (Supplementary Fig. 4f). Despite comparable total UHRF1 protein levels in wt and Dppa3KO ESCs (Supplementary Fig. 4h), loss of DPPA3 completely abolished the cytoplasmic fraction of UHRF1 (Supplementary Fig. 4f).
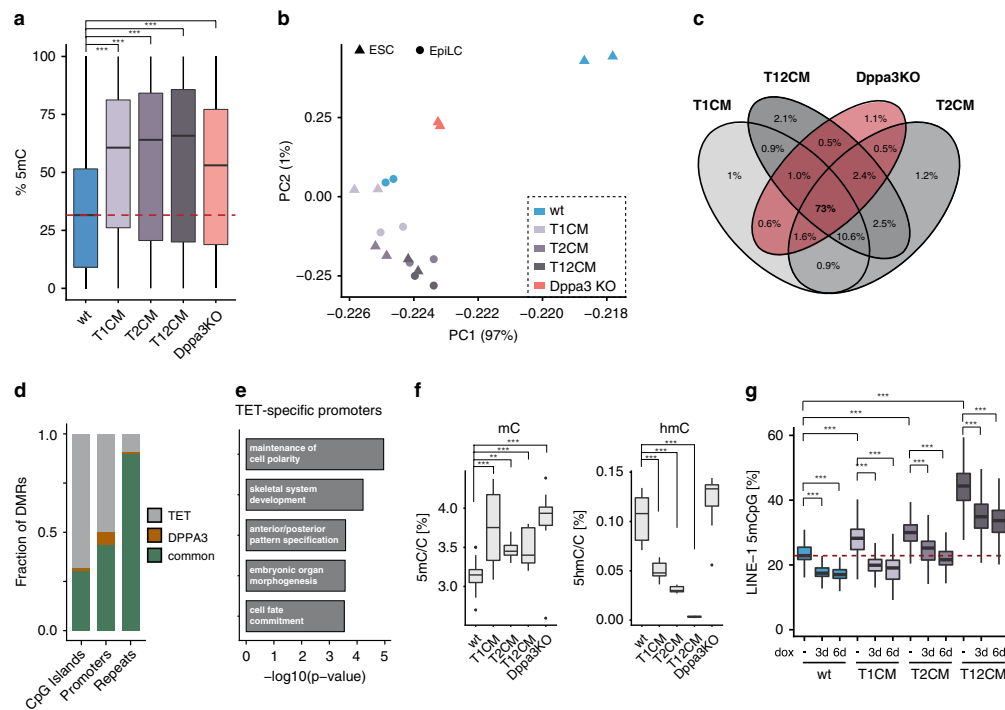
ARTICLE

Fig. 3 DPPA3 acts downstream of TET1 and TET2 to establish and preserve global hypomethylation. a *Dppa3* loss results in global hypermethylation. Percentage of total 5mC as measured by RRBS using $n = 2$ biologically independent samples per condition. **b** *Dppa3* prevents the premature acquisition of a primed methylome. Principal component (PC) analysis of RRBS data from wt, T1CM, T2CM, and T12CM ESCs and EpiLCs and Dppa3KO ESCs. **c** DPPA3 and TET proteins promote demethylation of largely similar targets. Venn Diagram depicting the overlap of hypermethylated sites among T1CM, T2CM, T12CM, and Dppa3KO ESCs. **d** *Dppa3* protects mostly repeats from hypermethylation. Fraction of hypermethylated genomic elements classified as TET-specific (only hypermethylated in TET mutant ESCs), DPPA3-specific (only hypermethylated in Dppa3KO ESCs), or common (hypermethylated in TET mutant and Dppa3KO ESCs). **e** Gene ontology (GO) terms associated with promoters specifically dependent on TET activity; adjusted p-values calculated using a two-sided Fisher's exact test followed by Benjamini-Hochberg correction for multiple testing. **f** TET activity remains unaffected in Dppa3KO ESCs. DNA modification levels for 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) as measured by mass spectrometry (LC-MS/MS) in wt ($n = 24$), T1CM ($n = 8$), T2CM ($n = 12$), T12CM ($n = 11$), Dppa3KO ($n = 12$) mESC biological replicates. **g** *Dppa3* expression can rescue the hypermethylation in TET mutant ESCs. DNA methylation levels at LINE-1 elements (%) as measured by bisulfite sequencing 0, 3, or 6 days after doxycycline (dox) induction of *Dppa3* expression using $n = 2$ replicates per condition. The dashed red line indicates the median methylation level of wt ESCs. In the boxplots in (**a**, **f** and **g**), horizontal black lines within boxes represent median values, boxes indicate the upper and lower quartiles, and whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge. In (**a**, **f**, and **g**), p-values were calculated using Welch's two-sided t-test: ***$p < 2e-16$. Source data are provided as a Source Data file.

As maintenance DNA methylation critically depends on the correct targeting and localization of UHRF1 within the nucleus[82–85], we asked whether TET-dependent regulation of DPPA3 might affect the subnuclear distribution of UHRF1. To this end, we tagged endogenous UHRF1 with GFP in wild-type (U1G/wt) as well as Dppa3KO and T12CM ESCs (U1G/Dppa3KO and U1G/T12CM, respectively) enabling us to monitor UHRF1 localization dynamics in living cells (Supplementary Fig. 4b, d). Whereas UHRF1-GFP localized to both the nucleus and cytoplasm of wt ESCs, UHRF1-GFP localization was solely nuclear in Dppa3KO and T12CM ESCs (Supplementary Fig. 4i, j). In addition, UHRF1 appeared to display a more diffuse localization in wt ESCs compared to Dppa3KO and T12CM ESCs, in which we observed more focal patterning of UHRF1 particularly at heterochromatic foci

(Supplementary Fig. 4i). To quantify this observation, we calculated the coefficient of variation (CV) of nuclear UHRF1-GFP among wt, Dppa3KO, and T12CM ESCs. The CV of a fluorescent signal correlates with its distribution, with low CV values reflecting more homogenous distributions and high CV values corresponding to more heterogeneous distributions[86,87]. Indeed, the pronounced focal accumulation of UHRF1-GFP observed in Dppa3KO and T12CM ESCs corresponded with a highly significant increase in the CV values of nuclear UHRF1-GFP compared with wt ESCs (Supplementary Fig. 4i, j).

To assess whether these differences in nuclear UHRF1 distribution reflected altered chromatin binding, we used fluorescence recovery after photobleaching (FRAP) to study the dynamics of nuclear UHRF1-GFP in wt, Dppa3KO, and T12CM ESCs. Our FRAP analysis revealed markedly increased UHRF1
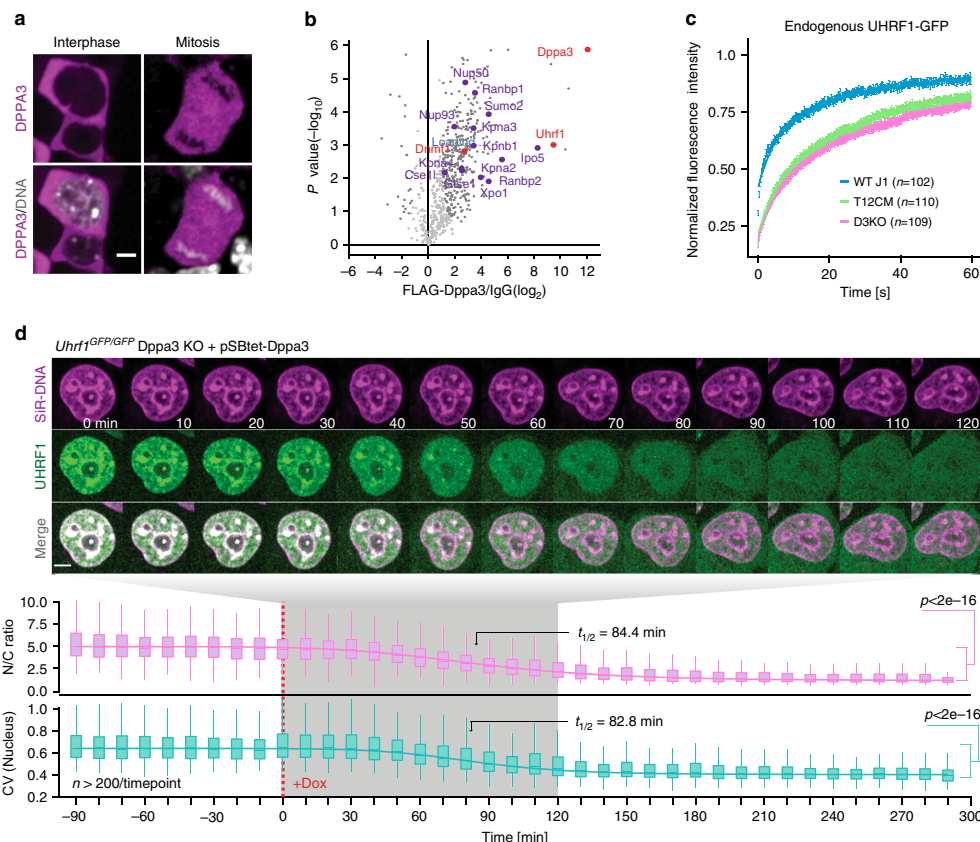
ARTICLE

Fig. 4 TET-dependent expression of DPPA3 alters UHRF1 localization and chromatin binding in naïve ESCs. a Localization of endogenous DPPA3-HALO in live ESCs counterstained with SiR-Hoechst (DNA). Representative result, $n \geq 4$. Scale bar: 5 µm. **b** Volcano plot from DPPA3-FLAG pulldowns in ESCs. Dark gray dots: significantly enriched proteins. Red dots: proteins involved in DNA methylation regulation. Purple dots: proteins involved in nuclear transport. anti-FLAG antibody: $n = 3$ biological replicates, IgG control antibody: $n = 3$ biological replicates. Statistical significance determined by performing a Student's t test with a permutation-based FDR of 0.05 and an additional constant S0 = 1. **c** FRAP analysis of endogenous UHRF1-GFP. Each genotype comprises the combined single-cell data from two independent clones acquired in two independent experiments. **d** Localization dynamics of endogenous UHRF1-GFP in response to *Dppa3* induction in U1G/D3KO + pSBtet-D3 ESCs with confocal timelapse imaging over 8 h (10 min intervals). $t = 0$ corresponds to start of *Dppa3* induction with doxycycline (+Dox). (*top panel*) Representative images of UHRF1-GFP and DNA (SiR-Hoechst stain) throughout confocal timelapse imaging. Scale bar: 5 µm. (middle panel) Nucleus to cytoplasm ratio (N/C ratio) of endogenous UHRF1-GFP signal. (bottom panel) Coefficient of variance (CV) of endogenous UHRF1-GFP intensity in the nucleus. (*middle and bottom panel*) N/C ratio and CV values: measurements in $n > 200$ single cells per time point (precise values can be found in the Source Data file), acquired at $n = 16$ separate positions. Curves represent fits of four parameter logistic (4PL) functions to the N/C ratio (pink line) and CV (green line) data. Live-cell imaging was repeated three times with similar results. In (**c**), the mean fluorescence intensity of n cells (indicated in the plots) at each timepoint are depicted as shaded dots. Error bars indicate mean ± SEM. Curves (solid lines) indicate double-exponential functions fitted to the FRAP data. In the boxplots in (**d**), darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge. *P*-values based on Welch's two-sided t test. Source data are provided as a Source Data file.

chromatin binding in both Dppa3KO and T12CM ESCs as demonstrated by the significantly slower recovery of UHRF1-GFP in these cell lines compared to wt ESCs (Fig. 4c and Supplementary Fig. 4k, l). These data confirmed the notion that the more pronounced focal patterning of nuclear UHRF1 observed in Dppa3KO and T12CM ESCs (Supplementary Fig. 4i, j) was indeed a consequence of increased UHRF1 chromatin binding. Interestingly, although strongly reduced compared to wt

ESCs, UHRF1 mobility was slightly higher in T12CM ESCs than Dppa3KO ESCs, consistent with a severe but not total loss of *Dppa3* in the absence of TET activity (Supplementary Fig. 4m). Induction of ectopic *Dppa3* rescued the cytoplasmic fraction of UHRF1 (N/C ratio: Fig. 4d) as well as the diffuse localization of nuclear UHRF1 in Dppa3KO ESCs (CV: Fig. 4d), which reflected a striking increase in the mobility of residual nuclear UHRF1-GFP as assessed by FRAP (Supplementary Figs. 4n and 5a, b).

ARTICLE

Our analysis also revealed that the nuclear export of UHRF1 and the inhibition of UHRF1 chromatin binding caused by *Dppa3* induction occur with almost identical kinetics (N/C $t_{1/2} = 84.4$ min; CV $t_{1/2} = 82.8$) (Fig. 4d). UHRF1 is required for the proper targeting of DNMT1 to DNA replication sites and therefore essential for DNA methylation maintenance[80,81]. We observed a marked reduction of both UHRF1 and DNMT1 at replication foci upon induction of *Dppa3*, indicating that DPPA3 promotes hypomethylation in naïve ESCs by impairing DNA methylation maintenance (Supplementary Fig. 5c, d). Ectopic expression of DPPA3 not only altered the subcellular distribution of endogenous UHRF1 in mouse ESCs (Fig. 4d and Supplementary Fig. 5e) but also in human ESCs suggesting evolutionary conservation of this mechanism among mammals (Supplementary Fig. 5f, g). Collectively, our results demonstrate that TET proteins control both the subcellular localization and chromatin binding of UHRF1 in naïve ESCs via the regulation of DPPA3 levels. Furthermore, these data show that DPPA3 is both necessary and sufficient for ensuring the nucleocytoplasmic translocation, diffuse nuclear localization, and attenuated chromatin binding of UHRF1 in ESCs.

DPPA3-mediated demethylation is achieved via inhibition of UHRF1 chromatin binding and attenuated by nuclear export. Our results demonstrated that *Dppa3* induction causes UHRF1 to be released from chromatin and exported to the cytoplasm near simultaneously (Fig. 4d, Supplementary Figs. 4n and 5a, b). In principle, either a reduction in the nuclear concentration of UHRF1 or the impairment of UHRF1 chromatin binding alone would suffice to compromise effective maintenance DNA methylation[84,88]. To dissect the contribution of these distinct modes of disrupting UHRF1 activity to DPPA3-mediated DNA demethylation in naïve ESCs, we generated inducible *Dppa3*-mScarlet expression cassettes (Supplementary Fig. 6a) harboring mutations to residues described to be critical for its nuclear export (ΔNES)[61] and the interaction with UHRF1 (KRR and R107E)[62], as well as truncated forms of DPPA3 found in zygotes, 1-60 and 61-150[78] (Fig. 5a). After introducing these *Dppa3* expression cassettes into U1GFP/ Dppa3KO ESCs, we used live-cell imaging to track each DPPA3 mutant's localization and ability to rescue the Dppa3KO phenotype (Fig. 5b). DPPA3-ΔNES and DPPA3 61-150, which both lacked a functional nuclear export signal, were retained in the nucleus (Fig. 5b). In contrast DPPA3-WT as well as the DPPA3-KRR, DPPA3-R107E, and DPPA3 1-60 mutants localized primarily to the cytoplasm (Fig. 5b), closely mirroring the localization of endogenous DPPA3 in naïve ESCs (Fig. 4a). However, all tested DPPA3 mutants failed to efficiently reestablish nucleocytoplasmic translocation of UHRF1 (Fig. 5b and Supplementary Fig. 6b), indicating that the DPPA3-UHRF1 interaction and nuclear export of DPPA3 are both required for the shuttling of UHRF1 from the nucleus to the cytoplasm in naïve ESCs.

Nevertheless, DPPA3-ΔNES and DPPA3 61-150 managed to significantly disrupt the focal pattern and heterochromatin association of UHRF1 within the nucleus, with DPPA3-ΔNES causing a more diffuse localization of nuclear UHRF1 than DPPA3-WT (Fig. 5b and Supplementary Fig. 6c). In contrast, the loss or mutation of residues critical for its interaction with UHRF1 compromised DPPA3's ability to effectively restore the diffuse localization of nuclear UHRF1 (Fig. 5b and Supplementary Fig. 6c). FRAP analysis revealed that the disruption or deletion of the UHRF1 interaction interface (DPPA3-KRR, DPPA3-R107E, DPPA3 1-60) severely diminished the ability of DPPA3 to release UHRF1 from chromatin (Fig. 5c and Supplementary Fig. 6f–k). On the other hand, the C-terminal half of DPPA3, lacking a nuclear export signal but retaining

UHRF1 interaction, came close to fully restoring the mobility of UHRF1 (Fig. 5c and Supplementary Fig. 6i–k). DPPA3-ΔNES mobilized UHRF1 to a greater extent than DPPA3-WT (Fig. 5c and Supplementary Fig. 6d, e, j, k), suggesting that active nuclear export might antagonize DPPA3-mediated inhibition of UHRF1 chromatin binding. Supporting this notion, chemical inhibition of nuclear export using leptomycin-B (LMB) significantly enhanced the inhibition of UHRF1 chromatin binding in U1G/D3KO ESCs expressing DPPA3-WT (Supplementary Fig. 5h–k). Taken together, our data show that the efficiency of DPPA3-dependent release of UHRF1 from chromatin requires its interaction with UHRF1 but not its nuclear export.

To further address the question whether the nucleocytoplasmic translocation of UHRF1 and impaired UHRF1 chromatin binding both contribute to DPPA3-mediated inhibition of DNA methylation maintenance, we assessed the ability of each DPPA3 mutant to rescue the hypermethylation of LINE-1 elements in Dppa3KO ESCs (Fig. 5d). Strikingly, DPPA3-ΔNES fully rescued the hypermethylation and achieved a greater loss of DNA methylation than DPPA3-WT, whereas DPPA3 mutants lacking the residues important for UHRF1 binding failed to restore low methylation levels (Fig. 5d). Overall, the ability of each DPPA3 mutant to reduce DNA methylation levels closely mirrored the extent to which each mutant impaired UHRF1 chromatin binding (Fig. 5c and Supplementary Fig. 6d–k). In line with the high mobility of UHRF1 achieved by the DPPA3-ΔNES, (Fig. 5c, Supplementary Figs. 5h–k and 6d, e, j, k), nuclear export is not only dispensable for DPPA3-mediated demethylation, but attenuates the ability of DPPA3 to inhibit maintenance methylation (Fig. 5d). Collectively, our findings demonstrate the inhibition of UHRF1 chromatin binding, as opposed to nucleocytoplasmic translocation of UHRF1, to be the primary mechanism by which DPPA3 drives hypomethylation in naïve ESCs.

DPPA3 binds nuclear UHRF1 with high affinity prompting its release from chromatin in ESCs. Next, we set out to investigate the mechanistic basis of DPPA3's ability to inhibit UHRF1 chromatin binding in naïve ESCs. DPPA3 has been reported to specifically bind H3K9me2[77], a histone modification critical for UHRF1 targeting[84,89,90]. These prior findings led us to consider two possible mechanistic explanations for DPPA3-mediated UHRF1 inhibition in naïve ESCs: (1) DPPA3 blocks access of UHRF1 to chromatin by competing in binding to H3K9me2, (2) DPPA3 directly or indirectly binds to UHRF1 and thereby prevents it from accessing chromatin.

To simultaneously assess the dynamics of both UHRF1 and DPPA3 under physiological conditions in live ES cells, we employed raster image correlation spectroscopy with pulsed interleaved excitation (PIE-RICS) (Fig. 6a). RICS is a confocal imaging method that measures the diffusive properties of fluorescently labeled molecules, and thereby also their binding, in living cells. Using images acquired on a laser scanning confocal microscope, spatiotemporal information of fluorescently labeled proteins can be extracted from the shape of the spatial autocorrelation function (SACF). A diffusive model is fitted to the SACF which yields the average diffusion coefficient, the concentration, and the fraction of quickly diffusing and slowly diffusing (in this case, bound) molecules[91]. If two proteins are labeled with distinct fluorophores and imaged simultaneously with separate detectors, the extent of their interaction can be extracted from the cross-correlation of their fluctuations using cross-correlation RICS (ccRICS) (Fig. 6a)[92].

We first measured the mobility of DPPA3-mScarlet variants expressed in U1GFP/D3KOs (Supplementary Fig. 7a, b). The RICS analysis revealed that, over the timescale of the
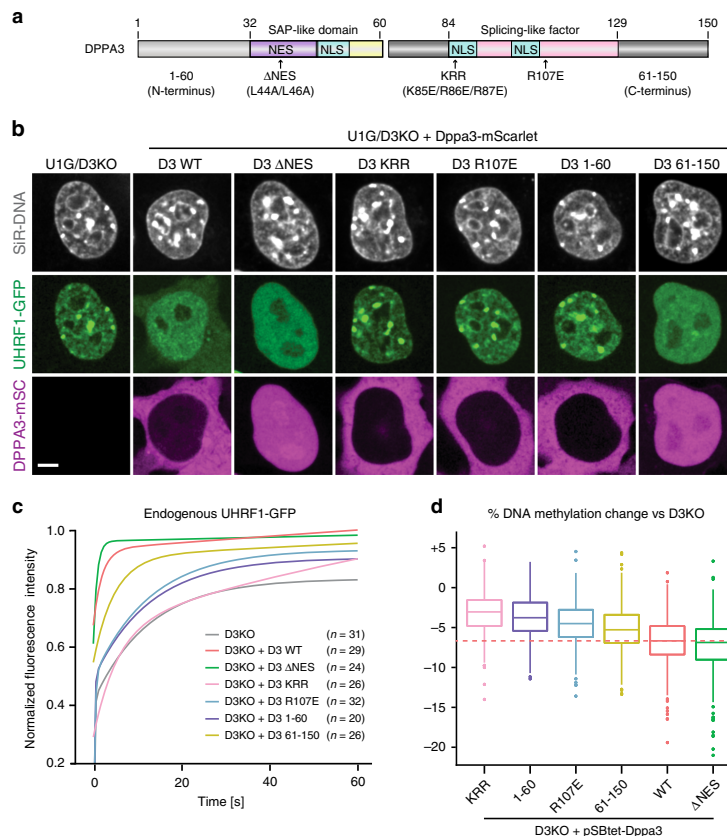
ARTICLE

Fig. 5 DPPA3-mediated demethylation is achieved via inhibition of UHRF1 chromatin binding and attenuated by nuclear export. a Schematic illustration of murine DPPA3 with the nuclear localization signals (NLS), nuclear export signal (NES), and predicted domains (SAP-like and splicing factor-like[210]) annotated. For the DPPA3 mutant forms used in this study, point mutations are indicated with arrows (ΔNES, KRR, R107E) and the two truncations are denoted by the middle break (1–60, left half; 61–150, right half). **b**, **c** Nuclear export and the N-terminus of DPPA3 are dispensable for disrupting focal UHRF1 patterning and chromatin binding in ESCs. **b** Representative confocal images illustrating the localization of endogenous UHRF1-GFP and the indicated mDPPA3-mScarlet fusions in live U1G/D3KO + pSB-D3-mSC ESCs after doxycycline induction. DNA counterstain: SiR-Hoechst. Scale bar: 5 μm. **c** FRAP analysis of endogenous UHRF1-GFP in U1G/D3KO ESCs expressing the indicated mutant forms of DPPA3. FRAP Curves (solid lines) indicate double-exponential functions fitted to the FRAP data acquired from *n* cells (shown in the plots). For single-cell FRAP data and additional quantification, see Supplementary Fig. 6d–k. **d** DPPA3-mediated inhibition of UHRF1 chromatin binding is necessary and sufficient to promote DNA demethylation. Percentage of DNA methylation change at LINE-1 elements (%) in D3KO ESCs after induction of the indicated mutant forms of *Dppa3* as measured by bisulfite sequencing of *n* = 4 biological replicates. In the boxplot in (**d**), horizontal lines within boxes represent median values, boxes indicate the upper and lower quartiles, whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge, and dots indicate outliers. Source data are provided as a Source Data file.

measurements, nuclear DPPA3-WT was predominantly unbound from chromatin and freely diffusing through the nucleus at a rate of 7.18 ± 1.87 μm²/s (Supplementary Fig. 7f). The fraction of mobile mDPPA3-mScarlet molecules was measured to be $88.4 \pm 5.2\%$ (Fig. 6f), validating the globally weak binding inferred from ChIP-Seq profiles[76]. These mobility parameters were largely unaffected by disruption of the UHRF1 interaction, with the DPPA3-KRR mutant behaving similarly to wild-type DPPA3 (Fig. 6f and Supplementary Fig. 7f). To rule out a potential competition between UHRF1 and DPPA3 for H3K9me2 binding, we next used RICS to determine if DPPA3 dynamics are altered

in the absence of UHRF1. For this purpose, we introduced the DPPA3-WT-mScarlet cassette into Uhrf1KO (U1KO) ESCs[93], in which free eGFP is expressed from the endogenous *Uhrf1* promoter (Supplementary Fig. 7c). However, neither the diffusion rate nor the mobile fraction of DPPA3 were appreciably altered in cells devoid of UHRF1, suggesting the high fraction of unbound DPPA3 to be unrelated to the presence of UHRF1 (Fig. 6f and Supplementary Fig. 7f). Overall, our RICS data demonstrate that, in contrast to zygotes[77], DPPA3 in ESCs lacks a strong capacity for chromatin binding, and, as such, is not engaged in competition with UHRF1 for chromatin binding.
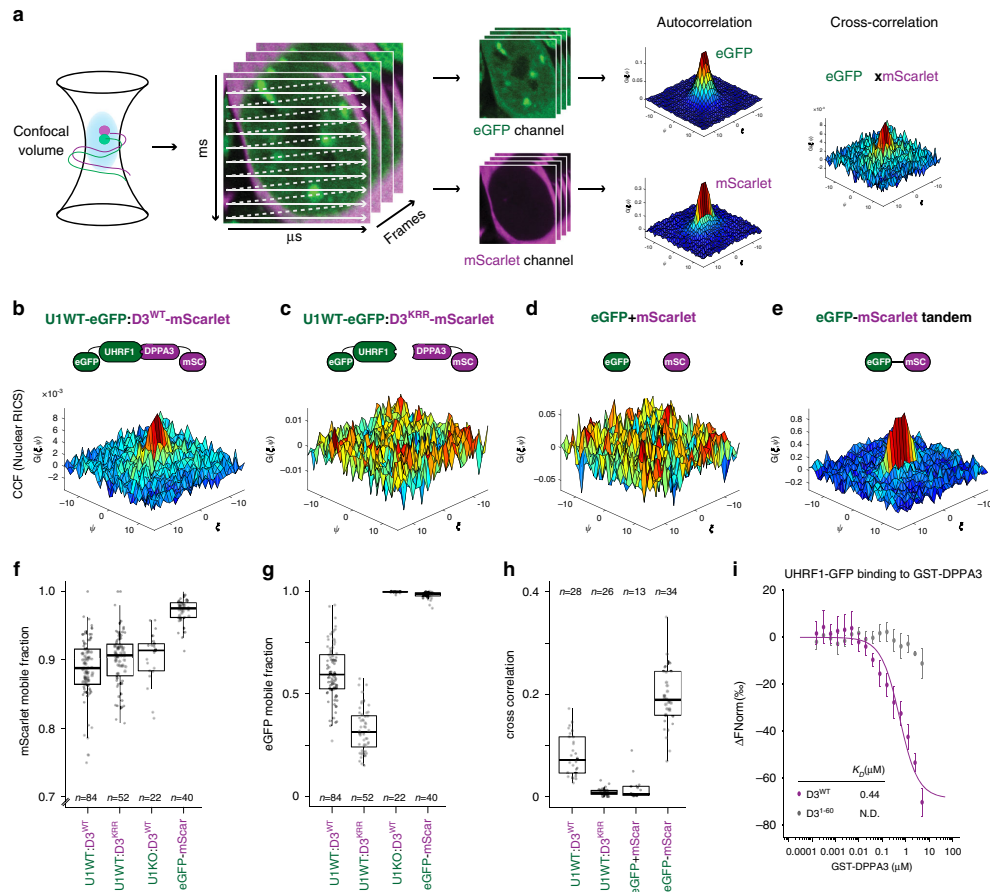
9

ARTICLE

Fig. 6 DPPA3 binds nuclear UHRF1 with high affinity prompting its release from chromatin in naïve ESCs. a Overview of RICS and ccRICS. Confocal image series are acquired on a laser scanning confocal microscope, containing spatiotemporal fluorescence information on the microsecond and millisecond timescales. A spatial autocorrelation function (SACF) is calculated from the fluorescence image and fit to a diffusive model. The cross-correlation of intensity between two channels is used to estimate the co-occurrence of two fluorescent molecules in live cells. The mean cross-correlation of the fluctuations is calculated and shown in the 3D plot color-coded according to the correlation value. **b–e** Representative plots of the spatial cross-correlation function (SCCF) between the depicted fluorescent molecules in cells from each cell line measured: (**b**) wild-type (U1WT:D3^{WT}) and (**c**) K85E/R85E/R87E DPPA3 mutant (U1WT:D3^{KRR}), and control ESCs expressing (**d**) free eGFP, free mScarlet (eGFP + mScarlet) and (**e**) an eGFP-mScarlet tandem fusion (eGFP-mScarlet). **f, g** Mobile fraction of (**f**) mScarlet and (**g**) eGFP species in the cell lines depicted in (**b**, **c**, and **e**) and in Uhrf1KO ESCs expressing free eGFP and wild-type DPPA3-mScarlet (U1KO:D3^{WT}). The mobile fraction was derived from a two-component model fit of the autocorrelation function. Data are pooled from three (U1WT:D3^{WT}, U1WT:D3^{KRR}) or two (U1KO:D3^{WT}, eGFP-mScar) independent experiments. **h** Mean cross-correlation values of mobile eGFP and mScarlet measured in the cell lines depicted in (**b–e**). The spatial lag in the x-dimension (sensitive to fast fluctuations) is indicated by ξ, and the spatial lag in the y-dimension (sensitive to slower fluctuations) is indicated by ψ. Data are pooled from two independent experiments. **i** Microscale thermophoresis measurements of UHRF1-eGFP binding to GST-DPPA3 WT (D3^{WT}) or GST-DPPA3 1–60 (D3$^{1-60}$). Error bars indicate the mean ± SEM of $n = 2$ technical replicates from $n = 4$ independent experiments. In (**f–h**), each data point represents the measured and fit values from a single cell where n = number of cells measured (indicated in the plots). In the boxplots, darker horizontal lines within boxes represent median values. The limits of the boxes indicate the upper and lower quartiles; whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge. Source data are provided as a Source Data file.

We next used RICS to analyze the dynamics of UHRF1-GFP in response to DPPA3 induction (Fig. 6a). In cells expressing DPPA3-KRR, RICS measurements revealed that only 32.4 ± 10% of UHRF1 is mobile, indicating that the majority of UHRF1 is chromatin-bound (Fig. 6g). In contrast, expression of wild-type DPPA3 leads to a dramatic increase in the mobile fraction of UHRF1 (60.6 ± 13.7% mobile fraction for UHRF1) (Fig. 6g and Supplementary Fig. 7g, h). Furthermore, the mobile fraction of

ARTICLE

UHRF1 increased as a function of the relative abundance of nuclear DPPA3 to UHRF1 (Supplementary Fig. 7i), thereby indicating a stoichiometric effect of DPPA3 on UHRF1 chromatin binding, consistent with physical interaction. Thus, these results demonstrate that DPPA3 potently disrupts UHRF1 chromatin binding in live ESCs and suggest its interaction with UHRF1 to be critical to do so.

To determine whether such an interaction is indeed present in the nuclei of live ESCs, we performed cross-correlation RICS (ccRICS) (Fig. 6a). We first validated ccRICS in ESCs by analyzing live cells expressing a tandem eGFP-mScarlet fusion (Fig. 6e and Supplementary Fig. 7d), or expressing both freely diffusing eGFP and mScarlet (Fig. 6d and Supplementary Fig. 7e). For the tandem eGFP-mScarlet fusion, we observed a clear positive cross-correlation indicative of eGFP and mScarlet existing in the same complex (Fig. 6e, h), as would be expected for an eGFP-mScarlet fusion. On the other hand, freely diffusing eGFP and mScarlet yielded no visible cross-correlation (Fig. 6d, h), consistent with two independent proteins that do not interact. Upon applying ccRICS to nuclear UHRF1-GFP and DPPA3-mScarlet, we observed a prominent cross-correlation between wild-type DPPA3 and the primarily unbound fraction of UHRF1 (Fig. 6b, h), indicating that mobilized UHRF1 exists in a high affinity complex with DPPA3 in live ESCs. In marked contrast, DPPA3-KRR and UHRF1-GFP failed to exhibit detectable cross-correlation (Fig. 6c, h), consistent with the DPPA3-KRR mutant's diminished capacity to bind[62] and mobilize UHRF1 (Fig. 5c and Supplementary Fig. 6f, j, k). Overall, these findings demonstrate that nuclear DPPA3 interacts with UHRF1 to form a highly mobile complex in naïve ESCs which precludes UHRF1 chromatin binding.

To determine whether the DPPA3-UHRF1 complex identified in vivo (Fig. 6h) corresponds to a high affinity direct interaction, we performed microscale thermophoresis (MST) measurements using recombinant UHRF1-GFP and DPPA3 proteins. MST analysis revealed a direct and high affinity (K_D: 0.44 μM) interaction between the DPPA3 WT and UHRF1 (Fig. 6i). No binding was observed for DPPA3 1-60, lacking the residues essential for interaction with UHRF1 (Fig. 6i). In line with the results obtained by ccRICS, these data support the notion that DPPA3 directly binds UHRF1 in vivo. Interestingly, the affinity of the UHRF1-DPPA3 interaction was comparable or even greater than that reported for the binding of UHRF1 to H3K9me3 or unmodified H3 peptides, respectively[94,95].

To better understand how UHRF1 chromatin loading is impaired by its direct interaction with DPPA3, we applied a fluorescent-three-hybrid (F3H) assay to identify the UHRF1 domain bound by DPPA3 in vivo (Supplementary Fig. 7j, k). In short, this method relies on a cell line harboring an array of *lac* operator binding sites in the nucleus at which a GFP-tagged "bait" protein can be immobilized and visualized as a spot. Thus, the extent of recruitment of an mScarlet-tagged "prey" protein to the nuclear GFP spot offers a quantifiable measure of the interaction propensity of the "bait" and "prey" proteins in vivo (Supplementary Fig. 7k)[96]. Using UHRF1-GFP domain deletions as the immobilized bait (Supplementary Fig. 7j, k), we assessed how the loss of each domain affected the recruitment of DPPA3-mScarlet to the GFP spot. In contrast to the other UHRF1 domain deletions, removal of the PHD domain essentially abolished recruitment of DPPA3 to the *lac* spot, demonstrating DPPA3 binds UHRF1 via its PHD domain in vivo (Supplementary Fig. 7l, m). The PHD of UHRF1 is essential for its recruitment to chromatin[88,95,97], ubiquitination of H3 and recruitment of DNMT1 to replication foci[82,83]. Thus, our in vivo results suggest that the high affinity interaction of DPPA3 with UHRF1's PHD domain precludes UHRF1 from binding

chromatin in ESCs, which is also supported by a recent report demonstrating that DPPA3 specifically binds the PHD domain of UHRF1 to competitively inhibit H3 tail binding in vitro[98].

DPPA3 can inhibit UHRF1 function and drive global DNA demethylation in distantly related, non-mammalian species. Whereas UHRF1 and TET proteins are widely conserved throughout plants and vertebrates[99,100], both early embryonic global hypomethylation[101] and the *Dppa3* gene are unique to mammals. Consistent with UHRF1's conserved role in maintenance DNA methylation, a multiple sequence alignment of UHRF1's PHD domain showed that the residues critical for the recognition of histone H3 are completely conserved from mammals to invertebrates (Fig. 7a). This prompted us to consider the possibility that DPPA3 might be capable of modulating the function of distantly related UHRF1 homologs outside of mammals. To test this hypothesis, we used amphibian (*Xenopus laevis*) egg extracts to assess the ability of mouse DPPA3 (mDPPA3) to interact with a non-mammalian form of UHRF1. Despite the 360 million years evolutionary distance between mouse and *Xenopus*[102], mDPPA3 not only bound *Xenopus* UHRF1 (xUHRF1) with high affinity (Fig. 7b, c and Supplementary Fig. 8a, b) it also interacted with xUHRF1 specifically via its PHD domain (Supplementary Fig. 8c–e). Moreover, the first 60 amino acids of DPPA3 were dispensable for its interaction with UHRF1 (Supplementary Fig. 8a, b). Interestingly, mutation to R107, reported to be critical for DPPA3's binding with mouse UHRF1[62], diminished but did not fully disrupt the interaction (Supplementary Fig. 8b, e). The R107E mutant retained the ability to bind the xUHRF1-PHD domain but exhibited decreased binding to xUHRF1-PHD-SRA under high-salt conditions (Supplementary Fig. 8e), suggesting that R107E changes the binding mode of mDPPA3 to xUHRF1, rather than inhibiting the complex formation. Considering the remarkable similarity between DPPA3's interaction with mouse and *Xenopus* UHRF1, we reasoned that the ability of DPPA3 to inhibit UHRF1 chromatin binding and maintenance DNA methylation might be transferable to *Xenopus*. To address this, we took advantage of a cell-free system derived from interphase *Xenopus* egg extracts to reconstitute DNA maintenance methylation[82]. Remarkably, recombinant mDPPA3 completely disrupted chromatin binding of both *Xenopus* UHRF1 and DNMT1 without affecting the loading of replication factors such as xCDC45, xRPA2, and xPCNA (Fig. 7d). We determined that the inhibition of xUHRF1 and xDNMT1 chromatin loading only requires DPPA3's C-terminus (61-150 a.a.) and is no longer possible upon mutation of R107 (R107E) (Supplementary Fig. 8h), in line with our results in mouse ESCs (Fig. 5d). Moreover, DPPA3-mediated inhibition of xUHRF1 chromatin loading resulted in the severe perturbation of histone H3 dual-monoubiquitylation (H3Ub2), which is necessary for the recruitment of DNMT1[82,83,103] (Supplementary Fig. 8f). To determine whether mDPPA3 can displace xUHRF1 already bound to chromatin, we first depleted *Xenopus* egg extracts of xDNMT1 to stimulate the hyper-accumulation of xUHRF1 on chromatin[82,104] and then added recombinant mDPPA3 after S-phase had commenced (Supplementary Fig. 8g). Under these conditions, both wild-type mDPPA3 and the 61-150 fragment potently displaced xUHRF1 from chromatin, leading to suppressed H3 ubiquitylation (Supplementary Fig. 8g).

We next assessed the effect of DPPA3 on *Xenopus* maintenance DNA methylation. Consistent with the severe disruption of xDNMT1 chromatin loading, both DPPA3 wild-type and 61–150 effectively abolished replication-dependent DNA methylation in *Xenopus* egg extracts (Fig. 7e). In contrast, DPPA3 1-60 and DPPA3 R107E, which both failed to suppress xUHRF1 and
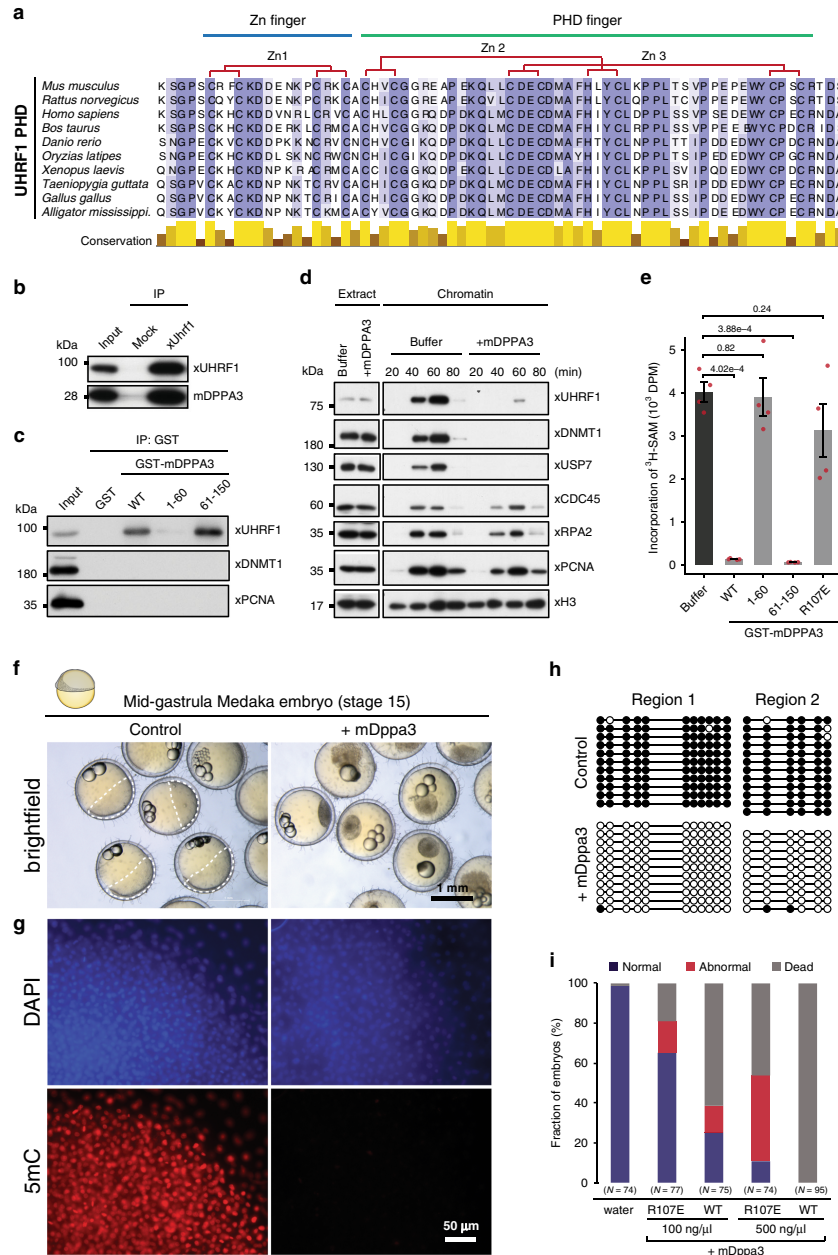
xDNMT1 binding, did not significantly alter maintenance DNA methylation activity (Fig. 7e and Supplementary Fig. 8d, e). Taken together, our data demonstrate DPPA3 to be capable of potently inhibiting maintenance DNA methylation in a non-mammalian system.

These findings raised the question whether a single protein capable of inhibiting UHRF1 function like DPPA3 could establish a mammalian-like global hypomethylation during the early embryonic development of a non-mammalian organism. To explore this possibility we turned to the biomedical model fish

                                                                    ARTICLE

Fig. 7 DPPA3 evolved in boreoeutherian mammals but also functions in lower vertebrates. a Protein sequence alignment of the PHD domain of the UHRF1 family. **b** Endogenous xUHRF1 binds mDPPA3. IPs were performed on *Xenopus* egg extracts incubated with FLAG-mDPPA3 using either a control (Mock) or anti-xUHRF1 antibody and then analyzed by immunoblotting using the indicated antibodies. Representative of $n = 3$ independent experiments. **c** GST-tagged mDPPA3 wild-type (WT), point mutant R107E, and truncations (1–60 and 61–150) were immobilized on GSH beads and incubated with *Xenopus* egg extracts. Bound proteins were analyzed using the indicated antibodies. Representative of $n = 3$ independent experiments. **d** Sperm chromatin was incubated with interphase *Xenopus* egg extracts supplemented with buffer (+buffer) or GST-mDPPA3 (+mDPPA3). Chromatin fractions were isolated and subjected to immunoblotting using the antibodies indicated. Representative of $n = 3$ independent experiments. **e** The efficiency of maintenance DNA methylation was assessed by the incorporation of radiolabelled methyl groups from S-[methyl-^{3}H]-adenosyl-L-methionine (^{3}H-SAM) into DNA purified from egg extracts. Disintegrations per minute (DPM). Error bars indicate mean ± SD calculated from $n = 4$ independent experiments. Depicted p-values based on Welch's two-sided t-test. **f** Representative images of developing mid-gastrula stage embryos (control injection) and arrested, blastula stage embryos injected with *mDppa3*. Injections were performed on one-cell stage embryos and images were acquired ~18 h after fertilization. **g** Immunofluorescence staining of 5mC in control and *mDppa3*-injected medaka embryos at the late blastula stage (~8 h after fertilization). Images are representative of $n = 3$ independent experiments. DNA counterstain: DAPI,4′,6-diamidino-2-phenylindole. **h** Bisulfite sequencing of two intergenic regions (Region 1: chr20:18,605,227-18,605,449, Region 2: chr20:18,655,561-18,655,825) in control and *mDppa3*-injected medaka embryos at the late blastula stage. **i** Percentage of normal, abnormal, or dead medaka embryos. Embryos were injected with wild-type *mDppa3* (WT) or *mDppa3* R107E (R107E) at two different concentrations (100 ng/μl or 500 ng/μl) or water at the one-cell stage and analyzed ~18 h after fertilization. N = number of embryos from $n = 3$ independent injection experiments. Source data are provided as a Source Data file.

medaka (*Oryzias latipes*), which does not exhibit genome-wide erasure of DNA methylation[105] and diverged from mammals 450 million years ago[102]. We injected medaka embryos with *Dppa3* mRNA at the one-cell stage and then tracked their developmental progression. Remarkably, medaka embryos injected with *Dppa3* failed to develop beyond the blastula stage (Fig. 7f) and exhibited a near-complete elimination of global DNA methylation as assessed by immunofluorescence and bisulfite sequencing (Fig. 7g, h). DPPA3-mediated DNA methylation loss was both dose dependent and sensitive to the R107E mutation, which induced only partial demethylation (Supplementary Fig. 8i). Interestingly, medaka embryos injected with DPPA3 R107E showed far fewer developmental defects than those injected with wild-type DPPA3 (Fig. 7i), suggesting that the embryonic arrest resulting from DPPA3 expression is truly a consequence of the global loss of DNA methylation. Taken together, these results demonstrate that mammalian DPPA3 can inhibit UHRF1 to drive passive demethylation in distant, non-mammalian contexts.

Discussion
In this study, we aimed to identify the mechanistic basis for the formation of genome-wide DNA hypomethylation unique to mammals. As the role of TET enzymes in active demethylation is well documented[106], we investigated their contribution to the hypomethylated state of naïve ESCs. Mutation of the catalytic core of TET enzymes caused—as expected—a genome-wide increase in DNA methylation but mostly at sites where TET proteins do not bind suggesting a rather indirect mechanism. Among the few genes depending on TET activity for expression in naïve ESCs and downregulated at the transition to EpiLCs was *Dppa3*. Demethylation at the *Dppa3* locus coincides with TET1 and TET2 binding and TDG-dependent removal of oxidized cytosine residues via base excision repair. DPPA3 in turn binds and displaces UHRF1 from chromatin and thereby prevents the recruitment of DNMT1 and the maintenance of DNA methylation in ESCs (see graphic summary in Fig. 8).

Despite long recognized as a marker of naïve ESCs resembling the inner cell mass[63,107], we provide, to our knowledge, the first evidence that DPPA3 directly promotes the genome-wide DNA hypomethylation characteristic of mammalian naïve pluripotency. This unique pathway, in which TET proteins indirectly cause passive demethylation, is based upon two uniquely mammalian innovations: the expression of TET genes in pluripotent cell types[53,79,108] and the evolution of the novel *Dppa3* gene, which is positioned within a conserved pluripotency gene cluster[109] and dependent on TET activity for expression. In support

of this novel pathway for passive demethylation, we found that TET mutant ESCs show a similar phenotype as Dppa3KO cells with respect to UHRF1 chromatin binding and hypermethylation and can be rescued by ectopic expression of *Dppa3*.

Our findings also provide the missing link to reconcile previous, apparently conflicting results. To date, three distinct mechanisms have been proposed for the global hypomethylation accompanying naïve pluripotency: TET-mediated active demethylation[51,54,58], impaired maintenance DNA methylation[58], and PRDM14-dependent suppression of methylation[50,51,71]. As a downstream target of both TETs and PRDM14 as well as a direct inhibitor of maintenance DNA methylation, DPPA3 mechanistically connects and integrates these three proposed pathways of demethylation (see graphic summary in Fig. 8).

Our mechanistic data showing DPPA3 to displace UHRF1 and DNMT1 from chromatin provide a conclusive explanation for the previous observation that global hypomethylation in naïve ESCs was accompanied by reduced levels of UHRF1 at replication foci[58]. The hypomethylated state of naïve ESCs has also been reported to be dependent on PRDM14[50,71], which has been suggested to promote demethylation by repressing de novo DNA methyltransferases[50,54,71,73]. However, recent studies have demonstrated that the loss of de novo methylation only marginally affects DNA methylation levels in mouse and human ESCs[58,110]. Interestingly, while the loss of *Prdm14* leads to global hypermethylation, it also causes downregulation of *Dppa3*[71,73,111]. Our results suggest that the reported ability of PRDM14 to promote hypomethylation in naïve ESCs largely relies on its activation of the *Dppa3* gene ultimately leading to an inhibition of maintenance methylation.

Of note, other epigenetic pathways such as suppression of H3K9me2 by MAD2L2 as well as eRNA dependent enhancer regulation also have been shown to positively regulate the transcription of *Dppa3*[109,112], and silencing of *Dppa3* has been shown to depend on Lin28a, TBX3, and intact DNA methylation maintenance[113–115]. Taken together, these findings suggest that *Dppa3* is regulated by a complex network of pathways to ensure proper timing of its expression in order to prevent unwanted global DNA demethylation.

The comparison of TET catalytic mutants and Dppa3KO ESCs allows us to distinguish TET-dependent passive DNA demethylation mediated by DPPA3 from *bona fide* active demethylation. We show that TET activity is indispensable for the active demethylation of a subset of promoters in naïve ESCs, especially those of developmental genes. These findings uncover two evolutionary and mechanistically distinct functions of TET catalytic activity.
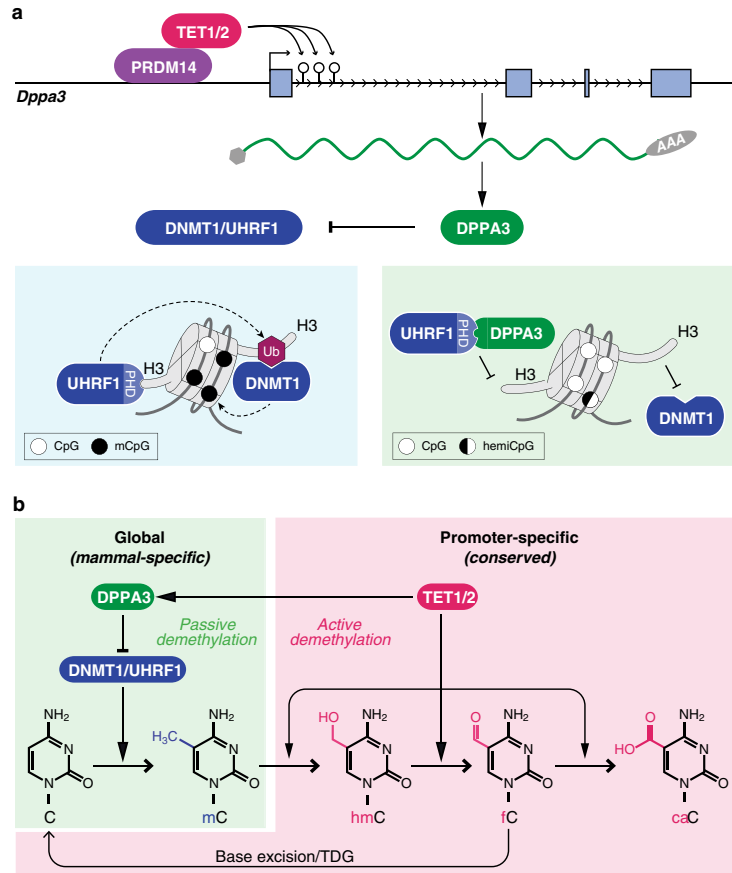
Fig. 8 Recent evolution of a TET-controlled and DPPA3-mediated pathway of DNA demethylation in boreoeutherian mammals. a In mammals, TET1 and TET2 are recruited by PRDM14 to the promoter of *Dppa3* where they promote active DNA demethylation and transcription of *Dppa3*. In most cellular contexts, high fidelity maintenance DNA methylation is guaranteed by the concerted activities of UHRF1 and DNMT1 at newly replicated DNA. Both the recruitment and activation of DNMT1 critically depend on the binding and ubiquitination of H3 tails by UHRF1. In naïve pluripotent cells, DPPA3 is expressed and inhibits maintenance DNA methylation by directly binding UHRF1 via its PHD domain and releasing it from chromatin. **b** TET1 and TET2 control DNA methylation levels by two evolutionary and mechanistically distinct pathways. TET-mediated active demethylation regulates focal DNA methylation states e.g. developmental genes and is evolutionarily conserved among vertebrates. The use of TET proteins to promote global demethylation appears to be specific to mammalian pluripotency and mediated by the recently evolved *Dppa3*.

Whereas TET-mediated active demethylation of developmental genes is evolutionarily conserved among vertebrates[79,116–118], the use of TET proteins to promote global demethylation appears to be specific to mammalian pluripotency[51,54,58] and mediated by the recently evolved *Dppa3* (Figs. 2c, 8).

In contrast to our findings in TET catalytic mutant ESCs, TET knockout ESCs do not appear to exhibit global hypermethylation[58]. This discrepancy might be explained by recent findings demonstrating that TET proteins influence global DNA methylation not only via their catalytic activity but also by their genomic binding[119,120]. Knockout of TET proteins results in a seemingly paradoxical loss of DNA methylation at repetitive elements like LINEs and LTRs due to a global redistribution of DNMT3A from heterochromatin to euchromatic sites previously

occupied by TETs. In contrast to TET KOs, disruption of TET catalytic activity would not be expected to affect global TET occupancy, presumably leaving DNMT3A genomic occupancy intact. Thus, the extensive hypermethylation occurring upon TET inactivation, but not TET knockout, could be attributable to both the preservation of TET binding as well as the enhanced loading of the DNA methylation machinery on chromatin in TET CM ESCs.

To date, our understanding of DPPA3's function in the regulation of DNA methylation has been clouded by seemingly conflicting reports from different developmental stages and cell types. DPPA3's ability to modulate DNA methylation was first described in the context of zygotes[61], where it was demonstrated to specifically protect the maternal genome from TET3-dependent

ARTICLE

demethylation[29,74,77]. In contrast, DPPA3 was later shown to prevent aberrant DNA hypermethylation during PGC specification[59], iPSC reprogramming[75], and oocyte maturation[62,121]. Whereas DPPA3 was shown to disrupt UHRF1 function by sequestering it to the cytoplasm in oocytes[62], we demonstrate that DPPA3-mediated nucleocytoplasmic translocation of UHRF1 is not only dispensable but actually attenuates DPPA3's promotion of hypomethylation in ESCs. Another example of development- and context-specific function of DPPA3 is its role in the regulation of imprinting. While DPPA3 has no impact on ICR methylation in oocytes[61,62], it is required to prevent the loss of both paternal and maternal imprints in zygotes[77]. In naïve ESCs, we found that the Dppa3 KO results in a gain of DNA methylation at ICRs. Although contrary to its zygotic role in protecting imprints from demethylation, our data is consistent with previous findings examining the effect of Dppa3 loss on iPSC generation, where imprints also became hypermethylated[75].

In light of our data from naïve ESCs, *Xenopus*, and medaka, DPPA3's capacity to directly bind UHRF1's PHD domain and thereby inhibit UHRF1 chromatin binding appears to be its most basal function. Considering that DPPA3 localization is highly dynamic during the different developmental time periods at which it is expressed[59,78,122], it stands to reason that its role in modulating DNA methylation might also be dynamically modulated by yet-to-be determined regulatory mechanisms. For example, immediately following fertilization, full-length DPPA3 is cleaved and its C-terminal domain is specifically degraded[78]. Interestingly, we identified this exact C-terminal stretch of DPPA3 to be necessary and sufficient for DPPA3's inhibition of maintenance DNA methylation. Thus, the precisely timed destruction of this crucial domain might offer an explanation for the differing roles of DPPA3 in regulating DNA methylation between oocytes and zygotes[62,74,77,121].

As the most basic and evolutionarily conserved function of DNA methylation is the repression of TEs[6], the post-fertilization wave of DNA demethylation found in mammals raises several fundamental questions. Considering the mutational risks associated with TE activity, why have mammals come to dispense with such a central genomic defense mechanism during early development? Whereas derepression of TEs leads to genomic instability and ultimately cell death in most cell types[9,10,13,14], TE activity is not only tolerated but increasingly appreciated to fulfill key roles in early mammalian development[123–129]. The activation of TEs, in particular endogenous retroviruses (ERVs), appears to be a conserved feature of early mammalian embryos[130], beginning after fertilization and continuing for the duration of gestation in the cells of the trophoblast and the placenta[131,132]. During mammalian evolution, the placenta emerged in the common ancestor of therian mammals, after the divergence from the egg-laying monotremes[133,134]. Accumulating evidence suggests ERVs facilitated the complex, network level changes necessary for the evolutionary emergence and diversification of placental viviparity[135–137]. By enabling embryos to directly regulate the allocation of maternal resources, placental viviparity creates unique evolutionary challenges absent in egg-laying species[138]. At the fetal–maternal interface, the interests of the mother and her offspring as well as those of the paternal and maternal genomes within the embryo are brought into conflict, unleashing a coevolutionary arms race for control of maternal resources and provisioning[139]. The existence of such an evolutionary struggle is perhaps best exemplified by the emergence of genomic imprinting, or parent-of-origin-specific gene expression, in therian mammals[140]. Transposons, particularly ERVs, have played an important role in the evolution of genomic imprinting as an adaption to

parental conflict; many of the cis-elements controlling imprinting status and, in some cases, even the imprinted genes themselves are derived from ERV insertions[141–143]. The retroviral origins of genomic imprinting are further illustrated by the use of conserved vertebrate host defense systems, namely DNA methylation and KRAB-ZFPs, to maintain imprint status[144,145]. In agreement with the parental conflict hypothesis, the evolution of more elaborated and invasive placentation has been accompanied by the expansion of genomic imprinting, with only 6 genes imprinted in marsupials compared with >100 in eutherians[146]. Indeed, the progressive co-option of retrotransposons over evolutionary time appears to have been a key driver in the transformation of a marsupial-like reproductive mode to the invasive and extended pregnancy of eutherians by facilitating the emergence of many of the unique, defining features of eutherian development such as the early allocation of the trophoblast cell lineage, invasive placentation, and suppression of the maternal immune response provoked by implantation[124,147–150]. Despite the importance of ERVs in eutherian development, the majority of ERV-derived regulatory elements, genes, and cis-elements controlling genomic imprinting are the result of evolutionarily recent and largely species-specific insertions[123,125,128,151–153].

How did eutherians come to rely on ERVs for so many aspects of their unique development? Such prolific ERV co-option among eutherians is proposed to have been a consequence of the evolution of precocious zygotic genome activation (ZGA) and an epigenetically permissive environment during early embryonic development[154,155]. It is tempting to speculate that post-fertilization demethylation was an important event in Eutherian evolution that contributed to the emergence and expansion of ERV/TE-based developmental regulation, including genomic imprints. Once ERV-derived genes and, in particular, regulatory networks acquired essential roles, mammalian preimplantation and placental development would have become "addicted" to the active transcription of ERVs[156]. Likewise, proper host development would require the establishment and maintenance of epigenetic states permissive for global ERV activity. In both mice and humans, the onset of ERV-dependent regulation coincides with a wave of genome-wide DNA demethylation, which commences upon fertilization and reaches its nadir in the ICM and trophectoderm of the blastocyst[19,40,157]. Whereas ERVs are silenced in the cells of the embryo proper by the wave of global de novo DNA methylation accompanying implantation, ERV activity and DNA hypomethylation persist in the trophoblast lineage throughout development[123,126,157–161]. Indeed, hypomethylation of the placenta relative to somatic cells appears to be conserved throughout Eutheria, despite dramatic differences in the embryonic and placental development among taxa[162].

As genome-wide DNA methylation is static throughout the lifecycle of most vertebrates, the evolution of novel mechanisms would have been required for the emergence of global DNA methylation erasure in the early embryonic development of eutherian mammals. DPPA3 may have arisen as a means to facilitate the early embryonic exposure of ERVs by neutralizing the host defense system of an ancestral eutherian mammal. In line with this notion, mouse embryos lacking *Dppa3* exhibit extensive genome-wide hypermethylation and undergo developmental arrest before the blastocyst stage as a result of impaired ERV activation and ZGA failure[62,76]. As *Dppa3* orthologs exhibit similar patterns of early embryonic expression in mice, humans, marmosets, cows, sheep, and pigs[163–167], it is plausible that function of DPPA3 during development is broadly conserved among mammals. However, our analysis identified *Dppa3* orthologs to be present in only a single clade of placental mammals, namely Boreoeutheria (Fig. 2c).

This raises the question whether eutherian lineages that lack DPPA3 also erase their methylomes and if so how? Pre-implantation DNA demethylation has been documented in every boreoeutherian species tested to date (e.g. mice, humans, monkeys, pigs, cows, sheep, rabbits)[19,36–43], however early embryonic DNA methylation dynamics have not been investigated in Eutherian lineages other than Boreoeutheria, i.e Afrotheria and Xenarthra, not to mention the more distant marsupial and monotreme groups. Likewise, the functional importance of ERV activity in early developmental and placental gene expression programs has also only been demonstrated in boreoeutherian species. Thus, it is currently wholly unclear whether global DNA demethylation and ERV-dependent regulatory networks are even present, let alone important for early embryonic and trophoblast development outside of Boreoeutheria. Follow-up studies that investigate the origins of *Dppa3* and whether a similar ERV-based rewiring of early development may have occurred in other, not yet studied branches of vertebrates, are needed to understand how global DNA demethylation shaped the evolution of placental mammals.

Methods

Cell culture. Naïve J1 mouse ESCs were cultured and differentiated into EpiLCs using an established protocol[168,169]. In brief, for both naïve ESCs and EpiLCs defined media was used, consisting of N2B27: 50% neurobasal medium (Life Technologies), 50% DMEM/F12 (Life Technologies), 2 mM L-glutamine (Life Technologies), 0.1 mM β-mercaptoethanol (Life Technologies), N2 supplement (Life Technologies), B27 serum-free supplement (Life Technologies), 100 U/mL penicillin, and 100 µg/mL streptomycin (Sigma). Naïve ESCs were maintained on flasks treated with 0.2% gelatin in defined media containing 2i (1 µM PD032591 and 3 µM CHIR99021 (Axon Medchem, Netherlands)), 1000 U/mL recombinant leukemia inhibitory factor (LIF, Millipore), and 0.3% BSA (Gibco) for at least three passages before commencing differentiation. To differentiate naïve ESCs into Epiblast-like cells (EpiLCs), flasks were first pre-treated with Geltrex (Life Technologies) diluted 1:100 in DMEM/F12 (Life Technologies) and incubated at 37 °C overnight. Naïve ESCs were plated on Geltrex-treated flasks in defined medium containing 10 ng/mL Fgf2 (R&D Systems), 20 ng/mL Activin A (R&D Systems) and 0.1× Knockout Serum Replacement (KSR) (Life Technologies). Media was changed after 24 h and EpiLCs were harvested for RRBS and RNA-seq experiments after 48 h of differentiation.

For CRISPR-assisted cell line generation, mouse ESCs were maintained on 0.2% gelatin-coated dishes in Dulbecco's modified Eagle's medium (Sigma) supplemented with 16% fetal bovine serum (FBS, Sigma), 0.1 mM ß-mercaptoethanol (Invitrogen), 2 mM L-glutamine (Sigma), 1× MEM Non-essential amino acids (Sigma), 100 U/mL penicillin, 100 µg/mL streptomycin (Sigma), homemade recombinant LIF tested for efficient self-renewal maintenance, and 2i (1 µM PD032591 and 3 µM CHIR99021 (Axon Medchem, Netherlands)).

Human ESCs (line H9) were maintained in mTeSR1 medium (05850, STEMCELL Technologies) on Matrigel-coated plates (356234, Corning) prepared by 1:100 dilution, and 5 ml coating of 10 cm plates for 1 h at 37 °C. Colonies were passaged using the gentle cell dissociation reagent (07174, StemCell Technologies).

All cell lines were regularly tested for Mycoplasma contamination by PCR.

Sleeping beauty constructs. To generate the sleeping beauty donor vector with an N-terminal 3xFLAG tag and a fluorescent readout of doxycycline induction, we first used primers with overhangs harboring SfiI sites to amplify the IRES-DsRed-Express from pIRES2-DsRed-Express (Clontech)(Supplementary Data 5). This fragment was then cloned into the NruI site in pUC57-GentR via cut-ligation to generate an intermediate cloning vector pUC57-SfiI-IRES-DsRed-Express-SfiI. A synthesized gBlock (IDT, Coralville, IA, USA) containing Kozak-BIO-3XFLAG-AsiSI-NotI-V5 was cloned into the Eco47III site of the intermediate cloning vector via cut-ligation. The luciferase insert from pSBtet-Pur[170] (Addgene plasmid #60507) was excised using SfiI. The SfiI-flanked Kozak-BIO-3XFLAG-AsiSI-NotI-V5-IRES-DsRed-Express cassette was digested out of the intermediate cloning vector using SfiI and ligated into the pSBtet-Pur vector backbone linearized by SfiI. The end result was the parental vector, pSBtet-3xFLAG-IRES-DsRed-Express-PuroR. The pSBtet-3x-FLAG-mScarlet-PuroR vector was constructed by inserting a synthesized gBlock (IDT, Coralville, IA, USA) containing the SfiI-BIO-3XFLAG-AsiSI-NotI-mScarlet sequence into the SfiI-linearized pSBtet-Pur vector backbone using Gibson assembly[171]. For *Dppa3* expression constructs, the coding sequence of wild-type and mutant forms of *Dppa3* were synthesized as gBlocks (IDT, Coralville, IA, USA) and inserted into the pSBtet-3xFLAG-IRES-DsRed-Express-PuroR vector (linearized by AsiSI and NotI) using Gibson assembly. To produce the *Dppa3*-mScarlet fusion expression constructs, wild-type and mutant forms of *Dppa3* were amplified from pSBtet-3xFLAG-Dppa3-IRES-DsRed-Express-PuroR constructs using primers with overhangs homologous to the AsiSI and NotI

restriction sites of the pSBtet-3x-FLAG-mScarlet-PuroR vector (Supplementary Data 5). Wild-type and mutant *Dppa3* amplicons were subcloned into the pSBtet-3x-FLAG-mScarlet-PuroR vector (linearized with AsiSI and NotI) using Gibson assembly.

For experiments involving the SBtet-3xFLAG-*Dppa3* cassette, all inductions were performed using 1 µg/mL doxycycline (Sigma-Aldrich). The DPPA3-WT construct was able to rescue the cytoplasmic localization and chromatin association of UHRF1 indicating that C-terminally tagged DPPA3 remains functional (Fig. 5b–d).

CRISPR/Cas9 genome engineering. For the generation of *Tet1*, *Tet2*, and *Tet1/Tet2* catalytic mutants, specific gRNAs targeting the catalytic center of *Tet1* and *Tet2* (Supplementary Data 5) were cloned into a modified version of the SpCas9-T2A-GFP/gRNA plasmid (px458[172], Addgene plasmid #48138), where we fused a truncated form of human Geminin (hGem) to SpCas9 in order to increase homology-directed repair efficiency[173] generating SpCas9-hGem-T2A-GFP/gRNA.

To generate *Tet1* and *Tet2* catalytic mutant targeting donors, 200 bp single-stranded DNA oligonucleotides carrying the desired HxD mutations (*Tet1*: H1652Y and D1654A, *Tet2*: H1304Y and D1306A) and ~100 bp homology arms were synthesized (IDT, Coralville, IA, USA) (Supplementary Data 5). For targetings in wild-type J1 ESCs, cells were transfected with a 4:1 ratio of donor oligo and SpCas9-hGem-T2A-GFP/gRNA construct. Positively transfected cells were isolated based on GFP expression using fluorescence-activated cell sorting (FACS) and plated at clonal density in ESC media 2 days after transfection. After 5–6 days, single colonies were picked and plated on 96-well plates. These plates were then duplicated 2 days later and individual clones were screened for the desired mutation by PCR followed by restriction fragment length polymorphism (RFLP) analysis. Cell lysis in 96-well plates, PCR on lysates, and restriction digests were performed as previously described[169]. The presence of the desired *Tet1* and/or *Tet2* catalytic mutations in putative clones was confirmed by Sanger sequencing.

As C-terminally tagged GFP labeled UHRF1 transgenes were shown to be able to rescue U1KO[83], the tagging of endogenous *Uhrf1* was also performed at the C-terminus. For insertion of the HALO or eGFP coding sequence into the endogenous *Dppa3* and *Uhrf1* loci, respectively, *Dppa3* and *Uhrf1* specific gRNAs were cloned into SpCas9-hGem-T2A-Puromycin/gRNA vector, which is a modified version of SpCas9-T2A-Puromycin/gRNA vector (px459;[172] Addgene plasmid #62988) similar to that described above. To construct the homology donors plasmids, gBlocks (IDT, Coralville, IA, USA) were synthesized containing either the HALO or eGFP coding sequence flanked by homology arms with ~200-400 bp homology upstream and downstream of the gRNA target sequence at the *Dppa3* or *Uhrf1* locus, respectively, and then cloned into the NruI site of pUC57-GentR via cut-ligation. ESCs were transfected with equimolar amounts of gRNA and homology donor vectors. Two days after transfection, cells were plated at clonal density and subjected to a transient puromycin selection (1 µg/mL) for 40 h. After 5-6 days, ESCs positive for HALO or eGFP integration were isolated via fluorescence-activated cell sorting (FACS) and plated again at clonal density in ESC media. After 4–5 days, colonies were picked and plated on Optical bottom µClear 96-well plates and re-screened for the correct expression and localization of eGFP or HALO using live-cell spinning-disk confocal imaging. Clones were subsequently genotyped using the aforementioned cell lysis strategy and further validated by Sanger sequencing[169].

To generate *Dppa3* knockout cells, the targeting strategy entailed the use of two gRNAs with target sites flanking the *Dppa3* locus to excise the entire locus on both alleles. gRNA oligos were cloned into the SpCas9-T2A-PuroR/gRNA vector (px459) via cut-ligation (Supplementary Data 5). ESCs were transfected with an equimolar amount of each gRNA vector. Two days after transfection, cells were plated at clonal density and subjected to a transient puromycin selection (1 µg/mL) for 40 h. Colonies were picked 6 days after transfection. The triple PCR strategy used for screening is depicted in Supplementary Fig. 3a. Briefly, PCR primers 1F and 4R were used to identify clones in which the *Dppa3* locus had been removed, resulting in the appearance of a ~350 bp amplicon. To identify whether the *Dppa3* locus had been removed from both alleles, PCRs were performed with primers 1F and 2R or 3F and 4R (Supplementary Data 5) to amplify upstream or downstream ends of the *Dppa3* locus, which would only be left intact in the event of mono-allelic locus excision. Removal of the *Dppa3* locus was confirmed with Sanger sequencing and loss of *Dppa3* expression was assessed by qRT-PCR.

For CRISPR/Cas gene editing, all transfections were performed using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's instructions. All DNA oligos used for gene editing and screening are listed in Supplementary Data 5.

Bxb1-mediated recombination and Sleeping Beauty transposition. To generate stable mESC lines carrying doxycycline-inducible forms of *Dppa3* or *Dppa3*-*mScarlet*, mES cells were first transfected with equimolar amounts of the pSBtet-3xFLAG-Dppa3-IRES-DsRed-PuroR or pSBtet-3xFLAG-Dppa3-mScarlet-PuroR and the Sleeping Beauty transposase, pCMV(CAT)T7-SB100[174] (Addgene plasmid #34879) vector using Lipofectamine 3000 (Thermo Fisher Scientific) according to manufacturer's instructions. Two days after transfection, cells were plated at clonal density and subjected to puromycin selection (1 µg/mL) for 5–6 days. To ensure

ARTICLE

comparable levels of *Dppa3* induction, cells were first treated for 18 h with doxycycline (1 μg/mL) and then sorted with FACS based on thresholded levels of DsRed or mScarlet expression, the fluorescent readouts of successful induction. Post sorting, cells were plated back into media without doxycycline for 7 days before commencing experiments.

To generate stable doxycycline-inducible *Dppa3* hESC lines, hES cells were first transfected with equimolar amounts of the pSBtet-3xFLAG-Dppa3-IRES-DsRed-PuroR and Sleeping Beauty transposase pCMV(CAT)T7-SB100[175] (Addgene plasmid #34879) vector using using the P3 Primary Cell 4D-NucleofectorTM Kit (V4XP-3012 Lonza) and the 4D-Nucleofector™ Platform (Lonza), program CB-156. Two days after nucleofection, cells were subjected to puromycin selection (1 μg/mL) for subsequent two days, followed by an outgrowth phase of 4 days. At this stage, cells were sorted with FACS based on thresholded levels of DsRed expression to obtain two bulk populations of positive stable hESC lines with inducible *Dppa3*.

For the generation of the *Uhrf1*[GFP/GFP] cell line, we used our previously described ESC line with a C-terminal MIN-tag (*Uhrf1*[attP/attP]; Bxb1 *attP* site) and inserted the GFP coding sequence as described previously[169]. Briefly, attB-GFP-Stop-PolyA (Addgene plasmid #65526) was inserted into the C-terminal of the endogenous *Uhrf1*[attP/attP] locus by transfection with equimolar amounts of Bxb1 and attB-GFP-Stop-PolyA construct, followed by collection of GFP-positive cells with FACS after 6 days.

Cellular fractionation. Cell fractionation was performed as described previously with minor modifications[175]. Approximately 1×10^7 ESCs were resuspended in 250 μL of buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl₂, 0.34 M sucrose, 10% glycerol, 0.1% Triton X-100, 1 mM DTT, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1x mammalian protease inhibitor cocktail (PI; Roche)) and incubated for 5 min on ice. Nuclei were collected by centrifugation (4 min, 1300 × *g*, 4 °C) and the cytoplasmic fraction (supernatant) was cleared again by centrifugation (15 min, 20,000 × *g*, 4 °C). Nuclei were washed once with buffer A, and then lysed in buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, 1 mM PMSF, 1× PI). Insoluble chromatin was collected by centrifugation (4 min, 1700 × *g*, 4 °C) and washed once with buffer B. Chromatin fraction was lysed with 1× Laemmli buffer and boiled (10 min, 95 °C).

Western blot. Western blots were performed as described previously[82,169]. The following antibodies were used:

Rabbit anti-UHRF1 (polyclonal; 1:250; Cell Signalling, D6G8E), mouse anti-alpha-Tubulin (monoclonal; 1:500; Sigma, T9026), rabbit anti-H3 (polyclonal; 1:1000; Abcam, ab1791), mouse anti-GFP (monoclonal; 1:1000; Roche), mouse anti-FLAG M2 (monoclonal; 1:1000; Sigma, F3165), rabbit anti-xDNMT1 (polyclonal;[82]), rabbit anti-xUHRF1 (polyclonal;[82]), rabbit anti-USP7 (polyclonal; Bethyl Lab., A300-033A), rabbit anti-H3 (polyclonal; Abcam, ab1791), rat anti-TET1 (monoclonal; 1:10;[176]), rat anti-alpha-Tubulin (monoclonal; 1:250; Abcam, ab6160). goat anti-rat HRP (polyclonal; 1:1000; Jackson ImmunoResearch), goat anti-rabbit HRP (polyclonal; 1:1000; BioRad), mouse anti-xCDC45 (monoclonal;[177]), mouse anti-xRPA2 (monoclonal;[178]), and mouse anti-PCNA (monoclonal; Santa Cruz, sc56). Uncropped and unprocessed scans of blots can be found in the Source Data file.

Quantitative real-time PCR (qRT-PCR) analysis. Total RNA was isolated using the NucleoSpin Triprep Kit (Macherey-Nagel) according to the manufacturer's instructions. cDNA synthesis was performed with the High-Capacity cDNA Reverse Transcription Kit (with RNase Inhibitor; Applied Biosystems) using 500 ng of total RNA as input. qRT-PCR assays with oligonucleotides listed in Supplementary Data 5 were performed in 8 μL reactions with 1.5 ng of cDNA used as input. FastStart Universal SYBR Green Master Mix (Roche) was used for SYBR green detection. The reactions were run on a LightCycler480 (Roche).

LC-MS/MS analysis of DNA samples. Isolation of genomic DNA was performed according to earlier published work[57]. 1.0–5 μg of genomic DNA in 35 μL H₂O were digested as follows: An aqueous solution (7.5 μL) of 480 μM ZnSO₄, containing 18.4 U nuclease S1 (Aspergillus oryzae, Sigma-Aldrich), 5 U Antarctic phosphatase (New England BioLabs) and labeled internal standards were added ([¹⁵N₂]-cadC 0.04301 pmol, [¹⁵N₂,D₂]-hmdC 7.7 pmol, [D₃]-mdC 51.0 pmol, [¹⁵N₅]-8-oxo-dG 0.109 pmol, [¹⁵N₂]-fdC 0.04557 pmol) and the mixture was incubated at 37 °C for 3 h. After addition of 7.5 μl of a 520 μM [Na]₂-EDTA solution, containing 0.2 U snake venom phosphodiesterase I (Crotalus adamanteus, USB corporation), the sample was incubated for 3 h at 37 °C and then stored at −20 °C. Prior to LC-MS/MS analysis, samples were filtered by using an AcroPrep Advance 96 filter plate 0.2 μm Supor (Pall Life Sciences).

Quantitative UHPLC-MS/MS analysis of digested DNA samples was performed using an Agilent 1290 UHPLC system equipped with a UV detector and an Agilent 6490 triple quadrupole mass spectrometer. Natural nucleosides were quantified with the stable isotope dilution technique. An improved method, based on earlier published work[57,179] was developed, which allowed the concurrent analysis of all nucleosides in one single analytical run. The source-dependent parameters were as follows: gas temperature 80 °C, gas flow 15 L/min (N₂), nebulizer 30 psi, sheath gas heater 275 °C, sheath gas flow 15 L/min (N₂), capillary voltage 2,500 V in the

positive ion mode, capillary voltage −2,250 V in the negative ion mode and nozzle voltage 500 V. The fragmentor voltage was 380 V/ 250 V. Delta EMV was set to 500 V for the positive mode. Chromatography was performed by a Poroshell 120 SB-C8 column (Agilent, 2.7 μm, 2.1 mm × 150 mm) at 35 °C using a gradient of water and MeCN, each containing 0.0085% (v/v) formic acid, at a flow rate of 0.35 mL/min: 0 → 4 min; 0 → 3.5% (v/v) MeCN; 4 → 6.9 min; 3.5 → 5% MeCN; 6.9 → 7.2 min; 5 → 80% MeCN; 7.2 → 10.5 min; 80% MeCN; 10.5 → 11.3 min; 80 → 0% MeCN; 11.3 → 14 min; 0% MeCN. The effluent up to 1.5 min and after 9 min was diverted to waste by a Valco valve. The autosampler was cooled to 4 °C. The injection volume amounted to 39 μL. Data were processed according to earlier published work[57].

RNA-seq library preparation. Digital gene expression libraries for RNA-seq were prepared using the single-cell RNA barcoding sequencing (SCRB-seq) method as described previously[180–182], with minor modifications to accommodate bulk cell populations. In brief, RNA was extracted and purified from ~1 × 10⁶ cells using the NucleoSpin Triprep Kit (Machery-Nagel) according to the manufacturer's instructions. In the initial cDNA synthesis step, purified, bulk RNA (70 ng) from individual samples were subjected to reverse transcription in 10 μL reactions containing 25 units of Maxima H Minus reverse transcriptase (ThemoFisher Scientific), 1× Maxima RT Buffer (ThemoFisher Scientific), 1 mM dNTPs (Thermo-Fisher Scientific), 1 μM oligo-dT primer with a sample-specific barcode (IDT), and 1 μM template-switching oligo (IDT). Reverse transcription reactions were incubated 90 min at 42 °C. Next, the barcoded cDNAs from individual samples were pooled together and then purified using the DNA Clean & Concentrator-5 Kit (Zymo Research) according to the manufacturer's instructions. Purified pooled cDNA was eluted in 18 μL DNase/RNase-Free Distilled Water (Thermo Fisher) and then, to remove residual primers, incubated with 1 μL Exonuclease I Buffer (NEB) and 1 μL Exonuclease I (NEB) (final reaction volume: 20 μL) at 37 °C for 30 min followed by heat-inactivation at 80 °C for 20 min. Full-length cDNA was then amplified via PCR using KAPA HiFi HotStart ReadyMix (KAPA Biosystems) and SINGV6 primer (IDT). The pre-amplification PCR was performed using the following conditions: 3 min at 98 °C for initial denaturation, 10 cycles of 15 s at 98 °C, 30 s at 65 °C, and 6 min at 68 °C, followed by 10 min at 72 °C for final elongation. After purification using CleanPCR SPRI beads (CleanNA), the pre-amplified cDNA pool concentration was quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher). A Bioanalyzer run using the High-sensitivity DNA Kit (Agilent Technologies) was then performed to confirm the concentration and assess the size distribution of the amplified cDNA pool (Agilent Technologies). Next, 0.8 ng of the pure, amplified cDNA pool was used as input for generating a Nextera XT DNA library (Illumina) following the Manufacturer's instructions with the exception that a custom P5 primer (P5NEXTPT5) (IDT) was used to pre-ferentially enrich for 3′ cDNA ends in the final Nextera XT Indexing PCR[180–182]. After an initial purification step using a 1:1 ratio of CleanPCR SPRI beads (CleanNA), the amplified Nextera XT Library the 300–800 bp range of the library was size-selected using a 2% E-Gel Agarose EX Gels (Life Technologies) and then extracted from the gel using the MinElute Gel Extraction Kit (Qiagen, Cat. No. 28606) according to manufacturer's recommendations. The final concentration, size distribution, and quality of Nextera XT library were assessed with a Bioanalyzer (Agilent Technologies) using a High-sensitivity DNA Kit (Agilent Technologies). The Nextera XT RNA-seq library was paired-end sequenced using a high output flow cell on an Illumina HiSeq 1500.

Reduced representation bisulfite sequencing (RRBS) library preparation. For RRBS library preparation, genomic DNA was isolated using the QIAamp DNA Mini Kit (QIAGEN), after an overnight lysis and proteinase K treatment. RRBS library preparation was performed as described previously[183], with slight modifications. In brief, once purified, genomic DNA (100 ng) from each sample was used as starting material and first digested with 60 units of MspI (New England Biolabs) in a 30 μl reaction volume at 37 °C overnight. Digested DNA ends were then repaired and A-tailed by adding a 2 μl of a mixture containing 10 mM dATP, 1 mM dCTP, 1 mM dGTP and Klenow fragment (3′→5′ exo-) (New England Biolabs) to the unpurified digestion reaction and incubated first at 30 °C for 20 min followed by 37 °C for 20 min. Individual end-repaired and A-tailed DNA samples were purified using a 2:1 ratio of CleanPCR SPRI beads (CleanNA) and eluted in 20 μl elution buffer (10 mM Tris-HCl, pH 8.5). Next, barcoded adapters were ligated to the eluted DNA fragments in a 30 μl reaction containing 1× T4 Ligase Buffer (New England Biolabs), 2000 units of T4 Ligase (New England Biolabs), and 0.8 μM sample-specific TruSeq adapters (Illuminas) and incubated at 16 °C overnight. After adapter ligation, individual samples were first pooled before being purified with a 2:1 ratio of CleanPCR SPRI beads (CleanNA) and then eluted using 4 μl elution buffer times the number of samples in the pool. Pooled samples were then bisulfite converted using the EZ DNA Methylation-Gold™ Kit (Zymo Research) according to the manufacturer's instructions with the exception that libraries were eluted 2 × 20 μL M-elution buffer (Zymo Research). After bisulfite conversion, libraries were amplified in a 200 μl large-scale PCR reaction containing, 1x PfuTurbo Cx Reaction Buffer (Agilent Technologies), 10 units of PfuTurbo Cx Hotstart DNA Polymerase (Agilent Technologies), 1 mM dNTPs (New England Biolabs), 0.3 μM TruSeq Primers (Illumina), and 20 μl of pooled, bisulfite-converted DNA samples. After dividing the reaction into 4 wells of a 96-well plate

ARTICLE

(each containing 50 µl), the PCR was performed using the following cycling conditions: 2 min at 95 °C for initial denaturation and Polymerase activation, 16 cycles of 30 s at 95 °C, 30 s at 65 °C, and 45 s at 72 °C, followed by 7 min at 72 °C for final elongation. After amplification, the samples are pooled together again, subjected to a final round of purification using a 1.2:1 ratio of CleanPCR SPRI beads (CleanNA), and eluted in 40 µl of elution buffer. For an initial assessment of quality and yield, purified RRBS libraries were first analyzed on 2% E-Gel Agarose EX Gels (Life Technologies) and the concentrations then measured using the Quant-iT™ PicoGreen™ dsDNA Assay-Kit (ThermoFisher). The final concentration, size distribution, and quality of each RRBS library was then assessed with a Bioanalyzer (Agilent Technologies) using a High-sensitivity DNA Kit (Agilent Technologies). RRBS libraries were then sequenced on an Illumina HiSeq 1500.

Targeted bisulfite amplicon (TaBA) sequencing. Genomic DNA was isolated from 10^6 cells using the PureLink Genomic DNA Mini Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. The EZ DNA Methylation-Gold Kit (Zymo Research) was used for bisulfite conversion according to the manufacturer's instructions but with the following alterations: 500 ng of genomic DNA was used as input and bisulfite converted DNA was eluted in 2×20 µL Elution Buffer (10 mM Tris-HCl, pH 8.5).

TaBA-seq library preparation entailed two sequential PCRs to first amplify a specific locus and then index sample-specific amplicons. For the first PCR, the locus specific primers were designed with Illumina TruSeq and Nextera compatible overhangs (Supplementary Data 5). The amplification of bisulfite converted DNA was performed in 25 µL PCR reaction volumes containing 0.4 µM each of forward and reverse primers, 2 mM Betaiinitialne (Sigma-Aldrich, B0300-1VL), 10 mM Tetramethylammonium chloride solution (Sigma-Aldrich T3411-500ML), 1x MyTaq Reaction Buffer, 0.5 units of MyTaq HS (Bioline, BIO-21112), and 1 µL of the eluted bisulfite converted DNA (~12.5 ng). The following cycling parameters were used: 5 min for 95 °C for initial denaturation and activation of the polymerase, 40 cycles (95 °C for 20 s, 58 °C for 30 s, 72 °C for 25 s) and a final elongation at 72 °C for 3 min. Agarose gel electrophoresis was used to determine the quality and yield of the PCR. For purifying amplicon DNA, PCR reactions were incubated with 1.8× volume of CleanPCR beads (CleanNA, CPCR-0005) for 10 min. Beads were immobilized on a DynaMag™-96 Side Magnet (Thermo Fisher, 12331D) for 5 min, the supernatant was removed, and the beads washed 2× with 150 µL 70% ethanol. After air drying the beads for 5 min, DNA was eluted in 15 µL of 10 mM Tris-HCl pH 8.0. Amplicon DNA concentration was determined using the Quant-iT™ PicoGreen™ dsDNA Assay Kit (Thermo Fisher, P7589) and then diluted to 0.7 ng/µL.

Thereafter, indexing PCRs were performed in 25 µL PCR reaction volumes containing 0.08 µM (1 µL of a 2 µM stock) each of i5 and i7 Indexing Primers (Supplementary Data 5), 1x MyTaq Reaction Buffer, 0.5 units of MyTaq HS (Bioline, BIO-21112), and 1 µL of the purified PCR product from the previous step. The following cycling parameters were used: 5 min for 95 °C for initial denaturation and activation of the polymerase, 40 cycles (95 °C for 10 s, 55 °C for 30 s, 72 °C for 40 s) and a final elongation at 72 °C for 5 min. Agarose gel electrophoresis was used to determine the quality and yield of the PCR. An aliquot from each indexing reaction (5 µL of each reaction) was then pooled and purified with CleanPCR magnetic beads as described above and eluted in 1 µL × Number of pooled reactions. Concentration of the final library was determined using PicoGreen and the quality and size distribution of the library was assessed with a Bioanalyzer. Dual indexed TaBA-seq libraries were sequenced on an Illumina MiSeq in 2 × 300 bp output mode.

RNA-seq processing and analysis. RNA-seq libraries were processed and mapped to the mouse genome (mm10) using the zUMIs pipeline[184]. UMI count tables were filtered for low counts using HTSFilter[185]. Differential expression analysis was performed in R using DESeq2[186] and genes with an adjusted $P < 0.05$ were considered to be differentially expressed. Hierarchical clustering was performed on genes differentially expressed in TET mutant ESCs respectively, using k-means clustering ($k = 4$) in combination with the ComplexHeatmap (v 1.17.1) R-package[187]. Principal component analysis was restricted to genes differentially expressed during wild-type differentiation and performed using all replicates of wild-type, TET mutant, and Dppa3KO ESCs.

RRBS alignment and analysis. Raw RRBS reads were first trimmed using Trim Galore (v.0.3.1) with the "-rrbs" parameter. Alignments were carried out to the mouse genome (mm10) using bsmap (v.2.90) using the parameters "-s 12 -v 10 -r 2 -I 1". Summary statistics of the RRBS results are provided in Supplementary Data 6 and sample reproducibility information is shown in Supplementary Fig. 9. CpG-methylation calls were extracted from the mapping output using bsmaps methratio.py. Analysis was restricted to CpG with a coverage >10. methylKit[188] was used to identify differentially methylated regions between the respective contrasts for the following genomic features: (1) all 1-kb tiles (containing a minimum of three CpGs) detected by RRBS; (2) Repeats (defined by Repbase); (3) gene promoters (defined as gene start sites −2 kb/+2 kb); and (4) gene bodies (defined as longest isoform per gene) and CpG islands (as defined by Ilingworth et al.[189]). Differentially methylated regions were identified as regions with $P < 0.05$ and a difference in

methylation means between two groups greater than 20%. Principal component analysis of global DNA methylation profiles was performed on single CpGs using all replicates of wild-type, T1KO and T1CM ESCs and EpiLCs.

Chromatin immunoprecipitation (ChIP) and Hydroxymethylated-DNA immunoprecipitation (hMeDIP) alignment and analysis. ChIP-seq reads for TET1 binding in ESCs and EpiLCs were downloaded from GSE57700[67] and PRJEB19897[66], respectively. hMeDIP reads for wild-type ESCs and T1KO ESCs were download from PRJEB13096[66]. Reads were aligned to the mouse genome (mm10) with Bowtie (v.1.2.2) with parameters "-a -m 3 -n 3 -best -strata". Subsequent ChIP-seq analysis was carried out on data of merged replicates. Peak calling and signal pileup was performed using MACS2 callpeak[190] with the parameters "-extsize 150" for ChIP, "-extsize 220" for hMeDIP, and "-nomodel -B -nolambda" for all samples. Tag densities for promoters and 1 kb Tiles were calculated using the deepTools2 computeMatrix module[191]. TET1 bound genes were defined by harboring a TET1 peak in the promoter region (defined as gene start sites −2 kb/+2 kb).

Immunofluorescence staining. For immunostaining, naïve ESCs were grown on coverslips coated with Geltrex (Life Technologies) diluted 1:100 in DMEM/F12 (Life Technologies), thereby allowing better visualization of the cytoplasm during microscopic analysis. All steps during immunostaining were performed at room temperature. Coverslips were rinsed two times with PBS (pH 7.4; 140 mM NaCl, 2.7 mM KCl, 6.5 mM Na_2HPO_4, 1.5 mM KH_2PO_4) prewarmed to 37 °C, cells fixed for 10 min with 4% paraformaldehyde (pH 7.0; prepared from paraformaldehyde powder (Merck) by heating in PBS up to 60 °C; store at −20 °C), washed three times for 10 min with PBST (PBS, 0.01% Tween20), permeabilized for 5 min in PBS supplemented with 0.5% Triton X-100, and washed two times for 10 min with PBS. Primary and secondary antibodies were diluted in blocking solution (PBST, 4% BSA). Coverslips were incubated with primary and secondary antibody solutions in dark humid chambers for 1 h and washed three times for 10 min with PBST after primary and secondary antibodies. For DNA counterstaining, coverslips were incubated 6 min in PBST containing a final concentration of 2 µg/mL DAPI (Sigma-Aldrich) and washed three times for 10 min with PBST. Coverslips were mounted in antifade medium (Vectashield, Vector Laboratories) and sealed with colorless nail polish.

The following antibodies were used: rabbit anti-DPPA3 (polyclonal; 1:200; Abcam, ab19878), mouse anti-UHRF1 (monoclonal; 1:250; Santa Cruz, sc373750), goat anti-mouse A488 (polyclonal; 1:500; used in IF; Invitrogen, A11029), donkey anti-rabbit Dylight594 (polyclonal; 1:500; Dianova, 711-516-152), anti-GFP-Booster ATTO488 (1:200; Chromotek), mouse anti-5mC (monoclonal; 1:200; Active Motif, 39649), donkey anti-anti-rabbit A555 (polyclonal; 1:500; Invitrogen, A31572), and donkey anti-anti-rabbit A488 (polyclonal; 1:500; Dianova, 711-547-003).

Immunofluorescence and Live-cell imaging. For immunofluorescence, stacks of optical sections were collected on a Nikon TiE microscope equipped with a Yokogawa CSU-W1 spinning-disk confocal unit (50 µm pinhole size), an Andor Borealis illumination unit, Andor ALC600 laser beam combiner (405 nm/488 nm/561 nm/640 nm), Andor IXON 888 Ultra EMCCD camera, and a Nikon 100×/1.45 NA oil immersion objective. The microscope was controlled by software from Nikon (NIS Elements, ver. 5.02.00). DAPI or fluorophores were excited with 405 nm, 488 nm, or 561 nm laser lines and bright-field images acquired using Nikon differential interference contrast optics. Confocal image z-stacks were recorded with a step size of 200 nm, 16-bit image depth, 1×1 binning, a frame size of 1024×1024 pixels, and a pixel size of 130 nm. Within each experiment, cells were imaged using the same settings on the microscope (camera exposure time, laser power, and gain) to compare signal intensities between cell lines.

For live-cell imaging, cells were plated on Geltrex-coated glass bottom 2-well imaging slides (Ibidi). Both still and timelapse images were acquired on the Nikon spinning-disk system described above equipped with an environmental chamber maintained at 37 °C with 5% CO_2 (Oko Labs), using a Nikon 100x/1.45 NA oil immersion objective and a Perfect Focus System (Nikon). Images were acquired with the 488, 561, and 640 nm laser lines, full-frame (1024 × 1024) with 1×1 binning, and with a pixel size of 130 nm. Transfection of a RFP-PCNA vector[192] was used to identify cells in S-phase. For DNA staining in live cells, cells were exposed to media containing 200 nM SiR-DNA (Spirochrome) for at least 1 h before imaging. For imaging endogenous DPPA3-HALO in live cells, cells were treated with media containing 50 nM HaloTag-TMR fluorescent ligand (Promega) for 1 h. After incubation, cells were washed 3× with PBS before adding back normal media. Nuclear export inhibition was carried out using media containing 20 nM leptomycin-B (Sigma-Aldrich). Live-cell imaging data was acquired with NIS Elements ver. 4.5 (Nikon). NIS Elements ver. 5.02.00 (Nikon) and Volocity (PerkinElmer) were used for acquiring FRAP data. RICS measurements were acquired using FABSurf (v 1.0).

Image analysis. For immunofluorescence images, Fiji software (ImageJ 1.51j)[193,194] was used to analyze images and create RGB stacks. For analysis of live-cell imaging data, CellProfiler Software (version 3.0)[195] was used to quantify fluorescence intensity

     ARTICLE

in cells stained with SiR-DNA. CellProfiler pipelines used in this study are available upon request. In brief, the SiR-DNA signal was used to segment ESC nuclei. Mean fluorescence intensity of GFP was measured both inside the segmented area (nucleus) and in the area extending 4–5 pixels beyond the segmented nucleus (cytoplasm). GFP fluorescence intensity was normalized by subtracting the experimentally-determined mean background intensity and background-subtracted GFP intensities were then used for all subsequent quantifications shown in Fig. 4 and Supplementary Figs. 4h, 5h, and 6b, c.

Fluorescence recovery after photobleaching (FRAP). For FRAP assays, cells cultivated on Geltrex-coated glass bottom 2-well imaging slides (Ibidi) were imaged in an environmental chamber maintained at 37 °C with 5% CO₂ either using the Nikon system mentioned above equipped with a FRAPPA photobleaching module (Andor) or on an Ultraview-Vox spinning-disk system (Perkin-Elmer) including a FRAP Photokinesis device mounted to an inverted Axio Observer D1 microscope (Zeiss) equipped with an EMCCD camera (Hamamatsu) and a 63x/1.4 NA oil immersion objective, as well as 405, 488 and 561 nm laser lines.

For endogenous UHRF1-GFP FRAP, eight pre-bleach images were acquired with the 488 nm laser, after which an area of 4 × 4 pixels was irradiated for a total of 16 ms with a focused 488 nm laser (leading to a bleached spot of ~1 μm) to bleach a fraction of GFP-tagged molecules within cells, and then recovery images were acquired every 250 ms for 1-2 min. Recovery analysis was performed in Fiji. Briefly, fluorescence intensity at the bleached spot was measured in background-subtracted images, then normalized to pre-bleach intensity of the bleached spot, and normalized again to the total nuclear intensity in order to account for acquisition photobleaching. Images of cells with visible drift were discarded.

***Xenopus* egg extracts**. The interphase extracts (low-speed supernatants (LSS)) were prepared as described previously[82]. After thawing, LSS were supplemented with an energy regeneration system (5 μg/ml creatine kinase, 20 mM creatine phosphate, 2 mM ATP) and incubated with sperm nuclei at 3000–4000 nuclei per μl. Extracts were diluted 5-fold with ice-cold CPB (50 mM KCl, 2.5 mM MgCl₂, 20 mM HEPES-KOH, pH 7.7) containing 2% sucrose, 0.1% NP-40 and 2 mM NEM, overlaid onto a 30% sucrose/CPB cushion, and centrifuged at 15,000 *g* for 10 min. The chromatin pellet was resuspended in SDS sample buffer and analyzed by SDS-PAGE. GST-mDPPA3 was added to egg extracts at 50 ng/μl at final concentration.

Monitoring DNA methylation in *Xenopus* egg extracts. DNA methylation was monitored by the incorporation of S-[methyl-³H]-adenosyl-L-methionine, incubated at room temperature, and the reaction was stopped by the addition of CPB containing 2% sucrose up to 300 μl. Genomic DNA was purified using a Wizard Genomic DNA purification kit (Promega) according to the manufacturer's instructions. Incorporation of radioactivity was quantified by liquid synchillation counter.

Plasmid construction for recombinant mDPPA3. To generate GST-tagged mDPPA3 expression plasmids, mDPPA3 fragment corresponding to full-length protein was amplified by PCR using mouse DPPA3 cDNA and specific primers (Supplementary Data 5). The resulting DNA fragment was cloned into pGEX4T-3 vector digested with EcoRI and SalI using an In-Fusion HD Cloning Kit.

Protein expression and purification. For protein expression in *Escherichia coli* (BL21-CodonPlus), the mDPPA3 genes were transferred to pGEX4T-3 vector as described above. Protein expression was induced by the addition of 0.1 mM Iso-propyl β–D-1-thiogalactopyranoside (IPTG) to media followed by incubation for 12 h at 20 °C. For purification of Glutathione S transferase (GST) tagged proteins, cells were incubated in Lysis buffer (20 mM HEPES-KOH (pH 7.6), 0.5 M NaCl, 0.5 mM EDTA, 10% glycerol, 1 mM DTT) supplemented with 0.5% NP40 and protease inhibitors, and were then disrupted by sonication on ice. After centrifugation, the supernatant was applied to Glutathione Sepharose (GSH) beads (GE Healthcare) and rotated at 4 °C for 2 h. Beads were then washed three times with Wash buffer 1 (20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% TritionX-100, 1 mM DTT) three times and with Wash buffer 2 (100 mM Tris-HCl (pH 7.5), 100 mM NaCl) once. Bound proteins were eluted in Elution buffer (100 mM Tris-HCl (pH 7.5), 100 mM NaCl, 5% glycerol, 1 mM DTT) containing 42 mM reduced Glutathione and purified protein was loaded on PD10 desalting column equili-brated with EB buffer (10 mM HEPES/KOH at pH 7.7, 100 mM KCl, 0.1 mM CaCl₂, 1 mM MgCl₂) containing 1 mM DTT, and then concentrated by Vivaspin (Millipore).

Data collection for the presence of TET1, UHRF1, DNMT1, and DPPA3 throughout metazoa. Reference protein sequences of TET1 (Human Q8NFU7, Mouse Q3URK3, *Naegleria gruberi* D2W6T1), DNMT1 (Rat Q9Z330, Human P26358, Mouse P13864, Chicken Q92072, Cow Q92072), UHRF1 (Mouse Q8VDF2, Rat Q7TPK1, Zebra fish E7EZF3, Human Q96T88, Cow A7E320, Xenopus laevis F6UA42) and DPPA3 (Mouse Q8QZY3, Human Q6W0C5, Cow A9Q1J7) were downloaded from the Universal Protein Resource (UniProt). Orthologous were identified with *hmmsearch* of the HMMER (http://hmmer.org/)

toolkit using default parameters. Presence of the proteins throughout metazoa was visualized using iTOL[196].

Chromatin immunoprecipitation coupled to Mass Spectrometry and Pro-teomics data analysis. For Chromatin immunoprecipitation coupled to Mass Spectrometry (ChIP-MS), whole cell lysates of the doxycycline-inducible *Dppa3*-FLAG mES cells were used by performing three separate immunoprecipitations with an anti-FLAG antibody and three samples with a control IgG. Trypsinized cells were washed twice by PBS and subsequently diluted to 15*10⁶ cells per 10 mL PBS. Paraformaldehyde (PFA) was added to a final concentration of 1% and crosslinking was performed at room temperature on an orbital shaker for 10 min. Free PFA was quenched by 125 mM Glycine for 5 min and crosslinked cells were washed twice by ice-cold PBS before cell lysis. Proteins were digested on the beads after the pulldown and desalted subsequently on StageTips with three layers of C18[197]. Here, peptides were separated by liquid chromatography on an Easy-nLC 1200 (Thermo Fisher Scientific) on in-house packed 50 cm columns of ReproSil-Pur C18-AQ 1.9-μm resin (Dr. Maisch GmbH). Peptides were then eluted suc-cessively in an ACN gradient for 120 min at a flow rate of around 300 nL/min and were injected through a nanoelectrospray source into a Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific). After mea-suring triplicates of a certain condition, an additional washing step was scheduled. During the measurements, the column temperature was constantly kept at 60 °C while after each measurement, the column was washed with 95% buffer B and subsequently with buffer A. Real time monitoring of the operational parameters was established by SprayQc[198] software. Data acquisition was based on a top10 shotgun proteomics method and data-dependent MS/MS scans. Within a range of 400-1650 m/z and a max. injection time of 20 ms, the target value for the full scan MS spectra was 3×10^6 and the resolution at 60,000.

The raw MS data was then analyzed with the MaxQuant software package (version 1.6.0.7)[199]. The underlying FASTA files for peak list searches were derived from Uniprot (UP000000589_10090.fasta and UP000000589_10090 additional. fasta, version June 2015) and an additional modified FASTA file for the FLAG-tagged *Dppa3* in combination with a contaminants database provided by the Andromeda search engine[200] with 245 entries. During the MaxQuant-based analysis the "Match between runs" option was enabled and the false discovery rate was set to 1% for both peptides (minimum length of 7 amino acids) and proteins. Relative protein amounts were determined by the MaxLFQ algorithm[201], with a minimum ratio count of two peptides.

For the downstream analysis of the MaxQuant output, the software Perseus[202] (version 1.6.0.9) was used to perform two-sided Student's *t*-test with a permutation-based FDR of 0.05 and an additional constant S0 = 1 in order to calculate fold enrichments of proteins between triplicate chromatin immunoprecipitations of anti-FLAG antibody and control IgG. The result was visualized in a scatter plot. The complete catalog of proteins interacting with FLAG-DPPA3 in ESCs including statistics can be found in Supplementary Data 3.

For GO analysis of biological processes the Panther classification system was used[203]. For the analysis, 131 interactors of DPPA3 were considered after filtering the whole amount of 303 significant interactors for a *p*-value of at least 0.0015 and 3 or more identified peptides. The resulting GO groups (determined by a two-sided Fisher's exact test) were additionally filtered for a fold enrichment of observed over expected amounts of proteins of at least 4 and a *p*-value of 5.30 E−08. The result can be found in Supplementary Data 4.

***Dppa3* overexpression in medaka embryos and immunostaining**. Medaka d-rR strain was used. Medaka fish were maintained and raised according to standard protocols. Developmental stages were determined based on a previous study[204]. *Dppa3* and mutant *Dppa3* (R107E) mRNA were synthesized using HiScribe T7 ARCA mRNA kit (NEB, E2060S), and purified using RNeasy mini kit (QIAGEN, 74104). *Dppa3* or mutant *Dppa3* (R107E) mRNA was injected into the one-cell stage (stage 2) medaka embryos. After 7 h of incubation at 28 °C, the late blastula (stage 11) embryos were fixed with 4% PFA in PBS for 2 h at room temperature, and then at 4 °C overnight. Embryos were dechorionated, washed with PBS, and permeabilized with 0.5% Triton X-100 in PBS for 30 min at room temperature. DNA was denatured in 4 M HCl for 15 min at room temperature, followed by neutralization in 100 mM Tris-HCl (pH 8.0) for 20 min. After washing with PBS, embryos were blocked in blocking solution (2% BSA, 1%DMSO, 0.2% Triton X-100 in PBS) for 1 h at room temperature, and then incubated with 5-methylcytosine antibody (1:200; Active Motif #39649) at 4 °C overnight. The embryos were washed with PBSDT (1% DMSO, 0.1% Triton X-100 in PBS), blocked in blocking solution for 1 h at room temperature, and incubated with Alexa Fluor 555 goat anti-mouse 2nd antibody (1:500; ThermoFisher Scientific #A21422) at 4 °C overnight. After washing with PBSDT, cells were mounted on slides and examined under a fluor-escence microscope.

Fluorescence three hybrid (F3H) assay. The F3H assay was performed as described previously[96]. In brief, BHK cells containing multiple lac operator repeats were transiently transfected with the respective GFP- and mScarlet-constructs on coverslips using PEI and fixed with 3.0% formaldehyde 24 h after transfection. For DNA counterstaining, coverslips were incubated in a solution of DAPI (200 ng/ml)

ARTICLE

in PBS-T and mounted in Vectashield. Images were collected using a Leica TCS SP5 confocal microscope. To quantify the interactions within the lac spot, the following intensity ratio was calculated for each cell in order to account for different expression levels: mScarlet$_{spot}$ − mScarlet$_{background}$)/(GFP$_{spot}$ − GFP$_{background}$).

Microscale thermophoresis (MST). For MST measurements, mUHRF1 C-terminally tagged with GFP- and 6xHis-tag was expressed in HEK 293 T cells and then purified using Qiagen Ni-NTA beads (Qiagen #30230). Recombinant mDPPA3 WT and 1-60 were purified as described above. Purified UHRF1 (200 nM) was mixed with different concentrations of purified DPPA3 (0.15 nM to 5 μM) followed by a 30 min incubation on ice. The samples were then aspirated into NT.115 Standard Treated Capillaries (NanoTemper Technologies) and placed into the Monolith NT.115 instrument (NanoTemper Technologies). Experiments were conducted with 80% LED and 80% MST power. Obtained fluorescence signals were normalized (F$_{norm}$) and the change in F$_{norm}$ was plotted as a function of the concentration of the titrated binding partner using the MO. Affinity Analysis software version 2.1 (NanoTemper Technologies). For fluorescence normalization (F$_{norm}$ = F$_{hot}$/F$_{cold}$), the manual analysis mode was selected and cursors were set as follows: F$_{cold}$ = −1 to 0 s, F$_{hot}$ = 10 to 15 s. The Kd was obtained by fitting the mean F$_{norm}$ of eight data points (four independent replicates, each measured as a technical duplicate).

RICS. Data for Raster Image Correlation Spectroscopy (RICS) was acquired on a home-built laser scanning confocal setup equipped with a 100x NA 1.49 NA objective (Nikon) pulsed interleaved excitation (PIE) as used elsewhere[205]. Samples were excited using pulsed lasers at 470 (Picoquant) and 561 nm (Toptica Photonics), synchronized to a master clock, and then delayed ~20 ns relative to one another to achieve PIE. Laser excitation was separated from descanned fluorescence emission by a Di01-R405/488/561/635 polychroic mirror (Semrock, AHF Analysentechnik) and eGFP and mScarlet fluorescence emission was separated by a 565 DCXR dichroic mirror (AHF Analysentechnik) and collected on avalanche photodiodes, a Count Blue (Laser Components) and a SPCM-AQR-14 (Perkin-Elmer) with 520/40 and a 630/75 emission filters (Chroma, AHF Analysentechnik). Detected photons were recorded by time-correlated single-photon counting.

The alignment of the system was verified prior to each measurement session by performing FCS with PIE on a mixture of Atto-488 and Atto565 dyes excited with pulsed 470 and 561 nm lasers set to 10 μW (measured in the collimated space before entering the galvo-scanning mirror system), 1 μm above the surface of the coverslip[206]. Cells were plated on Ibidi two-well glass bottom slides, and induced with doxycycline overnight prior to measurements. Scanning was performed in cells maintained at 37 ˚C using a stage top incubator, with a total field-of-view of 12 μm × 12 μm, composed of 300 pixels × 300 lines (corresponding to a pixel size of 40 nm), a pixel dwell time of 11 μs, a line time of 3.33 ms, at one frame per second, for 100–200 s. Pulsed 470 and 561 nm lasers were adjusted to 4 and 5 μW, respectively.

Image analysis was done using the Pulsed Interleaved Excitation Analysis with Matlab (PAM) software[207]. Briefly, time gating of the raw photon stream was performed by selecting only photons collected on the appropriate detector after the corresponding pulsed excitation, thereby allowing cross-talk free imaging for each channel. Then, using the Microtime Image Analysis (MIA) analysis program, slow fluctuations were removed by subtracting a moving average of 3 frames and a region of interest corresponding to the nucleus was selected, excluding nucleoli and dense aggregates. The spatial autocorrelation and cross-correlation functions (SACF and SCCF) were calculated as done previously[208] using arbitrary region RICS:

$$G(\xi, \psi) = \frac{\langle I_{\text{RICS},1}(x,y) I_{\text{RICS},2}(x+\xi, y+\psi)\rangle_{XY}}{\langle I_{\text{RICS},1}\rangle_{XY}\langle I_{\text{RICS},2}\rangle_{XY}} \quad (1)$$

where ξ and ψ are the correlation lags in pixel units along the x- and y-axis scan directions. The correlation function was then fitted to a two-component model (one mobile and one immobile component) in MIAfit:

$$G_{\text{fit}}(\xi, \psi) = A_{\text{mob}} G_{\text{fit, mob}}(\xi, \psi) + A_{\text{imm}} \exp\left(-\delta r^2 \omega_{\text{imm}}^{-2}\left(\xi^2 + \psi^2\right)\right) + y_0, \quad (2)$$

where:

$$G_{\text{fit, mob}}(\xi, \psi) = \left(1 + \frac{4D(\tau_p \xi + \tau_l \psi)}{\omega_r^2}\right)^{-1}\left(1 + \frac{4D(\tau_p \xi + \tau_l \psi)}{\omega_z^2}\right)^{-1/2} \\ \cdot \exp\left(-\frac{\delta r^2(\xi^2 + \psi^2)}{\omega_r^2 + 4D(\tau_p \xi + \tau_l \psi)}\right) \quad (3)$$

which yields parameters such as the diffusion coefficient (D) and the amplitudes of the mobile and immobile fractions (A$_{mob}$ and A$_{imm}$). The average number of mobile molecules per excitation volume on the RICS timescale was determined by

$$N_{\text{mob}} = \left(\frac{\gamma}{A_{\text{mob}}}\right)\left(\frac{2\Delta F}{2\Delta F + 1}\right), \quad (4)$$

where γ is a factor pertaining to the 3D Gaussian shape of the PSF, and 2ΔF/(2ΔF + 1) is a correction factor when using a moving average subtraction prior to calculating the SACF. The immobilized molecules (i.e. bound fraction) is the contribution of particles that remain visible without significant motion during the acquisition of 5–10 lines of

the raster scan, corresponding to ~30 ms. The cross-correlation model was fitted to the cross-correlation function and the extent of cross-correlation was calculated from the amplitude of the mobile fraction of the cross-correlation fit divided by the amplitude of the mobile fraction of the autocorrelation fit of DPPA3-mScarlet.

Statistics and reproducibility. No statistical methods were used to predetermine sample size, the experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment. Blinding was not implemented in this study as analysis was inherently objective in the overwhelming majority of experiments. For microscopy analysis, where possible, experimenter bias was avoided by selecting fields of view (or individual cells) for acquisition of UHRF1-GFP or DNMT1-GFP signal using the DNA stain (or another marker not being directly assessed in the experiment e.g. DsRed/mScarlet as a readout of Dppa3 induction or RFP-PCNA). To further reduce bias, imaging analysis was subsequently performed indiscriminately on all acquired images using semi-automated analysis pipelines (either with CellProfiler or Fiji scripts). All experimental findings were reliably reproduced in independent experiments as indicated in the Figure legends. In general, all micrographs from immunofluorescence and live cell imaging, immunoblots, and DNA gel images depicted in this study are representative of $n \geq 2$ independent experiments. The number of replicates used in each experiment are described in the figure legends and/or in the Methods section, as are the Statistical tests used. P values or adjusted P values are given where possible. Unless otherwise indicated, all statistical calculations were performed using R Studio 1.2.1335. Next-generation sequencing experiments include at least two independent biological replicates. RNA-seq experiments include $n = 4$ biological replicates comprised of $n = 2$ independently cultured samples from two clones (for T1CM, T2CM, T12CM ESCs and EpiLCs) or four independently cultured samples (for wild-type ESCs and EpiLCs). For RRBS experiments, data are derived from $n = 2$ biological replicates. For bisulfite sequencing of LINE-1 elements $n = 2$ biological replicates were analyzed from two independent clones for T1CM, T2CM, T12CM, and Dppa3KO ESCs or two independent cultures for wt ESCs. LC-MS/MS quantification was performed on at least four biological replicates comprising at least two independently cultured samples (usually even more) from $n = 2$ independent clones (T1CM, T2CM, T12CM, and Dppa3KO ESCs) or four independently cultured samples (wild-type ESCs and cell lines shown in Fig. 5d).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Sequencing data reported in this paper are available at ArrayExpress (EMBL-EBI) under accessions "E-MTAB-6785" (wild-type and Tet catalytic mutants RRBS), "E-MTAB-6797" (RNA-seq), "E-MTAB-6800" (Dppa3KO RRBS), "E-MTAB-9654" (TaBA-seq of Tet catalytic mutants during Dppa3 induction) and "E-MTAB-9653" (TaBA-seq of Dppa3KO cells expressing *Dppa3* mutant constructs). The raw mass spectrometry proteomics data from the FLAG-DPPA3 pulldown have been deposited at the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier "PXD019794". Publically available data sets used in this study can be found here: "GSE77420" (RRBS of TET triple knockout ESCs), "GSE42616" (PRDM14 ChIP-seq), "GSE46111" (5caC-DIP in TDK knockout ESCs), "GSE57700" (TET1 and TET2 ChIP-seq).

Supplementary Data 1 contains the entire list of differentially methylated promoters classified as either "TET-specific", "DPPA3-specific" or "common", which are summarized in Supplementary Fig. 3i. Supplementary Data 2 contains the extended gene ontology analysis of TET-specific promoters with the five most significant terms displayed in Fig. 3e. Supplementary Data 3 contains the complete catalog of proteins interacting with FLAG-DPPA3 in ESCs, which are plotted in Fig. 4b. Supplementary Data 4 contains the full gene ontology analysis of significant DPPA3 interactors. Source data are provided with this paper.

Code availability

The PAM and MIA software is available as source code, requiring MATLAB, or as a precompiled, standalone distribution for Windows or MacOS at http://www.cup.uni-muenchen.de/ pc/lamb/software/pam.html or hosted in Git repositories under http://www.gitlab.com/PAM-PIE/PAM and http://www.gitlab.com/PAM-PIE/PAMcompiled.

References

1. Ladstätter, S. & Tachibana, K. Genomic insights into chromatin reprogramming to totipotency in embryos. *J. Cell Biol.* **218**, 70–82 (2019).

ARTICLE

2.   Warren, I. A. et al. Evolutionary impact of transposable elements on genomic diversity and lineage-specific innovation in vertebrates. *Chromosome Res.* **23**, 505–531 (2015).

3.   Arkhipova, I. R. Neutral theory, transposable elements, and eukaryotic genome evolution. *Mol. Biol. Evol.* **35**, 1332–1337 (2018).

4.   Jacobs, F. M. J. et al. An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature* **516**, 242–245 (2014).

5.   Friedli, M. & Trono, D. The developmental control of transposable elements and the evolution of higher species. *Annu. Rev. Cell Dev. Biol.* **31**, 429–451 (2015).

6.   Schmitz, R. J., Lewis, Z. A. & Goll, M. G. DNA methylation: shared and divergent features across eukaryotes. *Trends Genet.* **35**, 818–827 (2019).

7.   Schübeler, D. Function and information content of DNA methylation. *Nature* **517**, 321–326 (2015).

8.   Rowe, H. M. & Trono, D. Dynamic control of endogenous retroviruses during development. *Virology* **411**, 273–287 (2011).

9.   Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**, 116–117 (1998).

10.   Jackson-Grusby, L. et al. Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat. Genet.* **27**, 31–39 (2001).

11.   Chernyavskaya, Y. et al. Loss of DNA methylation in zebrafish embryos activates retrotransposons to trigger antiviral signaling. *Development* **144**, 2925–2939 (2017).

12.   Iida, A. et al. Targeted reduction of the DNA methylation level with 5-azacytidine promotes excision of the medaka fish Tol2 transposable element. *Genet. Res.* **87**, 187–193 (2006).

13.   Chiappinelli, K. B. et al. Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. *Cell* **169**, 361 (2017).

14.   Roulois, D. et al. DNA-demethylating agents target colorectal cancer cells by inducing viral mimicry by endogenous transcripts. *Cell* **162**, 961–973 (2015).

15.   Veenstra, G. J. & Wolffe, A. P. Constitutive genomic methylation during embryonic development of Xenopus. *Biochim. Biophys. Acta* **1521**, 39–44 (2001).

16.   Stancheva, I., El-Maarri, O., Walter, J., Niveleau, A. & Meehan, R. R. DNA methylation at promoter regions regulates the timing of gene activation in Xenopus laevis embryos. *Dev. Biol.* **243**, 155–165 (2002).

17.   Ortega-Recalde, O., Day, R. C., Gemmell, N. J. & Hore, T. A. Zebrafish preserve global germline DNA methylation while sex-linked rDNA is amplified and demethylated during feminisation. *Nat. Commun.* **10**, 3053 (2019).

18.   Skvortsova, K. et al. Retention of paternal DNA methylome in the developing zebrafish germline. *Nat. Commun.* **10**, 3054 (2019).

19.   Monk, M., Boubelik, M. & Lehnert, S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**, 371–382 (1987).

20.   Sanford, J. P., Clark, H. J., Chapman, V. M. & Rossant, J. Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev.* **1**, 1039–1046 (1987).

21.   Lee, H. J., Hore, T. A. & Reik, W. Reprogramming the methylome: erasing memory and creating diversity. *Cell Stem Cell* **14**, 710–719 (2014).

22.   Tahiliani, M. et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).

23.   He, Y.-F. et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).

24.   Cortellino, S. et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146**, 67–79 (2011).

25.   Ito, S. et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011).

26.   Rougier, N. et al. Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev.* **12**, 2108–2113 (1998).

27.   Howlett, S. K. & Reik, W. Methylation levels of maternal and paternal genomes during preimplantation development. *Development* **113**, 119–127 (1991).

28.   Carlson, L. L., Page, A. W. & Bestor, T. H. Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev.* **6**, 2536–2541 (1992).

29.   Wossidlo, M. et al. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.* **2**, 241 (2011).

30.   Gu, T.-P. et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–610 (2011).

31.   Wang, L. et al. Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**, 979–991 (2014).

32.   Guo, F. et al. Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* **15**, 447–459 (2014).

33.   Shen, L. et al. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell Stem Cell* **15**, 459–471 (2014).

34.   Amouroux, R. et al. De novo DNA methylation drives 5hmC accumulation in mouse zygotes. *Nat. Cell Biol.* **18**, 225–233 (2016).

35.   Iyer, L. M., Abhiman, S. & Aravind, L. Natural history of eukaryotic DNA methylation systems. *Prog. Mol. Biol. Transl. Sci.* **101**, 25–104 (2011).

36.   Dean, W. et al. Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *Proc. Natl Acad. Sci. USA* **98**, 13734–13738 (2001).

37.   Beaujean, N. et al. Effect of limited DNA methylation reprogramming in the normal sheep embryo on somatic cell nuclear transfer. *Biol. Reprod.* **71**, 185–193 (2004).

38.   Fulka, H., Mrazek, M., Tepla, O. & Fulka, J. Jr. DNA methylation pattern in human zygotes and developing embryos. *Reproduction* **128**, 703–708 (2004).

39.   Chen, T. et al. The DNA methylation events in normal and cloned rabbit embryos. *FEBS Lett.* **578**, 69–72 (2004).

40.   Smith, Z. D. et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).

41.   Guo, H. et al. The DNA methylation landscape of human early embryos. *Nature* **511**, 606–610 (2014).

42.   Gao, F. et al. De novo DNA methylation during monkey pre-implantation embryogenesis. *Cell Res.* **27**, 526–539 (2017).

43.   Ivanova, E. et al. DNA methylation changes during preimplantation development reveal inter-species differences and reprogramming events at imprinted genes. *Clin. Epigenetics* **12**, 64 (2020).

44.   Macleod, D., Clark, V. H. & Bird, A. Absence of genome-wide changes in DNA methylation during development of the zebrafish. *Nat. Genet.* **23**, 139–140 (1999).

45.   Bogdanović, O., Long, S. W. & van Heeringen, S. J. Temporal uncoupling of the DNA methylome and transcriptional repression during embryogenesis. *Genome Res.* **21**, 1313–1327 (2011).

46.   Andersen, I. S., Reiner, A. H., Aanes, H., Alestrøm, P. & Collas, P. Developmental features of DNA methylation during activation of the embryonic zebrafish genome. *Genome Biol.* **13**, R65 (2012).

47.   de Mendoza, A. et al. Convergent evolution of a vertebrate-like methylome in a marine sponge. *Nat. Ecol. Evol.* **3**, 1464–1473 (2019).

48.   Xu, X. et al. Evolutionary transition between invertebrates and vertebrates via methylation reprogramming in embryogenesis. *Natl Sci. Rev.* **6**, 993–1003 (2019).

49.   Marks, H. et al. The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590–604 (2012).

50.   Leitch, H. G. et al. Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* **20**, 311–316 (2013).

51.   Hackett, J. A. et al. Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. *Stem Cell Rep.* **1**, 518–531 (2013).

52.   Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* **30**, 733–750 (2016).

53.   Ito, S. et al. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).

54.   Ficz, G. et al. FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* **13**, 351–359 (2013).

55.   Boroviak, T. et al. Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis. *Dev. Cell* **35**, 366–382 (2015).

56.   Habibi, E. et al. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* **13**, 360–369 (2013).

57.   Pfaffeneder, T. et al. Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat. Chem. Biol.* **10**, 574–581 (2014).

58.   von Meyenn, F. et al. Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. *Mol. Cell* **62**, 848–861 (2016).

59.   Nakashima, H. et al. Effects of dppa3 on DNA methylation dynamics during primordial germ cell development in mice. *Biol. Reprod.* **88**, 125 (2013).

60.   Funaki, S. et al. Inhibition of maintenance DNA methylation by Stella. *Biochem. Biophys. Res. Commun.* **453**, 455–460 (2014).

61.   Nakamura, T. et al. PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat. Cell Biol.* **9**, 64–71 (2007).

62.   Li, Y. et al. Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. *Nature* **564**, 136–140 (2018).

63.   Hayashi, K., Lopes, S. M. C., de, S., Tang, F. & Surani, M. A. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* **3**, 391–401 (2008).

64.   Kalkan, T. et al. Tracking the embryonic stem cell transition from ground state pluripotency. *Development* **144**, 1221–1234 (2017).

ARTICLE

65. Auclair, G., Guibert, S., Bender, A. & Weber, M. Ontogeny of CpG island methylation and specificity of DNMT3 methyltransferases during embryonic development in the mouse. *Genome Biol.* **15**, 545 (2014).

66. Khoueiry, R. et al. Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nat. Genet.* **49**, 1061–1072 (2017).

67. Xiong, J. et al. Cooperative action between SALL4A and TET proteins in stepwise oxidation of 5-methylcytosine. *Mol. Cell* **64**, 913–925 (2016).

68. Shen, L. et al. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* **153**, 692–706 (2013).

69. Sharif, J. et al. Activation of endogenous retroviruses in Dnmt1−/− ESCs involves disruption of SETDB1-mediated repression by NP95 Binding to Hemimethylated DNA. *Cell Stem Cell* **19**, 81–94 (2016).

70. Magnúsdóttir, E. et al. A tripartite transcription factor network regulates primordial germ cell specification in mice. *Nat. Cell Biol.* **15**, 905–915 (2013).

71. Yamaji, M. et al. PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* **12**, 368–382 (2013).

72. Okashita, N. et al. PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. *Development* **141**, 269–280 (2014).

73. Grabole, N. et al. Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. *EMBO Rep.* **14**, 629–637 (2013).

74. Han, L. et al. Embryonic defects induced by maternal obesity in mice derive from Stella insufficiency in oocytes. *Nat. Genet.* **50**, 432–442 (2018).

75. Xu, X. et al. Dppa3 expression is critical for generation of fully reprogrammed iPS cells and maintenance of Dlk1-Dio3 imprinting. *Nat. Commun.* **6**, 6008 (2015).

76. Huang, Y. et al. Stella modulates transcriptional and endogenous retrovirus programs during maternal-to-zygotic transition. *Elife* **6**, e22345 (2017).

77. Nakamura, T. et al. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415–419 (2012).

78. Shin, S.-W., John Vogt, E., Jimenez-Movilla, M., Baibakov, B. & Dean, J. Cytoplasmic cleavage of DPPA3 is required for intracellular trafficking and cleavage-stage development in mice. *Nat. Commun.* **8**, 1643 (2017).

79. Bogdanović, O. et al. Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nat. Genet.* **48**, 417–426 (2016).

80. Sharif, J. et al. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908–912 (2007).

81. Bostick, M. et al. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).

82. Nishiyama, A. et al. Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* **502**, 249–253 (2013).

83. Qin, W. et al. DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. *Cell Res.* **25**, 911–929 (2015).

84. Rothbart, S. B. et al. Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat. Struct. Mol. Biol.* **19**, 1155–1160 (2012).

85. Zhao, Q. et al. Dissecting the precise role of H3K9 methylation in crosstalk with DNA maintenance methylation in mammals. *Nat. Commun.* **7**, 12464 (2016).

86. Weihs, F. et al. Heterogeneous localisation of membrane proteins in Staphylococcus aureus. *Sci. Rep.* **8**, 3657 (2018).

87. Osswald, M., Santos, A. F. & Morais-de-Sá, E. Light-Induced Protein Clustering for Optogenetic Interference and Protein Interaction Analysis in Drosophila S2 Cells. *Biomolecules* **9**, 61 (2019).

88. Rothbart, S. B. et al. Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation. *Genes Dev.* **27**, 1288–1298 (2013).

89. Citterio, E. et al. Np95 is a histone-binding protein endowed with ubiquitin ligase activity. *Mol. Cell. Biol.* **24**, 2526–2535 (2004).

90. Karagianni, P., Amazit, L., Qin, J. & Wong, J. ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation. *Mol. Cell. Biol.* **28**, 705–717 (2008).

91. Digman, M. A. et al. Measuring fast dynamics in solutions and cells with a laser scanning microscope. *Biophys. J.* **89**, 1317–1327 (2005).

92. Digman, M. A., Wiseman, P. W., Horwitz, A. R. & Gratton, E. Detecting protein complexes in living cells from laser scanning confocal image sequences by the cross correlation raster image spectroscopy method. *Biophys. J.* **96**, 707–716 (2009).

93. Karg, E. et al. Ubiquitome analysis reveals PCNA-associated factor 15 (PAF15) as a specific ubiquitination target of UHRF1 in embryonic stem cells. *J. Mol. Biol.* **429**, 3814–3824 (2017).

94. Fang, J. et al. Hemi-methylated DNA opens a closed conformation of UHRF1 to facilitate its histone recognition. *Nat. Commun.* **7**, 11197 (2016).

95. Harrison, J. S. et al. Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1. *Elife* **5**, e17101 (2016).

96. Herce, H. D., Deng, W., Helma, J., Leonhardt, H. & Cardoso, M. C. Visualization and targeted disruption of protein interactions in living cells. *Nat. Commun.* **4**, 2660 (2013).

97. Arita, K. et al. Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. *Proc. Natl Acad. Sci. USA* **109**, 12950–12955 (2012).

98. Du, W. et al. Stella protein facilitates DNA demethylation by disrupting the chromatin association of the RING finger-type E3 ubiquitin ligase UHRF1. *J. Biol. Chem.* **294**, 8907–8917 (2019).

99. Feng, S. et al. Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl Acad. Sci. USA* **107**, 8689–8694 (2010).

100. Iyer, L. M., Tahiliani, M., Rao, A. & Aravind, L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* **8**, 1698–1710 (2009).

101. Wu, S. & Zhang, Y. Active DNA demethylation: many roads lead to Rome. *Nat. Rev. Mol. Cell Biol.* **11**, 607–620 (2010).

102. Kumar, S. & Hedges, S. B. A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920 (1998).

103. Ishiyama, S. et al. Structure of the Dnmt1 reader module complexed with a unique two-mono-ubiquitin mark on histone H3 reveals the basis for DNA methylation maintenance. *Mol. Cell* **68**, 350–360.e7 (2017).

104. Yamaguchi, L. et al. Usp7-dependent histone H3 deubiquitylation regulates maintenance of DNA methylation. *Sci. Rep.* **7**, 55 (2017).

105. Walter, R. B., Li, H.-Y., Intano, G. W., Kazianis, S. & Walter, C. A. Absence of global genomic cytosine methylation pattern erasure during medaka (Oryzias latipes) early embryo development. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **133**, 597–607 (2002).

106. Wu, X. & Zhang, Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat. Rev. Genet.* https://doi.org/10.1038/nrg.2017.33 (2017).

107. Singer, Z. S. et al. Dynamic heterogeneity and DNA methylation in embryonic stem cells. *Mol. Cell* **55**, 319–331 (2014).

108. Almeida, R. D. et al. 5-hydroxymethyl-cytosine enrichment of non-committed cells is not a universal feature of vertebrate development. *Epigenetics* **7**, 383–389 (2012).

109. Blinka, S., Reimer, M. H. Jr, Pulakanti, K. & Rao, S. Super-enhancers at the nanog locus differentially regulate neighboring pluripotency-associated genes. *Cell Rep.* **17**, 19–28 (2016).

110. Liao, J. et al. Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat. Genet.* **47**, 469–478 (2015).

111. Ma, Z., Swigut, T., Valouev, A., Rada-Iglesias, A. & Wysocka, J. Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. *Nat. Struct. Mol. Biol.* **18**, 120–127 (2011).

112. Rahjouei, A., Pirouz, M., Di Virgilio, M., Kamin, D. & Kessel, M. MAD2L2 promotes open chromatin in embryonic stem cells and derepresses the Dppa3 locus. *Stem Cell Rep.* **8**, 813–821 (2017).

113. Sang, H. et al. Dppa3 is critical for Lin28a-regulated ES cells naïve-primed state conversion. *J. Mol. Cell Biol.* **11**, 474–488 (2019).

114. Mochizuki, K., Tachibana, M., Saitou, M., Tokitake, Y. & Matsui, Y. Implication of DNA demethylation and bivalent histone modification for selective gene regulation in mouse primordial germ cells. *PLoS ONE* **7**, e46036 (2012).

115. Waghray, A. et al. Tbx3 controls Dppa3 levels and exit from pluripotency toward mesoderm. *Stem Cell Rep.* **5**, 97–110 (2015).

116. Li, X. et al. Tet proteins influence the balance between neuroectodermal and mesodermal fate choice by inhibiting Wnt signaling. *Proc. Natl Acad. Sci. USA* **113**, E8267–E8276 (2016).

117. Dai, H.-Q. et al. TET-mediated DNA demethylation controls gastrulation by regulating Lefty–Nodal signalling. *Nature* **538**, 528 (2016).

118. Verma, N. et al. TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nat. Genet.* **50**, 83–95 (2018).

119. Gu, T. et al. DNMT3A and TET1 cooperate to regulate promoter epigenetic landscapes in mouse embryonic stem cells. *Genome Biol.* **19**, 88 (2018).

120. López-Moyado, I. F. et al. Paradoxical association of TET loss of function with genome-wide DNA hypomethylation. *Proc. Natl Acad. Sci. USA* **116**, 16933–16942 (2019).

121. Han, L., Ren, C., Zhang, J., Shu, W. & Wang, Q. Differential roles of Stella in the modulation of DNA methylation during oocyte and zygotic development. *Cell Disco.* **5**, 9 (2019).

122. Payer, B. et al. Generation of stella-GFP transgenic mice: a novel tool to study germ cell development. *Genesis* **44**, 75–83 (2006).

123. Macfarlan, T. S. et al. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).

124. Dupressoir, A., Lavialle, C. & Heidmann, T. From ancestral infectious retroviruses to bona fide cellular genes: role of the captured syncytins in placentation. *Placenta* **33**, 663–671 (2012).

125. Chuong, E. B., Rumi, M. A. K., Soares, M. J. & Baker, J. C. Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat. Genet.* **45**, 325–329 (2013).

B.3 Paper 3: Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive DNA demethylation in mammals

B APPENDED PAPERS

ARTICLE

126. Grow, E. J. et al. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature* **522**, 221–225 (2015).

127. Fu, B., Ma, H. & Liu, D. Endogenous retroviruses function as gene expression regulatory elements during mammalian pre-implantation embryo development. *Int. J. Mol. Sci.* **20**, 790 (2019).

128. Hendrickson, P. G. et al. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).

129. Percharde, M. et al. A LINE1-nucleolin partnership regulates early development and ESC identity. *Cell* **174**, 391–405.e19 (2018).

130. Rodriguez-Terrones, D. & Torres-Padilla, M.-E. Nimble and ready to mingle: transposon outbursts of early development. *Trends Genet* **34**, 806–820 (2018).

131. Gill, T. J. III, Johnson, P. M., Lyden, T. W. & Mwenda, J. M. Endogenous retroviral expression in the human placenta. *Am. J. Reprod. Immunol.* **23**, 115–120 (1990).

132. Harris, J. R. Placental endogenous retrovirus (ERV): structural, functional, and evolutionary significance. *Bioessays* **20**, 307–316 (1998).

133. Wildman, D. E. et al. Evolution of the mammalian placenta revealed by phylogenetic analysis. *Proc. Natl Acad. Sci. USA* **103**, 3203–3208 (2006).

134. Selwood, L. & Johnson, M. H. Trophoblast and hypoblast in the monotreme, marsupial and eutherian mammal: evolution and origins. *Bioessays* **28**, 128–145 (2006).

135. Emera, D. & Wagner, G. P. Transposable element recruitments in the mammalian placenta: impacts and mechanisms. *Brief. Funct. Genomics* **11**, 267–276 (2012).

136. Frank, J. A. & Feschotte, C. Co-option of endogenous viral sequences for host cell function. *Curr. Opin. Virol.* **25**, 81–89 (2017).

137. Villarreal, L. P. Persistent virus and addiction modules: an engine of symbiosis. *Curr. Opin. Microbiol.* **31**, 70–79 (2016).

138. Zeh, D. W. & Zeh, J. A. Reproductive mode and speciation: the viviparity-driven conflict hypothesis. *Bioessays* **22**, 938–946 (2000).

139. Jangam, D., Feschotte, C. & Betrán, E. Transposable element domestication as an adaptation to evolutionary conflicts. *Trends Genet.* **33**, 817–831 (2017).

140. Renfree, M. B., Hore, T. A., Shaw, G., Graves, J. A. M. & Pask, A. J. Evolution of genomic imprinting: insights from marsupials and monotremes. *Annu. Rev. Genomics Hum. Genet.* **10**, 241–262 (2009).

141. Pask, A. J. et al. Analysis of the platypus genome suggests a transposon origin for mammalian imprinting. *Genome Biol.* **10**, R1 (2009).

142. Kaneko-Ishino, T. & Ishino, F. The role of genes domesticated from LTR retrotransposons and retroviruses in mammals. *Front. Microbiol.* **3**, 262 (2012).

143. Bogutz, A. B. et al. Evolution of imprinting via lineage-specific insertion of retroviral promoters. *Nat. Commun.* **10**, 5674 (2019).

144. Takahashi, N. et al. ZNF445 is a primary regulator of genomic imprinting. *Genes Dev.* **33**, 49–54 (2019).

145. Imbeault, M., Helleboid, P.-Y. & Trono, D. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature* **543**, 550–554 (2017).

146. Edwards, C. A., Takahashi, N., Corish, J. A. & Ferguson-Smith, A. C. The origins of genomic imprinting in mammals. *Reprod. Fertil. Dev.* **31**, 1203–1218 (2019).

147. Lynch, V. J., Leclerc, R. D., May, G. & Wagner, G. P. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nat. Genet.* **43**, 1154–1159 (2011).

148. Lynch, V. J. et al. Ancient transposable elements transformed the uterine regulatory landscape and transcriptome during the evolution of mammalian pregnancy. *Cell Rep.* **10**, 551–561 (2015).

149. Frankenberg, S. R., de Barros, F. R. O., Rossant, J. & Renfree, M. B. The mammalian blastocyst. *Wiley Interdiscip. Rev.: Developmental Biol.* **5**, 210–232 (2016).

150. Chavan, A. R., Griffith, O. W. & Wagner, G. P. The inflammation paradox in the evolution of mammalian pregnancy: turning a foe into a friend. *Curr. Opin. Genet. Dev.* **47**, 24–32 (2017).

151. Kunarso, G. et al. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat. Genet.* **42**, 631–634 (2010).

152. Jacques, P.-É., Jeyakani, J. & Bourque, G. The majority of primate-specific regulatory sequences are derived from transposable elements. *PLoS Genet.* **9**, e1003504 (2013).

153. Esnault, C., Cornelis, G., Heidmann, O. & Heidmann, T. Differential evolutionary fate of an ancestral primate endogenous retrovirus envelope gene, the EnvV syncytin, captured for a function in placentation. *PLoS Genet.* **9**, e1003400 (2013).

154. Haig, D. A. Going retro: transposable elements, embryonic stem cells, and the mammalian placenta. *Bioessays* **37**, 1154 (2015).

155. Haig, D. Transposable elements: Self-seekers of the germline, team-players of the soma. *Bioessays* **38**, 1158–1166 (2016).

156. Chuong, E. B. Retroviruses facilitate the rapid evolution of the mammalian placenta. *Bioessays* **35**, 853–861 (2013).

157. Smith, Z. D. et al. DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611–615 (2014).

158. Sanford, J. P., Chapman, V. M. & Rossant, J. DNA methylation in extraembryonic lineages of mammals. *Trends Genet.* **1**, 89–93 (1985).

159. Smith, Z. D. et al. Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. *Nature* **549**, 543–547 (2017).

160. Peaston, A. E. et al. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* **7**, 597–606 (2004).

161. Meyer, T. J., Rosenkrantz, J. L., Carbone, L. & Chavez, S. L. Endogenous retroviruses: with us and against us. *Front Chem.* **5**, 23 (2017).

162. Schroeder, D. I. et al. Early developmental and evolutionary origins of gene body DNA methylation patterns in mammalian placentas. *PLoS Genet.* **11**, e1005442 (2015).

163. Sato, M. et al. Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. *Mech. Dev.* **113**, 91–94 (2002).

164. Thélie, A. et al. Differential regulation of abundance and deadenylation of maternal transcripts during bovine oocyte maturation in vitro and in vivo. *BMC Dev. Biol.* **7**, 125 (2007).

165. Masala, L. et al. Delay in maternal transcript degradation in ovine embryos derived from low competence oocytes. *Mol. Reprod. Dev.* **85**, 427–439 (2018).

166. Wasielak, M., Więsak, T., Bogacka, I., Jalali, B. M. & Bogacki, M. Zygote arrest 1, nucleoplasmin 2, and developmentally associated protein 3 mRNA profiles throughout porcine embryo development in vitro. *Theriogenology* **86**, 2254–2262 (2016).

167. Boroviak, T. et al. Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. *Development* **145**, dev.167833 (2018).

168. Hayashi, K. & Saitou, M. Generation of eggs from mouse embryonic stem cells and induced pluripotent stem cells. *Nat. Protoc.* **8**, 1513–1524 (2013).

169. Mulholland, C. B. et al. A modular open platform for systematic functional studies under physiological conditions. *Nucleic Acids Res.* **43**, e112 (2015).

170. Kowarz, E., Löscher, D. & Marschalek, R. Optimized Sleeping Beauty transposons rapidly generate stable transgenic cell lines. *Biotechnol. J.* **10**, 647–653 (2015).

171. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

172. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).

173. Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G. F. & Chin, L. Post-translational regulation of Cas9 during G1 enhances homology-directed repair. *Cell Rep.* **14**, 1555–1566 (2016).

174. Mátés, L. et al. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat. Genet.* **41**, 753–761 (2009).

175. Méndez, J. & Stillman, B. Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol. Cell. Biol.* **20**, 8602–8612 (2000).

176. Bauer, C. et al. Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). *J. Biol. Chem.* **290**, 4801–4812 (2015).

177. Shintomi, K. & Hirano, T. Releasing cohesin from chromosome arms in early mitosis: opposing actions of Wapl–Pds5 and Sgo1. *Genes Dev.* 23, 2224–2236 (2009).

178. Françon, P. et al. A hypophosphorylated form of RPA34 is a specific component of pre-replication centers. *J. Cell Sci.* **117**, 4909–4920 (2004).

179. Wagner, M. et al. Age-dependent levels of 5-methyl-, 5-hydroxymethyl-, and 5-formylcytosine in human and mouse brain tissues. *Angew. Chem. Int. Ed. Engl.* **54**, 12511–12514 (2015).

180. Soumillon, M., Cacchiarelli, D., Semrau, S., van Oudenaarden, A. & Mikkelsen, T. S. Characterization of directed differentiation by high-throughput single-cell RNA-Seq. *bioRxiv* https://doi.org/10.1101/003236 (2014).

181. Ziegenhain, C. et al. Comparative analysis of single-cell RNA sequencing methods. *Mol. Cell* **65**, 631–643.e4 (2017).

182. Bagnoli, J. W. et al. Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat. Commun.* **9**, 2937 (2018).

183. Boyle, P. et al. Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biol.* **13**, R92 (2012).

184. Parekh, S., Ziegenhain, C., Vieth, B., Enard, W. & Hellmann, I. zUMIs—a fast and flexible pipeline to process RNA sequencing data with UMIs. *Gigascience* **7**, giy059 (2018).

185. Rau, A., Gallopin, M., Celeux, G. & Jaffrézic, F. Data-based filtering for replicated high-throughput transcriptome sequencing experiments. *Bioinformatics* **29**, 2146–2152 (2013).

186. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

187. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

ARTICLE

188. Akalin, A. et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* **13**, R87 (2012).
189. Illingworth, R. S. et al. Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet.* **6**, e1001134 (2010).
190. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
191. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
192. Sporbert, A., Domaing, P., Leonhardt, H. & Cardoso, M. C. PCNA acts as a stationary loading platform for transiently interacting Okazaki fragment maturation proteins. *Nucleic Acids Res.* **33**, 3521–3528 (2005).
193. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
194. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
195. McQuin, C. et al. CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol.* **16**, e2005970 (2018).
196. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128 (2007).
197. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
198. Scheltema, R. A. & Mann, M. SprayQc: a real-time LC–MS/MS quality monitoring system to maximize uptime using off the shelf components. *J. Proteome Res.* **11**, 3458–3466 (2012).
199. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367 (2008).
200. Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
201. Cox, J., Hein, M. Y., Luber, C. A., Paron, I. & Nagaraj, N. MaxLFQ allows Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).
202. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731 (2016).
203. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
204. Iwamatsu, T. Stages of normal development in the medaka Oryzias latipes. *Mech. Dev.* **121**, 605–618 (2004).
205. Hendrix, J. et al. Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-interacting Gag oligomers. *J. Cell Biol.* **210**, 629–646 (2015).
206. Müller, B. K., Zaychikov, E., Bräuchle, C. & Lamb, D. C. Pulsed interleaved excitation. *Biophys. J.* **89**, 3508–3522 (2005).
207. Schrimpf, W., Barth, A., Hendrix, J. & Lamb, D. C. PAM: a framework for integrated analysis of imaging, single-molecule, and ensemble fluorescence data. *Biophys. J.* **114**, 1518–1528 (2018).
208. Hendrix, J., Dekens, T., Schrimpf, W. & Lamb, D. C. Arbitrary-region raster image correlation spectroscopy. *Biophys. J.* **111**, 1785–1796 (2016).
209. Wang, H.-Q., Tuominen, L. K. & Tsai, C.-J. SLIM: a sliding linear model for estimating the proportion of true null hypotheses in datasets with dependence structures. *Bioinformatics* **27**, 225–231 (2011).
210. Payer, B. et al. Stella is a maternal effect gene required for normal early development in mice. *Curr. Biol.* **13**, 2110–2117 (2003).

Author contributions
C.B.M. and S.B. designed and conceived the study. S.B. and H.L. supervised the study. C.B.M., S.B., and H.L. prepared the manuscript with the help of M.D.B. C.B.M. performed cellular and molecular experiments. C.B.M. generated cell lines with help from M.Y., C.B.M. performed RRBS and RNA-Seq with help and supervision from C.Z., S.B., and W.E., J.R. and C.B.M. performed live-cell microscopy and photobleaching analyses. I.G., J.R., and C.B.M. performed MST and F3H assays. A.N. performed Xenopus experiments under the supervision of M.N., R.N. performed the experiments in medaka embryos under the supervision of H.T. M.D.B., and P.S. helped with cell line validation and performed fluorescence microscopy analysis. W.Q. performed the biochemical analyses with assistance from A.A., M.M. performed hESC experiments. E.U. conducted proteomics experiments and analyses under the guidance of M.W. F.R.T., and E.P. quantified modified cytosines by LC-MS/MS with the supervision by T.C. S.B. performed data analysis. All authors read, discussed, and approved the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-19603-1.

Correspondence and requests for materials should be addressed to S.B. or H.L.

Peer review information *Nature Communications* thanks Maxim Greenberg, Michael Weber and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

SUPPLEMENTARY FIGURES for:

**Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive demethylation in mammals**

Authors:

Christopher B. Mulholland[1], Atsuya Nishiyama[2†], Joel Ryan[1†], Ryohei Nakamura[3], Merve Yiğit[1], Ivo M. Glück[4], Carina Trummer[1], Weihua Qin[1], Michael D. Bartoschek[1], Franziska R. Traube[5], Edris Parsa[5], Enes Ugur[1,8], Miha Modic[7], Aishwarya Acharya[1], Paul Stolz[1], Christoph Ziegenhain[6], Michael Wierer[8], Wolfgang Enard[6], Thomas Carell[5], Don C. Lamb[4], Hiroyuki Takeda[3], Makoto Nakanashi[2], Sebastian Bultmann*[1§] and Heinrich Leonhardt*[1§]

Correspondence: bultmann@bio.lmu.de or h.leonhardt@lmu.de

This PDF includes:

1
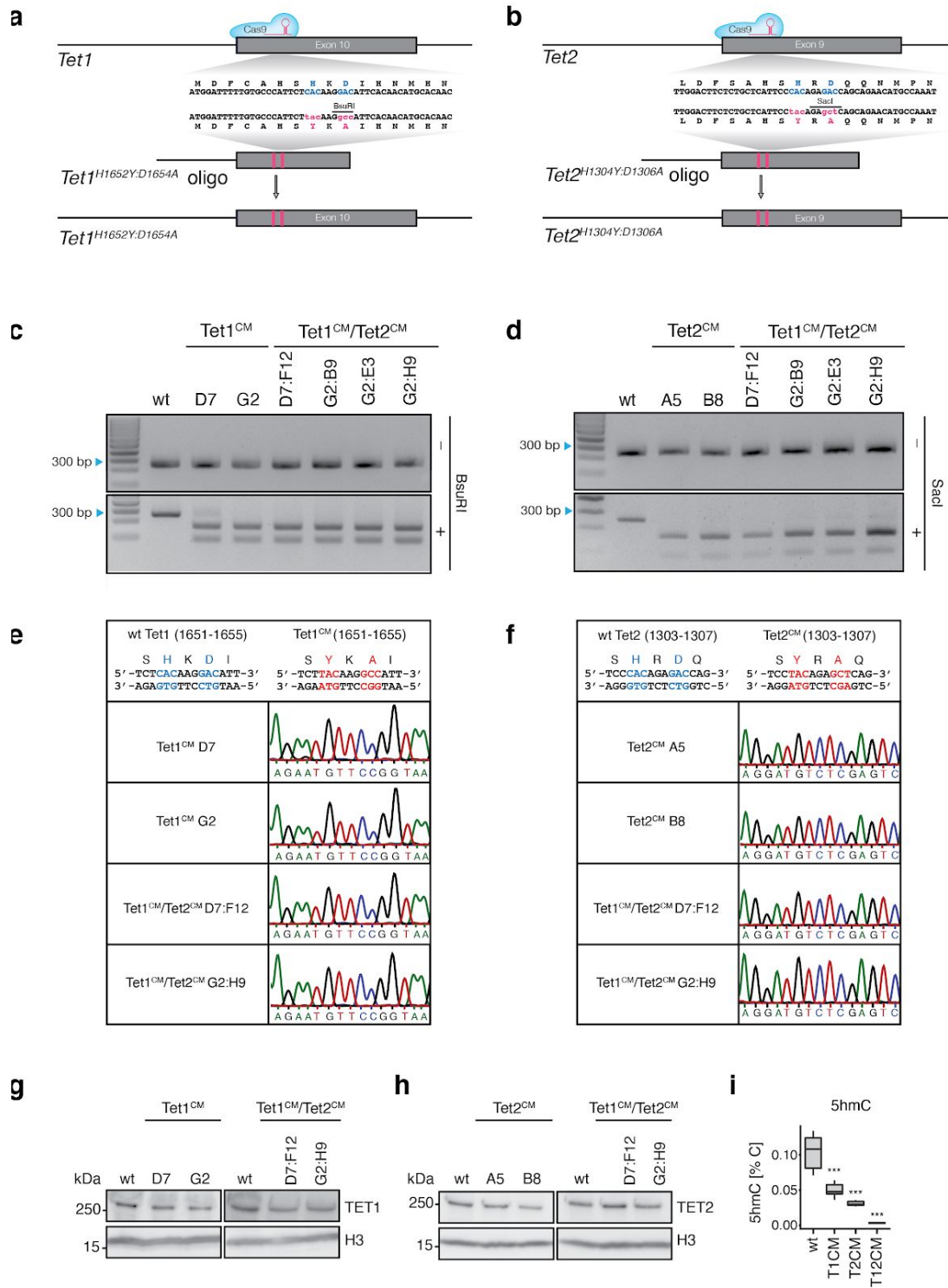
Supplementary Figure 1: Generation and characterization of T1CM, T2CM, and T12CM mESCs

a,b, Schematic representation of the CRISPR/Cas9 gene editing strategies used to mutate the catalytic center (HxD iron-chelating motif) of *Tet1* (**a**) and *Tet2* (**b**). gRNA target sequences and restriction enzyme recognition sites for restriction fragment length polymorphism (RFLP) screening are shown (See Supplementary Data 5). **c,d,** Genotyping using RFLP analysis of *Tet1* (**c**) and *Tet2* (**d**) loci. Data in c and d are representative of n = 2 independent experiments. **e,f,** DNA sequencing traces confirming the successful mutation of the HxD motif to YxA in *Tet1* (**e**) and *Tet2* (**f**). The top boxes depict the DNA and corresponding amino acid sequences of the HxD motif before (blue) and after (red) gene editing. **g,h,** Immunoblot detection of endogenous TET1 (**g**) and TET2 (**h**) protein levels in T1CM, T2CM, and T12CM mESC. Data in g and h are representative of n = 3 independent experiments. **i,** DNA modification levels as percentage of total cytosines measured in wt (n = 24), T1CM (n = 8), T2CM (n = 12), and T12CM (n = 11) mESC biological replicates. Depicted are mean values ± standard deviation; ***P < 2e-16 to wt as determined using a one-way ANOVA followed by a post-hoc Tukey HSD test. Source data are provided as a Source Data file.

3

4
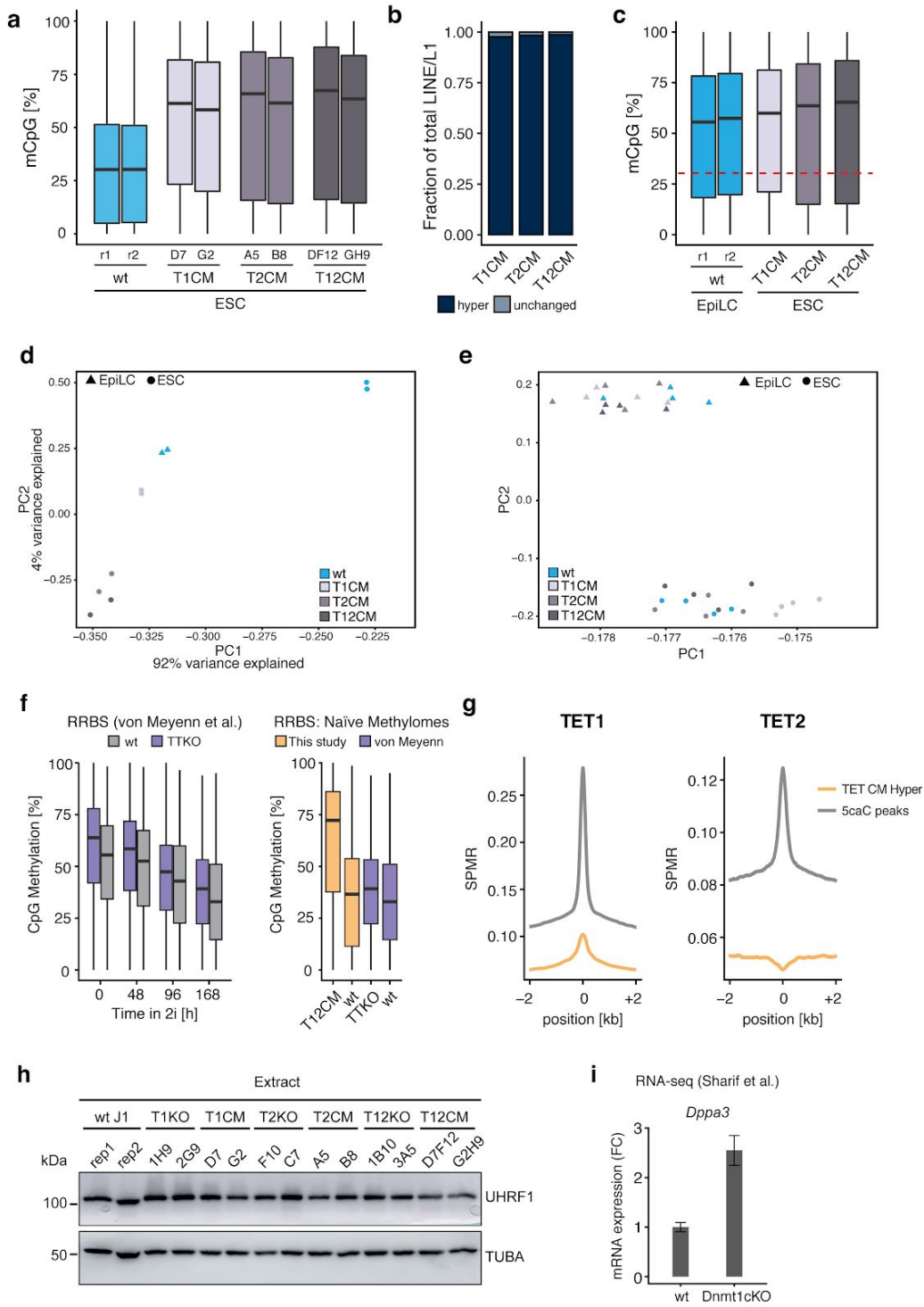
Supplementary Figure 2: Methylome and transcriptome analysis of T1CM, T2CM, and T12CM ESCs

a, Percentage of total 5mC as measured by RRBS in all analyzed wt, T1, T2, and T12 CM ESC replicates from n = 2 independent cell lines per genotype. **b,** Relative proportion of DNA hypermethylation (q value < 0.05; absolute methylation difference > 20%) at LINE1/L1 elements in T1CM, T2CM, and T12CM ESCs compared to wt ESCs. For each genotype, n = 2 independent clones. **c,** Loss of TET catalytic activity in ESCs results in similar or higher DNA methylation levels than in wt EpiLCs. Comparison of the total CpG DNA methylation (5mCpG) as measured by RRBS between individual wt EpiLC samples (n = 1 for each depicted cell line) and T1CM, T2CM, and T12CM ESCs (n = 2 independent cell lines per genotype). Dashed red line indicates the median mCpG methylation in wt EpiLCs. **d,** Principal component (PC) analysis of RRBS data from wt, T1CM, T2CM, and T12CM ESCs and wt EpiLCs. **e,** PC analysis of RNA-seq data from wt, T1CM, T2CM, and T12CM ESCs during EpiLC differentiation. **f,** Boxplots comparing total 5mC levels in Tet catalytic mutants (T12CM) and Tet triple (TET1, TET2, TET3) knockout ESCs (TTKO)[1] based on RRBS measurements from n = 2 biological replicates per genotype. **g,** TET1[2] and TET2[3] occupancy over 1 kb tiles hypermethylated (orange) in T1CM and T2CM ESCs or regions of 5caC enrichment (grey)[4]. (SPMR: Signal per million reads). **h,** UHRF1 protein levels are unaffected by inactivation or loss of TET proteins. Whole-cell extracts from wt J1, TET catalytic mutant (T1CM, T2CM, and T12CM) and TET knockout (*Tet1* KO (T1KO), *Tet2* KO (T2KO), and *Tet1/2* KO (T12KO))[5] ESCs analyzed by immunoblot detection using the indicated antibodies. Immunoblot depicts *n* = 2 biological replicates (wt J1: rep1 and rep2) or independent clones (TET CM and TET KO ESC lines) for each genotype. **i,** *Dppa3* mRNA levels in wild type and conditional *Dnmt1* knockout cells (Dnmt1cKO)[6]. Error bars indicate the mean ± SD calculated from n = 3 biological replicates. For the boxplots in **a, c,** and **f,** horizontal black lines within boxes represent median values, boxes indicate the upper and lower quartiles, and whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge. Source data are provided as a Source Data file.
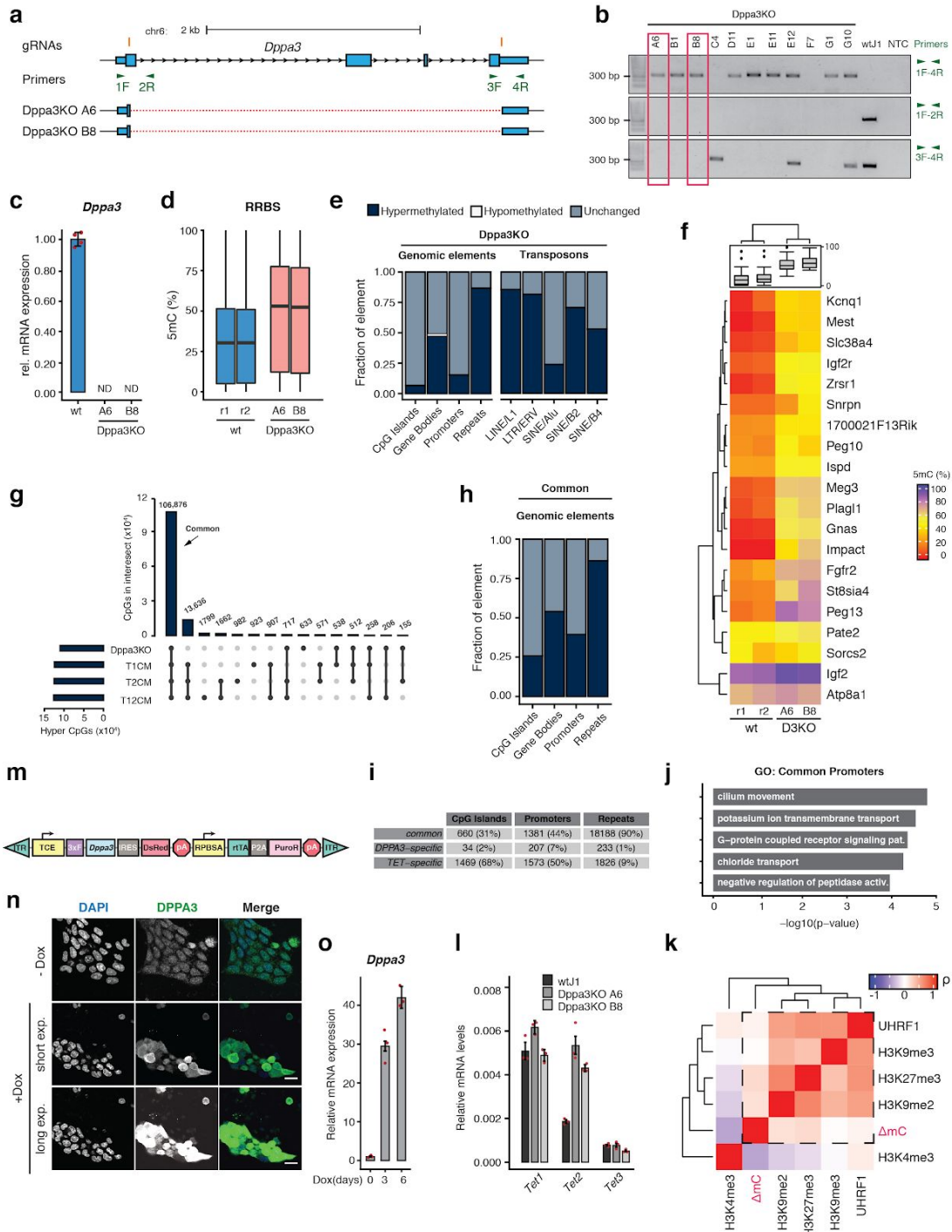
Supplementary Figure 3: Generation, characterization, and methylome profiling of Dppa3KO ESCs

a, Schematic representation of the CRISPR/Cas9 gene editing strategy used to excise the entire *Dppa3* locus. The position of the two locus-flanking gRNAs are shown in orange. PCR primers for determination of locus removal and zygosity are indicated in green. **b,** Results of PCR using the primers indicated in (**a**). *Dppa3* knockout (Dppa3KO) clones, A6 and B8, chosen for further experiments are indicated by the red boxes. The wild-type (wt) and no template control (NTC) reactions are depicted on the right. Data are representative of n = 2 independent experiments. The sequence alignment of the amplicon generated using primers 1 and 4 for Dppa3KO clones A6 and B8 is provided in the lower portion of (**a**) with solid boxes and the dashed lines representing successfully aligned sequences and genomic sequences not found in the sequenced amplicons respectively. **c,** *Dppa3* transcript levels of the two Dppa3KO clones A6 and B8 are depicted relative to mRNA levels in wt J1 ESCs after normalization to *Gapdh*. Error bars indicate mean ± SD calculated from technical triplicate reactions from *n* = 4 biological replicates. N.D., expression not detectable. **d,** Percentage of total 5mC as measured by RRBS in wt and Dppa3KO ESC replicates (n = 2 independent cell lines per genotype). Horizontal black lines within boxes represent median values, boxes indicate the upper and lower quartiles, and whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge. **e,** Relative proportion of DNA hypermethylation occurring at each genomic element or retrotransposon class in Dppa3KO ESCs. **f,** Heatmap depicting DNA methylation levels over ICRs in wild type and Dppa3KO ESCs. **g,** Upset plot depicting the overlap of hypermethylated CgGs between DPPAKO, T1CM, T2CM, and T12CM ESCs. Vertical bar plots indicate numbers of overlapping CpGs for each intersection. Intersections are indicated by filled circles connected by black lines. Horizontal bar plots indicate total numbers of hypermethylated CpGs for each cell line. **h,** Relative proportion of DNA hypermethylation common to T1CM, T2CM, and T12CM ESCs and Dppa3KO ESCs at each class of genomic element. Dark blue indicates common hypermethylated elements. **i,** Summary of differentially methylated regions either unique to TET mutants (TET-specific) and Dppa3KO ESCs (DPPA3-specific) or shared among TET mutants and Dppa3KO ESCs (common). **j,** Gene ontology (GO) terms associated with promoters specifically dependent on TET activity; adjusted p-values calculated using two-sided Fisher's exact test. **k,** Cross-correlation analysis between DNA hypermethylation (ΔmC) in Dppa3KO ESCs and genomic occupancy of histone modifications and UHRF1 binding. Correlations are calculated over 1kb tiles. Positive correlations with ΔmC are bounded by a dashed-line square. **l,** Hypermethylation in Dppa3 KO ESCs is not associated with *Tet* downregulation. Expression of *Tet* genes in wt Dppa3KO ESC clones depicted as mRNA levels relative to *Gapdh*. Error bars indicate mean ± SD calculated from technical triplicate reactions from *n* = 2 biological replicates per genotype depicted. **m,** Schematic representation of the pSBtet-D3 (pSBtet-3xFLAG-Dppa3-IRES-DsRed) cassette for the Sleeping Beauty transposition-mediated generation of doxycycline (Dox) inducible *Dppa3* ESC lines. Abbreviations: inverted terminal repeat (ITR), tetracycline response element plus minimal CMV (TCE), 3xFLAG tag (3xF), internal ribosomal entry site (IRES), polyA signal (pA), constitutive RPBSA promoter (RPBSA), reverse tetracycline-controlled transactivator (rtTA), self-cleaving peptide P2A (P2A), puromycin resistance (PuroR). **n,** Confirmation of DPPA3 protein induction as assessed by immunofluorescence in uninduced ESCs (-Dox) or after 24 h of doxycycline treatment (+Dox). To illustrate the increase in DPPA3 protein levels, the same acquisition settings used to detect DPPA3, including a long exposure, in uninduced cells were applied for detection of DPPA3 after induction (+Dox (long exp.); bottom panel) leading to a saturated signal. Shorter exposure settings were also applied to the induced cells (+Dox (short exp.); middle panel) to better resolve the localization of DPPA3 after induction. The staining was repeated in 3 biological replicates with similar results. Scale bar: 10 μm. **o,** *Dppa3* expression before induction and after 3 or 6 days of doxycycline treatment in pSBtet-D3 ESCs. mRNA levels of *Dppa3* are shown relative to those in uninduced wt J1 ESCs after normalization to *Gapdh*. Error bars indicate mean ± SD calculated from technical triplicate reactions from *n* = 3 biological replicates.

In **e, g and h,** hypermethylation is defined as a gain in 5mC compared to wt ESCs (q value < 0.05; absolute methylation difference > 20%) with q-values calculated using a two-sided Wald test followed by p-value adjustment using SLIM. Source data are provided as a Source Data file.

7
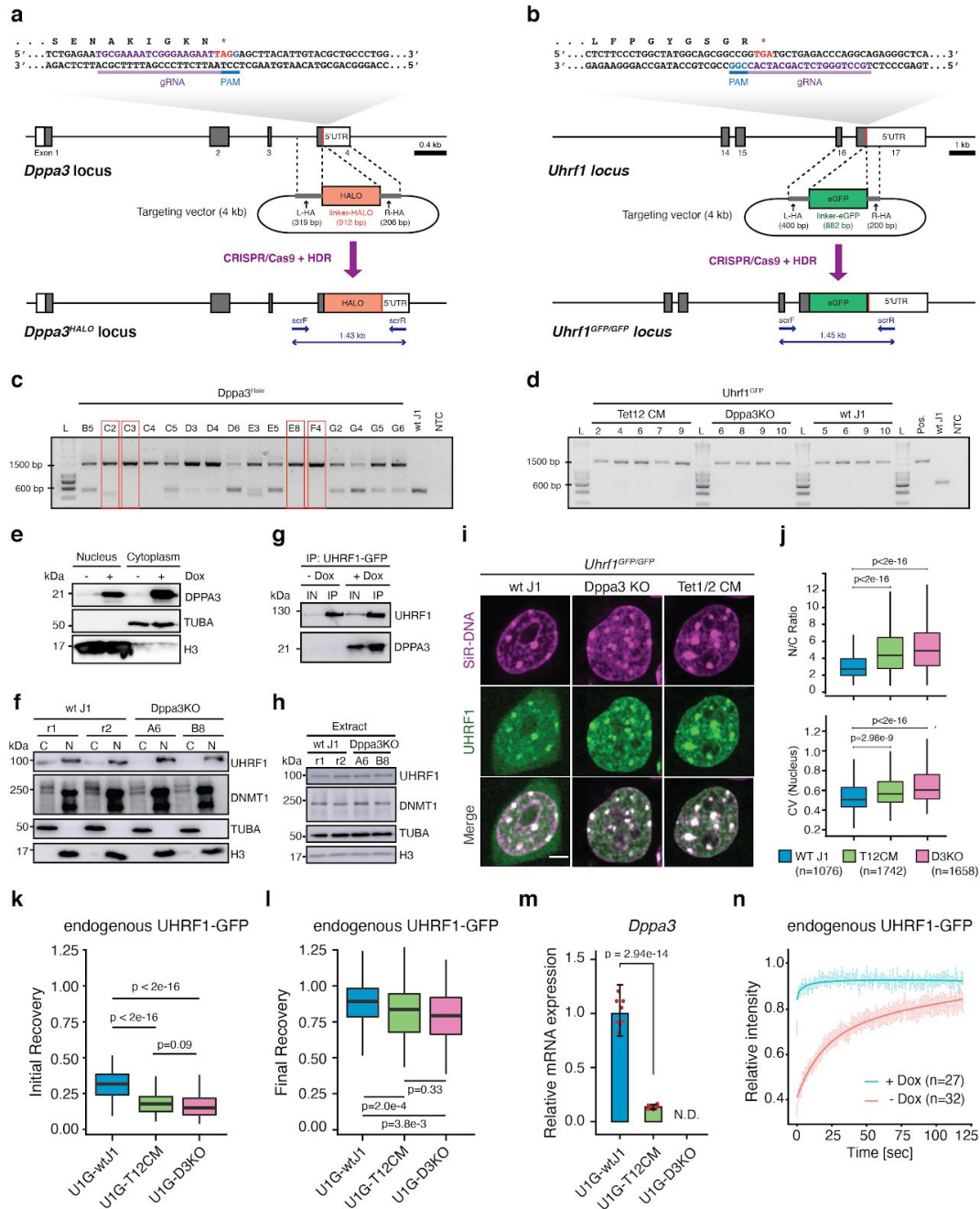
**Supplementary Figure 4: TET proteins control UHRF1 chromatin association in mESCs by regulating DPPA3 levels**

a,b, Schematic of the CRISPR/Cas9 targeting strategy used to insert fluorescent protein tags at the (**a**) *Dppa3* and (**b**) *Uhrf1* loci. Exons are depicted as boxes, with the coding sequences shaded in grey. Stop codons are indicated with a red line. The nucleotide and amino acid sequence encompassing the gRNA target site are enlarged for detail. gRNAs target sequences are colored purple and their respective PAMs blue. Donor constructs harboring the coding sequence for (**a**) HALO or (**b**) eGFP are flanked by homology arms (L-HA, left homology arm and R-HA, right homology arm), which are indicated with grey rectangles. Genotyping PCR primers used in (**c,d**) are shown as arrows along with the size of the amplicon confirming successful integration of the donor sequence. **c,d,** Results of genotyping PCRs for confirming the successful integration of (**c**) HALO into the *Dppa3* locus of wt J1 ESCs and (**d**) eGFP into the *Uhrf1* locus of wt J1, Dppa3KO, and T12CM ESCs. Wild-type (wt) and a negative control without DNA template (NTC) are depicted on the right. **c,** Clones with Sanger-sequence validated homozygous insertion of the HALO coding sequence are indicated with red boxes. **d,** PCR results after pre-screening via microscopy. All depicted clones have a correct homozygous insertion of eGFP as validated by Sanger-sequencing. PCRs were performed using the primers depicted in (**a,b**). **e,** DPPA3 is primarily but not exclusively localized to the cytoplasm in ESCs. Cytoplasmic and nuclear fractions from doxycycline (Dox) inducible FLAG-*Dppa3* ESC lines in the presence (+) or absence of Dox analyzed by immunoblot detection using the indicated antibodies or, in the case of DPPA3, an anti-FLAG antibody. **f,** DPPA3 regulates subcellular UHRF1 distribution. Cytoplasmic, "C", and nuclear, "N", fractions from wt and Dppa3KO ESCs analyzed by immunoblot detection using the indicated antibodies. The fractionation and immunoblot were repeated 2 times with similar results obtained. **g,** Immunoblot detection of anti-GFP immunoprecipitated material (IP) and the corresponding input (IN) extracts from U1G/wt + pSBtet-D3 ESCs treated before (-Dox) and after (+Dox) 24 h induction of *Dppa3* with doxycycline. Immunoblot was repeated 2 times with similar results obtained. **h,** UHRF1 and DNMT1 protein levels are unaffected by DPPA3 loss. Whole-cell extracts from wt and Dppa3KO ESCs analyzed by immunoblot detection using the indicated antibodies. Immunoblot depicts n = 2 biological replicates. **i,j,** DPPA3 alters the localization and nuclear patterning of endogenous UHRF1-GFP. **i,** Representative confocal images of UHRF1-GFP in live U1G-wt, U1G-T12CM, U1G-Dppa3KO ESCs with DNA counterstain (SiR-Hoechst). Scale bar: 5 μm. **j,** Quantification of endogenous UHRF1-GFP (top panel) nucleus to cytoplasm ratio (N/C Ratio) and (bottom panel) coefficient of variance (CV) within the nucleus of cells (number indicated in the plot) from n = 3 biological replicates per genotype. The assay was repeated 2 times with similar results. **k,l,** Endogenous DPPA3 prevents excessive UHRF1 chromatin binding in ESCs. Further analysis of the single cell FRAP data presented in Fig. 4c. showing the (**k**) initial and (**l**) final relative recovery of endogenous UHRF1-GFP. Intensity measurements 1 s and 60 s after photobleaching were used for calculating (**k**) initial recovery and (**l**) final recovery, respectively. **m,** *Dppa3* is significantly downregulated in U1G-T12CM ESCs. *Dppa3* transcript levels in U1G-wt, U1G-T12CM, U1G-Dppa3KO ESCs are depicted relative to mRNA levels in U1G-wt J1 ESCs after normalization to Gapdh. Error bars indicate mean ± SD calculated from technical triplicate reactions from n = 4 biological replicates (from n= 2 independent clones per genotype). N.D., expression not detectable. **n,** DPPA3 expression abolishes UHRF1 chromatin binding. FRAP analysis of endogenous UHRF1-GFP before (-Dox) and after 48 h of *Dppa3* induction (+Dox) in U1G-wt + pSBtet-D3 ESCs. The FRAP experiment was repeated at least 3 times with similar results obtained. In the boxplots in **j-l**, darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge. P-values in **j-m**; Welch's two-sided t-test. In **n**, the mean fluorescence intensity of n cells (indicated in the plots) at each timepoint depicted as a shaded dot. Error bars indicate mean ± SEM. Curves (solid lines) indicate double-exponential functions fitted to the FRAP data. Source data are provided as a Source Data file.

9
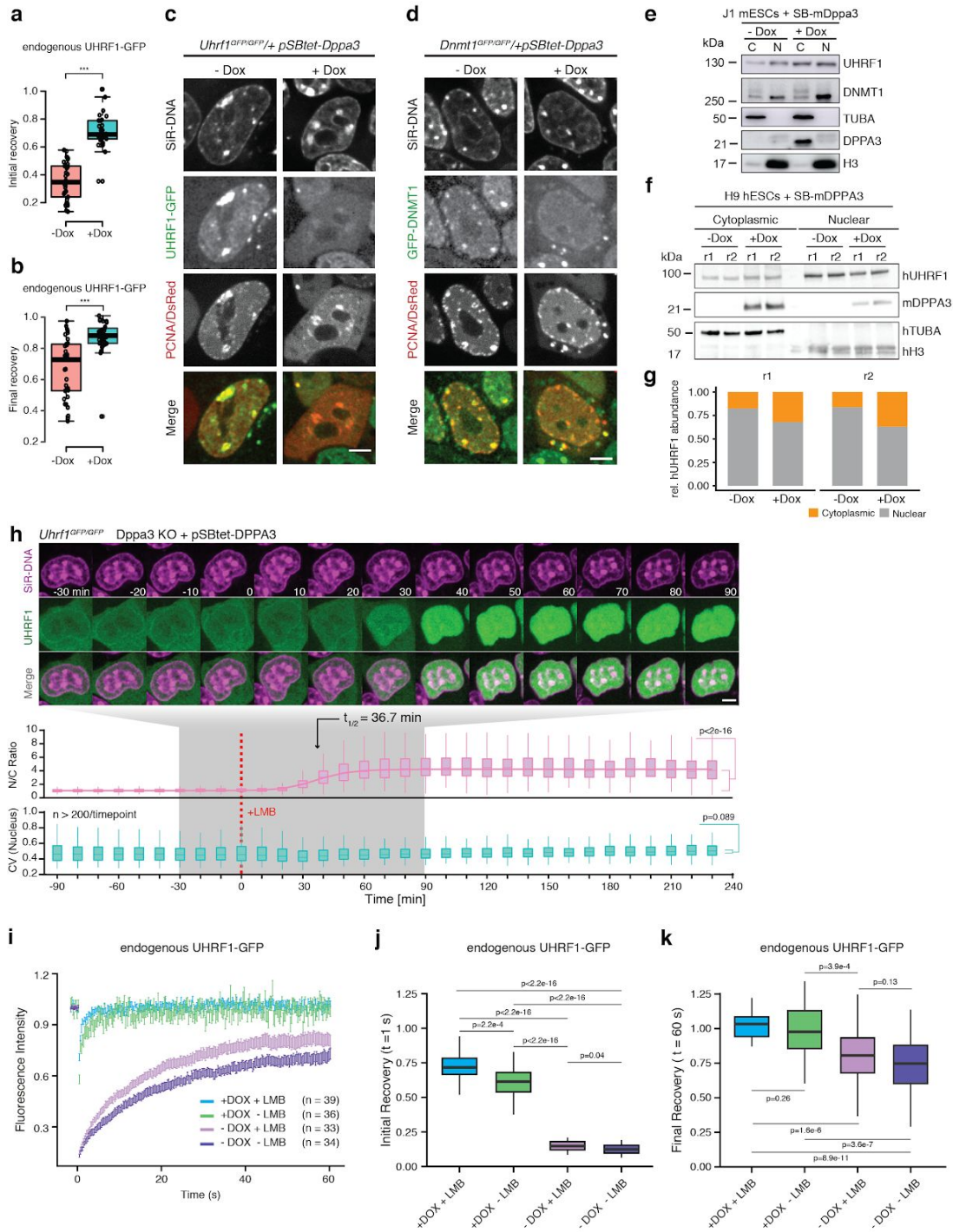
Supplementary Figure 5: DPPA3 alters the localization and chromatin binding of endogenous UHRF1 and DNMT1 in ESCs

a,b, DPPA3 releases UHRF1 from chromatin. Initial (1 s post-bleach) (**a**) and final (120 s post-bleach) (**b**) relative recovery of UHRF1-GFP in *Uhrf1^{GFP/GFP}/SBtet-3xFLAG-Dppa3* ESCs before (-Dox) and after 48 h of *Dppa3* induction (+Dox) (+Dox: *n* = 27; -Dox: *n* = 32) (data from FRAP experiment in Supplementary Fig. 4m). **c,d,** DPPA3 disrupts the recruitment of UHRF1 and DNMT1 to replication foci. Live cell imaging illustrating the late S-phase localization of (**c**) UHRF1-GFP in U1G/wt + pSBtet-D3 ESCs or (**d**) GFP-DNMT1 in live *Dnmt1^{GFP/GFP}* + pSBtet-D3 ESCs before (-Dox) and after 48 h of *Dppa3* induction. Transfected RFP-PCNA marks sites of active DNA replication within the nucleus. In both c and d, expressed free DsRed is a marker of doxycycline induction (see cytoplasmic signal in +Dox PCNA/DsRed panels and Supplementary Fig. 3m). DNA was stained in live cells using a 30 min treatment of SiR-DNA (SiR-Hoechst). Data are representative of n = 2 independent experiments. Scale bar: 5 μm. **e-g,** DPPA3 alters the subcellular distribution of endogenous UHRF1 in mouse ESCs and human ESCs. **e,f,** Immunoblots of nuclear, "N", and cytoplasmic, "C", fractions from (**e**) U1G/wt + pSBtet-D3 mouse ESCs and (**f**) Human H9 ESCs + SBtet-3xFLAG-Dppa3 before (-Dox) and after 24 hours of *Dppa3* induction (+Dox) using the indicated antibodies. An anti-FLAG antibody was used for detection of FLAG-DPPA3. **g,** Quantification of the relative abundance of hUHRF1 in the cytoplasmic and nuclear fractions shown in (**f**) with the results of two independent biological replicates (r1 and r2) displayed. **h,** Nuclear export is dispensable for DPPA3-mediated inhibition of UHRF1 chromatin association. Localization dynamics of endogenous UHRF1-GFP in response to inhibition of nuclear export using leptomycin-B (LMB) after *Dppa3* induction in U1G/D3KO + pSBtet-D3 ESCs with confocal time-lapse imaging over 5.5 h (10 min intervals). *t = 0* corresponds to start of nuclear export inhibition (+LMB). (*top panel*) Representative images of UHRF1-GFP and DNA (SiR-Hoechst stain) throughout confocal time-lapse imaging. Scale bar: 5 μm. (middle panel) Nucleus to cytoplasm ratio (N/C Ratio) of endogenous UHRF1-GFP signal. (bottom panel) Coefficient of variance (CV) of endogenous UHRF1-GFP intensity in the nucleus. (*middle and bottom panel*) N/C Ratio and CV values: measurements in *n* > 200 single cells per time point (precise values can be found in the Source Data file), acquired at *n* = 20 separate positions . Curves represent fits of four parameter logistic (4PL) functions to the N/C Ratio (pink line) and CV (green line) data. The experiment was repeated 2 times with similar results. **i-k,** Nuclear export is not only dispensable for but attenuates DPPA3-mediated inhibition of UHRF1-GFP chromatin binding. **i,** FRAP analysis of endogenous UHRF1-GFP within the nucleus of U1G/D3KO + pSBtet-D3 ESCs before (-Dox) and after 48 h of *Dppa3* induction and before (-LMB) and after (+LMB) 3 h of leptomycin-B (LMB) treatment. The mean fluorescence intensity of *n* cells (indicated in the plots) at each timepoint is depicted as a shaded dot. Error bars indicate mean ± SEM. The FRAP experiment was repeated 3 times with similar results. **j,k,** Boxplots showing the (**j**) initial (0.25 s post-bleach) and (**k**) final (57.75 s post-bleach) relative recovery of endogenous UHRF1-GFP in the FRAP experiment plotted in (**i**). For each condition, recovery intensity was calculated from the same n cells depicted in (**i**).

For the boxplots in **a, b, h, j, k,** darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge. All p-values were calculated using Welch's two-sided t-test and are depicted in the figure or as follows *** P < 2e-16. Source data are provided as a Source Data file.

11
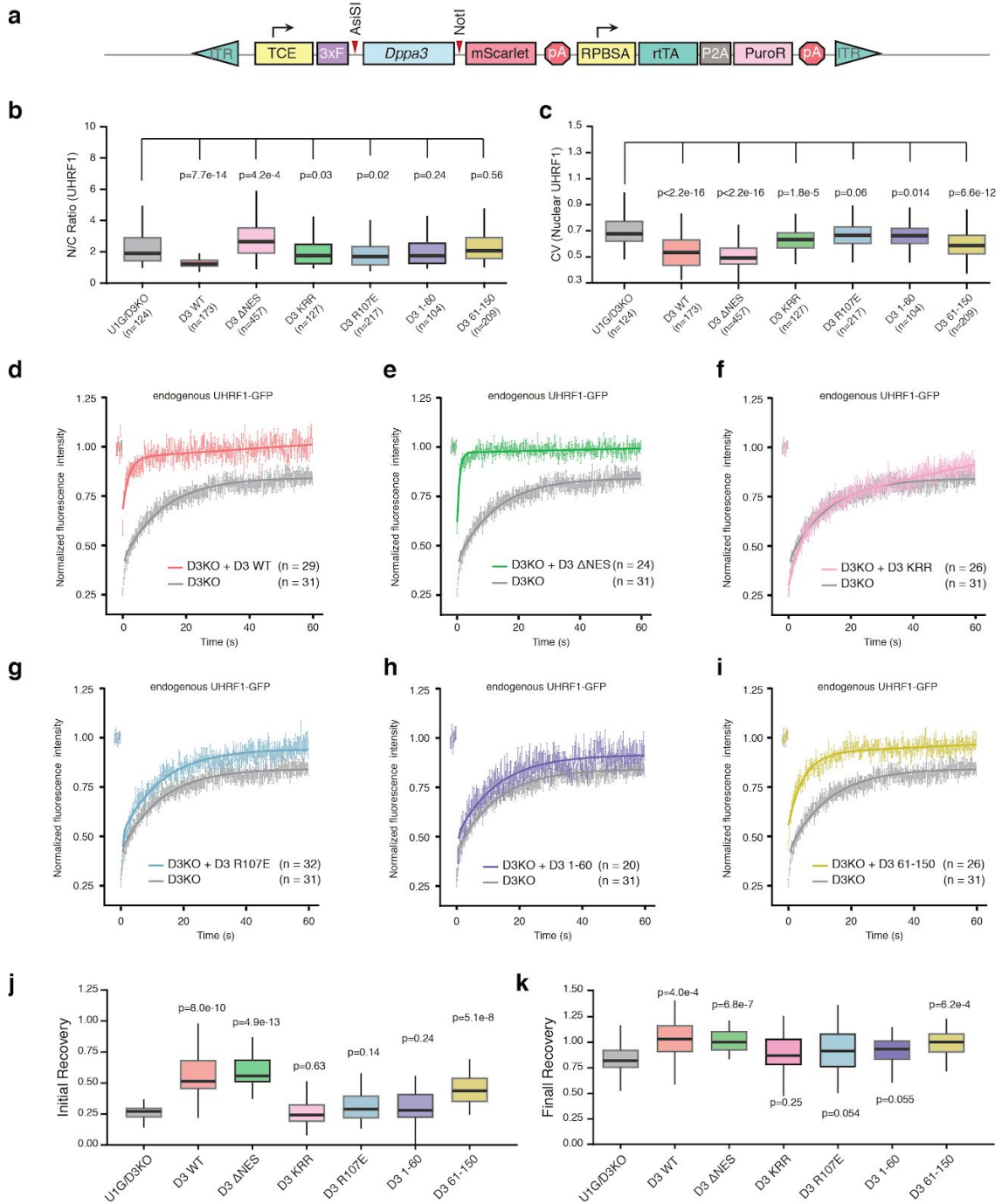
Supplementary Figure 6: Further analysis of endogenous UHRF1-GFP localization and chromatin binding in response to mutant forms of mDPPA3

a, Schematic representation of the pSBtet-D3-mSC (pSBtet-3xFLAG-Dppa3-mScarlet) cassette for the Sleeping Beauty transposition-mediated generation of ESC lines with doxycycline (Dox) inducible expression of DPPA3-mScarlet fusions. Abbreviations: inverted terminal repeat (ITR), tetracycline response element plus minimal CMV (TCE), 3xFLAG tag (3xF), internal ribosomal entry site (IRES), polyA signal (pA), constitutive RPBSA promoter (RPBSA), reverse tetracycline-controlled transactivator (rtTA), self-cleaving peptide P2A (P2A), puromycin resistance (PuroR). Restriction sites used for exchanging the *Dppa3* coding sequences are indicated with red arrows. **b,c,** DPPA3 alters the nuclear localization of UHRF1 independently of promoting UHRF1 nucleocytoplasmic translocation. Quantification of endogenous UHRF1-GFP (**b**) nucleus to cytoplasm ratio (N/C Ratio) and (**c**) coefficient of variance (CV) within the nucleus of U1G/D3KO + pSBtet-D3 ESCs expressing the indicated mutant forms of *Dppa3*. The data depicted for each mutant in (b) and (c) was derived from analysis of the same *n* cells (indicated in the plot). **d-i** Nuclear export and the N-terminus of DPPA3 are dispensable for its inhibition of UHRF1 chromatin binding. FRAP analysis of endogenous UHRF1-GFP within the nucleus of U1G/D3KO + pSBtet-D3 ESCs expressing the following forms of DPPA3: (**d**) wild-type (D3 WT), (**e**) nuclear export mutant L44A/L46A (D3-ΔNES), (**f**) K85E/R85E/K87E mutant (D3-KRR), (**g**) R107E mutant (D3-R107E), (**h**) N-terminal 1-60 fragment (D3 1-60), (**i**) C-terminal 61-150 fragment (D3 61-150). In **d-i**, the mean fluorescence intensity of *n* cells (indicated in the plots) at each timepoint is depicted as a shaded dot. Error bars indicate mean ± SEM. Curves (solid lines) indicate double-exponential functions fitted to the FRAP data. Individual same fits correspond to those in Fig. 5c. FRAP of UHRF1 in cells expressing *Dppa3* mutants was repeated 2 times. **j,k,** Boxplots showing the (**j**) initial (0.25 s post-bleach) and (**k**) final (57.75 s post-bleach) relative recovery of endogenous UHRF1-GFP in the FRAP experiments plotted in **(d-i).** For each mutant, recovery intensity was calculated from the same n cells depicted in the corresponding plots **(d-i).**

In the boxplots in **b, c, j,** and **k,** darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge. P-values were calculated using Welch's two-sided t-tests. Source data are provided as a Source Data file.

13
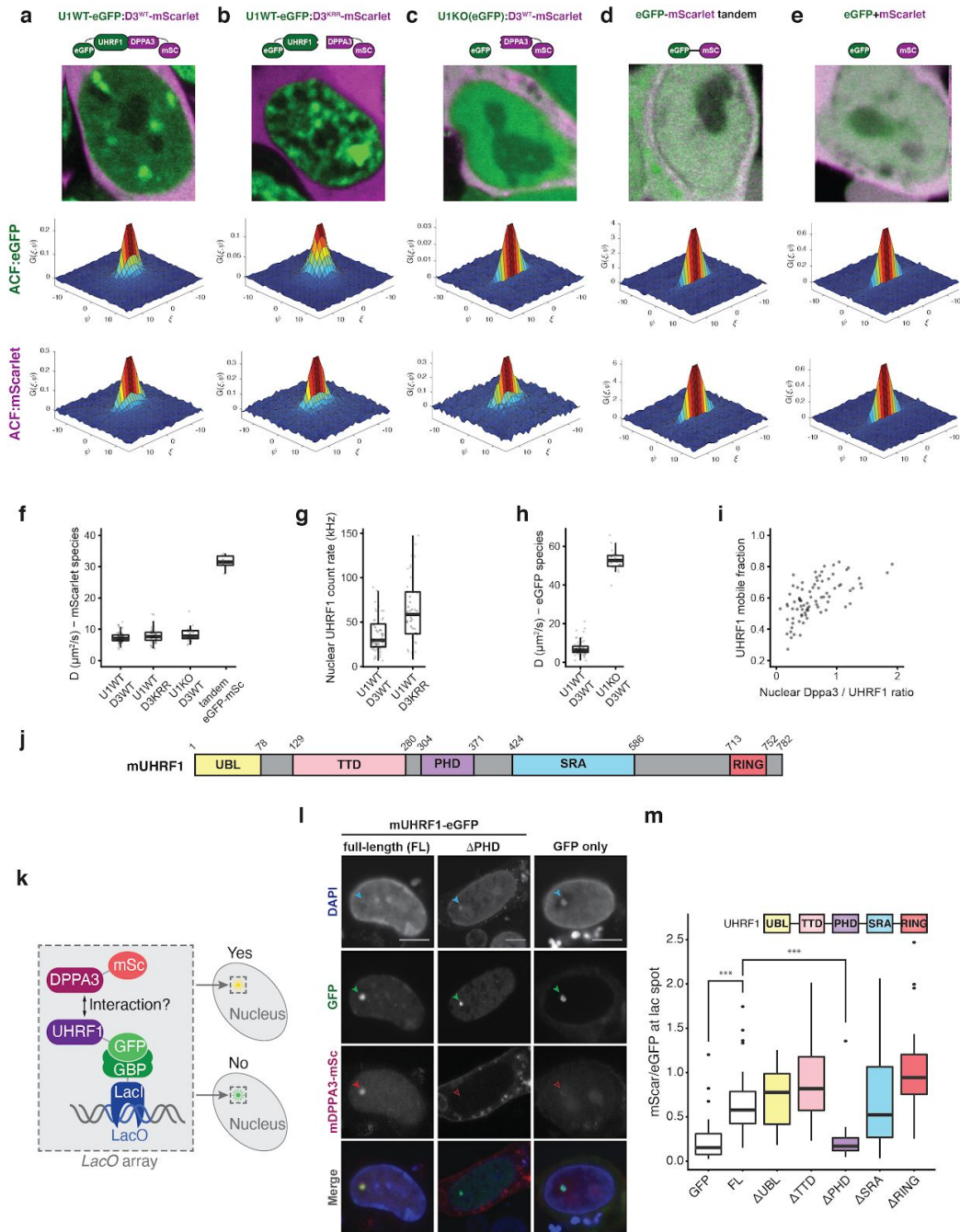
Supplementary Fig. 7: DPPA3 binds UHRF1 via its PHD domain to form a mobile complex in ESCs

a-e, Representative dual-color live-cell confocal images (top row) and associated 3D plots of the mean spatial autocorrelation (SACF) values of eGFP- (middle row) and mScarlet-labeled (bottom row) species in U1G/D3KO + pSBtet-D3 ESCs expressing the following forms of DPPA3-mScalet: (**a**) wild-type (U1WT:D3^{WT}) and (**b**) K85E/R85E/K87E mutant (U1WT:D3^{KRR}), in (**c**) Uhrf1KO ESCs expressing free eGFP and wild-type *Dppa3*-mScarlet (U1KO:D3^{WT}), and in control ESCs expressing (**d**) an eGFP-mScarlet tandem fusion (eGFP-mScarlet) and (**e**) free eGFP and free mScarlet (eGFP + mScarlet). In **a-e**, each image is the result of merging the sum projections of 250-frame image series acquired simultaneously from the eGFP (green) and mScarlet (magenta) channels. Autocorrelation plots are color-coded to indicate the mean correlation value. The spatial lag in the x-dimension (sensitive to fast timescales) is indicated by ξ, and the spatial lag in the y-dimension (sensitive to slower timescales) is indicated by ψ. **f,** The calculated diffusion coefficient of the mScarlet species in different cell types (those shown in (**a-d**)), derived from a two-component fit of mScarlet autocorrelation functions. Data are pooled from three independent experiments, except for the tandem eGFP-mScarlet, which was from two independent experiments. **g,** Photon count rates of UHRF1-eGFP in the nuclei of U1G/D3KOs expressing either DPPA3-WT (U1WT:D3^{WT}) or the DPPA3-KRR mutant (U1WT:D3^{KRR}). Data are pooled from three independent experiments. **h,** Calculated diffusion coefficients of UHRF1-GFP (in U1WT:D3^{WT}) or free eGFP (in U1KO:D3^{WT}), derived as those in C. Data are pooled from three independent experiments. **i,** A scatter plot showing the relationship between the mobile fraction of UHRF1-eGFP and the ratio of DPPA3-mScarlet:UHRF1-eGFP photon count rates in the nucleus. Data are pooled from three independent experiments. **j,** A schematic diagram illustrating the domains of mUHRF1: ubiquitin-like (UBL), tandem tudor (TTD), plant homeodomain (PHD), SET-and-RING-associated (SRA), and really interesting new gene (RING). **k,** Overview of the F3H assay used to find the domain of UHRF1 that binds to DPPA3. **l,m,** The PHD domain of UHRF1 is necessary for mediating the interaction with DPPA3. **l,** Representative confocal images of full-length UHRF1-GFP, ΔPHD UHRF1-GFP, and free GFP constructs (**j**) immobilized at the lacO array (indicated with green arrows). Efficient or failed recruitment of DPPA3-mScarlet to the lacO spot are indicated by solid or unfilled red arrows respectively. Similar results were obtained from n = 2 independent experiments. Scale bar: 5 μm. **m,** The efficiency of DPPA3-mScarlet recruitment to different UHRF1-eGFP deletion constructs immobilized at the lacO array is given as the fluorescence intensity ratio of mScarlet (DPPA3) to eGFP (UHRF1 constructs) at the nuclear LacO spot. Data represent the results of *n* = 25 cells analyzed from two independent experiments for each construct. P-values were calculated using a two-sided Welch's t-test: : *** P < 0.2e-16. In **f-h**, each data point indicates the measured and fit values from a single cell. In the box plots **f-h** and **m**, darker horizontal lines within boxes represent median values. The limits of the boxes indicate upper and lower quartiles, and whiskers extend to the most extreme value within 1.5 x the interquartile range from each hinge. Source data are provided as a Source Data file.

15
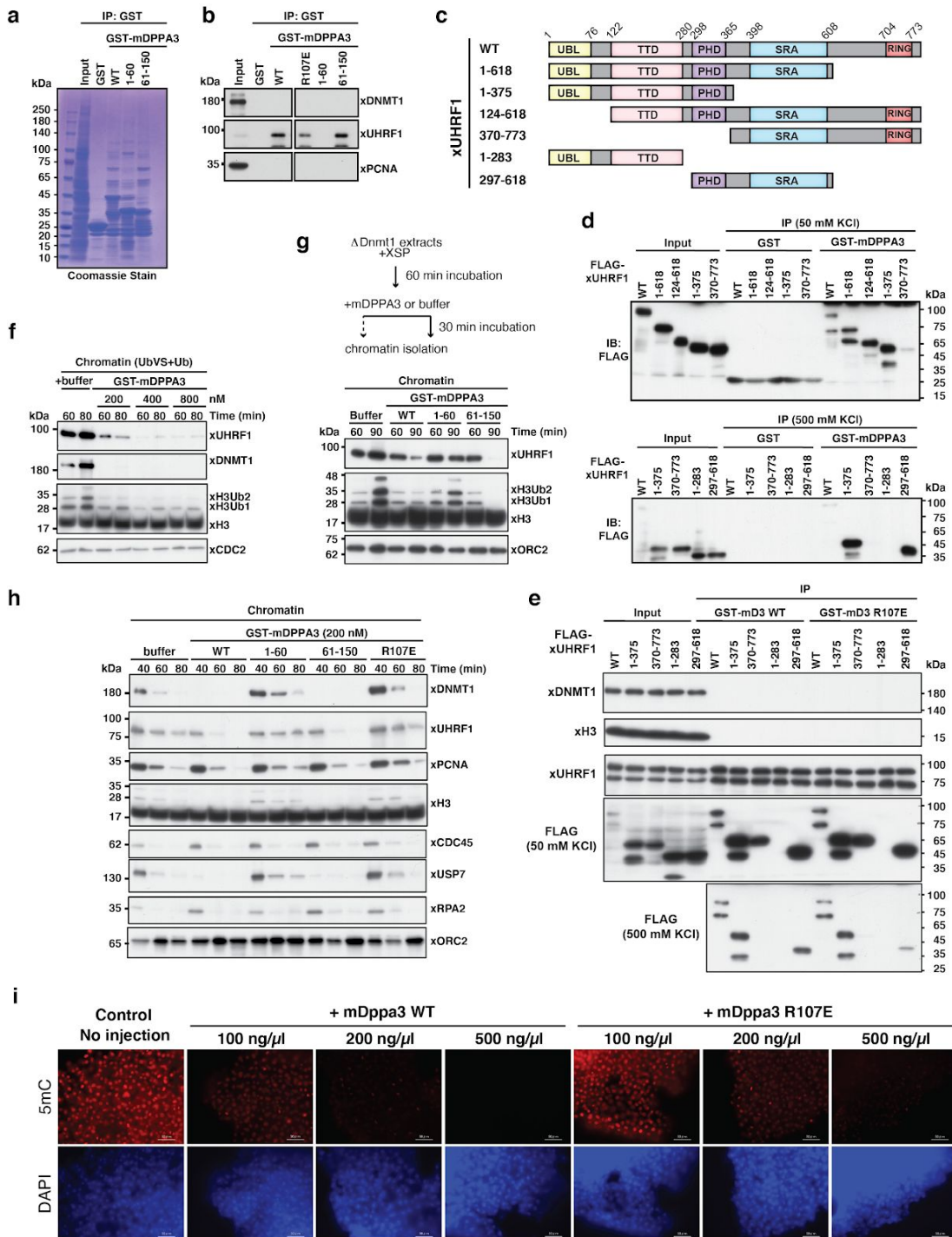
Supplementary Fig. 8: DPPA3 evolved in boreoeutherian mammals but its function is transferable to lower vertebrates

a, mDPPA3 C-terminus is sufficient for xUHRF1 binding. GST-tagged mDPPA3 wild-type (WT) and truncations (1-60 and 61-150) were immobilized on GSH beads and incubated with *Xenopus* egg extracts. The samples shown in Fig. 7c were subjected to SDS-PAGE and stained with Coomassie Blue. Representative of *n* = 3 independent experiments. **b,** High affinity interaction between mDPPA3 and xUHRF1 is weakened by R107E. GST-tagged mDPPA3 wild-type (WT), the point mutant R107E, and truncations (1-60 and 61-150) were immobilized on GSH beads and incubated with *Xenopus* egg extracts . Pull-downs were subjected to stringent 500 mM KCl washing. Bound proteins were analyzed using the indicated antibodies. Representative of *n* = 3 independent experiments. **c,** Schematic diagram illustrating the xUHRF1 deletion constructs used in the pull-downs in (**d,e**). **d,e,** xUHRF1 binds mDPPA3 via its PHD domain. *In vitro* translated xUHRF1 fragments were added to interphase *Xenopus* egg extracts. (**d**) GST and GST-mDPPA3 wild-type (WT) or (**e**) GST-tagged mDPPA3 wild-type (WT) and GST-tagged mDPPA3 R107E were immobilized on GSH beads and then used for GST-pulldowns on egg extracts containing the indicated recombinant xUHRF1 fragments. Bound proteins were analyzed using the denoted antibodies. Pull-downs were subjected to either 50 mM KCl or more stringent 500 mM KCl washing as indicated. Blots in d and e are representative of *n* = 3 independent experiments. **f,** mDPPA3 disrupts xUHRF1-dependent ubiquitylation of H3. As dual-monoubiquitylation of H3 (H3Ub2) is hard to detect given its quick turnover[7], we specifically enhanced ubiquitylation by simultaneous treatment of extracts with ubiquitin vinyl sulfone (UbVS), a pan-deubiquitylation enzyme inhibitor[8] and free ubiquitin (+Ub) as described previously[7]. Buffer or the displayed concentration of recombinant mDPPA3 were then added to the extracts. After the indicated times of incubation, chromatin fractions were isolated and subjected to immunoblotting using the antibodies indicated.Representative of *n* = 2 independent experiments. **g,** mDPPA3 displaces chromatin-bound xUHRF1. To stimulate xUHRF1 accumulation on hemi-methylated chromatin, *Xenopus* extracts were first immuno-depleted of xDNMT1 as described previously[9]. After addition of sperm chromatin, extracts were incubated for 60 min to allow the accumulation of xUHRF1 on chromatin during S-phase and then the indicated form of recombinant mDPPA3 (or buffer) was added. Chromatin fractions were isolated either immediately (60 min) or after an additional 30 min incubation (90 min) and subjected to immunoblotting using the antibodies indicated. Representative of *n* = 2 independent experiments. **h,** The region 61-150 of mDPPA3 is sufficient but requires R107 to inhibit xUHRF1 chromatin binding. Sperm chromatin was incubated with interphase *Xenopus* egg extracts supplemented with buffer (+buffer) and GST-mDPPA3 wild-type (WT), the point mutants R107E, truncations (indicated). Chromatin fractions were isolated after the indicated incubation time and subjected to immunoblotting using the antibodies indicated. Representative of *n* = 2 independent experiments. **i,** mDPPA3-mediated DNA demethylation in medaka. Representative 5mC immunostainings in late blastula stage (~8 h after fertilization) medaka embryos injected with wild-type *mDppa3* (WT) or *mDppa3* R107E (R107E) at three different concentrations (100 ng/µl, 200 ng/µl, or 500 ng/µl). Scale bars represent 50 µm. DNA counterstain: DAPI,4′,6-diamidino-2-phenylindole. Source data are provided as a Source Data file.

17
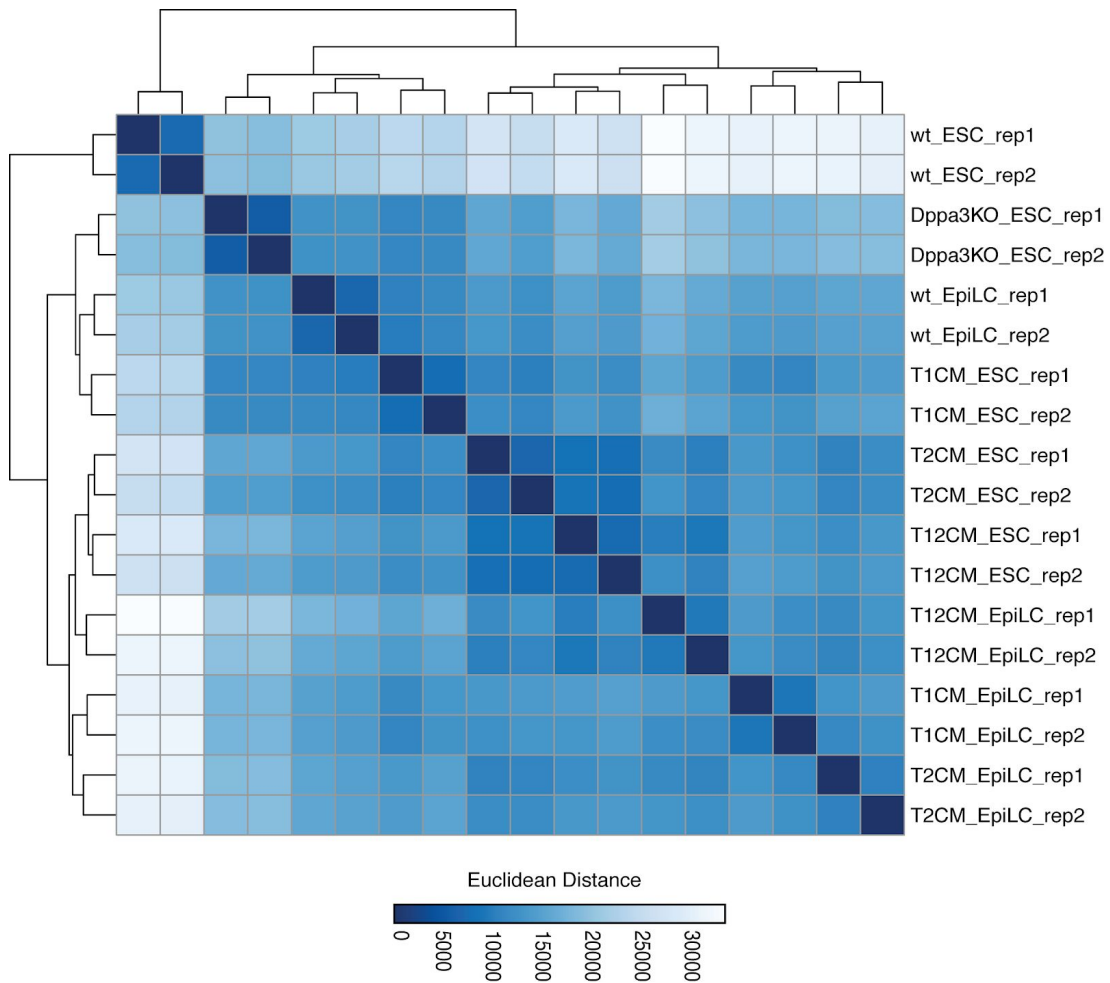
Supplementary Fig. 9: Methylome profiling of wild-type, T1CM, T2CM, T12CM, and Dppa3KO ESCs with RRBS

Heatmap depicting hierarchical clustering of the euclidean distance matrix generated from all RRBS samples used in this study.

18

References

1. von Meyenn, F. *et al.* Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. *Mol. Cell* **62**, 848–861 (2016).

2. Khoueiry, R. et al. Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nat. Genet.* **49**, 1061–1072 (2017).

3. Xiong, J. *et al.* Cooperative Action between SALL4A and TET Proteins in Stepwise Oxidation of 5-Methylcytosine. *Mol. Cell* **64**, 913–925 (2016).

4. Shen, L. *et al.* Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* **153**, 692–706 (2013).

5. Mulholland, C. B. *et al.* Distinct and stage-specific contributions of TET1 and TET2 to stepwise cytosine oxidation in the transition from naive to primed pluripotency. *Sci. Rep.* **10**, 12066 (2020).

6. Sharif, J. *et al.* Activation of Endogenous Retroviruses in Dnmt1−/− ESCs Involves Disruption of SETDB1-Mediated Repression by NP95 Binding to Hemimethylated DNA. *Cell Stem Cell* **19**, 81–94 (2016).

7. Yamaguchi, L. *et al.* Usp7-dependent histone H3 deubiquitylation regulates maintenance of DNA methylation. *Sci. Rep.* **7**, 55 (2017).

8. Borodovsky, A. *et al.* A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. *EMBO J.* **20**, 5187–5196 (2001).

9. Nishiyama, A. *et al.* Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* **502**, 249–253 (2013).

19