Characterisation of interactions between the pluripotency transcription factors Nanog, Oct4 and Sox2

Tapan Kumar Mistri[1, 2*], David Kelly[3], John Mak[1], Douglas Colby[1], Nicholas Mullin[1], Ian Chambers[1#]

[1]*Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, Edinburgh EH16 4UU, UK*

[2]*Department of Chemistry, Lovely Professional University, Jalandhar - Delhi G.T. Road, Phagwara, Punjab–144411, India*

[3]*Wellcome Trust Centre for Cell Biology, University of Edinburgh, Michael Swann Building, Kings Buildings, Max Born Crescent, Edinburgh-EH9 3BF, UK*

Running head: *Dynamic interactions of pluripotency transcription factors*

Keywords: Nanog, Sox2, Oct4, FCS, FCCS, CBIM.

Final character count (with space) = 33,772

Address correspondence to:

*****Tapan Kumar Mistri, Department of Trans-disciplinary Research, Lovely Professional University, Jalandhar - Delhi G.T. Road, Phagwara, Punjab-144411, India, Email: tapachemo@gmail.com

[#]Ian Chambers, MRC Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, Edinburgh EH16 4UU, Scotland, UK Email: ichambers@ed.ac.uk

ABSTRACT

The pluripotency transcription factors (TFs) Nanog, Sox2, and Oct4 are at the centre of the gene regulatory network that controls cell identity in embryonic stem (ES) cells. However, the mechanisms by which these factors control cell fate, and their interactions with one another are not fully understood. Here we combine biophysical and novel biochemical assays to assess how these factors interact with each other quantitatively. A new confocal microscopy method to detect binding of a target protein to a fluorescently labelled partner (coimmunoprecipitation bead imaging microscopy [CBIM]) is presented and used to demonstrate homotypic binding of Nanog and heterotypic binding between Nanog and Sox2 and between Nanog and Oct4. Using fluorescence correlation spectroscopy we show that in solution, Nanog but not Oct4 or Sox2 can form homodimers. Fluorescence Cross Correlation Spectroscopy shows that the affinity of Nanog for dimer formation is in the order Sox2 > Nanog > Oct4. Importantly, live cell analysis demonstrate the existence of Nanog homodimers in vivo. While Sox2 and Oct4 bind one another in a DNA-dependent manner, Nanog appears not to bridge Sox2 and Oct4, even though Nanog binds both Sox2 and Oct4 individually. Together these findings extend understanding of the molecular interactions occurring between these central mediators of the pluripotency gene regulatory network at the single molecule level.

INTRODUCTION

ES cells originate from the inner cell mass (ICM) of the preimplantation mammalian embryo (Evans and Kaufman 1981) and are defined by their capacity for indefinite self-renewal in culture and their ability to differentiate into all three embryonic germ layers (Smith 2001). A network of core transcription factors (TFs) including Sox2, Oct4, and Nanog controls ES cell self-renewal efficiency (Boyer, Lee et al. 2005, Chambers 2005). In particular, Nanog can enhance ES cell self-renewal to the point of cytokine independence, when overexpressed (Chambers, Colby et al. 2003). A gene knock-out study for Nanog resulted in preimplantation lethality due to a failure to properly specify pluripotent cell identity (Mitsui, Tokuzawa et al. 2003, Silva, Nichols et al. 2009). In ES cells, conditional deletion of Nanog increases the propensity of cells for differentiation, but does not eliminate pluripotency (Chambers, Silva et al. 2007). These observations led to the hypothesis that Nanog mainly functions as a rheostat with the level of Nanog positively correlated to ES cell self-renewal (Chambers, Colby et al. 2003, Chambers, Silva et al. 2007).

Previous studies have demonstrated that Nanog functions as a dimer (Mullin, Yates et al. 2008, Wang, Levasseur et al. 2008), with dimerisation occuring through a low complexity domain referred to as the tryptophan repeat (WR), due to the presence of a tryptophan at every fifth position (Mullin, Yates et al. 2008). The dimerisation of Nanog may also be affected by additional Nanog partner proteins. While the residues involved in Nanog homodimerisation and heterodimerisation with Sox2 have been determined (Mullin, Yates et al. 2008, Gagliardi, Mullin et al. 2013, Mullin, Gagliardi et al. 2017), how Nanog interacts with additional partners remains to be shown. For example, while Nanog has been reported to bind Oct4 (Wang, Levasseur et al. 2008), other studies did not report an interaction between Nanog and Oct4 (van den Berg, Snoek et al. 2010, Gagliardi, Mullin et al. 2013). Therefore, the differential binding specificity of Nanog as a homodimer or as a heterodimer with specific partners requires further exploration.

In the current study, we generated and functionally characterised fluorescent GFP and mCherry tagged fusions with Nanog to investigate the molecular interactions of Nanog with Sox2 and Oct4 using FCS and FCCS (Chen, Tapan et al. 2012, Mistri, Devasia et al. 2015). We used

single-molecule, sensitive, advanced techniques including fluorescence correlation spectroscopy (FCS) and cross-correlation spectroscopy (FCCS) (Macháň and Wohland 2014) to help advance understanding of how TFs interact with their binding partners both in vitro and under physiological conditions in live cells. In addition, we used a novel technique termed coimmunoprecipitated beam imaging spectroscopy (CBIM) to directly visualise DNA-Protein interactions. Our results reveal that Nanog has a unique single molecular binding dynamics among master regulators like Sox2 and Oct4.

MATERIALS AND METHODS

Cell culture and transfection-Chinese hamster ovary (CHO) cells (K1 line, ATCC # CRL-61) were cultured in Glasgow's modified Eagle's medium (GMEM, Life Technologies), 10% fetal bovine serum (GIBCO), 1x nonessential amino acids, 2 mM L-glutamine/pyruvate, 112 µM β-mercaptoethanol at 37 ºC, 7% CO_2 and 95 % humidity. Transfection of plasmids was performed using Lipofectamine 3000 (Invitrogen) according to the manufacturer's protocol. 10 µg of plasmid was transfected into CHO cells on 75cm^2 plates. The cells were collected after 24 hours for nuclear lysate preparation. Mouse E14/T cells were cultured on gelatin-coated dishes in GMEM/β-mercaptoethanol/10% FCS/LIF as described earlier (Smith 1991).

Motif discovery- To identify putative binding patterns ("motifs") for Nanog, Sox2, and Pou5f1 (Oct4) we made use of publicly available ChIP-Seq data (Chen, Xu et al. 2008, Marson, Levine et al. 2008, Whyte, Orlando et al. 2013). We had previously reanalysed these data using the GeneProf software (Halbritter, Vaidya et al. 2012, Gagliardi, Mullin et al. 2013, Halbritter, Kousa et al. 2014). We downloaded the coordinates of putative binding sites ("peaks") from GeneProf (accession gpXP_001309: http://www.geneprof.org/GeneProf/show?id=gpXP_001309) and then followed the same procedure as previously published to search for enriched binding motifs (Mistri, Devasia et al. 2015). Briefly, we used a local installation of XXMotif (version 1.6, parameters: --zoops --XXmasker --type ALL -g3 --merge-motif-threshold LOW --batch --revcomp) on the peak sequences. We employed DNaseI hypersensitive sites from the ENCODE project (http://genome.ucsc.edu/ENCODE) as a negative control set (parameter: --negSet) for eliminating non-specific sequence patterns.

Nuclear lysate preparation-Nuclear lysates were prepared from transfected CHO cells 24 hours post-transfection as described previously (Hutchins, Choo et al. 2013, Mistri, Devasia et al. 2015).

Co-immunoprecipitated-bead imaging microscopy (CBIM)- Coimmunoprecipitation **(**Co-IP) transfections were performed with plasmids encoding (flag)$_3$-tagged interaction partners or partners tagged with a fluorescent proteins using Lipofectamine 3000, (Life technologies) according to the manufacturer's protocol. The nuclear lysate was prepared 24 hours post-

transfection of 6 million cells in a final volume of 250 µl. The Co-IP was performed as described (Gagliardi, Mullin et al. 2013). After Co-IP, 30 µl of binding buffer was added to protein G sepharose beads. 10 µl of the bead solution was used for fluorescence detection by confocal microscopy. GFP and mCherry were imaged at a laser intensity of 31% with 488 nm and 561 nm lasers respectively. High laser power was used to eliminate background noise. For protein-DNA interactions, CoIP beads were incubated with Cy5-tagged DNA oligos, at a final concentration of 50nM in 100 µl of binding buffer for 1 hour at 4°C with rotation. The beads were washed with 1ml PBS buffer 5 times before imaging. Cy5 tagged beads were imaged with a 633 nm laser at 31% of laser intensity.

Fluorescence protein-based electrophoretic mobility shift assay (FP-EMSA)- FP-EMSAs using CHO nuclear lysates were performed as described previously (Chen, Tapan et al. 2012, Hutchins, Choo et al. 2013, Mistri, Devasia et al. 2015).

Fluorescence correlation and cross-correlation spectroscopy (FCS/FCCS)- FCS was performed using a laser scanning confocal microscope (Leica SP5, Leica Microsystems UK) with a water immersion objective (63×, NA1.2, Leica) coupled to an in-built FCS module. A dichroic mirror (560DRLP) and bandpass emission filters (500-550nm for Alexa 488/GFP and 607-683nm for Atto 565/mCherry, (Chroma Technology, VT) were used to separate the excitation light coming from the fluorescence emission. A standard calibration approach was used to determine the absolute concentration of the fusion proteins in the crude nuclear lysate (Mistri, Devasia et al. 2015, Mistri, Arindrarto et al. 2018). Appropriately chosen dyes with known diffusion coefficients were employed to determine the confocal volume accurately. To measure the concentration of GFP tagged proteins, calibration was performed with 2 nM of Alexa488 (diffusion constant, $D = 3.80\times 10^{-6}\ cm^2s^{-1}$) (Ramadurai, Holt et al. 2009) triggered by a 488 nm laser line at 30 µW (15% laser intensity). For mCherry-Sox2, the calibration was performed using a 561nm laser line at 30 µW (10% laser power) using 2nM of Atto565 (diffusion constant, $D = 2.59 \times 10^{-6}\ cm^2s^{-1}$) (Pan, Aw et al. 2006). For live-cell studies, GFP and mCherry localised to the nucleus were excited at 10 µW by 488 nm and 561 nm lasers respectively. Correlations in fluorescence fluctuations were computed online by a hardware correlator and were fitted using a theoretical model for free diffusion in solution and a triplet using Vista 3.6 (ISS, Inc Champaign IL) or using a in-house

fitting module based on Igor Pro 6.0 as described previously (Wohland, Friedrich et al. 1999). Absolute concentrations were calculated in nM.

RESULTS

To assess the molecular interactions between pluripotency transcription factors (TFs), fluorescent protein fusions were prepared. Sox2 and Oct4 fluorescent protein fusions have previously been shown to be functional (Mistri, Devasia et al. 2015). For Nanog, fusions to mCherry or eGFP were prepared and their functionality tested following transfection of E14/T ESCs. This shows that both mCherry-Nanog and Nanog-GFP retain the defining capacity of Nanog to drive LIF-independent ESC self-renewal (Fig. S1).

Nanog forms homo-multimers and binds to Sox2 and Oct4 in CHO nuclear lysate

To assess the ability of fluorescent partners to interact with Nanog, we developed a bead based coimmunoprecipitation assay termed coimmunoprecipitation bead imaging microscopy (CBIM). This uses confocal microscopy to directly image binding of a fluorescent molecule to a non-fluorescent partner protein immobilised on protein G-sepharose beads (Fig. S2). Nuclear lysates from cells transfected with fluorescent protein expression constructs were used as a source of fluorescent proteins. CHO cells were used as the transfection target cell as these cells do not express components of the pluripotency gene regulatory network and therefore lack proteins that interact with the fluorescent protein fusions in a cell-specific manner. When beads coated with anti-FLAG antibody were incubated with a nuclear lysate from untransfected CHO cells, no fluorescence was detected on the beads by confocal microscopy (Fig. 1A). Likewise, when beads bound by anti-FLAG antibody were incubated with nuclear lysates from CHO cells transfected with separate constructs expressing $(FLAG)_3$-Nanog and GFP, no fluorescence was apparent (Fig 1B). However, incubation with lysate expressing GFP-Nanog showed fluorescence localised on the bead surface, indicative of dimerization between $(FLAG)_3$-Nanog bound to the anti-FLAG conjugated beads and GFP-Nanog from the CHO nuclear lysate (Fig 1B). In contrast, when a GFP-Nanog fusion in which the tryptophan residues in the tryptophan repeat (WR; Nanog-W10A) was expressed in CHO cells alongside $(FLAG)_3$-Nanog, incubation of nuclear lysate with anti-FLAG-conjugated beads yielded no fluorescence (Fig 1B). This is consistent with previous reports of a critical role for WR tryptophan residues in homodimerisation of Nanog (Mullin, Yates et al. 2008, Wang, Levasseur et al. 2008, Mullin, Gagliardi et al. 2017). In addition, incubation of beads with $(FLAG)_3$-Nanog and GFP-Sox2, resulted in localisation of fluorescence to the bead surface (Fig 1B). This is consistent with the known heterodimerisation of Nanog with Sox2 (Gagliardi, Mullin

et al. 2013). We further assessed the heterodimerisation of Nanog and Sox2 by first immobilising $(FLAG)_3$-Sox2 onto anti-FLAG/proteinG-beads. Incubation with GFP-Nanog, but not GFP-NanogW10A resulted in fluorescent beads (Fig 1C), consistent with a central role of WR tryptophans in mediating heterodimerisation of Nanog with Sox2 (Gagliardi, Mullin et al. 2013). We next assessed the utility of mCherry fusion proteins in CBIM (Fig 1D). This established that interactions between partner proteins can be detected independent of the fluorescent fusion partner. These assays also showed that binding of Oct4 to Nanog could be detected by CBIM (Fig 1D). Previously, Oct4 has been reported to co-immunoprecipitate with Nanog (Wang, Levasseur et al. 2008). However, this result not uncontentious, since Oct4 is not detectable by Nanog-affinity purification and mass spectrometry (Wang, Levasseur et al. 2008) nor is Nanog present in the Oct4 interactome (van den Berg, Snoek et al. 2010). The current results confirm that Nanog and Oct4 can interact and suggest that CBIM may provide a sensitive means of detecting protein-protein interactions.

We next assessed the ability of CBIM to detect interactions between immobilised proteins and labelled DNA. When anti-FLAG conjugated beads were incubated with $(FLAG)_3$-Sox2, Cy5-labelled dsDNA containing the Oct/Sox motif from the *Nanog* gene produced a fluorescent halo on the beads (Fig 2A). This DNA interaction is specific, since Cy5-labelled dsDNA from the *Tcf3* gene that lacks an Oct/Sox motif was not retained on beads bound by Sox2 (Fig 2A). Similar results were obtained when FLAG-Oct4 bound beads were incubated with the same Cy5-labelled oligonucleotides (Fig 2B). In addition, incubation of dsDNA from the *Tcf3* gene containing a Nanog binding site but lacking Oct4 or Sox2 binding sites (Jauch, Ng et al. 2008) with nuclear lysate from CHO cells co-transfected with $(Flag)_3$-Nanog shows that Flag-Nanog also binds DNA on beads (Fig. 2C).

Nanog, but not Sox2 or Oct4 can form homo-multimers

While the CBIM assay provides qualitative data on TF binding activities, it does not assess the binding strength of TF interactions. Fluorescence correlation spectroscopy (FCS) was therefore used to assess protein-protein complexes at single molecule resolution. Nuclear lysates from CHO cells transfected with fluorescent protein fusions were examined by FCS to measure the photon count rate per single molecule (CPM). Assessing the brightness of the molecules passing through

the detection beam enables extraction of information about their multimeric state. These analyses show that the molecular brightness of GFP, GFP-Oct4 and GFP-Sox2 are similar, suggesting that the multimerisation state of GFP is unaffected by fusion to Sox2 and Oct4 (Fig. 3A). In contrast, GFP-Nanog is brighter than GFP, with the GFP-Nanog/GFP brightness ratio close to 1.5 (Fig. 3A, Fig. S3A). These results suggest that Nanog is in equilibrium between dimeric and monomeric states. In contrast, GFP-Nanog-W10A has a brightness similar to GFP, consistent with Nanog-W10A existing as a monomer (Mullin, Yates et al. 2008) (Fig. 3A). Analysis of mCherry fusions to Sox2 and Nanog also suggest that Nanog but not Sox2 can exist as a dimer (Fig. 3B, Fig. S3B).

Comparative assessment of the binding strengths of Nanog homo- and hetero-dimers

The previous results show that Nanog is in equilibrium between monomers and homo- and hetero-dimers in solution. To further assess the properties of Nanog dimers, we performed fluorescence cross-correlation spectroscopy (FCCS) experiments using mCherry and GFP fusions to Nanog, Sox2 and Oct4 (Fig. 4). By analysing the cross correlation between fluorescent protein signals originating from distinct partners in a dimer, we were able to assess the complex %, a parameter indicating the proportion of complex present in equilibrium and hence the affinity of binding (Table 1). Analysing CHO nuclear lysates expressing GFP and mCherry (Fig. 4A) or a tandem GFP-mCherry (Fig. 4A) provided the baseline and maximum for fluorescent proteins that are either separate, or always in complex. The calculated complex % for the negative control was 1.3 ± 0.8 and for the positive control 46.2 ± 6.7 (Fig. 4A, Table 1). FCCS of lysate expressing mCherry-Nanog and GFP-Nanog gave a complex % of 19.8 ± 2.9 (Fig. 4B, Table 1). Similar analysis of lysates expressing mCherry-Nanog with either GFP-Sox2 or GFP-Oct4 gave complex percentages of 26.1 ± 7.7 and 12 ± 6.7, respectively (Fig. 4B, Table 1). In contrast, FCCS analysis of GFP-Nanog-W10A in combination with mCherry-Nanog, mCherry-Sox2 or mCherry-Oct4 gave complex percent values indistinguishable from the negative control (Fig. 4C, Table 1). These results suggest an order of binding strength of Nanog for partners of Sox2>Nanog> Oct4.

Nanog diffusion dynamics decrease in the presence of either Sox2 or Oct4

To obtain insight into the dynamics of the interaction of Nanog with Sox2 or Oct4 we used fluorescence correlation spectroscopy (FCS) to measure the diffusion coefficient of GFP-Nanog in the presence or absence of either $(Flag)_3$-Oct4 or $(Flag)_3$-Sox2 (Fig. 5). This showed that the

diffusion coefficient of GFP-Nanog decreases to a similar extent in the presence of either Oct4 or Sox2. As Sox2 and Oct4 have a similar Mr, this suggests that the Nanog complexes they form have a similar stoichiometry. In addition, the diffusion coefficient for GFP-Nanog-W10A is higher than GFP-Nanog consistent with monomer formation by GFP-Nanog-W10A. Finally, as the diffusion coefficient of GFP-Nanog-W10A is unaffected by Sox2 or Oct4, this suggests that Oct4 interacts with Nanog through the WR tryptophan residues.

Differential effects of DNA on TF partner interactions

We next extended the FCCS analysis to assess the Oct4-Sox2 interaction using nuclear lysates from CHO cells co-transfected with GFP-Oct4 and mCherry-Sox2 (Fig. 6A). This showed that while Sox2 does not bind to Oct4 in solution (Complex % =3.9 ± 3.1) (Table 1) (Chen, Tapan et al. 2012), addition of DNA containing a Sox/Oct motif increased the complex formation between GFP-Oct4 and mCherry-Sox2 (7.5 ± 4.5) (Fig. 6A, Table 1) (Chen, Tapan et al. 2012). However, addition of Nanog did not affect the extent of the interaction between Oct4 and Sox2, either in the presence or absence of DNA containing an Oct/Sox motif (Fig. 6A, Table 1). The observation suggests that Nanog neither assists nor impairs DNA binding by the Oct/Sox2 complex.

We next analysed the interaction between GFP-Oct4 and mCherry-Nanog in the presence or absence of $(Flag)_3$-Sox2 (Fig. 6B). FCCS showed that the complex % formed by GFP-Oct4 and mCherry-Nanog was reduced from 12.5 (± 6.7) to 5.0 (± 3.5) when Sox2 was added (Fig. 6B and Table 1). There are at least two possible interpretations of this data. Firstly, Sox2 might bind to Oct4 and thereby reduce the Oct4/Nanog binding. However, as Sox2 binding to Oct4 is DNA-dependent (Fig. 6A) (Chen, Tapan et al. 2012), this seems unlikely. An alternative is that Sox2 may compete with Oct4 for Nanog binding. This possibility is supported by higher affinity of interaction of Sox2 for Nanog, compared to Oct4 and by the fact that binding of Nanog to both Sox2 and Oct4 is reduced by mutation of the WR tryptophan residues (Fig. 5B).

We next performed an FCCS assay with CHO nuclear lysates containing co-transfected GFP-Sox2 and mCherry-Nanog in the presence and absence of DNA (Fig. 6C). This shows that

DNA addition had a minimal effect on the % complex formed between mCherry-Nanog and GFP-Sox2 (Fig. 6C, Table 1).

Analysis of Nanog dimer formation *in vivo*

The preceding analyses establish that NANOG can interact with partners on beads and in solution. To investigate whether similar interactions may occur in vivo, we assessed live cells, initially using FCS to analyse fluorescent protein brightness in CHO cells transfected with either GFP or GFP-Nanog (Fig. 7). Similar to the situation in vitro, the brightness of GFP-Nanog was higher than that of GFP, with a GFP-Nanog/GFP ratio of approximately 1.5 (Fig. 7B). These results are consistent with the idea that Nanog exists in cells in equilibrium between dimeric and monomeric states.

The preceding FCS analysis is consistent with the existence of NANOG dimers *in vivo*. However, they do not provide information about the proportion of molecules present in monomeric or multimeric forms in vivo. To assess this, we used fluorescence cross-correlation spectroscopy to analyse live CHO cells co-transfected with Nanog-GFP and mCherry-Nanog (Fig. 8, Fig. S4). In cells expressing GFP and mCherry as a negative control, the baseline complex %, a measure of binding strength, was 3.7 ± 2.6 whereas that for the GFP-mCherry fusion, the positive control, complex % was 47.4 ± 9.2 (Fig. 8A). These values then served as the lower and upper limits of the complex % for experiments in which mCherry-Nanog was expressed in CHO cells alongside GFP-Nanog or GFP-Nanog-W10A. FCCS data indicate that the cross correlation between mCherry-Nanog and either GFP-Nanog-W10A or GFP was indistinguishable (Fig. 8A, B, C), indicating no interaction between mCherry-Nanog and GFP-Nanog-W10A. In contrast the cross correlation between GFP-Nanog and mCherry-Nanog, indicated a complex % of 16.8 ± 7.2 (Fig. 8B, C), a value close to that obtained from in vitro analyses (Table 1). These results indicate that Nanog exists in cells in an equilibrium between monomeric and dimeric forms.

DISCUSSION

In this study, we assess the molecular dynamics of Nanog in solution at single molecule sensitivity. It has previously been shown that Nanog homodimerises and heterodimerises with Sox2 to form DNA independent protein-protein complexes (Mullin, Yates et al. 2008, Wang, Levasseur et al. 2008, Gagliardi, Mullin et al. 2013). However, the quantitative details of Nanog protein-protein complex formation have only partly been addressed. Our analysis of the diffusion dynamics of Nanog fluorescent protein fusions confirms that two Nanog molecules participate in Nanog homo-multimerisation and that dimers remain in equilibrium with monomers, confirming previous analytical ultracentrifugation studies (Mullin, Yates et al. 2008). Importantly, our FCCS analysis establishes the existence of Nanog homodimers in cells. Additional FCCS analysis shows that the capacity of Nanog for homo-multimerisation is not shared with Oct4 or Sox2.

We describe a new confocal microscopy method to detect binding of a target protein to a fluorescently labelled partner protein or DNA target (coimmunoprecipitation bead imaging microscopy [CBIM]). Using CBIM allowed us to establish that Nanog binds Oct4, a previously contentious issue (Wang, Levasseur et al. 2008, van den Berg, Snoek et al. 2010, Gagliardi, Mullin et al. 2013). The high local concentration of binding partners that can be achieved on the bead surface may make CBIM a highly sensitive detection method for detecting low affinity interactions. This idea is supported by our FCCS analysis of the capacity of Nanog to form homodimers, or heterodimers with Sox2 or Oct4. This showed that the affinity of Nanog for binding partners is in the order Sox2 > Nanog > Oct4.

By adding a non-fluorescent partner into the FCCS analysis of a pair of partners, it is possible to determine the effect of the addition on dimerization. This can be achieved using DNA and enabled us to show the stimulatory effect of DNA on complex formation between Oct4 and Sox2 (Chen, Tapan et al. 2012). This can also be done using protein. For example, addition of Nanog to the GFP-Oct4 and mCherry-Sox2 in the presence of DNA reduced the amount of the ternary Oct4/Sox2/DNA complex, likely through competition of Nanog for Sox2.

ChIP-Seq studies show that Nanog often shares similar genomic binding sites with Sox2 (Papatsenko, Darr et al. 2015). In addition, Sox2 and Oct4 synergistically bind the Sox-Oct *cis* co-

motif (Ambrosetti, Basilico et al. 1997, Mistri, Devasia et al. 2015). Genomic data therefore suggests a correlation between Nanog binding interactions with Sox2 as well as with Oct4 within the Sox2-Oct4-DNA ternary complex. The protein interactions described here may in part explain why motif enrichment from Nanog Chip-Seq is closer to the Sox2 binding motif than to the octamer motif. However, such indirect interaction with DNA through partner protein bridges, cannot fully account for the biological function of Nanog, since a mutation in the Nanog homeodomain abaltes Nanog function (Navarro, Festuccia et al. 2012).

In conclusion, we demonstrated a unique binding pattern of Nanog at the level of the single-molecule with its partner proteins Sox2 and Oct4. Future quantitative analyses will help further illuminate how TFs interact with other partner proteins and with DNA to deliver their function.

ACKNOWLEDGEMENTS

Research in the IC lab was supported by the Medical Research Council of the UK.

AUTHOR CONTRIBUTIONS

T.K.M. did experiments and wrote the manuscript, D. K. helped T.K.M in FCCS and FLIM-FRET experiments, J. M. performed diffusion study of Nanog, D. C. performed fusion Nanog functional tests, N.M. provided reagents, advice and assisted in manuscript preparation and I.C. supervised the project and wrote the manuscript.

ADDITIONAL INFORMATION

Please see the **Supplementary Section** for further information.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES:

Ambrosetti, D. C., C. Basilico and L. Dailey (1997). "Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites." Molecular and cellular biology **17**(11): 6321.

Boyer, L. A., T. I. Lee, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray and R. G. Jenner (2005). "Core transcriptional regulatory circuitry in human embryonic stem cells." Cell **122**(6): 947-956.

Chambers, I. (2005). "Mechanisms and factors in embryonic stem cell self-renewal." Rendiconti Lincei **16**(2): 83-97.

Chambers, I., D. Colby, M. Robertson, J. Nichols, S. Lee, S. Tweedie and A. Smith (2003). "Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells." Cell **113**(5): 643-655.

Chambers, I., J. Silva, D. Colby, J. Nichols, B. Nijmeijer, M. Robertson, J. Vrana, K. Jones, L. Grotewold and A. Smith (2007). "Nanog safeguards pluripotency and mediates germline development." Nature **450**(7173): 1230-1234.

Chen, S. L., K. M. Tapan, H. F. Yong, S. Thankiah, T. G. Hui, R. David, H. L. Leng, C. Chai, R. Paul and W. Thorsten (2012). "DNA-dependent Oct4-Sox2 interaction and diffusion properties characteristic of the pluripotent cell state revealed by fluorescence spectroscopy." Biochemical Journal **448**(1): 21-33.

Chen, X., H. Xu, P. Yuan, F. Fang, M. Huss, V. B. Vega, E. Wong, Y. L. Orlov, W. Zhang and J. Jiang (2008). "Integration of external signaling pathways with the core transcriptional network in embryonic stem cells." Cell **133**(6): 1106-1117.

Evans, M. J. and M. H. Kaufman (1981). "Establishment in culture of pluripotential cells from mouse embryos." Nature **292**(5819): 154-156.

Gagliardi, A., N. P. Mullin, Z. Ying Tan, D. Colby, A. I. Kousa, F. Halbritter, J. T. Weiss, A. Felker, K. Bezstarosti and R. Favaro (2013). "A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell selfâ€renewal." The EMBO journal **32**(16): 2231-2247.

Halbritter, F., A. I. Kousa and S. R. Tomlinson (2014). "GeneProf data: a resource of curated, integrated and reusable high-throughput genomics experiments." Nucleic acids research **42**(D1): D851-D858.

Halbritter, F., H. J. Vaidya and S. R. Tomlinson (2012). "GeneProf: analysis of high-throughput sequencing experiments." Nature methods **9**(1): 7.

Hutchins, A. P., S. H. Choo, T. K. Mistri, M. Rahmani, C. T. Woon, C. Keow Leng Ng, R. Jauch and P. Robson (2013). "Co-Motif Discovery Identifies an Esrrb-Sox2-DNA Ternary Complex as a Mediator of

Transcriptional Differences Between Mouse Embryonic and Epiblast Stem Cells." Stem Cells **31**(2): 269-281.

Jauch, R., C. K. L. Ng, K. S. Saikatendu, R. C. Stevens and P. R. Kolatkar (2008). "Crystal structure and DNA binding of the homeodomain of the stem cell transcription factor Nanog." Journal of molecular biology **376**(3): 758-770.

Macháň, R. and T. Wohland (2014). "Recent applications of fluorescence correlation spectroscopy in live systems." FEBS letters **588**(19): 3571-3584.

Marson, A., S. S. Levine, M. F. Cole, G. M. Frampton, T. Brambrink, S. Johnstone, M. G. Guenther, W. K. Johnston, M. Wernig and J. Newman (2008). "Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells." Cell **134**(3): 521-533.

Mistri, T. K., W. Arindrarto, W. P. Ng, C. Wang, L. H. Lim, L. Sun, I. Chambers, T. Wohland and P. Robson (2018). "Dynamic changes in Sox2 spatio-temporal expression promote the second cell fate decision through Fgf4/Fgfr2 signaling in preimplantation mouse embryos." Biochemical Journal **475**(6): 1075-1089.

Mistri, T. K., A. G. Devasia, L. T. Chu, W. P. Ng, F. Halbritter, D. Colby, B. Martynoga, S. R. Tomlinson, I. Chambers and P. Robson (2015). "Selective influence of Sox2 on POU transcription factor binding in embryonic and neural stem cells." EMBO reports **16**(9): 1177-1191.

Mitsui, K., Y. Tokuzawa, H. Itoh, K. Segawa, M. Murakami, K. Takahashi, M. Maruyama, M. Maeda and S. Yamanaka (2003). "The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells." Cell **113**(5): 631-642.

Mullin, N. P., A. Gagliardi, D. Colby, E. Hall-Ponsele, A. J. Rowe and I. Chambers (2017). "Distinct contributions of tryptophan residues within the dimerization domain to Nanog function." Journal of molecular biology **429**(10): 1544-1553.

Mullin, N. P., A. Yates, A. J. Rowe, B. Nijmeijer, D. Colby, P. N. Barlow, M. D. Walkinshaw and I. Chambers (2008). "The pluripotency rheostat Nanog functions as a dimer." Biochemical Journal **411**(2): 227-231.

Navarro, P., N. Festuccia, D. Colby, A. Gagliardi, N. P. Mullin, W. Zhang, V. Karwacki-Neisius, R. Osorno, D. Kelly and M. Robertson (2012). "OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells." The EMBO journal **31**(24): 4547-4562.

Pan, X., C. Aw, Y. Du, H. Yu and T. Wohland (2006). "Characterization of poly (acrylic acid) diffusion dynamics on the grafted surface of poly (ethylene terephthalate) films by fluorescence correlation spectroscopy." Biophysical Reviews and Letters **1**(04): 433-441.

Papatsenko, D., H. Darr, I. V. Kulakovskiy, A. Waghray, V. J. Makeev, B. D. MacArthur and I. R. Lemischka (2015). "Single-cell analyses of ESCs reveal alternative pluripotent cell states and molecular mechanisms that control self-renewal." Stem cell reports **5**(2): 207-220.

Ramadurai, S., A. Holt, V. Krasnikov, G. van den Bogaart, J. A. Killian and B. Poolman (2009). "Lateral diffusion of membrane proteins." Journal of the American Chemical Society **131**(35): 12650-12656.

Silva, J., J. Nichols, T. W. Theunissen, G. Guo, A. L. van Oosten, O. Barrandon, J. Wray, S. Yamanaka, I. Chambers and A. Smith (2009). "Nanog is the gateway to the pluripotent ground state." Cell **138**(4): 722-737.

Smith, A. G. (1991). "Culture and differentiation of embryonic stem cells." Journal of tissue culture methods **13**(2): 89-94.

Smith, A. G. (2001). "Embryo-derived stem cells: of mice and men." Annual review of cell and developmental biology **17**(1): 435-462.

van den Berg, D. L., T. Snoek, N. P. Mullin, A. Yates, K. Bezstarosti, J. Demmers, I. Chambers and R. A. Poot (2010). "An Oct4-centered protein interaction network in embryonic stem cells." Cell stem cell **6**(4): 369-381.

Wang, J., D. N. Levasseur and S. H. Orkin (2008). "Requirement of Nanog dimerization for stem cell self-renewal and pluripotency." Proceedings of the National Academy of Sciences **105**(17): 6326-6331.

Whyte, W. A., D. A. Orlando, D. Hnisz, B. J. Abraham, C. Y. Lin, M. H. Kagey, P. B. Rahl, T. I. Lee and R. A. Young (2013). "Master transcription factors and mediator establish super-enhancers at key cell identity genes." Cell **153**(2): 307-319.

Wohland, T., K. Friedrich, R. Hovius and H. Vogel (1999). "Study of ligand-receptor interactions by fluorescence correlation spectroscopy with different fluorophores: evidence that the homopentameric 5-hydroxytryptamine type 3As receptor binds only one ligand." Biochemistry **38**(27): 8671-8681.

FIGURE LEGENDS

Figure 1. CBIM demonstrates Nanog binding to Sox2 and Oct4. A. Confocal images of protein-G beads incubated with antibody alone or antibody and CHO lysate. **B**. Confocal images of protein-G beads after CoIP from nuclear lysate of CHO cells co-transfected with $(Flag)_3$-Nanog and GFP, $(Flag)_3$-Nanog and GFP-Nanog, $(Flag)_3$-Nanog and GFP-Nanog-W10A or $(Flag)_3$-Nanog and GFP-Sox2. **C**. Confocal images of beads after CoIP from nuclear lysate of CHO cells co-transfected with $(Flag)_3$-Sox2 and GFP-Nanog or $(Flag)_3$-Sox2 and GFP-Nanog-W10A. **D**. Confocal images of beads after CoIP from nuclear lysate of CHO cells co-transfected with $(Flag)_3$-Nanog and mCherry, $(Flag)_3$-Nanog and mCherry-Nanog, $(Flag)_3$-Nanog and mCherry-Oct4 or $(Flag)_3$-Nanog and mCherry-Sox2.

Figure 2. CBIM demonstrates DNA binding by TFs. A. Confocal images of protein-G beads incubated with CHO nuclear lysate from untransfected cells and a Cy5 labelled oligonucleotide encompassing the *Nanog* Sox/-Oct (NSO) element (upper) or with nuclear lysate from $(Flag)_3$-Sox2 transfections together with either the *Nanog* Sox/-Oct (NSO) element described above (middle) or a Cy5 labelled probe encompassing the *Tcf3* specific Nanog (TN) element (lower). **B**. Confocal images of protein-G beads incubated with CHO nuclear lysate from untransfected cells or from cells transfected with $(Flag)_3$-Oct4 together with the NSO or TN elements described above. **C**. Confocal images of protein-G beads incubated with Cy5 labelled TN DNA after immunoprecipitation from CHO nuclear lysates containing $(Flag)_3$-Nanog and GFP. All the DNA sequences used are shown in supplementary table 1.

Figure 3. FCS analysis of transcription factors in nuclear lysates. A. Count rate per species (CPS) of GFP, GFP-Oct4, GFP-Sox2, GFP-Nanog W10A and GFP-Nanog were measured by FCS **B**. CPS values for mCherry, mCherry-Sox2 and mCherry-Nanog. For all experiments, n = 12, mean ± SD.

Figure 4. FCCS analysis demonstrates interaction between Nanog, Sox 2 and Oct4. A. Autocorrelation function (ACF) and cross-correlation function (CCF) curves from FCCS assay of nuclear lysates containing mCherry-Nanog co-transfected with either GFP-Nanog, GFP-Sox2 or GFP-Oct4. **B**. Autocorrelation function and cross-correlation function curves from FCCS assay of

nuclear lysate containing GFP-Nanog-W10A co-transfected with mCherry-Nanog, mCherry-Sox2 or mCherry-Oct4. **C**. Autocorrelation function and cross-correlation function curves from FCCS assay of nuclear lysates containing co-transfected GFP and mCherry (left) or containing a tandem GFP-mCherry protein (right).

Figure 5. The diffusion dynamics of Nanog change in the presence of co-factors. A. Diffusion coefficients of GFP-Nanog and GFP-Nanog in the presence of Sox2 or Oct4 were measured by FCS. Statistical significance measured using an unpaired t-test. **B**. Diffusion coefficients of GFP-Nanog-W10A and GFP-Nanog. For all experiments n = 12, mean ± SD.

Figure 6. Nanog binds directly to Sox2 and Oct4. A. Autocorrelation function (ACF) and cross-correlation function (CCF) curves from FCCS with nuclear lysate from a co-transfection of GFP-Oct4 and mCherry-Sox2 in the presence of either $(Flag)_3$Nanog or DNA or in the presence of both $(Flag)_3$Nanog and DNA. **B**. Autocorrelation function and cross-correlation function curves from FCCS with nuclear lysate from a co-transfection of GFP-Oct4 and mCherry-Nanog in the presence of either $(Flag)_3$-Sox2 or DNA or in the presence of both $(Flag)_3$-Sox2 and DNA. **C**. Autocorrelation function and cross-correlation function curves from FCCS with nuclear lysate from a co-transfection of GFP-Sox2 and mCherry-Nanog in the presence of DNA. For all experiments, n = 12, mean ± SD. The percentage complex formation calculated form the above assays are shown in Table 1.

Figure 7. FCS analysis of transcription factors in live cells. A. Confocal images (left) and best-fit autocorrelation function (ACF) curves generated from FCS assay of cells expressing GFP (top) or GFP-Nanog (bottom). Scale bar = 10µm. **B**. Normalized CPS values for GFP and GFP-Nanog. For all experiments, n = 15, mean ± SD.

Figure 8. FCCS analysis of transcription factors in live cells. A. Autocorrelation functions (ACF) and cross-correlation functions (CCF) from FCCS assay with live CHO cells co-transfected with GFP and mCherry (left) or with tandem GFP-mCherry (right). **B**. Autocorrelation functions and cross-correlation functions from live CHO cells co-transfected with GFP-Nanog-W10A and

mCherry-Nanog (left panel) and GFP-Nanog and mCherry-Nanog (right panel). **C**. Percentage complex for species in A and B. For all experiments, cell number = 10, mean ± SD).

Table 1: Complex% values for different sets of *in-vitro* FCCS experiments from Figure 4 & 5.

Supplementary Figure 1. Functional test of Nanog fusion proteins in the self-renewal assay. E14/T ES cells were transfected with plasmids encoding Nanog fusion proteins and colonies stained with alkaline phosphatase 10 days post-transfection in the absence of LIF. Undifferentiated (ES), mixed and differentiated colonies were counted. n = 3, mean ± SD.

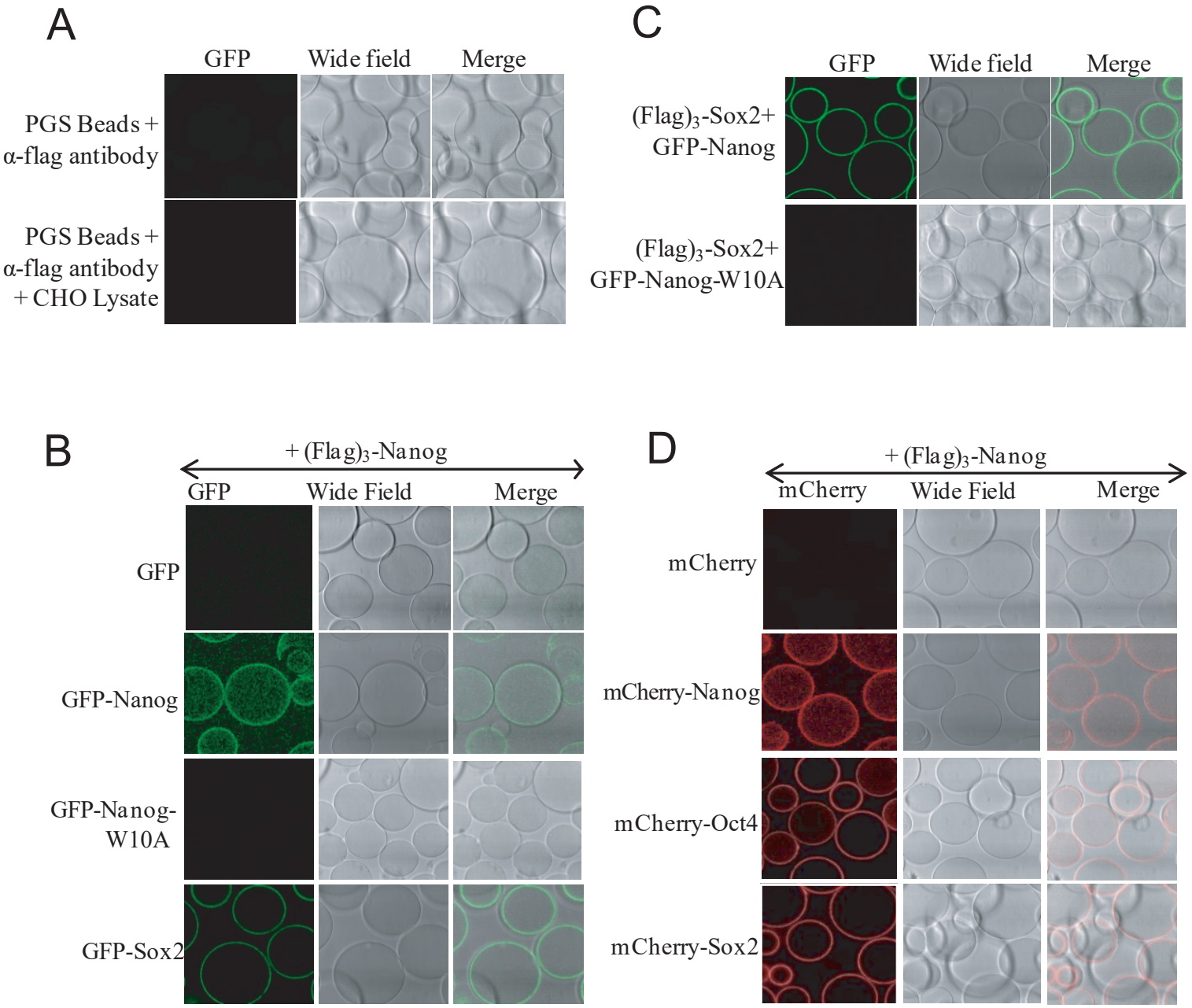
Supplementary Figure 2.CBIM diagram. Protein G beads are coated with anti-Flag antibody and used to immunoprecipitate proteins or complexes of interest. The protein of interest or a binding partner is fused to a fluorescent protein allowing interactions to be visualized by confocal microscopy. The assay can also be used to visualize the interaction between protein and a fluorescently tagged DNA target sequence.

Supplementary Figure 3. The best fit autocorrelation functions for fluorescent species from FCS. A. Representative best fit autocorrelation functions (ACFs) for GFP (left), GFP-Nanog (middle) and GFP-Nanog-W10A (right). For fitting ACFs of GFP-Nanog, a 3D-2P model was used. A 3D-1P model gives best fit for GFP and GFP-Nanog-W10A. **B**. Representative best-fit ACFs for mCherry (left panel, 3D-1P-1T model), mCherry-Nanog (middle panel, 3D-2P-1T) and mCherry-Sox2 (right panel, 3D-1P-1T) (Wohland, Friedrich et al. 1999, Mistri, Devasia et al. 2015).

Supplementary Figure 4. Confocal images of live cells used in FCCS analysis. (**A-B**). Representative confocal images for positive (tandem GFP-mCherry) (**A**) and negative (co-transfection of GFP and mCherry) controls (**B**). Scale bar is 10 µm. (**C-D**). Representative confocal images for CHO cells co-transfected cells with GFP-Nanog-W10A and mCherry-Nanog (**C**) and CHO cells co-transfected with GFP-Nanog and mCherry-Nanog (**D**). The scale bar is 3.5 µm. The intersection of the two lines refers to the position where FCCS was measured.

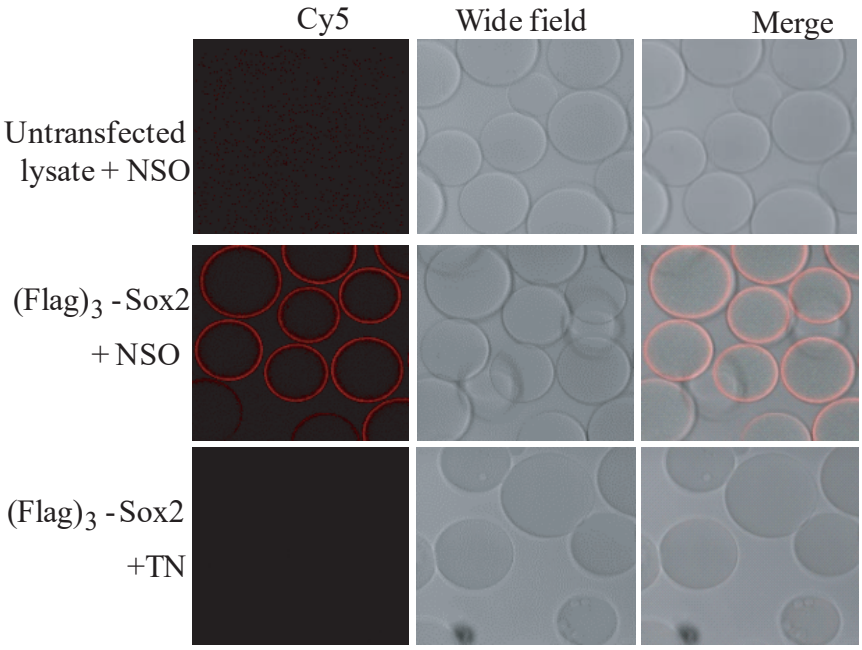
Supplementary Table 1. All the DNA sequences used in this paper are shown here.

Mistri et al
Figure 1

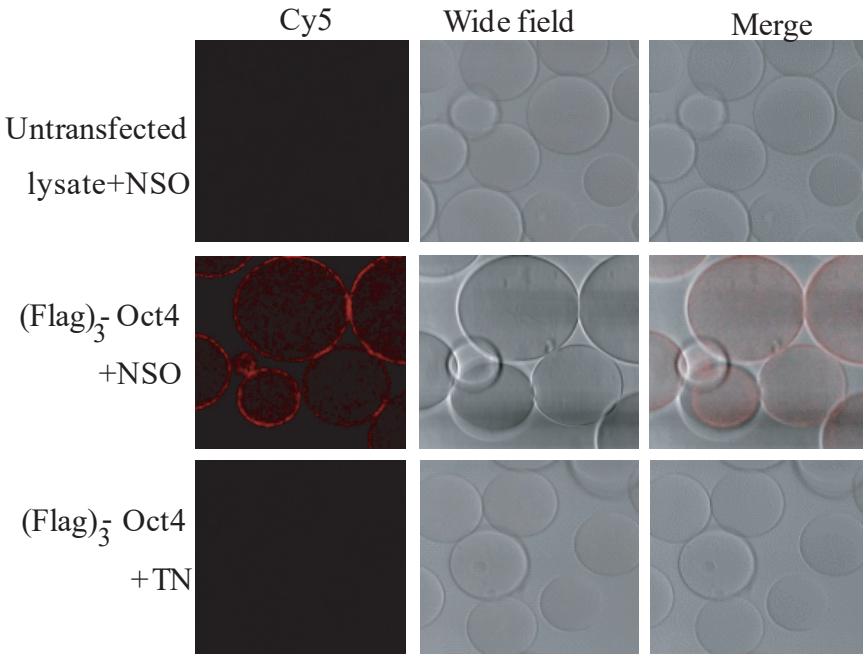


Mistri et al
Figure 2

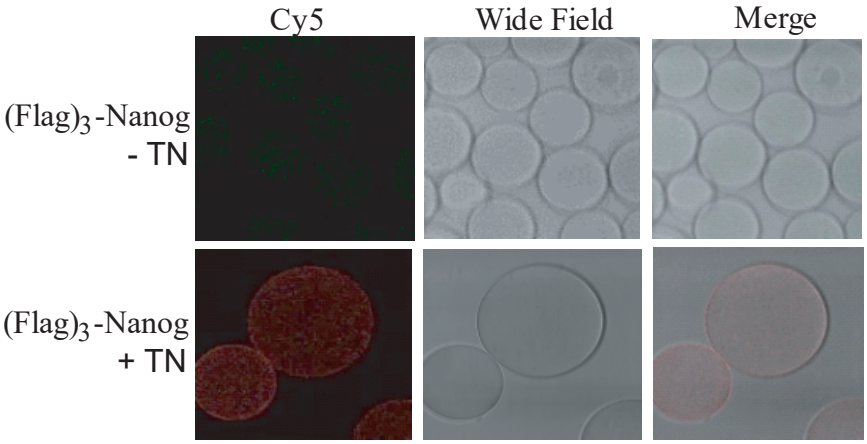
A



B

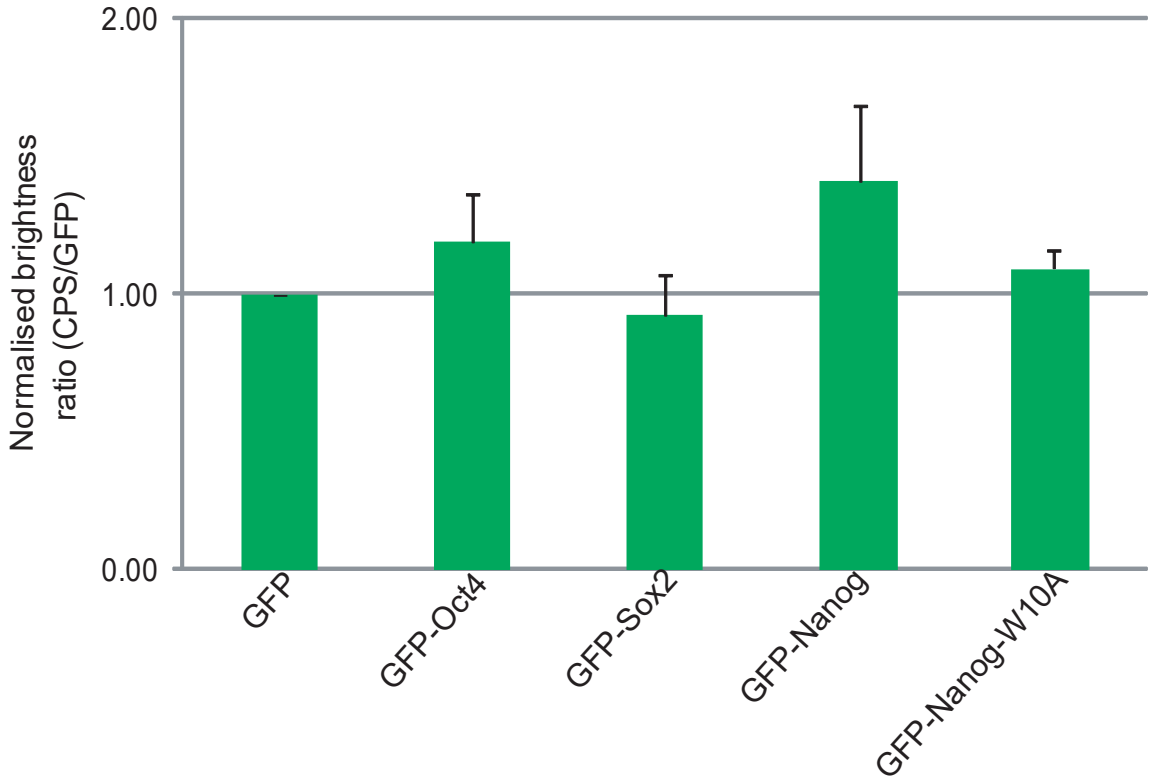


C

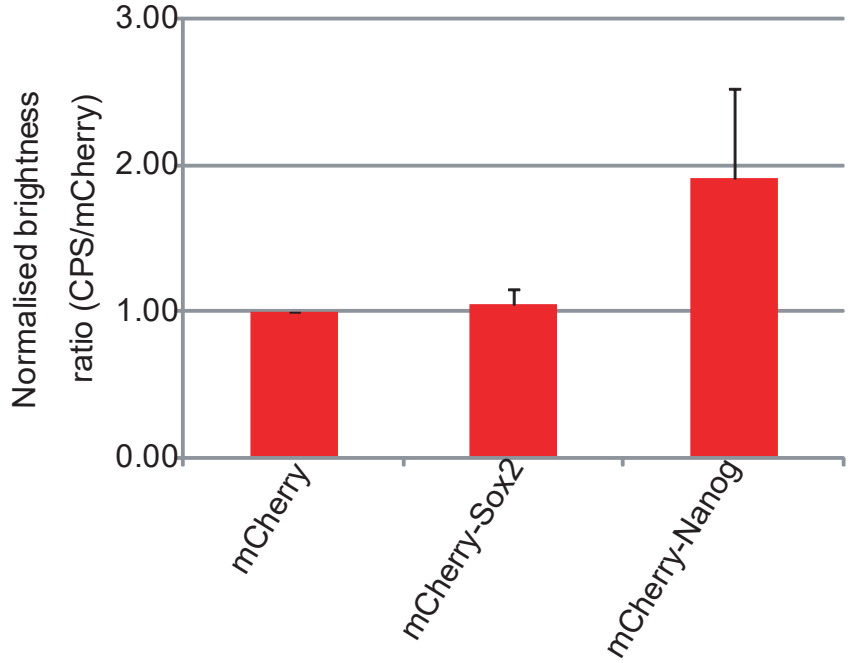


Mistri et al
Figure 3

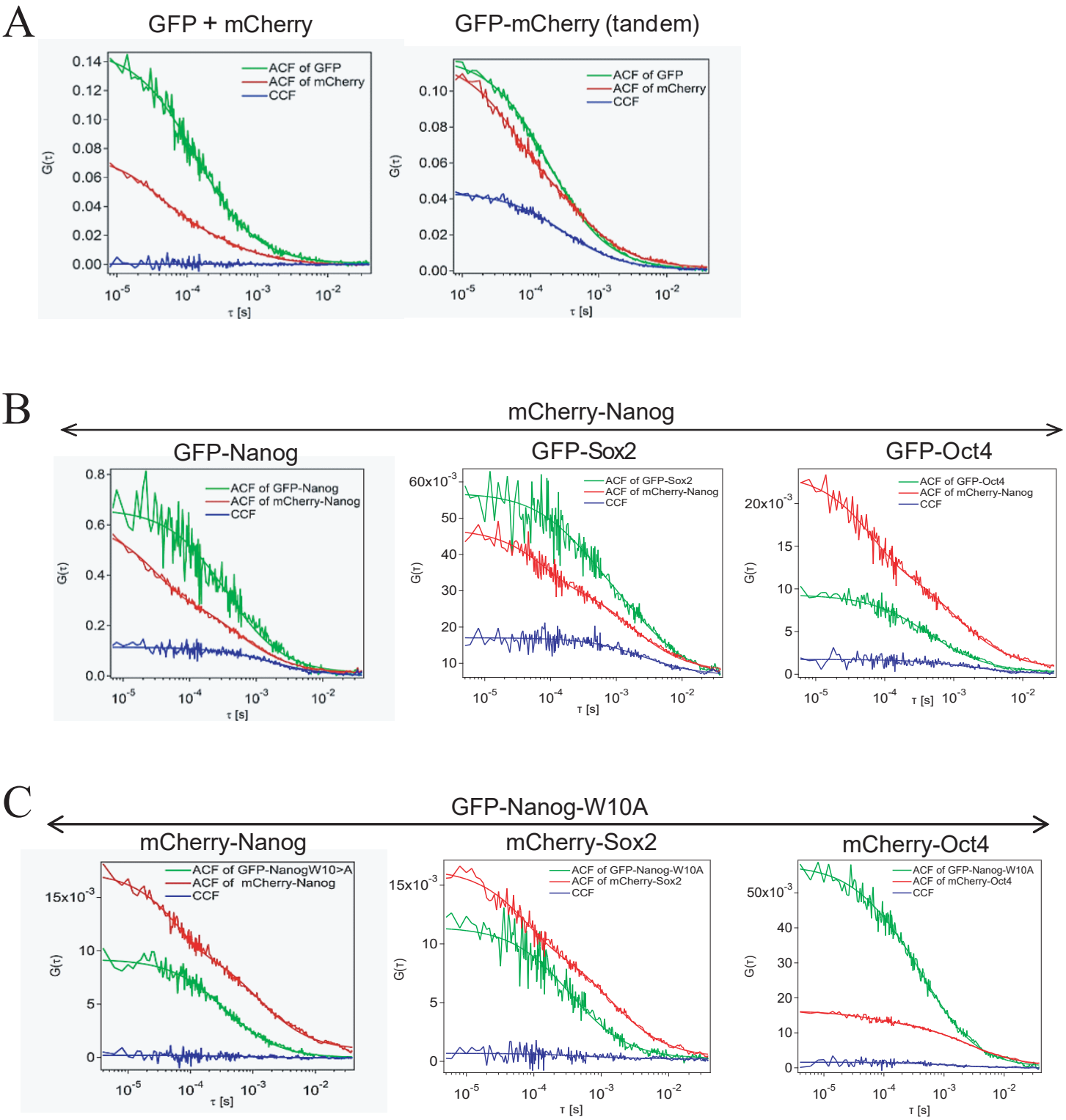
A



B

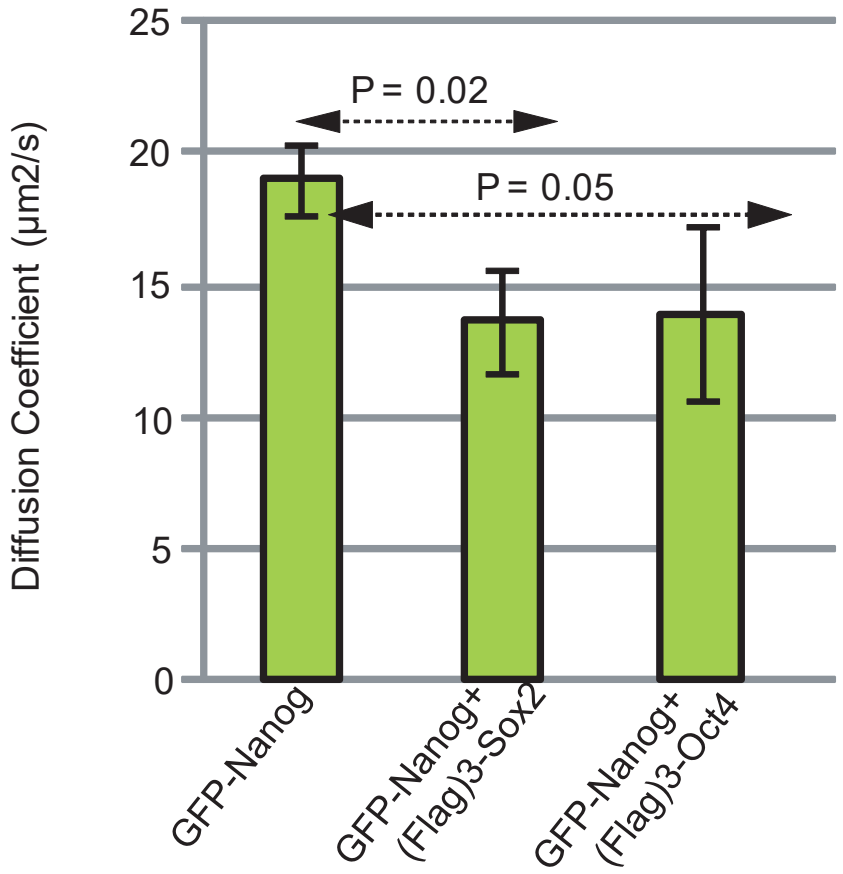


Mistri et al
Figure 4

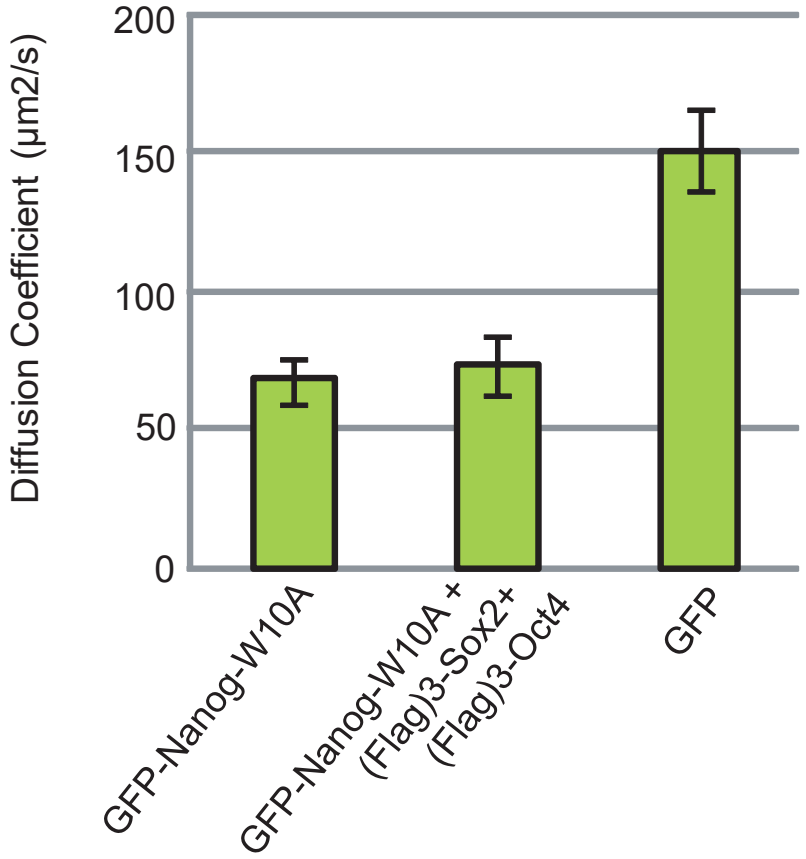


Mistri et al
Figure 5

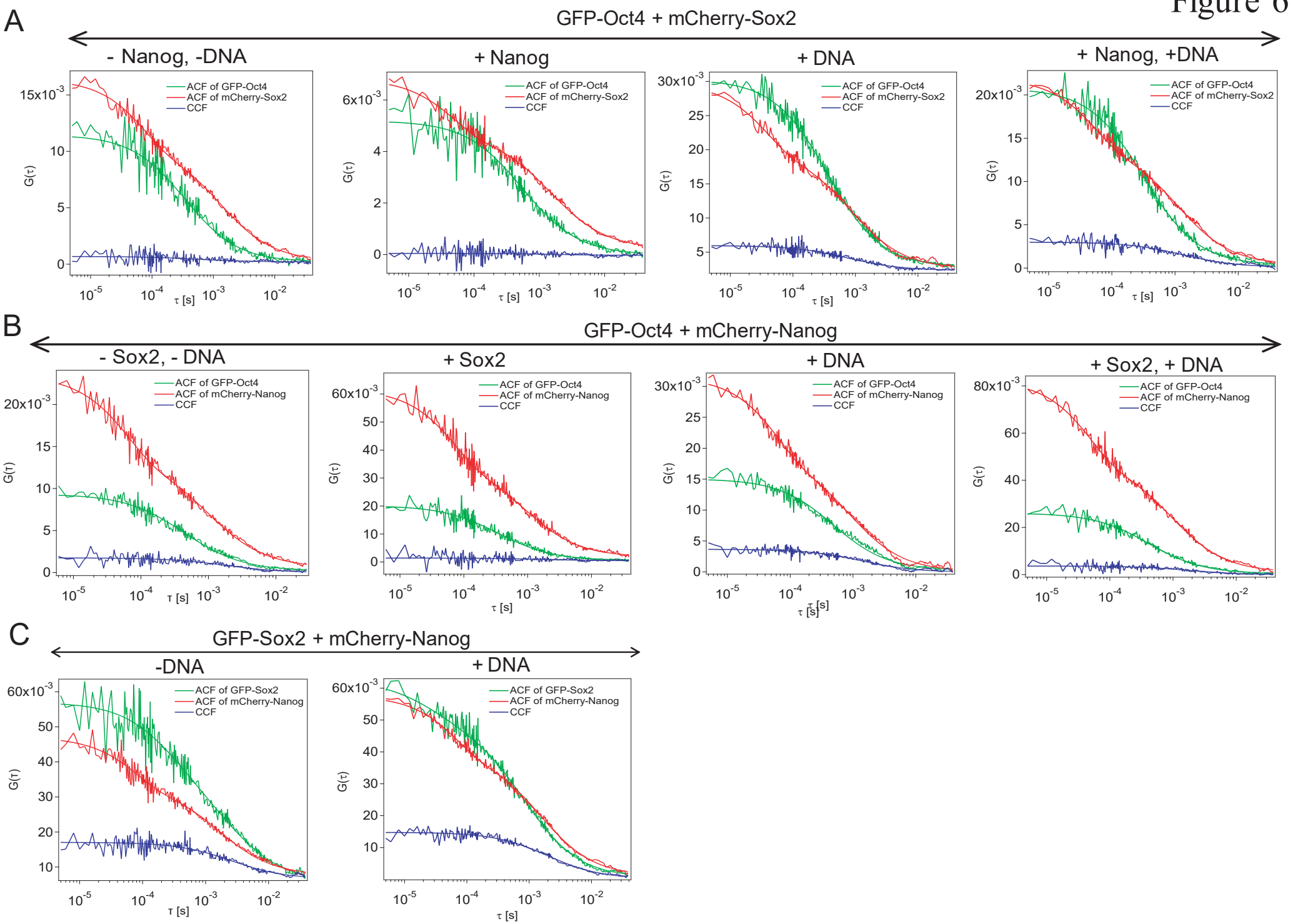
A



B

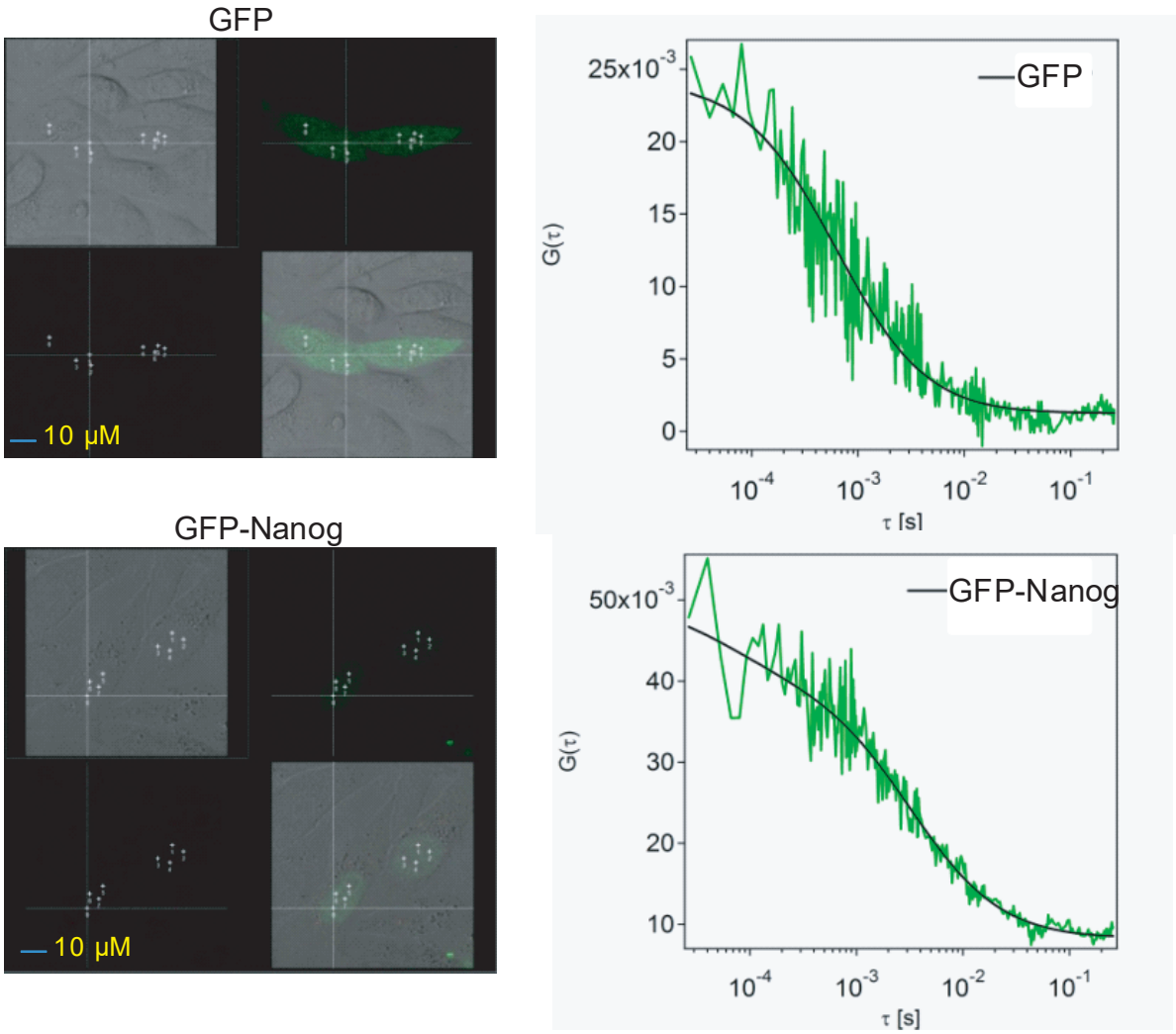


Mistri et al
Figure 6

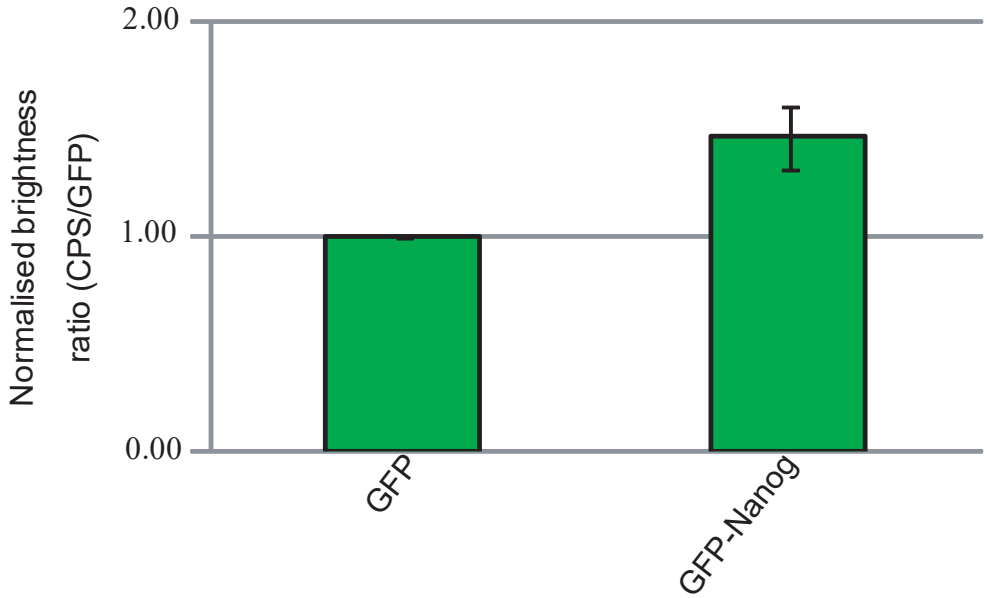


Mistri et al
Figure 7

A

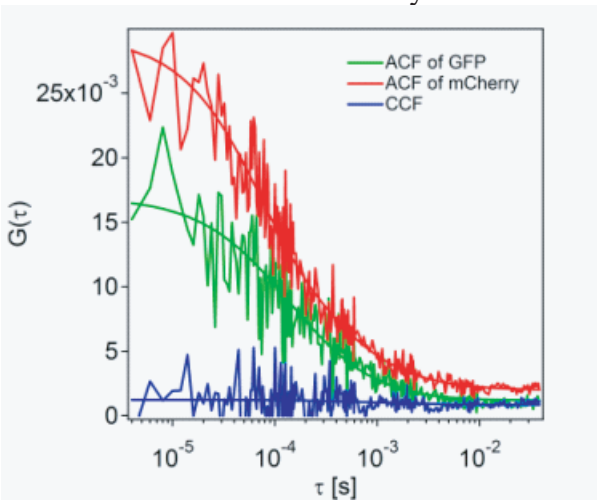


B

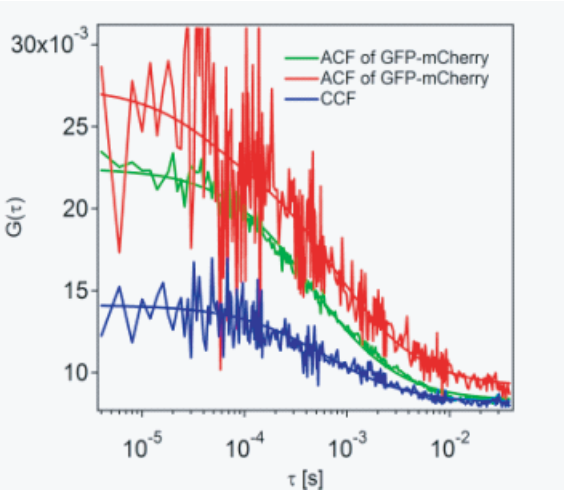


A

Negative control
GFP + mCherry

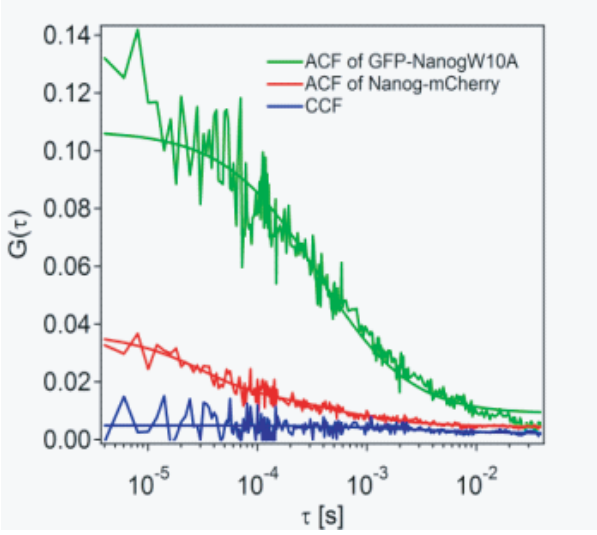


Positive control
GFP-mCherry

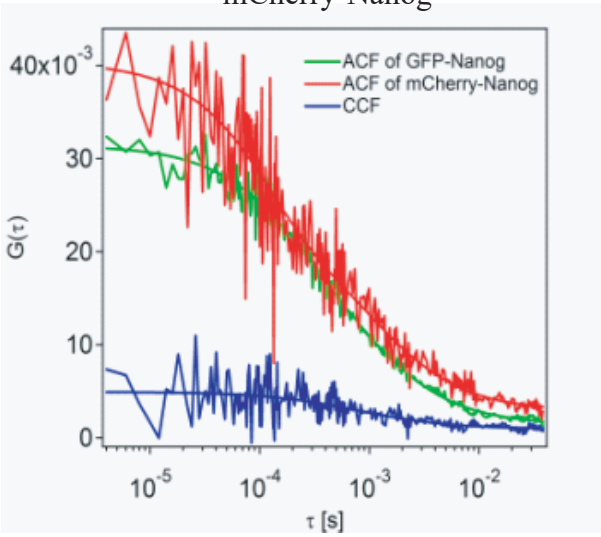


B

GFP -Nanog-W10A +
mCherry-Nanog



GFP -Nanog +
mCherry-Nanog



C

Samples	Complex %	n
GFP + mCherry	3.7 ± 2.6	15
GFP-mCherry	47.4 ± 9.2	15
GFP -Nanog-W10A + mCherry-Nanog	4.1 ± 2.8	15
GFP -Nanog + mCherry-Nanog	16.8 ± 7.2	15

Table 1

Interacting partners with additives	Complex %
GFP-Oct4 + mCherry-Sox2	
No additive	3.9 ± 3.1
+ Nanog	3.0 ± 2.7
+ DNA	7.5 ± 4.5
+ DNA, + Nanog	8.4 ± 5.0
GFP-Oct4 + mCherry-Nanog	
No additive	12.5 ± 6.7
+Sox2	5.0 ± 3.5
+DNA	8.5 ± 3.6
+DNA, + Sox2	6.3 ± 4.8
GFP-Sox2 + mCherry-Nanog	
No Additive	26.1 ± 7.7
+ DNA	15.8 ± 7.2
GFP-Nanog+ mCherry-Nanog	
No additives	19.8 ± 2.9
Positive and negative controls	
GFP-mCherry (tandem)	46.2 ± 6.7
GFP + mCherry	2.5 ± 1.9
mCherry-Nanog + GFP-Nanog-W10A	1.7 ± 1.4
mCherry-Oct4 + GFP-Nanog-W10A	3.6 ± 2.5
mCherry-Sox2 + GFP-Nanog-W10A	2.8 ± 1.7