CONCENTRATION AND DYNAMICS OF TWO MORPHOGENS, BCD AND CIC

CONCENTRATION AND DYNAMICS OF TWO EARLY FLY EMBRYO MORPHOGENS, BICOID AND CAPICUA, EXPLORED BY FLUORESCENCE CORRELATION SPECTROSCOPY

By

LILI ZHANG, B.Sc., M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree of

Doctor of Philosophy

McMaster University

Doctor of Philosophy (2021) McMaster University

(Department of Physics & Astronomy) Hamilton, Ontario, Canada

TITLE:	Concentration and dynamics of two early fly embryo morphogens, Bicoid and Capicua, explored by Fluorescence Correlation Spectroscopy
AUTHOR:	Lili Zhang B.Sc.(Nanjing University of Science & Technology) M.Sc.(Jinan University)
SUPERVISOR:	Dr. Cécile Fradin
NUMBER OF PAGES:	xvii, 138

Lay Abstract

Have you ever wondered how a single fertilized egg turns miraculouly into a beautifully organized living being, be it an insect, a cat, or a human? It turns out that an important group of molecules called morphogens govern the formation of body pattern. These molecules (usually proteins) form concentration gradients along the different body axes of that organism and influence gene expression. Abnormal distribution of morphogen can result in defects in embryo development and even death. Thus knowing how much morphogen is present in the early developing embryo, as well as how it forms gradients and how the morphogen concentration is translated into a pattern can help us better understand early embryo development. My thesis focuses on accurate measurements of morphogen concentrations and dynamics using fluorescence techniques. We were able to obtain concentration maps for two morphogens, the activator Bicoid and the repressor Capicua, in early developing fruit fly embryos. We also found that despite having opposite functions, the activator and the repressor have similar intranuclear dynamics, but drastically different internuclear mobility. Our findings provide clues to distinguish between multiple hypothetical models scientists have put forward to explain the mechanisms of transcription regulation.

Abstract

Morphogens (often acting as transcription activators or repressors) govern pattern formation and cell differentiation during early embryogenesis. Abnormal distributions of morphogens can result in developmental defects or even death. Oftentimes, thresholds of concentrations of morphogens behave like an ON/OFF switch for the activation or repression of downstream genes. Accurate measurements of morphogen concentration and mobility in space and time can help tackle the puzzle of how exactly cascades of hundreds of morphogens coordinate their targets precisely and promptly amidst crowded and complicated cellular environments. The research question at the centre of my thesis is that of the concentration and dynamics of two morphogens with opposite functions in the early fly embryo. In the work presented in this thesis, we use Fluorescence Correlation Spectroscopy (FCS) and confocal imaging to achieve extremely low ($\sim$ nM) concentration measurements in live *Drosophila* embryos expressing recombinant fluorescent morphogens, by carefully taking into account background noise and photobleaching effects. The dynamics of both Bicoid (Bcd) and Capicua (Cic), an activator and a repressor morphogens, were further studied using FCS, Fluorescence Recovery After Photobleaching (FRAP) and Monte Carlo simulation. We found that both types of morphogens are very mobile in nuclei, explaining how they are able to turn on or off gene expression in only a few minutes. However, these

two morphogens with opposite functions have drastically different nucleo-cytoplasmic transport behaviours, where the activator can pass through the nuclear envelop (NE) relatively freely while the repressor is jailed inside nuclei during interphase. These findings can provide clues to distinguish between several hypothetical models (including the newly proposed hub hypothesis) trying to explain the mechanisms of target gene search and transcription regulation.

In this thesis, a background introduction on transcription factors and morphogens is given in Chapter 1, with a focus on the two transcription factors (the activator Bicoid and the repressor Capicua) studied in this thesis. Next, experimental details such as fruit fly maintenance, and fluorescent techniques used to measure concentration and mobility are described in Chapter 2. From Chapter 3 to Chapter 5, three manuscripts from the thesis author, either published or in preparation for submission are presented in sequence. Chapter 3 introduces a new method to accurately measure protein concentration in the presence of noise and photobleahing in early *Drosophila* embryos using FCS. Chapter 4 contains the results of concentration and mobility measurements for Cic which contribute to the finding that Cic acts like a fast brake in transcription repression. Chapter 5 compares the similarities and differences of the dynamics of Bcd and Cic through multiple lenses. Finally, a conclusion and future outlook are given in Chapter 6.

Dedicated to

Mom and Dad

Acknowledgements

First and foremost, I'd like to express my gratitude towards my supervisor, Dr. Cécile Fradin for taking me on as her student and guiding my PhD thesis projects. Her constant patience, devotion, and trust in me as a researcher mean a lot to me and make me less shaky when research gets hard. Without you, I truly wouldn't be who I am and where I am today, for that I am forever grateful. I'm also grateful to Dr. Maikel Rheinstädter for leading me into biophysics and teaching me membrane biophysics using X-ray diffraction and MD simulations. I'd also like to thank my committee members, Drs. An-Chang Shi and Jose Moran-Mirabal, for their continuous support and feedbacks throughout the past five years. The professors teaching me courses in the Physics and Chemistry department have also been a source of inspiration so I also want to thank them: Profs. Bruce Gaulin, Takashi Imai, Catherine Kallin, An-Chang Shi, Paul Higgs and Jim Britten.

My lab members were very helpful and I'm glad we shared some experience together in this stage of our lives. Thanks Carmina, for teaching me the husbandry of fruit flies and preparing embryo samples. Thanks Markus, for teaching and sharing with me his preliminary confocal simulation programs. Thanks Michael, for helping me troubleshooting the more than 20 years old Instrument Insight. Thanks Sam and Mitchell, for your help with the FRAP instrument. Thanks to Ahmad and Liu for

being awesome labmates and friends, who always have so many fascinating stories and brilliant ideas to share. Thanks also Sebastian, Adree, Rick, Dylan, Alex, Renee for the encouragement and help during the early years and I did learn a lot from each of you.

The graduate students in the P&A department are a great source of support and the get-togethers they organized have been perfect break from research. Each one of them is so unique and fabulous and definitely made my PhD life worth remembering. I'd like to thank PIPA (previously GWIPA) led by Carmen, Sara, Sarah, Gwen, for it indeed gave me a sense of belonging and a feeling of being backed up. The department softball team, Liquid State, also provided lots of workout and fun. I'd also like to thank my dear friend and roommate Ivy, for always being there for me when I had life or school questions. I also won't forget to thank Rose, Tina and Mara, the secretaries of the department, for the help and support they kindly gave me.

Last but not least, I thank my parents, Xueqin Rong and Jiaxi Zhang, for allowing me to freely choose my life path and supporting me unconditionally, especially considering where we came from and what we had been through, this is incredibly wise and selfless of you to do.

Contents

List of Figures

List of Tables

Definitions, Abbreviations and Notation

Definitions

Morphogen A substance, usually a protein, which controls the morphology and pattern formation of a developing embryo or tissue through a concentration gradient

Transcription factor A protein that binds to specific DNA regions to regulate the transcription of genes

Activator A transcription factor that activates the expression of its target genes

Repressor A transcription factor that represses the expression of its target genes

Fluorescence Emission of visible light by a substance that has absorbed light or other electromagnetic radiation

Abbreviations

Bcd	Bicoid, a transcription activator and a morphogen; lower case and italic as in *bcd* represents the gene
Cic	Capicua, a transcription respressor and a morphogen; lower case and italic as in *cic* represents the gene
NLS	Nulcear Localization Signal, a short amino acid sequence that is recognized by nuclear transport proteins and lead to the protein carrying the NLS to be imported in the cell nucleus
TF	Transcription factor
eGFP	enhanced Green Fluorescent Protein
sfGFP	superfolder Green Fluorescent Protein
FCS	Fluorescence Correlation Spectroscopy
ACF	Autocorrelation Function
FRAP	Fluorescence Recovery After Photobleaching
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
AP	Anterior-Posterior
NC	Nuclear Cycle

Notation

Bcd-eGFP Recombinant protein where an eGFP molecule is attached to a Bcd protein

Cic-sfGFP Recombinant protein where a sfGFP molecule is attached to a Cic protein

NLS-eGFP Recombinant protein where an eGFP molecule is attached to a NLS

Chapter 1

Introduction

The proper development of a single, apolar fertilized cell into a polar, multicellular organism involves thousands of complicated biological processes, including cell division, cell differentiation, and cell proliferation on the cellular level. These processes are driven by gene transcription and translation on the molecular level (Fig. 1.1). Indeed, the central dogma of molecular biology (i.e. DNA transcripts to RNA, and RNA translates to protein), controls these important chains of reactions where DNA acts as the commander of the whole machinery of the organism. The execution of DNA transcription is controlled by many factors, and maybe the most important is transcription factors (TFs), which are proteins that bind to particular DNA regions to regulate transcription of particular genes. During the early embryonic development, the transcription of genes is regulated in space and time by molecules called morphogens (which are often transcription factors themselves), which control the pattern formation and morphology of the future adult into which the embryo turns (Shahbazi, 2020). Malfunction and abnormal distribution of these transcription factors can both result in developmental defects at early stages, or diseases and cancers later in life

(Latchman, 1996; Bushweller, 2019). Thus quantitative studies of morphogens *in vivo* in order to better understand how they function are an ongoing priority in the field.

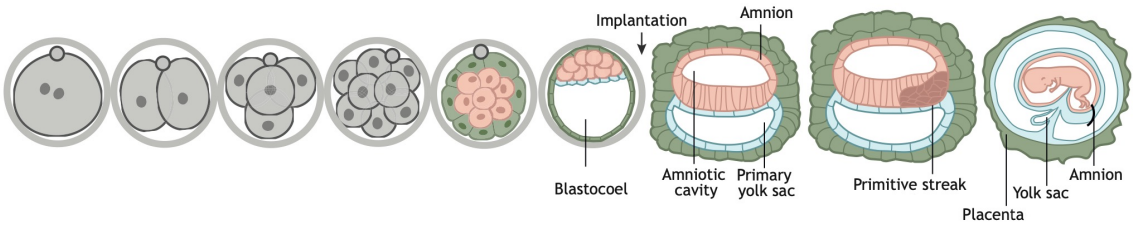

Figure 1.1: An illustration of the development of a human embryo. Upon fertilisation, the embryo undergoes a series of cellular divisions, and cell differentiation. Ultimately, a single apolar fertilized egg turns into a polar, multicellular organism. Reproduced with permission from (Shahbazi, 2020).

In the following sections of this chapter, first an introduction on transcription factors will be given. Then morphogens will be introduced, with emphasis on two aspects of their behaviours, concentration and dynamics. Third, the two morphogens studied in this thesis will be introduced as two representative morphogens having opposite transcriptional functions. Lastly, the use of *Drosophila melanogaster* as a model system is explained and justified.

1.1 Transcription factors

1.1.1 Transcription in general

As the first of the two steps in the central dogma of molecular biology, transcription is the process where the information stored in DNA is transferred to RNA. This process is regulated by transcription factors (TFs), which are proteins that can bind specific DNA sequences upstream of their target genes (the genes they help regulate) (Mitsis *et al.*, 2020). TFs include sequence-specific TFs, general transcription

factors (gTFs), and other core transcription machinery, such as RNA polymerase II, and chromatin remodelers (Lu and Lionnet, 2021). Sequence-specific TFs are called transcriptional activators when they activate the expression of their target genes and transcriptional repressors when they repress the expression of their target genes. The short ($100 \sim 1000$ base pairs) DNA base pair patterns in the non-coding DNA region upstream of target genes to which TFs bind are termed motifs, or *cis*-regulatory elements (CREs). CREs include promoters and sequences called enhancers in the case of transcriptional activation and silencers in the case of transcriptional repression. Transcriptional activators bind at enhancers to upregulate transcriptional activation, while repressors bind at silencers to downregulate gene expression (Jayavelu *et al.*, 2020). The promoters are bound by general TFs or other transcription machinery to initiate transcription (Fig. 1.2). The DNA-binding domains (DBDs) found on TFs include homeodomain (HD), helix-turn-helix (HTH), high-mobility group box (HMG) and zinc-finger (ZNF). TFs can be classified according to these different DBDs.

About 708 of the 14000 ($\sim 5\,\%$) protein-coding genes in *Drosophila* genome are transcription factors with characterized DNA-binding domains (Rhee *et al.*, 2014). Aberrant TFs numbers are closely connected to pathogenic abnormalities of gene expression, including either increased or decreased levels. In particular, reduced levels of some TFs may result in haploinsufficeint (HI) diseases, such as eye disorders, or facial and limb diseases (Auer *et al.*, 2020).

1.1.2 Transcription activation and repression

Even after decades of research in the field, the mechanism regarding transcription regulation is not fully understood, and multiple models have been proposed to answer the

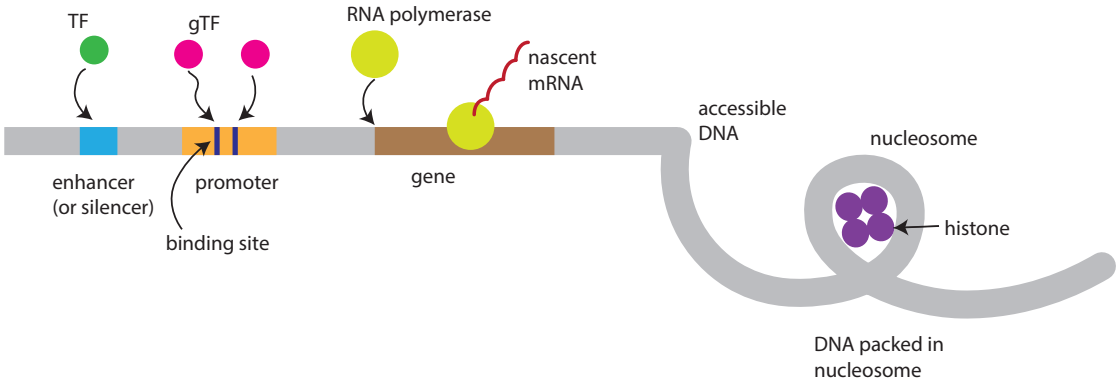

Figure 1.2: An illustration of the components in the transcription of DNA. Sequence-specific transcription factors (TFs) bind to either enhancer in the case of transcription activation or silencer in the case of transcription repression. General TFs (gTFs) bind to promoter, which is downstream of enhancer or silencer. RNA polymerase II bind to the start of the coding region of gene. DNA is wrapped around histones to form the nucleosome.

question of how exactly transcription regulation (including activation and repression) is enacted. In eukaryotic cell nuclei, histones are proteins that act as spools around which DNA winds. Histone acetylation has generally been associated with active transcription whereas deacetylation is associated with repression (Reynolds *et al.*, 2013). Acetylation of histone results in the unwinding of DNA side chains thus opening up the chromatin structure, allowing increased transcription. Histone deacetylation reverses this process thus leading to gene repression (Adcock and Caramori, 2009). This process is also referred to as Nucleosome Remodelling and deacetylation (NuRD). For example chromatin remodelling protein Mi-2 in *Drosophila melanogaster* is associated with gene repression.

For transcriptional activation, both a classical model and a more recent model exist regarding the mechanism of transcription regulation. The classical view supports an enhancer-promoter interaction where TFs bind to binding sites within the

enhancer and persist for a long periods due to low off rates (von Hippel, 1998). The emerging new view proposes the existence of hubs, or condensates, i.e. a high local concentration of TFs, which help maintain the occupancy during the binding and unbinding of TFs to the target sites (Kato *et al.*, 2012). Hubs of Bcd and Zelda have been observed in live *Drosophila* embryos and their activity are thought to coordinate active transcription (Mir *et al.*, 2017, 2018). The components of hubs are thought to be TFs, RNA polymerase II and other transcription machineries, such as mediators (Lim and Levine, 2021).

For transcriptional repression, two types of mechanisms exist: i) global repression and ii) gene-specific repression. The first case happens when repressor proteins modify RNA polymerase or pre-initiation protein. The second case happens when repressor proteins tamper with the concentration or the function of transcriptional activators. Gene-specific repression involves interaction with DNA directly or indirectly. Depending on whether the repressor proteins act distally or locally, they are sometimes called 'long-range' repressors or 'short-range' repressors (Gaston and Jayaraman, 2003) .

In *Drosophila*, transcription activators include Bicoid, and Hunchback, while transcription repressors include Hunchback, Nano, Krüppel, Giant, and Capicua. Some transcriptional factors can act as both activator and repressor, such as Torso and Hunchback.

1.2 Morphogens

1.2.1 Background information

Morphogens are substances that regulate the pattern formation in a developing embryo by 1) forming a concentration gradient along a body axis and 2) affecting gene expression in a concentration or activity-dependent manner. They exist in a wide range of organisms with various molecular types. Morphogens were first studied directly in fruit flies in late 1980s. The term "morphogen" was coined by Alan Turing (Turing, 1952), and the first morphogen, Bicoid, was discovered by Christiane Nüsslein-Volhard and Eric Wieschaus who were awarded the Nobel prize in Physiology or Medicine in 1995 for this discovery (Nüsslein-Volhard and Wieschaus, 1980). Morphogen gradients can form along different body axes of the embryo. For example, some morphogens have concentration gradients along the anterior-posterior (AP) axis of an embryo, such as Bicoid and Nanos, while others have gradients along the dorsal-ventral (DV) axis of the embryo, such as Dorsal. Besides Bicoid (AP gradient controlling head and thorax pattern), other morphogens in *Drosophila* include Dorsal (DV gradient controlling dorsoventral pattern), Nanos (AP gradient controlling formation of the abdomen), dpERK (AP gradient controlling terminal pattern), Decapentaplegic (Dpp, controlling tissue patterning) (Ferguson and Anderson, 1992), Hedgehog (Gallet, 2011) and Nodal (Schier, 2009) (controlling embryonic axis formation). The four maternal morphogens (i.e. the RNA or protiens are deposited in the fertilised egg by its mother), Bicoid, Dorsal, Nanos, and dpERK, sit at the top of the hierarchy of networks of gene expression thus playing critical role in pattern formation (Shvartsman *et al.*, 2008). Morphogens can be transcription factors in

which case they affect gene expression directly. A summary of morphogen family, the organism and function can be found in Table 1 in a recent review (Stapornwongkul and Vincent, 2021).

One of the first models trying to explain how morphogen gradients pattern developing tissues is the French Flag model (Wolpert, 1969). This model proposes that for each of the morphogen target genes, there is a single threshold concentration, above which the gene is expressed (or repressed), whereas it is not expressed (or repressed) below that concentration (Fig. 1.3). Because for each target gene the threshold concentration can be different, a single morphogen gradient can delimitate several regions, in which distinct sets of genes are expressed. In this way, an exponential decay concentration gradient of a morphogen can induce different cell fates and provide positional information to cells.

1.2.2 Concentration and dynamics of morphogens

One of the two defining characteristics of morphogens is their concentration gradients along a body axis, therefore quantitative exploration of the concentration profiles of morphogens *in vivo* is a natural component in the quest for better understanding the functions of morphogens. The concept of a threshold concentration has been proposed very early (Dalcq, 1938), even before the discovery of the first morphogen, to explain the phenomenon that particular cell fates emerge only above certain threshold concentrations. Later it has been found that tissue patterning is controlled not only by the concentrations but also the duration of exposure to morphogens, i.e. the dynamics of morphogens are also essential. As a result, concentration and dynamics are two important aspects of morphogens for a complete understanding of the functions of

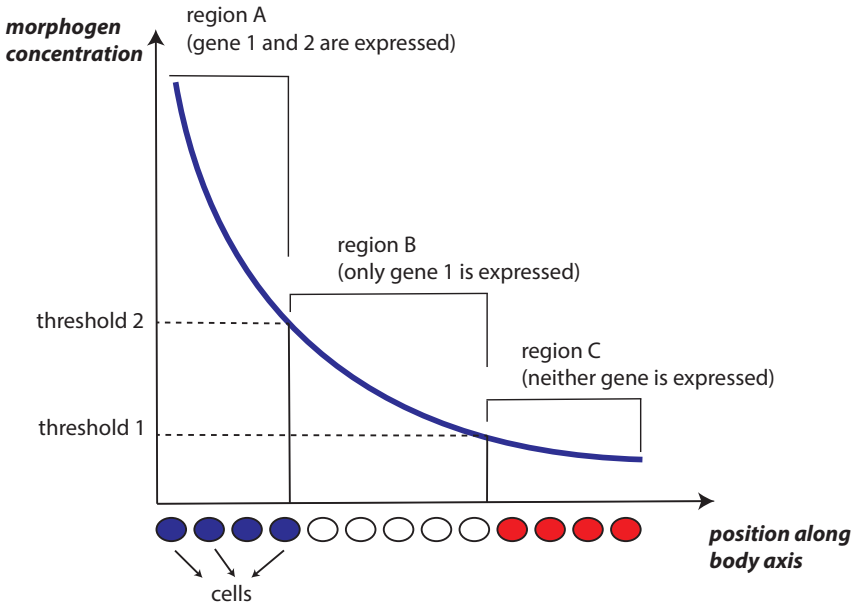

Figure 1.3: An illustration of the French Flag Model. In this illustration, two genes, gene 1 and gene 2, have different threshold concentrations, threshold 1 and 2. Above threshold 2, which is depicted as region A, both gene 1 and 2 are expressed. In between threshold 1 and 2, which is region B, only gene 1 is expressed. Below threshold 1, which is region C, neither gene is expressed. In this way, an exponential decay concentration gradient of a morphogen can induce different cell fates and provide positional information to cells.

morphogens. However, accurate concentration measurements of morphogens in a living organism is in itself a challenging task, *e.g.*, due to dynamic nature of embryos, and only after the advent of recombinant fluorescent protein fusion, did tackling this problem become possible. So far the concentration and dynamics of morphogens have been measured for only a few morphogens, including Bcd (Gregor *et al.*, 2007b; Abu-Arish *et al.*, 2009) and Dpp (Wartlick *et al.*, 2011).

1.3 Proteins studied in this thesis

Two proteins, Bicoid and Capicua, are studied in this thesis, which are both morphogens as well as transcription factors but with opposite functions. A third protein, NLS (Nuclear localization signal), which is neither a morphogen nor a transcription factor, is also examined to serve as a control for the first two proteins.

1.3.1 An activator morphogen: Bicoid (Bcd)

The first protein studied in this thesis is Bicoid. It is both a morphogen and a transcription factor, specifically a transcription activator. In the following, we will introduce Bicoid with regard to, first its morphogenic aspect, then its transcriptional aspect.

Morphogenic aspect

Bicoid (Bcd) is the first discovered (in 1988) and most studied morphogen (Nüsslein-Volhard and Wieschaus, 1980; Porcher and Dostatni, 2010; Fradin, 2017). The first characteristic for Bcd as a morphogen is that it forms an exponential decay concentration gradient along the Anterior-Posterior (AP) axis in the early ($1 \sim 3$ hrs after egg deposition) *Drosophila* embryo, with the maximum concentration at the anterior pole and minimum concentration at the posterior pole (Fig. 1.4a). The exponential gradient has the form: $c(x) = c_0 e^{-x/\lambda}$, where λ is the characteristic decay length. By using immunofluorescent staining or the recombinant fusion protein Bcd-eGFP, the characteristic decay length, λ, is found to be between 80 to 120 µm, about one fifth or one fourth of the embryo length (Abu-Arish *et al.*, 2010; Gregor *et al.*, 2007a). After taking into account the underestimation of the concentration measurement, due

to reasons like the maturation time for eGFP, the decay length is about one sixth of the embryo length (Tran *et al.*, 2020). The Bcd gradient is established rapidly ($\sim$ 1 hr after fertilization) with nuclear Bcd concentration quickly building up after each mitosis, to reach a nuclear to cytoplasmic concentration of $\sim$ 8 (Grimm and Wieschaus, 2010).

Bcd is a maternal gene where *bcd* mRNA is deposited and localized at the anterior pole of the embryo by its mother. During the very early embryo development (nuclear cycle (NC) 1 - 7), the embryo expresses only maternal genes, and Bcd protein is synthesized at the anterior pole. Previous experiments have shown that Bcd then diffuses along the anterior posterior axis (Abu-Arish *et al.*, 2009). Assuming the protein is degraded uniformly, leads to the prediction of an exponential decay profile along the AP axis with maximum concentration at the anterior and minimum concentration at the posterior as is observed experimentally (Fradin, 2017). This is the so called Synthesis Diffusion Degradation (SDD) model (Wolpert, 1969; Gregor *et al.*, 2005; Houchmandzadeh *et al.*, 2005). This model predicts $\lambda = \sqrt{D\tau}$, and that since $\tau \approx 25$ min and D was measured to be 7 µm^2/s in the cytoplasm, it predicts $\lambda = \sim 100$ µm, which agrees well with experimentally measured λ of 80 µm to 120 µm (Abu-Arish *et al.*, 2010; Gregor *et al.*, 2007a). Local production and uniform degradation in this SDD model ensures the stability of the Bcd profile. Halftime of Bcd is about 25 min, which is related to its degradation rate (Durrieu *et al.*, 2018). Other models, for example the ARTS (active RNA transport, synthesis), have also been brought up to explain the Bcd gradient formation (Baumgartner, 2018; Spirov *et al.*, 2009), however, they lack strong experimental back-up.

Transcriptional aspect

The second characteristic for Bcd as a morphogen is that it affects gene expression in a concentration-dependent manner. One of the genes being affected by Bcd is *hunchback* (*hb*), which is expressed above a certain threshold concentration of Bcd but not below (according to the French Flag Model) (Fig. 1.4b). Consequently, the smooth decay concentration gradient of Bcd results in a sharp step-like pattern of hunchback expression along the AP axis (Fig. 1.4). The border of the hunchback expression domain is formed at $\sim$ 45 % of the embryo length from the anterior pole (Hülskamp *et al.*, 1990; Holloway and Spirov, 2015). It was found that the time required for the formation of *hb* pattern controlled by Bcd gradient can be as fast as 3 min at NC 11 (Lucas *et al.*, 2018). Patterns of both the *hb* gene transcription and hb protein expression are step-like around the middle of the embryo body (Fig. 1.4 b,c). The precision of the readout of the *hb* gene according to the concentration of the Bcd protein is measured to be $\sim$ 99% (Dubuis *et al.*, 2013; Petkova *et al.*, 2019). Since decreasing or increasing the amount of Bcd results in a shift of the border of the steplike pattern of *hb*, it is thus proved that expression of *hb* depends on the concentration of Bcd (Driever and Nüsslein-Volhard, 1989). The regulation of *hb* transcription by Bcd is strongly nonlinear since the smooth exponential decay pattern of Bcd results in a sharp step-like pattern of *hb* expression. One model to explain this nonlinearity, based on the fact that there are multiple binding sites (5 $\sim$ 7) in each *hb* promoter (Tran *et al.*, 2018), is that Bcd binds cooperatively, as observed in experiments (Ma *et al.*, 1996; Lebrecht *et al.*, 2005).

It was shown that Bcd has at least 66 target genes (Chen *et al.*, 2012), one of them being *hunchback* (*hb*) gene. Hunchback itself is a transcription factor, which controls

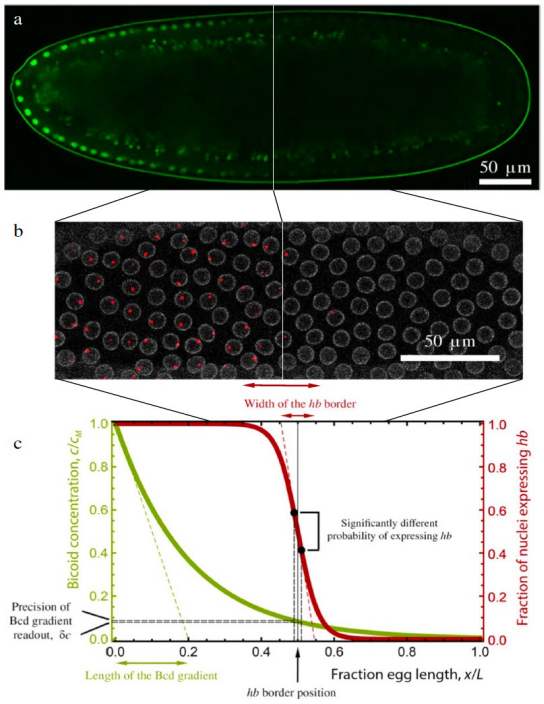

Figure 1.4: (a) Confocal image of a *Drosophila melanogaster* embryo expressing Bcd-eGFP at nuclear cycle 12. The round bright dots along the thin curve of the embryo membrane are nuclei containing Bcd-eGFP. The anterior is to the left side and the posterior pole to the right. (b) Transcription activity of the *hb* promoter, visualized by confocal microscopy using the MS2 system, where nascent *hb* mRNA is fluorescently labelled and shown in red. Each red dot represents an active *hb* promoter, grey circles show nuclear membranes. (c) Schematic profiles of Bcd protein concentration (green) and *hb* response (red). Reproduced with permission from (Fradin, 2017).

the expression of its own downstream targets, including *kruppel* (*kr*), *knirps* (*kni*), and *even skipped* (*eve*). The *hb* gene is also related to AP patterning. Its expression starts from NC 8, the onset of zygotic transcription (i.e. the point at which the embryo starts to transcribe its own genes instead of using maternal genes).

An ongoing question is how Bcd can reach its target so fast. Lattice Light sheet microscopy and single particle tracking shows that Bcd binds DNA transiently and with an average residence time of about 2 s (Mir *et al.*, 2018). Subnuclear hubs (i.e.

high local concentration of TFs) of Bcd have been observed and it is found that bicoid binding is enriched within Zelda hubs. The hubs are very dynamic and their location and intensity change rapidly, indicating an exchange of the TFs in the hubs with the rest of the nucleoplasm. From single particle tracking of Bcd-eGFP, Mir et al. found that $\sim$ 50% of Bcd are immobile (i.e. at any given time 50 % of bcd is transiently bound to DNA) in the nucleus (Mir *et al.*, 2018). It has been found in E. coli that the combination of 1D and 3D diffusion would result in a 100-fold faster search time compared with a pure 3D diffusion search (Hammar *et al.*, 2012), thus it has been suggested that Bcd might also be under this mode of searching for its binding site (Mirny *et al.*, 2009; Abu-Arish *et al.*, 2009; Porcher and Dostatni, 2010).

1.3.2 A repressor morphogen: Capicua (Cic)

Capicua (Cic) is the second protein studied in this thesis. It is also both a morphogen and a transcription factor, specifically, a transcription repressor. Cic has a flat concentration profile in the midbody of the embryo and minimum concentrations at the two terminal poles (Fig. 1.5). The *cic* gene is a maternal gene that is distributed uniformly in the embryo. The formation of the concentration profile for Cic is due to the degradation of Cic at the two poles induced by the Torso signal from the RTK (Receptor Tyrosine Kinase) and MAPK (mitogen-activated protein kinase) pathways (Fig. 1.5).

Cic was discovered in 2000 (Jiménez *et al.*, 2000). The name "capicua" is Catalan meaning "less amount at the two ends". As a transcription repressor, Cic has a high-mobility group (HMG) domain that binds to specific short DNA sequences within the regulatory regions to inhibit the expression of its target genes, including zygotic genes

tailless (*tll*) and *huckebein* (*hkb*), which are two genes responsible for the terminal pattern, as shown in Fig. 1.5A,C (Jiménez *et al.*, 2012). The base pair sequence of the binding region for Cic is TGAATGAA. Another difference from Bcd is that Cic is a conserved gene, meaning that it exists not only in invertebrates, but also in vertebrates, for example in human. *Drosophila* Cic is a homolog of Human Cic proteins (Cic-L and Cic-S) (Lam *et al.*, 2006). Recent findings have linked malfunctions of Cic to several types of cancers in humans, as Cic is thought to be a tumour repressor and mutations of *capicua* are linked to neoplasm such as brain, lung or gastric cancer (Tanaka *et al.*, 2017; Okimoto *et al.*, 2017).

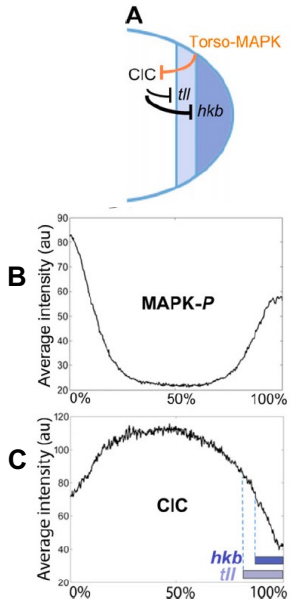

Figure 1.5: (A) Illustration of Capicua (Cic) downregulation by Torso-MARK at the pole of a *D. melanogaster* embryo, and repression of *tailless* (*tll*) and *huckebein* (*hkb*) in the middle region of the embryo. (B) Quantification of active, phosphorylated MAPK (MAPK-P) along the length of the embryo (0% and 100% represent the anterior and posterior poles, respectively); au, arbitrary units. (C) Quantification of Cic levels as in (B). Reproduced with permission from (Jiménez *et al.*, 2012).

Open questions about the two proteins

There are multiple open questions regarding the concentration and dynamics of the two proteins studied in this thesis. For example, what is the actual concentration of Bcd or Cic at the *hb* or *tll* expression border? How do Bcd and Cic find their target genes, especially during such a short time of only a few minutes? Is Bcd or Cic concentration uniform inside nuclei, or are they part of hubs or condensates? Are there any differences and/or similarities between transcription activation and repression by Bcd and Cic, respectively? My thesis work has been taken in the hope of answering those types of questions.

1.4 *Drosophila melanogaster* as a model system

The model system used in my work is *Drosophila melanogaster*, commonly known as fruit fly, which has been used in the field of genetics and embryogenesis since the 1900s (Allen, 1975). Significant advantages of using *Drosophila melanogaster* include the easiness with which it can be handled and maintained, its small size, fast reproduction and short life cycle ($\sim$ 28 days). Other important factors include the relative ease with which we can genetically modify flies (for example to make them express fluorescent protein fusions) and the fact that embryos are transparent, which makes fluorescence microscopy experiments possible. In my thesis, we study embryos of *Drosophila* that are 1 - 4 hours old. While mechanistic studies of human development are technically and ethically challenging, *Drosophila melanogaster* offers a window into the molecular biology of cell fate and tissue shape. Even though fruit flies do not appear to have similarities to human beings in terms of size and shape at

first glance, the molecular mechanism and mode of functions of many pathways are similar in flies and humans. Also, a study has found that 77% of the disease genes in human have matching cognate genes in *Drosophila* (Reiter *et al.*, 2001). Therefore, studies in flies can help us better understand molecular pathways in humans. For example, *D. melanogaster* has recently been used in the study of nanotoxicity (Ong *et al.*, 2015), cancer (Mirzoyan *et al.*, 2019), Alzheimer's disease (Prüßing *et al.*, 2013) and anti-aging drugs (Lee and Min, 2019). In this thesis, I used *Drosophila* as a model system to study the function of TFs in general. To complement *in vivo* experiments using fruit flies, *in vitro* experiments using organic dye and pure eGFP solutions were also carried out.

Chapter 2

Methods

The morphogen concentration in a *Drosophila melanogaster* embryo, which is about 500 µm in length and 200 µm in width, is usually in the nanomolar (nM, 10^{-9} mol/L) range. Fluorescence Correlation Spectroscopy (FCS) provides an ideal solution to measure concentrations in the nM range and dynamics in range of ms $\sim$ s. While FCS can provide dynamical information only for the mobile molecules, Fluorescence Recovery After Photobleaching (FRAP) can provide information for both mobile and immobile molecules, but on different times scales and length scales. To examine the kinetic models used for FCS and FRAP, dynamical simulations can be implemented as well.

In the following sections of this chapter, first an overview of the fly husbandry is given, including the characteristics and maintenance of the *Drosophila* flies, preparations of the embryos for imaging, as well as a description of the fluorescent protein fusions expressed in the flies. Since confocal imaging is a core component for both FCS and FRAP, a small section is spent on that. Subsequently, the three types of techniques used in this thesis, i.e. FCS, FRAP and simulations, are introduced and

representative results are shown.

2.1 Fly husbandry

2.1.1 An introduction to *Drosophila melanogaster*

Drosophila melanogaster is commonly known as the fruit fly, and can be easily found near rotted fruits. Due to its small size (3 mm $\times$ 2 mm for adults), short life cycle ($\sim$ 28 days) and ease of culture, *Drosophila* has been used for more than a hundred and twenty years by researchers in the field of developmental biology, molecular biology, genetics and so on (Allen, 1975). The life cycle of *Drosophila*, same as that of butterflies, includes an egg stage (embryo), larva stage (three stages of instars), pupa stage and finally flying adult stage (Fig. 2.1). Generally, the adult female is larger in size than the male, while the male has a larger black spot on its tail than the female.

Shortly after egg laying, the embryo of *Drosophila* starts to hatch, i.e. a series of nuclear divisions takes place. During the early embryo development (2 - 3 hours after egg deposition), *Drosophila* embryos have an unusual syncytial blastoderm, meaning that the nuclei divide without division of the cell membrane for 14 nuclear cycles (NC, i.e. between NC 1 - NC 14) (Fig. 2.2) (Tram *et al.*, 2001). After 14 rounds of nuclear cycles, the membrane starts to invaginate around the nuclei which are at the periphery of the embryo and the cellular blastoderm is formed. At nuclear cycle 14, there are about 6,000 nuclei in the embryo. Fruit fly body cells have 4 pairs of chromosomes (three autosomes, and one pair of sex chromosomes), in total 8 chromosomes.

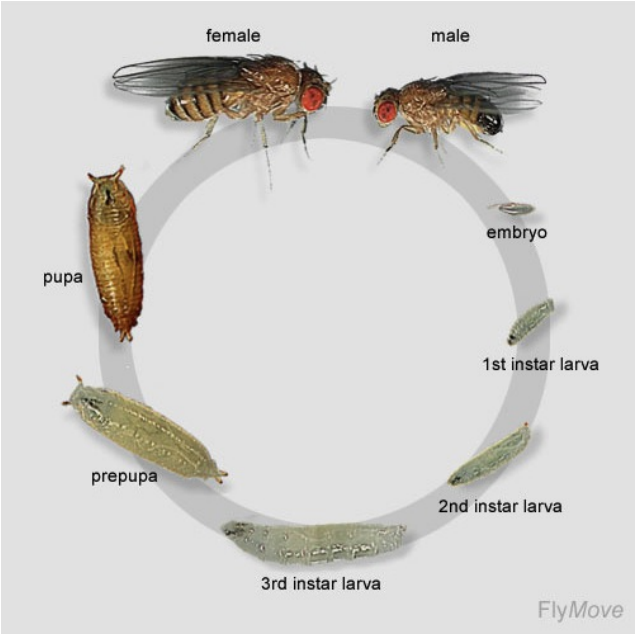

Figure 2.1: An illustration of the life cycle of a *Drosophila melanogaster*, which includes an egg stage (embryo), larva stage (three stages of instars), pupa stage and finally flying adult stage. The embryo hatches for one day, the three stages of instar take two days. After the roaming stage of the larva for two days, pupation occurs on the 6th day after egg laying. After another 5 or 6 days, the adult fly emerges from the pupa case. The life cycle of a fruit fly is about 28 days at 25°C. Reproduced with permission from (Abolaji *et al.*, 2013).

2.1.2 Stock keeping of live fly culture

The live fly stocks are cultured in plastic vials with moisturized food (Ward's, Rochester, NY, USA) at the bottom and a cotton plug on the top. These vials are stored in an incubator with temperature set at 25°C and an alternating lighting of 12 hours of darkness and 12 hours of brightness. The stocks of flies in each vial are transferred to a new vial with fresh food every 1 to 2 weeks. Though the fly stocks are relatively easy to maintain, fly keepers should watch out for mite infestation, and fungal or bacterial contamination in the culture. Useful resources on fly stock keeping can be

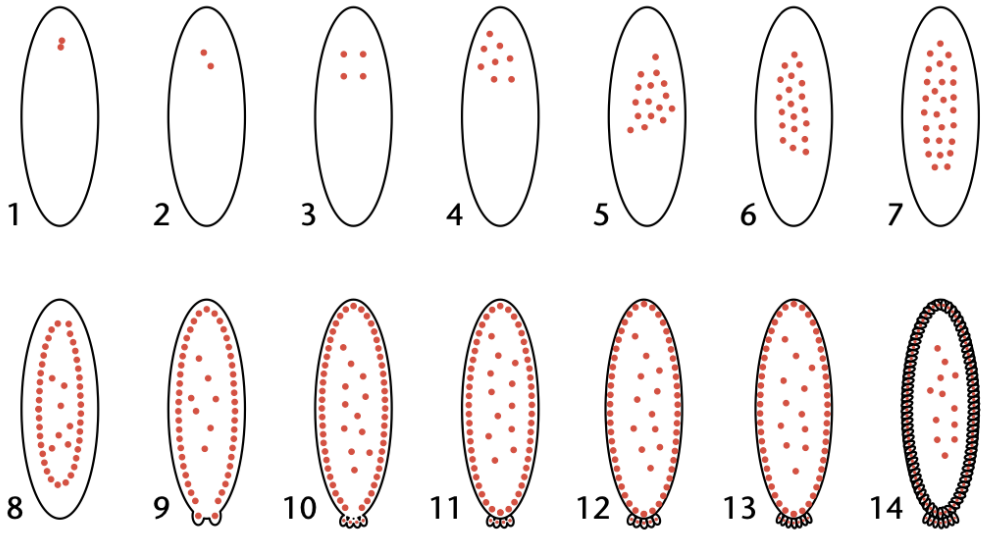

Figure 2.2: An illustration of the early nuclear division and migration during *Drosophila* embryogenesis. During the first 2 - 3 hours (i.e. NC 1 - NC 14) after egg deposition, the *Drosophila* embryo is a syncytium, meaning that only the nuclei divide while the cell membrane does not. Here the red dots represent nuclei and the oval black line represents the membrane of the developing embryo. The pole cells at the posterior side of the embryo start to appear at NC 9. Reproduced with permission from (Tram *et al.*, 2001).

found on the Bloomington Drosophila Stock Centre Web page[1].

2.1.3 Preparing embryos for imaging

To prepare embryos for experiments, either FCS or FRAP, the first step is to collect embryos. Embryos are collected by transferring a regular vial of flies to another empty vial with two open ends, where one end is stopped by a cotton plug, and the other put on an embryo collection plate (ECP) with yeast paste in the centre. After about 2 - 3 hours, the vial is removed and the embryos can be collected on the ECP with a tweezer tip by hand and placed on a double-sided tape. The second step is to remove the

[1]https://bdsc.indiana.edu/information/fly-culture.html

chorion of the embryo. This is done by gently rolling the embryo on the double-sided tape until the chorion sticks and remains on the tape. The last step is to transfer the dechorionated embryo to a 0.17 mm coverslip. The coverslip is prepared by drawing thin lines of heptane glue on it and then the embryo is transferred on one line by using the tip of the metal tweezer. Then a drop of halocarbon oil 700 (Sigma) is added to cover the embryo such that the water in the embryo won't evaporate whereas the oxygen can permeate the oil and be breathed by the embryo. In this manner, the embryo is kept alive during the subsequent experiments. A more detailed description regarding embryo preparation and related materials can be found in (Perez-Romero *et al.*, 2018).

2.1.4 Fluorescent protein fusions expressed in flies

To study morphogens *in vivo*, one needs to first identify and separate the protein of interest out of a sea of thousands of other molecules in the complex and crowded cellular environment of the living embryo. Thanks to the discovery and development of different fluorescent proteins, the protein of interest can be attached to a fluorescent protein such as the green fluorescent protein (GFP), first found in the jellyfish *Aequorea victoria*. This is done through genetic engineering. The resultant combination of the protein of interest and the fluorescent protein is called a protein fusion. Usually, the goal is to attach one fluorescent protein to one protein of interest in a 1-to-1 ratio, as is the case for the protein fusions studied in this thesis.

Fluorescence is a phenomenon where a photon is emitted as an atom or a molecule excited by the absorption of a photon of shorter wavelength goes back to its ground state. Depending on the wavelengths of excitation and emission, various types of

fluorescent proteins exist, such as red, magenta, yellow, blue, and green fluorescent proteins. Within the same color group, different types of fluorescent proteins exist as well. The two types of green fluorescent proteins used in my thesis projects are eGFP (enhanced Green Fluorescent Protein) and sfGFP (superfolder Green Fluorescent Protein).

eGFP (enhanced Green Fluorescent Protein)

The most widely used GFP is eGFP. It is a basic GFP published in 1996 (Cormack *et al.*, 1996) and derived from the GFP found in *Aequorea victoria*. eGFP is attached to Bcd and NLS (Nulcear Localization Signal, used as an *in vivo* control protein) to form Bcd-eGFP and NLS-eGFP. The *D. melanogaster* fly stocks expressing Bcd-eGFP and NLS-eGFP used in this thesis were kindly gifted to us by Dr. Eric Wieschaus (Gregor *et al.*, 2007b).

sfGFP (superfolder Green Fluorescent Protein)

sfGFP is another type of GFP, which is more recent compared to eGFP. It is a basic GFP published in 2006 (Pédelacq *et al.*, 2006) and also derived from the GFP found in *Aequorea victoria*. sfGFP is attached to Cic to form Cic-sfGFP. Our *D. melanogaster* fly stocks expressing Cic-sfGFP were kindly gifted to us by Dr. Stas Shvartsman (Patel *et al.*, 2021).

sfGFP is found to fold faster than eGFP when expressed as fusions with other proteins (Pédelacq *et al.*, 2006). This is an advantage when using it in a developing

organism where fluorescence needs to be observed early on. The fluorescence properties of eGFP and sfGFP are very similar: same excitation and emission wavelength, slightly different photostability where sfGFP is a bit more stable than eGFP under the same circumstances (Cranfill *et al.*, 2016; Scott *et al.*, 2018). A comparison of known parameters for these two proteins is listed in Table 2.1 [2] [3]. A new monomeric superfolder GFP as photostable as eGFP has been recently reported (Valbuena *et al.*, 2020).

Table 2.1: Comparison of eGFP and sfGFP (Cranfill *et al.*, 2016).

	eGFP	sfGFP
Molecular weight	26.9 kDa	26.8 kDa
Excitation wavelength	488 nm	485 nm
Emission wavelength	507 nm	510 nm
Maturation time	25 min	13.6 min
Photobleaching half time	179 s at 80 µW	208 s at 80 µW
Brightness	33.5 $mM^{-1}cm^{-1}$	54 $mM^{-1}cm^{-1}$
Quantum yield	0.6	0.65

2.2 Confocal Imaging

For both FCS and FRAP experiments, the setup is equipped with the functionality of confocal imaging, which is an indispensable step for acquiring fluorescence images

[2]Quantum yield (QY), also known as quantum efficiency (QE), i.e. the probability that an excitation of the electronic dipole of the chromophore leads to the emission of a photon.

[3]Brightness, is defined as the product of the QE and the fluorophore's molar extinction coefficient. Molar extinction coefficient (or molar absorptivity), is a measurement of how strongly a chemical species absorbs light at a given wavelength. The brightness here is different from molecular brightness appeared later.

of the sample before and after the measurements. The essence of confocal imaging is the application of a pinhole (an adjustable iris in the intermediate image plane, which is usually set as 1 Airy Unit) right before the detector to allow for emission light only from the focal plane of the sample, thus resulting in sharp images of the specimen. Fig. 2.3 shows confocal images of four different cross-sections of an embryo expressing Cic-sfGFP at NC 14.

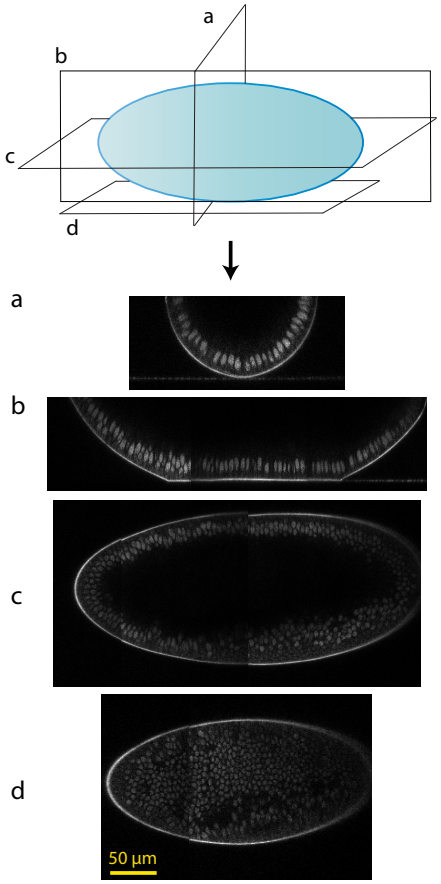

Figure 2.3: Confocal images of four different cross-sections of an embryo expressing Cic-sfGFP at NC 14. Panels **b**, **c**, and **d** contain composite images stitched together and scale bar for all four are the same, which is 5 µm.

2.3 FCS (Fluorescence Correlation Spectroscopy)

As one of the major biophysical techniques, Fluorescence Correlation Spectroscopy (FCS) has been widely applied to study molecular interactions and dynamics in the field of biology and medicine. It is minimally invasive and has high temporal and spatial resolutions. FCS was first utilized in the 1970s to measure diffusion and concentration (Elson and Magde, 1974; Magde *et al.*, 1974). It can measure concentrations of fluorescent molecules in extremely low range (a few hundred pM to a few hundred nM), and give dynamic information about the fluorescent molecules undergoing various types of motions including Brownian motion. FCS data are treated using a mathematical procedure called correlation to generate autocorrelation functions which contain dynamic and concentration information about the fluorescent molecules in the sample.

2.3.1 Principle of FCS

In an FCS experiment, an incident laser beam shines on the sample and illuminates (at its focus) a very small three dimensional volume ($\sim$ 1 fl), which is called the detection volume, V. Generally, this volume can be approximated as a 3D Gaussian. The fluorescent molecules in the detection volume are driven to one of their excited states and according to the energy minimization rule, these excited molecules first quickly relax to the lower energy excited states, then go back to the even lower energy ground state while emitting light of longer wavelength compared to that of the incident light. The fluorescence intensity, I, from the detection volume is recorded as a function of time, t, using a single photon detector. The recorded signal, $I(t)$, is then correlated with itself using a hardware correlator, which then results in the so

called autocorrelation function (ACF), $G(\tau)$, as shown in the following equation:

$$G(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I\rangle^2}. \tag{2.3.1}$$

The pointed brackets $\langle\rangle$ represent a temporal average of the quantity between them. The ACF represents the self-similarity of the fluorescence signal after a lag time of τ. Under certain conditions, the amplitude of the ACF, $G(0)$, is equal to the inverse of the number of fluorescent molecules, N, in the detection volume, V:

$$G(0) = \frac{1}{N}. \tag{2.3.2}$$

To obtain the concentration, knowledge of the volume V of the detection volume needs to be acquired. As mentioned before, the detection volume is usually modelled as a 3D Gaussian, which can be described mathematically as in the following equation:

$$W(\vec{r}) = I_0 \cdot \exp(-2\frac{x^2+y^2}{\omega_0^2} - 2\frac{z^2}{\omega_z^2}). \tag{2.3.3}$$

The above equation describes the spatial distribution of the incident laser light, where the intensity drops to $1/e^2$ at a distance ω_0 from the focal point in the lateral direction and at a distance ω_z in the axial direction. The relation between these two parameters is $\omega_z = S \cdot \omega_0$, where S is the aspect ratio ($S \approx 5-10$). The volume of detection can be calculates as (Schwille and Ries, 2011):

$$V = \pi^{\frac{3}{2}} \cdot \omega_0^2 \cdot \omega_z. \tag{2.3.4}$$

Then the concentration is $C = N/V$. An important parameter regarding dynamics is the characteristic diffusion time, τ_D, which represents the average time a fluorescent

molecule spends inside the detection volume before it moves out, and can be expressed in the following equation:

$$\tau_D = \frac{\omega_0^2}{4D}. \tag{2.3.5}$$

D is the diffusion coefficient. In an FCS experiment, the detection volume is routinely measured from calibration measurements using fluorescent dye with known diffusion coefficient D. In this thesis, calibration measurements are performed with solution of Alexa Fluor 488 (Invitrogen, now Life Technologies, Carlsbad, CA, USA) which has a known diffusion coefficient $D = 435$ µm^2/s at 22.5 °C (Petrášek and Schwille, 2008). The characteristic diffusion time, τ_D, can be obtained from fitting the ACF with an appropriate analytical model. Multiples theoretical models for analysing the ACF exist and will be listed and explained below.

2.3.2 Theoretical Models for FCS data analysis

Depending on different scenarios in the sample, i.e. whether there are one or two species of fluorescent molecules, whether taking into account the background noise and the photobleaching effect, or considering the binding and unbinding of molecules to the targets, there are four types of FCS analytical models used in this thesis. They are referred to as the One-Component Model, the Two-Component Model, the Stick-and-Diffuse Model, and the Modified Two-Component Model.

One-Component Model

The One-Component Model is the simplest model used here, which considers only one diffusing species, i.e. one population of molecules with the same molecular weight,

molecular brightness etc. Under this model, the autocorrelation function can be written as (Widengren *et al.*, 1995):

$$G_{1C}(\tau) = G(0)\frac{1 + T/(1-T)e^{-\tau/\tau_T}}{(1+\tau/\tau_D)(1+\tau/(S^2\tau_D))^{1/2}}. \tag{2.3.6}$$

The numerator represents the photophysics property of the fluorescent molecules (namely its blinking due to the existence of a dark triplet state). T is the fraction of molecules in the dark state, and τ_T is the relaxation time of the dark state. This model is useful in the case of *in vitro* experiments with organic dyes, such as Alexa Fluor 488, and with fluorescent proteins, such as purified eGFP.

Two-Component Model

If there are two species of diffusing molecules in the detection volume, the ACF is then characterized by the Two-Component Model, which is written as (Wachsmuth *et al.*, 2000):

$$G_{2C}(\tau) = G(0)\left(1 + T/(1-T)e^{-\tau/\tau_T}\right) \left[\frac{p}{(1+\tau/\tau_{\mathrm{D1}})(1+\tau/(S^2\tau_{\mathrm{D1}}))^{1/2}} + \frac{1-p}{(1+\tau/\tau_{\mathrm{D2}})(1+\tau/(S^2\tau_{\mathrm{D2}}))^{1/2}}\right]. \tag{2.3.7}$$

p is the fraction of the first diffusing species, with a diffusion characteristic time of τ_{D1}. The second species diffuse with a diffusing characteristic time of τ_{D2}. The reason why there may be two species of molecules in the sample could be that the first

species diffuse relative freely in the medium it is in without being obstructed by other structures, while the second species diffuse relatively slower, maybe because it is part of a larger molecular complex. This model is applied for *in vivo* experiments where things are generally more complicated than in solutions and we expect proteins of interest may interact with other molecules.

Stick-and-Diffuse Model

If there is a single diffusing fluorescent species in the solution, but this species can become transiently immobilized, for example through transient binding to an immobile or very slow structure, one can use the Stick-and-Diffuse Model to obtain the bound and unbound rate of the fluorescent molecule to its target. The ACF for the Stick-and-Diffuse Model for 3D diffusion can be expressed in the following equation (Yeung *et al.*, 2007; Abu-Arish *et al.*, 2009):

$$\begin{aligned} G_{SD}(\tau) = \frac{1}{N}(1 + \frac{T}{1-T}e^{-\frac{\tau}{\tau_T}})[\frac{e^{-k_{\text{off}}\tau}}{1+\frac{k_{\text{off}}}{k_{\text{on}}}} + \frac{1}{1+\frac{k_{\text{on}}}{k_{\text{off}}}}\frac{e^{-k_{\text{on}}\tau}}{(1+\tau/\tau_D)\sqrt{1+\tau/(S^2\tau_D)}} \\ + \frac{k_{\text{on}}k_{\text{off}}}{k_{\text{on}}+k_{\text{off}}}\sum_{n=1}^{\infty}\frac{1}{(n-1)!n!}\int_0^{\tau} ds \frac{e^{-k_{\text{off}}(\tau-s)-k_{\text{on}}s}}{(1+s/\tau_D)\sqrt{1+s/(S^2\tau_D)}}(2n \\ + k_{\text{off}}s + k_{\text{on}}(\tau-s))(k_{\text{on}}k_{\text{off}}s(\tau-s))^{n-1}] + C. \end{aligned} \tag{2.3.8}$$

k_{on} and k_{off} are the bound and unbound rates, respectively. It has been tested previously that n can be truncated at 7 in the Taylor series (Abu-Arish *et al.*, 2009) thus we set $n = 1, 2, ..., 7$ when fitting data with Eq. 2.3.8.

Effect of background noise and photobleaching on the ACF

When it comes to *in vivo* FCS measurements, compartments as well as the less photostable GFP utilized will render the effect of noise and photobleaching non-negligible. As illustrated in Fig. 2.4, the effect of the background noise and photobleaching on the fluorescence intensity (FI) as well as the autocorrelation function are prominent. The noise increases the average value of FI and thus decreases the amplitude of the ACF. The amplitude of the ACF is related to the amplitude of the fluctuations (δI) divided by the intensity (I), since $G(0) = \langle \delta I \rangle / \langle I^2 \rangle$. So increasing I (because of background noise) without increasing δI results in a decrease in $G(0)$. The photobleaching makes the FI gradually decrease, generally in an exponential manner, which induces a tail for the ACF in the longer lag time regime around lag times comparable to the characteristic time of the exponential decay. Detailed calculations for these modified ACFs are the object of Chapter 3, and will be developed in detail in that chapter.

- Contribution of noise

When the fluorescence signal decreases to a level that is comparable to the signal from the noise, the amplitude of the ACF is no longer in a simple inverse relationship with the number of molecules, N, but instead it is expressed as follows (Koppel, 1974; Rigler *et al.*, 1993):

$$G(0) = \frac{1}{N} \frac{1}{(1 + I_B / [I - I_B])^2} = \frac{1}{N} \frac{1}{(1 + m/N)^2}. \tag{2.3.9}$$

I_B represents the fluorescent intensity from the background noise. m is defined as

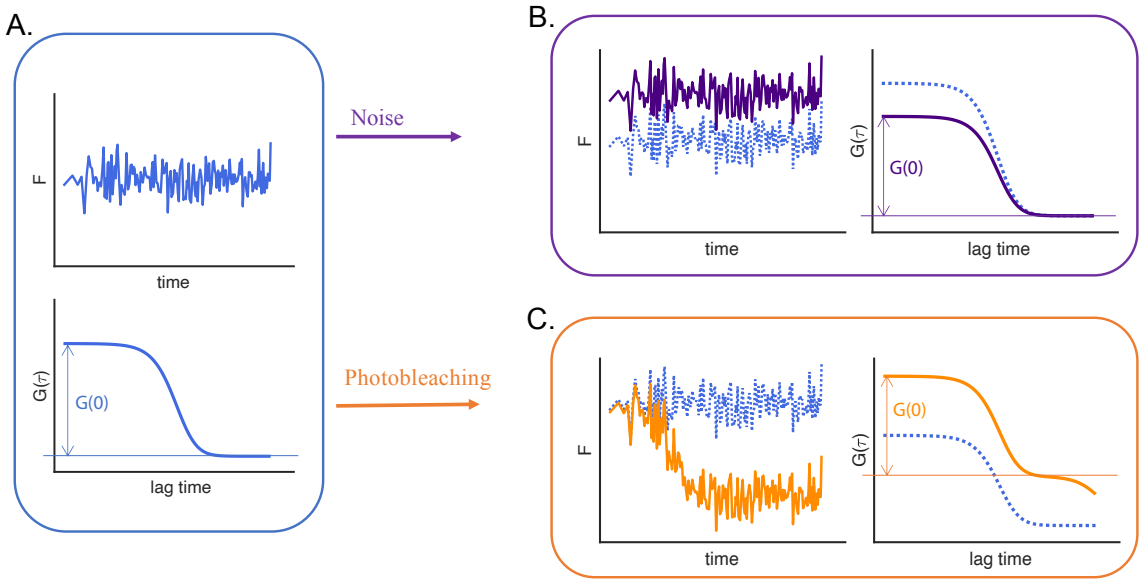

Figure 2.4: An illustration of the effects of background noise and photobleaching on the fluorescence intensity (FI), F, as well as the autocorrelation function (ACF), $G(\tau)$. (A) The original FI and ACF. (B) The effect of background noise on the FI and ACF, where the average value of FI increases and the amplitude of the ACF decreases. (C) The effect of photobleaching on the FI and ACF, where the FI gradually decreases roughly following an exponential decay, which induces a tail for the ACF in the longer lag time regime.

$m = I_B/(\gamma B)$ where B is the molecular brightness (usually in unit of kHz, and it describes the number of photons emitted from the fluorescent molecules per unit time) and γ a geometrical constant that has a value of $2^{3/2}$. The ratio N/m is a measure of the signal-to-noise.

- *Contribution of photobleaching*

The ACF is given by $G(\tau) = G_D(\tau) + G_P(\tau)$, where $G_D(\tau)$ is from Eq. 2.3.6, 2.3.7

or 2.3.8 (with $G(0)$ given by 2.3.9) and $G_P(\tau)$ is as follows (Zhang *et al.*, 2021):

$$G_P(\tau) = \left(\frac{t_M - \tau}{2\tau_P}\coth\left[\frac{t_M - \tau}{2\tau_P}\right] - 1\right) \Bigg/ \left(1 + \frac{m}{N_0}\frac{t_M - \tau}{\tau_P}\frac{1 + e^{\tau/\tau_P}}{1 - e^{-\frac{t_M - \tau}{\tau_P}}} + \left(\frac{m}{N_0}\right)^2\left(\frac{t_M - \tau}{\tau_P}\right)^2\frac{e^{\tau/\tau_P}}{1 - e^{-\frac{t_M - \tau}{\tau_P}}}\right). \quad (2.3.10)$$

τ_P is the photobleaching characteristic time and t_M the FCS measurement time (5 - 60 s for each measurement). N_0 represents the number of fluorescent molecules in the detection volume at $t = 0$.

Modified Two-Component Model

The Modified Two-Component Model was recently derived and reported by us (Zhang *et al.*, 2021) as will be described in Chapter 3. The highlight of this model is that it takes into account both the effects of noise and photobleaching explicitly, i.e. $G_{\mathrm{M2C}}(\tau) = G_{\mathrm{2D}}(\tau) + G_P(\tau)$ with $G(0)$ given by 2.3.9.

2.3.3 Experimental setup and procedures

Generally, an FCS setup can be conveniently converted from an inverted microscope. An illustration of a typical FCS setup is shown in Fig. 2.5, where a beam of incident laser light (represented in blue and entering from the right here), hits the dichroic mirror, then is deflected into the back aperture of a water-immersion objective, then passes through the coverslip and reaches the sample, which is generally in the state of an aqueous solution. The fluorescent molecules in the sample are excited to a higher energy state. The emitted fluorescence light (represented in green), enters through the objective and reaches the dichroic mirror. Due to its property, this same

dichroic mirror will allow the transmission of the emitted light. The emitted light passes through an emission filter and a pinhole (30 - 100 µm) to finally reach the APD (avalanche photodiode) detector, which is connected to a computer. The correlation of the fluorescence signal is realized by a hardware correlator. There are two types of sample holder for this setup, one of which can hold a 96-well plate, and the other can hold a single regular coverslip of 24 mm × 17 mm. Note that due to the settings of the objective and its collar, it is required to use a coverslip with a thickness of 0.17 mm on this setup (Insight, Evotec Technologies, Hamburg, Germany, now PerkinElmer, Waltham, MA, USA).

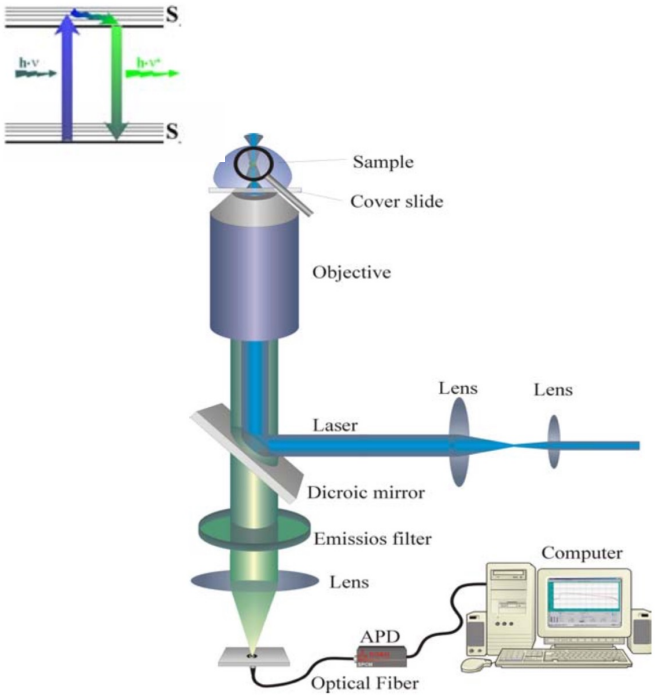

Figure 2.5: An illustration of a typical FCS setup. The excitation light is represented in blue. The emitted light from the sample is represented in green. See text for more detail. Reproduced with permission from (Schwille and Ries, 2011).

To perform an FCS experiment, one first turns on the laser with the required

wavelength (in the case of both *in vivo* and *in vitro* experiments, a 488 nm continuous wave laser (Sapphire 488-20/460-10, Coherent, Santa Clara, CA, USA)), and let it warm up for about half an hour (for the laser stability test, see Appendix). One then prepares the sample either on an individual coverslip if the sample is an embryo, or in a 96-well plate if the sample is aqueous solution of dye, then use the corresponding sample holder to mount the sample after putting a drop of deionized water (18 MΩ) on top of the water immersion objective (UApoN, ×40, 1.15 NA, Olympus Canada, Richmond Hill, ON, Canada). The proper excitation and emission filters are then selected and the required laser power is set. The sample can be viewed either using wide field imaging with a white light or using confocal imaging with the excitation laser. For this FCS setup, the confocal fluorescence imaging is viewed through the accompanying software (MIPSS (Evotec)). After adjusting the position of the sample to an optimal imaging condition where it is at its highest resolution as well as ideal field of view, one can turn on the detector and start FCS measurements by navigating the measurement software. Basically, a point that will be measured will be selected and the measurement time, generally 5 - 60 s, is set, then a series of single-point FCS measurements are acquired. The resulting files are automatically saved in a local folder. For the analysis of simple data from *in vitro* experiments, I used the FCS analysis software on the One- or Two-Component Models. For more complicated data from *in vivo* experiments, I wrote a program in Matlab to analyse FCS data using either the Modified Two Component Model or the Stick-and-Diffuse Model.

2.3.4 Examples of FCS results

As an example of an FCS measurement, the autocorrelation function (ACF), $G(\tau)$, measured from a 40 nM solution of Alexa Fluor 488 at a laser power of 75 µW is shown in Fig. 2.6. The blue line represents the original ACF data, and the magenta line represents the fit with the One-Component Model. It can be seen that the One-Component Model fits the data quite well and the residuals between the original and fitted data in the lower panel are only on the order of 10^{-3}. Fig. 2.6 shows an ACF obtained for Alexa Fluor 488, where the background noise and photbleaching effects are negligible due to the photostability of the dye as well as its relatively high concentration in solution.

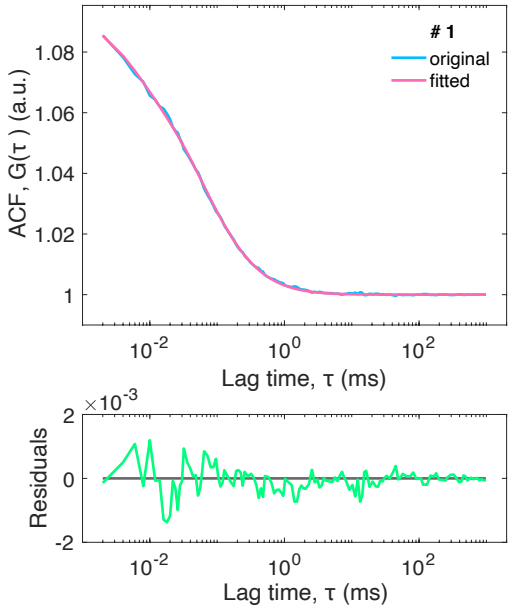

Figure 2.6: An example of an autocorrelation function obtained as the result of an FCS experiment. The autocorrelation function (ACF), $G(\tau)$, is the first one of a serial dilutions of Alexa Fluor 488, with concentration of about 40 nM and at laser power of 75 µW. The original ACF data (blue) can be well fitted with the One-Component Model (magenta) with small residuals (green) on the order of 10^{-3}.

2.4 FRAP (Fluorescence Recovery After Photobleaching)

FRAP is another powerful fluorescence technique introduced in the 1970s, which has been widely used in cell biology and related field to study protein diffusion and binding (Wachsmuth, 2014). FRAP experiments can be readily performed on a commercial confocal laser scanning microscope (developed in the 1980s). Initially, FRAP was not used as widely, due to the difficulties associated with purifying and labelling proteins and then injecting them into cells. Thanks to the revolutionary development of fluorescent protein technology, FRAP saw a tremendous boost in its applications in cell biological research (Houtsmuller, 2005). Nowadays, FRAP is often used in conjunction with FCS for cross-validation because they provide complementary information, as FRAP cannot resolve very fast motions, because its temporal resolution is limited by the duration of the photobleaching step while FCS cannot detect very slow or immobile particles, because these particles get photobleached. Computer modelling of FRAP experiments can also help analyse complex FRAP data (as done here in this thesis).

Photobleaching is a phenomenon where fluorescent molecules lose their capacity to fluoresce after a certain number of fluorescence cycles when exposed to an excitation light with a certain intensity. Generally, the higher the intensity of the excitation light, the stronger the photobleaching effect (meaning the shorter the time necessary to photobleach the molecule). It also depends on the photostability of the fluorescent molecules, as well as on pH, temperature, and buffer composition (Kremers *et al.*,

2007). In general, photobleaching is an unwanted side effect in microscopy experiments as it hampers time-lapse monitoring or imaging weakly fluorescent samples. However, FRAP makes use of this phenomenon to expose the dynamical behaviours of fluorescent molecules. In the following sections, the principle of FRAP will be explained, then two analytical models for FRAP will be introduced, and lastly an example of FRAP experiment result will be shown.

2.4.1 Principle of FRAP

The principle of FRAP is relatively easier than that of FCS as a FRAP experiment does not involve any complicated mathematical procedure such as the correlation of the signal. In essence, during a FRAP experiment, a laser beam of high power is focused on the region of the sample to be studied, and as a result the fluorescent molecules in the sample in that region are irreversibly photobleached, meaning that they become permanently non-fluorescent. Usually exposure to the photobleaching laser of hight intensity is very short (a few seconds) and the sample is then repeatedly imaged using a much reduced laser intensity for a longer period of time. Due to diffusion, all of the fluorescent (bright) and non-fluorescent (dark) molecules will move around and gradually, the bleached area which is dark right after the bleaching will gradually turn bright again (Fig. 2.7). Whether its brightness is the same as the unbleached region or is the same as the level from pre-bleach, depends on the properties of the sample, and specifically depends on the dynamics and motions of the fluorescent molecules in the sample.

Depending on the size of the photobleached area with respect to the overall size of the cell, the length of the photobleach step and the size of the region monitored

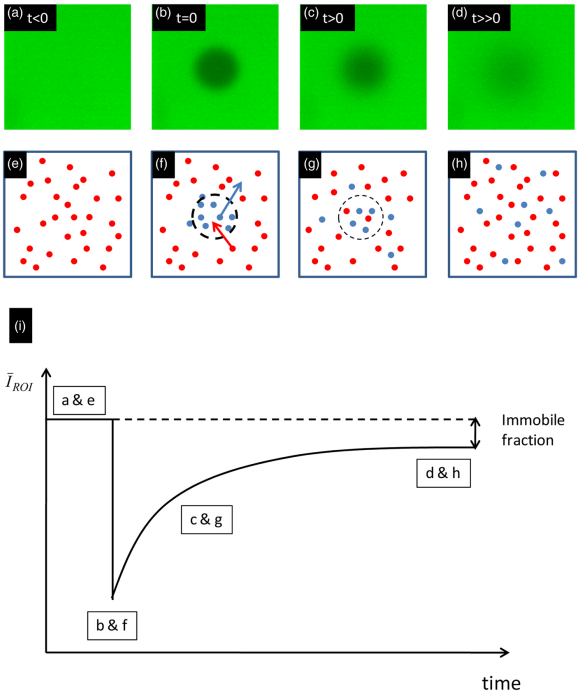

Figure 2.7: An illustration of a FRAP experiment, where the top row shows the fluorescent microscopy images of a sample and the middle row the schematics of the fluorescent (red) and non fluorescent (blue) particles in the field of view of the top images. The bottom panel shows the fluorescence intensity in the region of interest (ROI) represented by a circle in the above images. If the pool of molecules is not a limiting factor, the unrecovered fraction, i.e. the difference between the initial and the final plateau level of intensity represents the immobile fraction of the total fluorescent molecules in the sample. Reproduced with permission from (Lorén *et al.*, 2015).

after photobleaching, various types of FRAP exist, which share essentially the same principle (Houtsmuller, 2005).

2.4.2 Theoretical Models for FRAP data analysis

The theoretical models for FRAP experiments are derived from partial differential equations of motion with proper initial and boundary conditions (Sprague *et al.*, 2004; Mueller *et al.*, 2008). In deriving our analytical models for FRAP experiments,

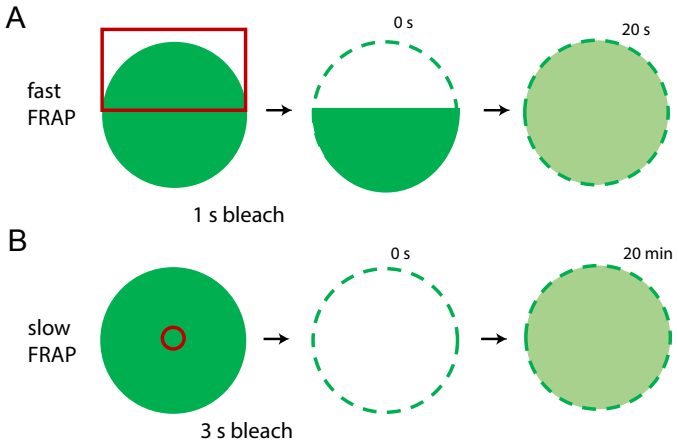

Figure 2.8: Illustrations of the two types of FRAP experiments performed in this thesis. (A) Fast FRAP, where a square region covering half of the nucleus is set as the bleaching region and photobleached for 1 s, and then monitored afterwards with a time interval of 1 s for a total duration of 20 s. (B) Slow FRAP, where a small circle at the center of the nucleus is set as the bleaching area and photobleached for 3 s, and then monitored with a time interval of 30 s for a total duration of 20 min.

what is taken in account are 1) diffusion of 1 species, and 2) nucleo-cytoplasmic transport.

There are two types of FRAP experiments that have been performed in this thesis (Fig. 2.8). One is termed fast FRAP where a square region covering half of the nucleus is set as the bleaching region and photobleached for 1 s, and then monitored afterwards with a time interval of 1 s for a total duration of 20 s. From fast FRAP experiments, information regarding intranuclear dynamics and presence of an immobile nuclear fraction can be obtained. The other is termed slow FRAP where a small circle at the center of the nucleus is set as the bleaching area and photobleached for 3 s, and then monitored with a time interval of 30 s for a total duration of 20 min. From slow FRAP, information regarding internuclear dynamics (i.e. kinetics of nucleo-cytoplasmic transport) can be extracted. The derivation of fitting models for fast and slow FRAP will be explained below.

Fast FRAP

It is assumed that there are two processes happening in the system after fast photobleaching of fluorescent proteins in half a nucleus, that is the exchange of molecules between the two halves of the nucleus (due for example to simple diffusion or to a mix of diffusion and binding, unbinding), and the exchange between nucleus and cytoplasm (due to nucleo-cytoplasmic transport). We assume that both processes give rise to an exponential recovery with different time scales (τ_f for the redistribution of fluorescent proteins within the nucleus, and τ_s for nucleo-cytoplasmic transport). Thus we have the following two relationships for the evolution of the fluorescence intensity for the bleached (I_b) and unbleached (I_u) halves of the nucleus:

$$I_b(t) = I_{\text{eq}} - a_b e^{-k_f t} - b e^{-k_s t}, \tag{2.4.1}$$

$$I_u(t) = I_{\text{eq}} + a_u e^{-k_f t} - b e^{-k_s t}. \tag{2.4.2}$$

Where all the variables ($I_{\text{eq}}, a_b, a_u, b, k_f$ and k_s) are positive, and $k_f = 1/\tau_f$, $k_s = 1/\tau_s$. In the case of Fast FRAP, $t \ll 1/k_s = \tau_s$, $e^{-k_s t} = 1 - k_s t$ (meaning that the experiment is stopped before significant exchange of molecules between the nucleus and cytoplasm takes place), so these two equations (Eqs. 2.4.1, 2.4.2) can be simplified to:

$$I_b(t) = (I_{\text{eq}} - b) - a_b e^{-k_f t} + b k_s t, \tag{2.4.3}$$

$$I_u(t) = (I_{\text{eq}} - b) + a_u e^{-k_f t} + b k_s t. \tag{2.4.4}$$

Since one cannot separate I_{eq} from b, and b from k_s, the equations can be written

as:

$$I_b(t) = I^*_{\text{eq}} - a_b e^{-k_f t} + b^* t, \tag{2.4.5}$$

$$I_u(t) = I^*_{\text{eq}} + a_u e^{-k_f t} + b^* t. \tag{2.4.6}$$

I^*_{eq} and b^* are constants.

Note that this model assumes that all the fluorescent proteins inside the nucleus are mobile, therefore at equilibrium $I_b(t) = I_u(t)$, and if exactly half the nucleus is bleached then $a_u = a_b$. If there is an immobile fraction, then these equations become:

$$I_b(t) = I^*_{\text{eq}} - a_b e^{-k_f t} - I_{\text{im}} + b^* t \ \text{(changed)}, \tag{2.4.7}$$

$$I_u(t) = I^*_{\text{eq}} + a_u e^{-k_f t} + b^* t \ \text{(unchanged)}. \tag{2.4.8}$$

The fraction of immobile fraction is then $p_{\text{im}} = I_m/a_b$, which can be calculated from I_u, the mean intensity of the unbleached half of nucleus, I_b, the mean intensity of the bleached half nucleus, and I_0, the mean intensity of a control nucleus that is not being bleached, when all three parameters reach their equilibrium values, as shown in Eq. 2.4.9.

$$p_{\text{im}} = (I_u - I_b)/I_0 = (c_u^{eq} - c_b^{eq})/c_0^{eq}. \tag{2.4.9}$$

Slow FRAP

In the case of Slow FRAP, $t \gg 1/k_f = \tau_f$ (meaning that the photobleaching step is longer than τ_f, so at any time the intensity across the nucleus is uniform), $e^{-k_f t} = 0$,

so these equations (Eqs. 2.4.1, 2.4.2) can be simplified to:

$$I_b(t) = I_u(t) = I_0 - I_{\text{im}} - be^{-k_s t} \tag{2.4.10}$$

$I(t)$ is the overall intensity of the whole nucleus. I_0, similarly as before, is the average fluorescence intensity from the faraway unbleached nucleus,which is the same for both halves of the nucleus (uniform intensity). I_{im} is the fluorescence intensity from the immobile fluorescent molecules in the nucleus and I_m from the mobile ones. The fluorescence intensity from the background, i.e. from the cytoplasm is represented by I_{bkg}. As a result, $I_0 = I_{\text{im}} + I_m + I_{\text{bkg}}$. For the slow FRAP we also need to include a term due to continuous photobleaching occurring while imaging since imaging is done over a long period of time. The continuous photobleaching can be modelled by a exponential decay of the overall fluorescence intensity:

$$I(t) = [I_0 - I_{\text{im}} - be^{-k_s t}] \times e^{-k_P t}. \tag{2.4.11}$$

The recovery percentage is calculated as follows:

$$\text{Recover}\% = \frac{I_0 - I_{\text{im}} - I_{\text{bkg}}}{I_0 - I_{\text{bkg}}} \times 100\%. \tag{2.4.12}$$

When $I_{\text{im}} = 0$, Recover% = 100%, and when $I_{\text{im}} = I_0 - I_{\text{bkg}}$, Recover% = 0%.

2.4.3 Experimental setup and procedures

The FRAP experiments can be carried out on confocal microscope that has the FRAP functionality. Same as for FCS experiment, the first thing to do is to turn on the

laser and let it warm up for about half an hour. Also, depending on the sample type, there are two types of sample holders to mount the sample to the instrument. In the Nikon Eclipe Ti microscope that was used in this thesis, the laser wavelength (488 nm) and power (25 - 30 µW), as well as the excitation and emission filters are selected on the NI Elements software. After mounting the sample with appropriate sample holder (note that the thickness of the coverslip is 0.17 mm as well), one then chooses the FRAP option in the software, in my case it is called ND stimulation, where one can set the intensity of the bleaching power and duration, and the time interval to record an image under lower power after photobleaching. Depending on how long one wants to monitor the sample after photobleaching, a FRAP experiment can take 30 s to 30 min. Depending on the software, one may need to manually save the FRAP results files. Commercial FRAP analysis software are available, here I used homebuilt program written in Matlab to analyse the FRAP data.

Preliminary data processing

The original output data from a FRAP experiment are usually a series of confocal images with the same field of view containing the photobleached region. To extract the fluorescence intensity from both the bleached and unbleached regions, one can make use of image analysis software, such as Ilastik (Berg *et al.*, 2019) and Fiji (Schindelin *et al.*, 2012). For the slow FRAP data, we used Ilastik, which utilizes machine learning algorithm to extract objects of certain features, to first obtain segmentation of nuclei, which then be used to obtain average intensity of each nucleus. For fast FRAP, we used Fiji to extract fluorescence intensity from both the bleached and unbleached halves of the nucleus.

2.4.4 Examples of FRAP results

Fig. 2.9 shows an example of a FRAP experiment result. It was performed on a *D. melanogaster* embryo expressing NLS-eGFP at NC14 where a whole nucleus was bleached and its fluorescent intensity was monitored for 20 min with an interval of 30 s. Three images from a total of 45 images in this FRAP experiment are shown in the first row, i.e the image before (A), right after (B), and 20 min after (C) the photobleach. Average intensity of each nucleus is obtained through a segmentation mask (D) with the help of the machine learning program Ilastik. The recovery curves (E) are obtained, subsequently, where blue symbols represent those for the bleached nucleus. The green symbols represent the average fluorescence intensity from the nearby nuclei closest to the bleached one while the orange symbols represent the average intensity from the faraway nuclei. Grey symbols represent the average intensity from the cytoplasm serving as a background reference. The blue solid line is a one component fit to the data. The fitted recovery time as well as the recovery percentage are indicated in the figure.

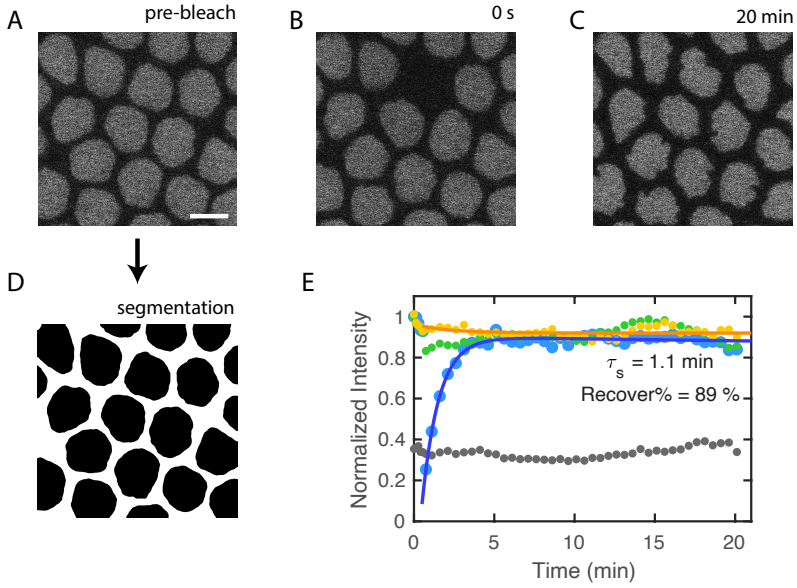

Figure 2.9: An example of a FRAP result. The FRAP experiment was performed on a *D. melanogaster* embryo expressing NLS-eGFP at NC14 where a whole nucleus was bleached and its fluorescent intensity was monitored for 20 min with an interval of 30 s. Three images from a total of 45 images in this FRAP experiment are shown in the first row,i.e the image before (A), right after (B), and 20 min after (C) the photobleach. Scale bar is 5 µm. Average intensity of each nucleus is obtained through a segmentation mask (D) with the help of the machine learning program Ilastik. The recovery curves (E) are obtained, subsequently, where blue symbols represent that for the bleached nucleus. The green symbols represent the average fluorescence intensity from the nearby nuclei closest to the bleached one while the orange symbols represent the average intensity from the faraway nuclei. Grey symbols represent the average intensity from the cytoplasm serving as a background reference. The blue solid line is a one component fit to the data.

2.5 Simulations

Despite having two complementary experimental methods to probe the dynamics of the morphogens in the samples, due to the potentially complicated nature of the motions of the morphogens, and the limitations of the dynamic models for those two

techniques, we decided to further test different dynamic models using Monte Carlo simulations. The core idea of the Monte Carlo simulation here is random walks, i.e. the particles in the simulation box perform random walks with steps drawn from a distribution, usually either a delta distribution or a Gaussian distribution, leading to same size steps or varying size steps, respectively. In the following section, the essential steps of the simulation of motion as well as the subsequent simulation of the acquisition of a confocal imaging will be discussed and examples of results in each case will be shown.

2.5.1 Simulation of protein diffusion

As we are interested in the dynamics of morphogens in the nuclei as well as the nucleocytoplasmic transport, we chose to use a cubic simulation box of size 8 µm containing a sphere with radius of 3 µm in its centre representing the nucleus. In this geometric space a number of particles can then undergo motions following various scenarios. For example, to simulate the fluorescence recovery after a whole nucleus photobleach (slow FRAP), a certain amount of particles (a few thousands to tens of thousands) are put in the cubic box, excluding the inside of the sphere. Once the simulation starts, the particles can move based on Brownian motion by adding a step taken in a random direction to each of its current coordinates, i.e. x, y, z. If this step leads the particle to cross the nuclear envelop (NE), the move is allowed to proceed only with a certain probability, representing the semi-permeability property of the nuclear membrane. This probability is directly related to the import rate of morphogen from the cytoplasm to the nucleus, k_{in}. Vice versa, when a particle inside the nucleus tries to cross the NE to go to the cytoplasm, it is allowed to proceed with

a different (lower) probability, which is related to the export rate of the morphogen from the nucleus to cytoplasm, k_{out}. Note that the steps in each direction were drawn from a Gaussian distribution with mean at 0 and variance $\sqrt{2D\delta t}$ where D is the diffusion coefficient and δt is the step time. An example of the particles in the simulation box at the start and end of a simulation is shown in Fig. 2.10. At the start, the particles are placed in the box excluding the sphere, and in the end, the particles can be observed to accumulate in the sphere, i.e the concentration of the particles in the nucleus is much higher than that in the cytoplasm, due to the effect of nucleo-cytoplasmic transport.

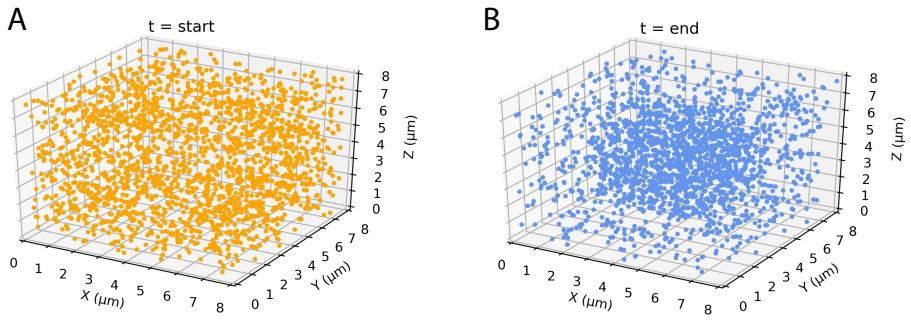

Figure 2.10: An example of the particles in the simulation box at the start (A) and end (B) of a simulation. At the start, the sphere of radius 3 µm is devoid of particles and particles are in the region outside the sphere. At the end of the simulation, i.e. after 60 s of simulation time with step time of $\delta t = 1$ ms, the particles accumulate in the sphere due to a higher nuclear import rate, in this case, $k_{in}/k_{out} = 4$, $N = 2000$, $D = 0.1$ µm^2/s.

The example shown above is only simple diffusion of one population of particles. We've also simulated one population of particles undergoing binding and unbinding, as in the Stick-and-Diffuse Model. We could build on that and simulate two populations with binding and unbinding, which represents closer to what actually happens for the morphogens in the nucleus.

2.5.2 Simulation of confocal imaging

A highlight of this simulation is that we also perform the simulation of confocal imaging in the hope that once the dynamic models are close to what happens in the real experiments, the simulated confocal image would be very similar to that obtained in real experiments. Remember that we have a box of moving fluorescent particles, and in a scanning confocal imaging measurement, the laser focus scans the field of view point by point starting by a row along the x direction and then moving to the next row in y direction and repeating the scanning point by point then line by line. The laser focus stays at each point, i.e. pixel, for a small duration of time, called the pixel dwell time, usually in the range of µs $\sim$ ms. During that pixel dwell time, $\delta\tau$, the detector collects photons emitted from the small volume being illuminated by the laser focus, which leads to a countrate in photons/pixel. As mentioned earlier, the volume that the laser focus illuminates is usually modelled as a 3D Gaussian, as expressed in Eq. 2.3.3. Note that the step time δt in the simulation and pixel dwell time $\delta\tau$ might be different, and in order to obtain accurate simulated image we need $\delta\tau \geq \delta t$ (Rose *et al.*, 2021). The simulation of protein motions is described in the previous section. In the example shown below Fig. 2.11, $\delta\tau = \delta t = 1$ ms, and simulated confocal images of the simulated system, obtained at the beginning (A) and at the end (B) of the simulation where the particles are diffusing. The number of pixels are 40 along each coordinate and pixel size is 0.2 µm/pixel. An example Python code for simulation of protein diffusion and confocal imaging can be found in the Appendix for simulation.

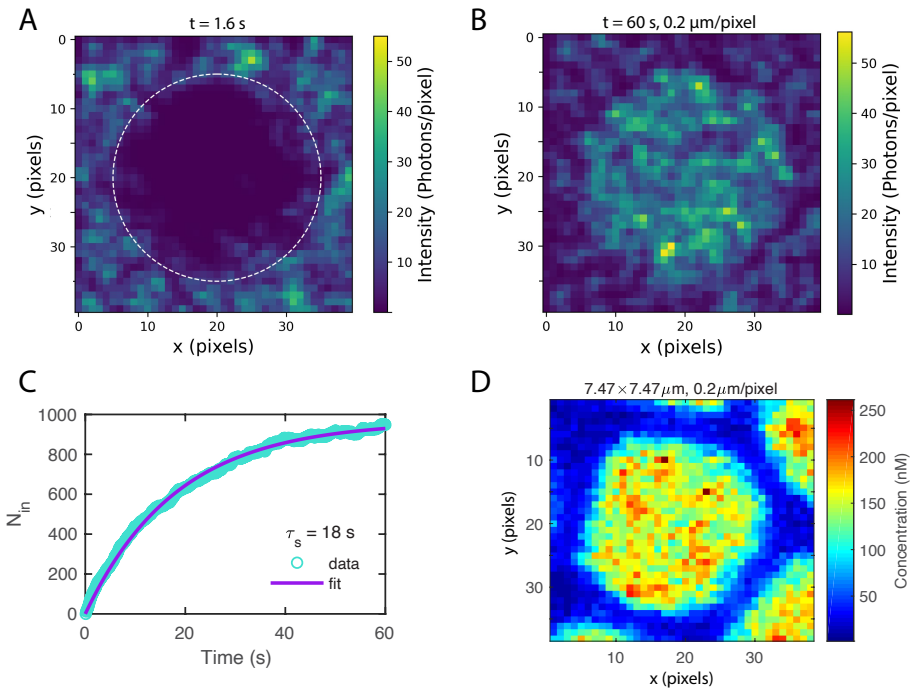

Figure 2.11: Examples of simulated confocal images of the simulated system, obtained at the beginning (A) and at the end (B) of the simulation where the particles are diffusing. (C) Time evolution of the number of particles inside the nucleus from the total 60 s of simulation is fitted with an exponential and the recovery time can be obtained. The number of pixel is 40 along each coordinate and pixel size is 0.2 µm/pixel. (D) A real experimental confocal image of a nucleus in a *Drosophila* embryo expressing Cic-sfGFP is also shown as a comparison.

Chapter 3

Paper 1: Accurate concentration measurements by FCS

The following manuscript in this chapter is a published article in Biophysical Journal:

Lili Zhang, Carmina Perez-Romero, Nathalie Dostatni, Cécile Fradin, Using FCS to accurately measure protein concentration in the presence of noise and photobleaching, Biophysical Journal, Volume 120, Issue 19, 2021, Pages 4230-4241, ISSN 0006-3495, https://doi.org/10.1016/j.bpj.2021.06.035.

Author contribution:

L.Z. designed the research, performed and analyzed all the experiments, and wrote the manuscript. C.P.-R. designed the research and performed preliminary experiments. N.D. designed the research. C.F. designed the research, performed analytical calculations, and wrote the manuscript.

Research Background:

FCS is in principle a great technique to measure fluorescent protein concentrations *in vivo*. However, often in living systems low protein concentrations ($\sim$ nM) accompanied by a high autofluorescence background result in a very low signal-to-noise ratio. In addition, the confinement of proteins in the small volume of cells or cellular compartments causes a regular decrease of the fluorescent protein population due to photobleaching. Both these phenomena, if not corrected for, lead to inaccurate concentration measurements by FCS.

Research Purpose:

The first project I set out to do was to establish how to precisely measure absolute concentration using FCS, taking into account the effects of background noise and photobleaching, so that we could precisely measure morphogen concentration in fly embryos. The method we came up with works by repeating FCS measurements for different fluorophore concentrations (something which is easy to do in samples prone to photobleaching) in order to retrieve two crucial parameters, molecular brightness and background noise. These two parameters then allow to turn confocal images into concentration maps. We first tested our method with *in vitro* experiments to measure the concentrations of dye and eGFP in solutions. We then applied this method *in vivo* to measure the concentrations of morphogens in live fruit fly embryos.

Methods: Fluorescence Correlation Spectroscopy (FCS) and Fluorescence Imaging.

Highlights of this project:

- We were able to measure fluorophore concentrations down to 1 pM (10^{-12} mol/L)

by taking into account and correcting for the effect of background noise.

- We derived an analytical expression for the ACF that can be directly applied to fit ACFs that have been affected by photobleaching.

- Using the simple relationship between average fluorescence intensity, background noise, molecular brightness and ACF amplitude, we proved that when fluorescent protein concentration is varied (either through dilution or continuous photobleaching), the molecular brightness of the fluorescent proteins and background noise can be obtained *in situ* and utilized to convert a fluorescence intensity image into a concentration map.

Significance: Fluorescent protein concentration, either *in vivo* or *in vitro*, can be accurately measured in the presence of noise and photobleaching using the method that has been demonstrated in this study.

Biophysical Journal
Article

Using FCS to accurately measure protein concentration in the presence of noise and photobleaching

Lili Zhang,[1] Carmina Perez-Romero,[1,2,3] Nathalie Dostatni,[2,3] and Cécile Fradin[1,4,*]
[1]Department of Physics and Astronomy, McMaster University, Hamilton, Ontario, Canada; [2]Institut Curie, PSL University, CNRS, Paris, France; [3]Nuclear Dynamics, Sorbonne University, Paris, France; and [4]Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada

ABSTRACT Quantitative cell biology requires precise and accurate concentration measurements, resolved both in space and time. Fluorescence correlation spectroscopy (FCS) has been held as a promising technique to perform such measurements because the fluorescence fluctuations it relies on are directly dependent on the absolute number of fluorophores in the detection volume. However, the most interesting applications are in cells, where autofluorescence and confinement result in strong background noise and important levels of photobleaching. Both noise and photobleaching introduce systematic bias in FCS concentration measurements and need to be corrected for. Here, we propose to make use of the photobleaching inevitably occurring in confined environments to perform series of FCS measurements at different fluorophore concentration, which we show allows a precise in situ measurement of both background noise and molecular brightness. Such a measurement can then be used as a calibration to transform confocal intensity images into concentration maps. The power of this approach is first illustrated with in vitro measurements using different dye solutions, then its applicability for in vivo measurements is demonstrated in *Drosophila* embryos for a model nuclear protein and for two morphogens, Bicoid and Capicua.

SIGNIFICANCE Many questions in cellular biology and biophysics would benefit from accurate measurements of protein concentration in vivo. For example, understanding how morphogen gradients are translated into target genes expression maps in developing embryos will necessitate determining absolute morphogen concentrations thatvary in space and time. Here, we propose a way to exploit fluorescence correlation spectroscopy data in samples prone to photobleaching to retrieve two crucial parameters, molecular brightness and background noise, which then allows one to turn confocal images into concentration maps. We demonstrate the efficacy of this method in fly embryos for three different nuclear proteins and suggest that it should be widely applicable to other types of eukaryotic systems.

INTRODUCTION

Many questions in cellular biophysics would benefit from accurate measurements of protein concentrations in live organisms (1,2), for example, understanding how morphogen concentration gradients are translated into expression domains of target genes necessitates determining morphogen concentration as it varies in space and time in developing embryos and tissues (3–10). In principle, fluorescence correlation spectroscopy (FCS), which allows measuring absolute concentrations of fluorescently tagged proteins noninvasively, should provide an ideal strategy to tackle this challenge. Single-point FCS is based on the quantification of the fluctuations in the fluorescence signal coming from a small confocal observation volume through the use of correlation functions (11,12). The signal is directly related to the Poisson-distributed number of observed mobile fluorophores, and knowledge of both its mean and standard deviation allows calculating absolute fluorophore concentration (13). This approach works very well in simple systems such as buffer solutions, but a number of issues arise when working with living systems.

Submitted April 5, 2021, and accepted for publication June 28, 2021.
*Correspondence: fradin@physics.mcmaster.ca
Editor: Stanislav Shvartsman.
https://doi.org/10.1016/j.bpj.2021.06.035

First, low biomolecule concentrations and high autofluorescence backgrounds often result in low signal/noise ratios in cells. This affects the ratio between the mean and the standard deviation of the signal and leads to a concentration overestimate. Correcting for this effect requires a precise measurement of the background fluorescence noise (14,15). Second, photobleaching of fluorescent proteins confined in the small volume of cells or cellular compartments causes a regular decrease of the protein population and fluorescent signal over time (16,17). This long-term photobleaching is an especially vexing issue because it not only leads to a concentration underestimate because of fluorophore depletion but also to the emergence of an additional timescale for the fluorescence fluctuations, which makes the interpretation of the FCS data more difficult. Photobleaching can be minimized by lowering excitation intensity; however, this comes at the cost of lowering the signal/noise ratio. Alternatively, the effect of long-term photobleaching in FCS data can be avoided by discarding or ignoring affected parts of the data (18–20) by considering only short time windows when correlating the signal (16,21–24) or by correcting the slow fluorescence decay with the help of an analytical function before correlating the signal (24,25). All of these methods, however, have drawbacks, and none of them allows analyzing uncorrected correlation functions, which is often the only type of data returned to the user by commercial FCS instruments.

Another important consideration is the heterogeneous and dynamic nature of living systems. Single-point FCS can give information about a few selected areas in the sample but cannot provide high throughput concentration data in space and time. FCS also fails to return proper concentrations if the environment is complex (e.g., because of the presence of membranes (26)) or if fluorophores form complexes or are immobile. One often-used workaround for these issues is to use single-point FCS to perform a calibration measurement in conditions in which the correct concentration can be recovered to obtain the molecular brightness B of a single fluorophore. Once B is known, confocal images acquired in the exact same conditions as the FCS measurement can be transformed into fluorophore concentration maps. This strategy has been used in a number of cases, for example, to measure the concentration of signaling proteins in *Escherichia coli* (27), histones in Hela cells (28), nuclear import factors at the nuclear pore complex (29), or morphogens in *Drosophila* embryos (5,18). A protocol detailing how to obtain FCS-calibrated concentration maps, including corrections for background and photobleaching, was recently published (30). A potential issue with this method, however, is that the molecular brightness of the fluorophore often has to be determined outside of the imaged area (e.g., buffer solution, cell with different expression levels, and different part of the cell or of the embryo) when it is known that B can vary a lot with environment (pH, buffer composition, and temperature (31,32)).

Building on these different ideas, we propose an original and direct way to obtain both fluorophore molecular brightness and background noise by using the artificial slow variation in fluorophore concentration due to photobleaching to build an in situ FCS calibration curve. These two quantities, in turn, make it possible to obtain very accurate FCS-calibrated concentration maps acquired just before performing the FCS calibration experiments. For this strategy to be successful, the issue of fitting correlation functions affected by slow photobleaching has to be tackled, which we explain how to do. We demonstrate the effectiveness of this strategy with a series of in vitro and in vivo experiments.

Theory

The autocorrelation function (ACF) of the fluorescence signal, $I(t)$, recorded during an FCS experiment is defined as follows:

$$G(\tau) = \langle I(t)I(t+\tau)\rangle / \langle I\rangle^2 - 1.$$

Simple form of the ACF

For a single fluorophore species with molecular brightness B and concentration $c(\vec{r},t)$, the recorded fluorescence signal is a function of the collection profile $W(\vec{r})$:

$$I(t) = \int BW\left(\vec{r}\right)c\left(\vec{r},t\right)d\vec{r}.$$

Both I and B are expressed in photons per second (or hertz). In a confocal instrument, $W(\vec{r})$ resembles a three-dimensional Gaussian ($1/e^2$ radius w, aspect ratio S), and the average detected signal is $I = \gamma BN$, where N is the average number of fluorophores in the effective detection volume $V = \pi^{3/2}Sw^3$, and $\gamma = 2^{-3/2}$ is a geometrical factor (33).

For a diffusive species (diffusion coefficient D) with a single dark state (exponential relaxation time τ_T, average fraction of dark molecules T), the ACF takes the following simple form (34):

$$G_D(\tau) = G(0)\frac{1 + T/(1-T)e^{-\tau/\tau_T}}{(1+\tau/\tau_D)\left(1+\tau/\left(S^2\tau_D\right)\right)^{1/2}}, \qquad (1)$$

where the characteristic diffusion time is $\tau_D = w^2/(4D)$ and the amplitude of the diffusive term is $G(0) = 1/N$.

If a second diffusive species is present, a second term needs to be added to this expression. As long as B is the same for both species (35), the following applies:where p

$$G_D(\tau) = G(0)\left(1+T\,/\,(1-T)e^{-\tau/\tau_T}\right)\left[\frac{p}{(1+\tau/\tau_{D1})\left(1+\tau/\left(S^2\tau_{D1}\right)\right)^{1/2}}+\frac{1-p}{(1+\tau/\tau_{D2})\left(1+\tau/\left(S^2\tau_{D2}\right)\right)^{1/2}}\right], \quad (2)$$

is the fraction of the first species and τ_{D1} and τ_{D2} are the diffusion coefficients of the first and second species, respectively. In this simple case (two species with same molecular brightness), the combined amplitude of both terms, $G(0)$, is related to the total number of fluorophores, N, as before: $G(0) = 1/N$.

Influence of background noise

In the presence of background noise with mean I_B, the average detected signal becomes the following:

$$I = \gamma BN + I_B = \gamma BN\left(1+\frac{m}{N}\right), \quad (3)$$

where we have defined the constant: $m = I_B/(\gamma B)$. The ratio N/m is a measure of the signal/noise.

As long as the background noise is uncorrelated, the ACF retains the same form as in the absence of noise (Eq. 1), but its amplitude decreases as N/m decreases (14,15):

$$G(0) = \frac{1}{N}\frac{1}{(1+I_B/[I-I_B])^2} = \frac{1}{N}\frac{1}{(1+m/N)^2}. \quad (4)$$

For a given amount of noise (that is a given value of m), $G(0)$ is maximum when $N = m$. If $N \gg m$ (high signal/noise ratio) we recover $G(0) \simeq 1/N$. However, in the limit where $N \ll m$, $G(0) \simeq N/m^2$ becomes proportional to N instead.

Equation 4 can be rewritten as a function of I, which in contrast to N is a quantity directly accessible through experiments:

$$G(0) = \gamma B\frac{I-I_B}{I^2}. \quad (5)$$

Influence of photobleaching

When fluorophores are confined to a small compartment, photobleaching may cause a slow decrease in the average number of fluorescent molecules. Consider the simple case of an exponential decrease of the fluorophore concentration:

$$\tilde{I}(t) = \gamma BN_0\left(e^{-\frac{t}{\tau_P}}+\frac{m}{N_0}\right), \quad (6)$$

where a tilde has been used to indicate averaging over a time much longer than the characteristic diffusion time, τ_D, yet much shorter than the characteristic photobleaching decay time, τ_P. N_0 is the average number of molecules in the detection volume at $t = 0$. The average signal between t_1 and t_2 is as follows:

$$\left\langle I\right\rangle_{t_1,t_2} = \gamma BN_0\left(\frac{\tau_P}{t_2-t_1}e^{-\frac{t_1}{\tau_P}}\left(1-e^{-\frac{t_2-t_1}{\tau_P}}\right)+\frac{m}{N_0}\right).$$

At any given time, the measured signal is the sum of $\tilde{I}(t)$ and of a fluctuation around this instantaneous average value due to fluorophore diffusion, $\delta I(t)$. Considering there is no correlation between these two contributions, we can write the following:

$$\langle I(t)I(t+\tau)\rangle = \left\langle \tilde{I}(t)\tilde{I}(t+\tau)\right\rangle + \langle \delta I(t)\delta I(t+\tau)\rangle. \quad (7)$$

The second term in Eq. 7 captures fluctuations in the number of fluorophores in the detection volume due to diffusion. As long as in-focus photobleaching is negligible, it gives rise to the same contribution as before in the ACF, given by Eq. 1. However, N now varies over the course of the measurement, and the amplitude of $G_D(\tau)$ is related to $\langle N\rangle$, the average value of N over the course of the measurement:

$$G(0) = \frac{1}{\langle N\rangle} = \frac{t_M/\tau_P}{N_0(1-e^{-t_M/\tau_P})}. \quad (8)$$

Importantly, because $\langle I\rangle = \gamma B\langle N\rangle$, the relationship between I and $G(0)$ (Eq. 5) is not modified by photobleaching.

The other term in Eq. 7 captures the slow decrease of the signal over time due to fluorophore photobleaching, resulting in a new term in the ACF, $G_P(\tau)$. Its exact expression depends on the normalization scheme used to calculate the denominator of the ACF (see Appendix). In the case of a symmetric normalization, the denominator is calculated using $\langle I\rangle_{0,t_M-\tau}\langle I\rangle_{\tau,t_M}$. This leads to the following expression for the photobleaching term of the ACF (in which the factor -1 present in the definition of the ACF has been included):

In the absence of background noise ($m = 0$), this expression reduces to that previously calculated by Bacia (36):

$$G_P(\tau) = \left(\frac{t_M-\tau}{2\tau_P}\coth\left[\frac{t_M-\tau}{2\tau_P}\right]-1\right)\Bigg/\left(1+\frac{m}{N_0}\frac{t_M-\tau}{\tau_P}\frac{1+e^{\tau/\tau_P}}{1-e^{-\frac{t_M-\tau}{\tau_P}}}+\left(\frac{m}{N_0}\right)^2\left(\frac{t_M-\tau}{\tau_P}\right)^2\frac{e^{\tau/\tau_P}}{1-e^{-\frac{t_M-\tau}{\tau_P}}}\right). \quad (9)$$

$$G_P(\tau) = \frac{t_M-\tau}{2\tau_P}\coth\left[\frac{t_M-\tau}{2\tau_P}\right]-1. \quad (10)$$

The ACF is given by $G(\tau) = G_D(\tau) + G_P(\tau)$.

From pixel intensity to concentration

The relationship between the fluorescence intensity i (in photons per pixel) measured in a confocal image and the fluorophore molar concentration c can be established considering that, by definition, $N = \mathcal{N}cV$ (where $\mathcal{N}$ is Avogadro's number) and that i is related to N through Eq. 3: $i/\delta = \gamma BN + I_B$ (where δ is the pixel dwell time). In the end, the result is as follows:

$$c(x,y) = \frac{i(x,y)/\delta - I_B}{\mathcal{N}\gamma BV}. \quad (11)$$

MATERIALS AND METHODS

Fluorophore solutions

Alexa Fluor 488 (AF488) was purchased from Invitrogen (now Life Technologies, Carlsbad, CA). It has a known diffusion coefficient $D = 435\ \mu m^2/s$ at 22.5°C (37). Solutions of AF488 were prepared in double-deionized water (resistance 18 MΩ). Purified enhanced green fluorescent protein (eGFP) was purchased from BioVision (catalog number: 4999; Milpitas, CA). This protein is labeled with two polyhistidine tags, and its molecular weight (32.7 kDa) is slightly larger than that of wild-type GFP. Solutions of eGFP were prepared in phosphate buffer saline.

Drosophila embryos

Drosophila embryos were prepared for imaging following the protocol described in (38). *Drosophila melanogaster* fly strains expressing nuclear localization signal (NLS)-eGFP, Bicoid (Bcd)-eGFP (a kind gift of Dr. Wieschaus) (5), or Capicua (Cic)-sfGPP (a kind gift of Dr. Shvartsman) (39) were stored and maintained in a 25°C incubator with alternating day-night lighting. To collect embryos for experiments, plastic tubes containing flies were inverted on an embryo collection plate with yeast paste in the center. After ~3 h, the embryos on the collection plate were transferred with a tweezer to a double-sided tape to remove the chorion. The dechorionated embryos were then transferred onto a thin layer of heptane glue on a 0.17-mm coverslip. The ventral side of the embryo was placed in contact with the glue such that as many nuclei as possible could be observed just above the coverslip. Lastly, a small drop of Halocarbon oil 700 (Sigma-Aldrich, St. Louis, MO) was added on top of the embryos to prevent evaporation while allowing oxygen permeation into the embryo.

FCS and confocal imaging

Single-point FCS data and confocal images were both recorded on an Insight Cell confocal microscope (Evotec Technologies, now PerkinElmer, Waltham, MA) using the same configuration. Fluorescence was excited with a 488-nm continuous wave solid state diode-pumped laser (Sapphire 488–20/460-10; Coherent, Santa Clara, CA). The excitation power was 75 μW when working with AF488 and 25 μW when working with eGFP in vitro. It was 20 μW when working with embryos expressing NLS-eGFP or Bcd-eGFP, and 25 μW when working with embryos expressing Cic-sfGFP. The excitation beam was set so as to underfill the back aperture of the water-immersion objective (UAPON, ×40, 1.15 NA; Olympus, Tokyo, Japan) and was used in conjunction with a 40-μm pinhole in the detection pathway. All experiments were performed at room temperature. Single-point FCS measurements of 5, 10, or 20 s were performed for 10, 20, or 40 repeats in each series. Fitting of the ACF obtained as a result of single-point FCS experiments was done either with the software FCS+plus Analyze (Evotec Technologies, now PerkinElmer, Waltham, MA) or with MATLAB (The MathWorks, Natick, MA).

RESULTS

We carried a series of FCS experiments, both in vitro and in vivo, to confirm the relationship between $G(0)$, the amplitude of the diffusive part of the ACF, and I, the mean count rate, which are both experimentally accessible quantities. This relationship is predicted to obey Eq. 5 and should depend on the values of the fluorophore molecular brightness (B) and the noncorrelated average background noise (I_B). Varying $G(0)$ and I, which can be achieved by systematically varying the fluorophore concentration, should allow one to retrieve the two crucial calibration parameters, B and I_B, and to calculate absolute fluorophore concentration directly from I.

In vitro: AF488 and eGFP

Single-point FCS experiments were performed on two types of in vitro samples: AF488 and eGFP solutions, in which fluorophore concentration was varied over several orders of magnitude by performing serial dilutions. Both samples serve as a model system for fluorophores undergoing free diffusion in the absence of photobleaching, giving the opportunity to explore the relationship between $G(0)$ and I in a simple system. Results from these experiments are shown in Fig. 1.

The ACFs obtained as a result of these experiments are, as expected, well fitted with a one-component model (Eq. 1),

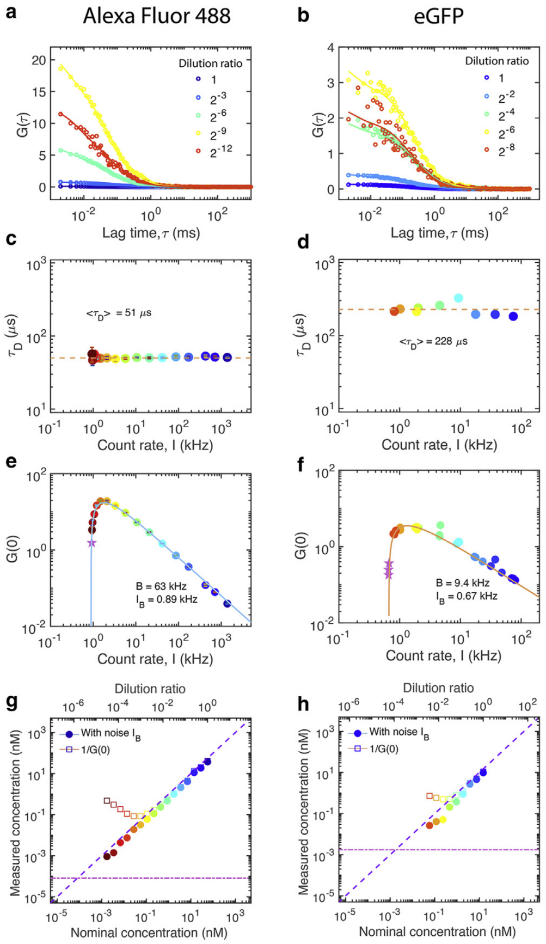

FIGURE 1 Single-point FCS concentration measurements in solution: AF488 (*left column*) and eGFP (*right column*). (*a* and *b*) Example of ACFs obtained at different concentrations (while the excitation power was kept constant). Solid lines are fitted with Eq. 1. (*c* and *d*) Characteristic relaxation time, τ_D, extracted from the ACFs and plotted as a function of count rate, *I*. Dashed line shows the mean τ_D. (*e* and *f*) Relationship between $G(0)$, the amplitude of the diffusive part of the ACF, and *I*, the mean count rate. The solid line is a fit with Eq. 5, which allows the extraction of both molecular brightness, *B*, and background noise, I_B. Pink stars represent data from the buffers for which very small amounts of contaminant gave rise to a detectable ACF. In (*f*), data points of the same color show the result from repeated measurements in the same sample. (*g* and *h*) Absolute concentration calculated from the estimated value of *N*. Open symbols show what happens when background noise is ignored (i.e., when assuming that $N = 1/G(0)$), whereas solid symbols show what happens when noise is taken into account (i.e., when using Eq. 4 to solve for *N* given the value of $m = I_B/(\gamma B)$ obtained from the fit of the dependency of $G(0)$ on *I* shown in *e* and *f*). The purple dashed line shows the nominal fluorophore concentration (provided by the suppliers), and the pink dotted dashed line shows the estimated value of the contaminant concentration in the buffer. To see this figure in color, go online.

assuming the presence of a single population of fluorophore in solution (Fig. 1, *a and b*). The characteristic diffusion times extracted from these fits are constant throughout the explored concentration range (Fig. 1, *c and d*). Their mean values ($\langle\tau_D\rangle = 51 \pm 3$ μs for AF488 and 228 ± 43 μs for eGFP, corresponding to diffusion coefficients of 435 and 97 $\mu m^2/s$, respectively) are in keeping with the respective molecular weight of the two fluorophores (0.72 kDa for AF488 and 32.7 kDa for eGFP). In sharp contrast to τ_D, the amplitude of the diffusive part of the ACF, $G(0)$, varied over two to three orders of magnitude as the fluorophore concentration was varied (Fig. 1, *e and f*).

It is often assumed when using FCS data to measure concentration that $G(0)$ is simply inversely related to the average number of fluorophores present in the confocal detection volume ($G(0) = 1/N$). In that case, we should see a monotonous increase of $G(0)$ as the fluorophore is diluted and as the count rate decreases, as is indeed observed at high fluorophore concentrations (high count rate). However, as *I* approaches the count rate measured for the buffer, $G(0)$ reaches a maximum and then sharply decreases (Fig. 1, *e and f*). This is what is expected in the presence of uncorrelated background noise, an effect that is captured in Eq. 5. Fitting of the data shows an excellent agreement with Eq. 5 and allows retrieving of two very important parameters—the fluorophore molecular brightness *B* and the background noise I_B. From these parameters, the ratio $m = I_B/(\gamma B)$ can be calculated, which gives an idea of how large background noise is compared with the effective brightness of a single molecule. We found that $m = 0.014$ for AF488 and 0.071 for eGFP, reflecting a large difference in molecular brightness for these two fluorophores.

Once the value of *m* is known for a particular sample and a particular set of experimental conditions, the actual relationship between $G(0)$ and *N* (Eq. 4) can be used to calculate *N* from the measured value of $G(0)$. The absolute fluorophore concentration can then be calculated using the value of *V* (observation volume) obtained from ACF measurements with a fluorophore with known diffusion coefficient (AF488). The concentrations measured for the AF488 and eGFP samples as a function of their dilution ratio are shown in Fig. 1, *g and h*. Strikingly, a linear relationship between these two quantities is obtained over the whole measurement range for both samples, showing that FCS allows precise concentration measurements in the sub-nanomolar range. Comparing these results (*solid symbols* in Fig. 1, *g and h*) with those obtained without taking into account background noise (*open symbols*) shows that the procedure described here extends the accessible concentration range by approximately two orders of magnitude. We note that for both fluorophores, the concentrations measured by FCS were about 30% smaller than the nominal concentrations indicated by the suppliers (*dashed lines* in Fig. 1, *g and h*), which is not overly surprising because fluorophores might improperly dissolve or adsorb on sample surfaces or photobleach.

Although the precision of the concentration measurements and their accessible range (down to 1 pM for AF488 and 20 pM for eGFP) are impressive for both fluorophores, it is noticeably lower for eGFP. This difference may be attributed in part to a certain instability of the eGFP sample (visible in the dispersion in the values of $G(0)$ measured for repeated measurements; Fig. 1 *f*) probably because of the presence of aggregates and to the interaction of the protein with the surfaces of the sample chamber. Mostly, it can be traced back to the relatively lower molecular brightness of eGFP and consequently larger m. We found that the lowest attainable concentration was about two orders of magnitude lower than that for which $G(0)$ reaches a maximum. Because the peak in the value of $G(0)$ is attained when $N = m$, the lower the value of m, the lower the concentration that can be directly measured by single-point FCS.

In vivo: NLS-eGFP, Bcd-eGFP, and Cic-sfGFP

In cells, systematic variations in apparent fluorophore concentration can be achieved through gradual photobleaching of the molecules present inside the cellular compartment where single-point FCS measurements are performed. We used this effect to establish the relationship between $G(0)$ and I inside the nuclei of live embryos. We used *D. melanogaster* embryos expressing different types of fluorescent protein fusion for which we can expect different dynamics, molecular brightness, and concentration: an NLS-eGFP, a transcription activator called Bcd-eGFP, and a transcription repressor called Cic-sfGFP. Both Bcd and Cic are important regulators of gene expression that are endogenously expressed during early fly development.

We performed series of 10–20 single-point FCS measurements (for durations of 5–20 s) at the center of nuclei in the midsection of embryos during nuclear cycle 13 or 14, as illustrated in Fig. 2 for NLS-eGFP. All three studied proteins are actively imported and accumulate into nuclei, resulting in the presence of brightly fluorescent nuclei in images of the cortical region of the embryo, where a single layer of nuclei is found at this stage of development (Fig. 2 *a*). The depletion of fluorescence due to continuous photobleaching during repeated FCS measurements in a single

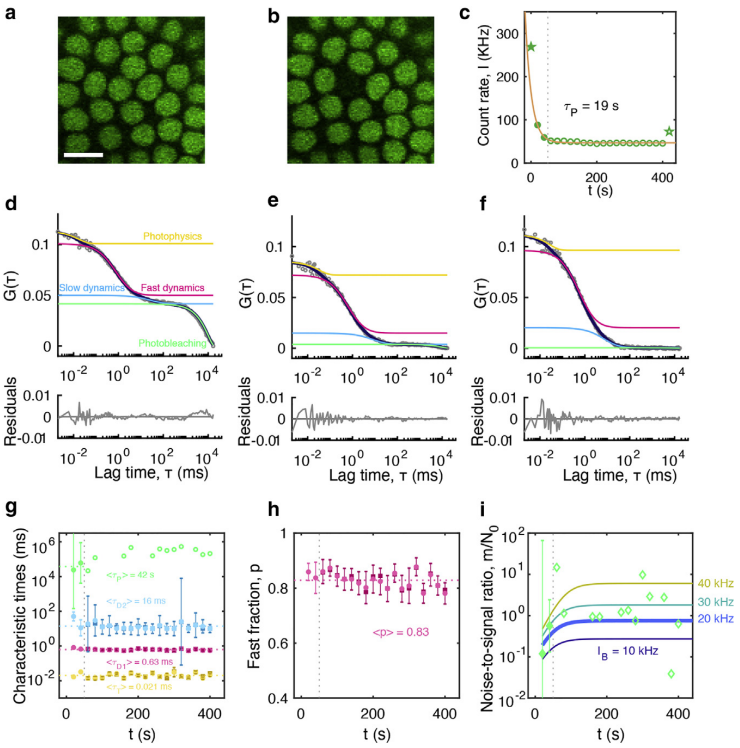

FIGURE 2 Single-point FCS measurements in *D. melanogaster* embryos expressing NLS-eGFP. (*a*) Representative confocal image of a field of nuclei in the cortical region of the embryo. Scale bar, 10 μm. (*b*) Same field of view, just after a series of 20 single-point FCS measurements. The nucleus where the FCS measurements were performed is only very faintly fluorescent. (*c*) Average count rate recorded for each of the 20 FCS measurements (*solid* and *open circles*, separated by *vertical dashed lines*, indicate measurements during which the count rate either significantly varied or was reasonably stable). The orange line is an exponential fit to the data (Eq. 6). The count rate recorded for the pixel at which the FCS measurements were performed is also shown for the image acquired just before ($t = 0$ s) and just after ($t = 420$ s) the FCS measurements (*star symbols*). (*d–f*) ACF obtained for the first (*d*), second (*e*), and 20th (*f*) measurements in this series of 20-s FCS measurements. Solid lines indicate a fit with a two-component model, taking account the possibility of photobleaching (Eqs. 2 and 9), and residuals are shown below. Lines of different colors indicate the different decays observed in the ACF due to triplet state relaxation (*orange*), fast diffusion (*magenta*), slow diffusion (*blue*), and photobleaching (*green*). (*g*) Different characteristic times, (*h*) fraction of fast molecules, and (*i*) noise/signal ratio obtained from the fit of the ACF for all the measurements in the series. Error bars in (*g*)–(*i*) correspond to 50% confidence intervals, except for m/N_0, for which the error bars correspond to 10% confidence intervals (values obtained from ACFs with a photobleaching decay too small for a reliable estimate are indicated by *open symbols*, and without error bars because they were out-of-range; a few fits did not converge properly for these parameters, in which case they were not shown). In (*g*) and (*h*), circles and squares correspond to values obtained from fits performed with and without the photobleaching term, respectively. Horizontal dashed lines indicate average values (for τ_P only reliable measurements, indicated by *solid symbols*, have been considered when calculating this average value). In (*i*), the solid lines indicate the predicted value of $m/N_0 = I_B/(I - I_B)$ for different values of I_B, approximating I by its fitted value (*orange line* in *c*). To see this figure in color, go online.

nucleus was evident in images taken immediately after these measurements, in which the fluorescence of the studied nucleus strongly diminished (Fig. 2 *b*). The decay in the fluorescence signal at the position of the FCS measurements, which can be attributed to the photobleaching of the nuclear fraction of NLS-eGFP, was well approximated by a single exponential. In a regular cell, the cytoplasmic fraction of the fluorescent protein would also eventually get photobleached over a time corresponding to nucleocytoplasmic exchange—resulting in an additional slower decay time. But because the fly embryo is a syncytium, the cytoplasmic concentration remains constant over experimental timescales, and only a single exponential decay is observed. The characteristic time associated with this decay is $\tau_P \approx$ 20 s in the conditions of our experiments (Fig. 2 *c*). Thus, after only a few measurements, an equilibrium is reached between the import of new fluorescent molecules into the nucleus and the photobleaching, allowing the fluorescence to stabilize. When excitation is stopped at the end of the series of measurements, the fluorescence immediately starts recovering because of nuclear import (*green star* in Fig. 2 *c*).

The strong photobleaching that occurred during the first few FCS measurements in a series resulted in a visible decay in the ACF around τ_P, i.e., at much larger lag times than the decay corresponding to the motion of the proteins (Fig. 2, *d and e*). In contrast, the ACFs corresponding to later measurements in the series, after stabilization of the fluorescence, did not show this large time decay (Fig. 2 *f*). We therefore fitted our data with a model that included a two-component diffusive term (Eq. 2, because most nuclear proteins show at least two mobile components (28,35,40,41)) and a term corresponding to long-term photobleaching (Eq. 9; see Theory for a derivation of this term). This model allowed adequately fitting of all measurements in a series (for all studied proteins), as shown for NLS-eGFP in (Fig. 2, *d–f*), and retrieving four different characteristic times (for photophysics, fast diffusion, slow diffusion, and photobleaching; Fig. 2 *g*) for each of them; the fraction of fast molecules p obtained from the relative amplitude of the fast and slow diffusion terms (Fig. 2 *h*); and an estimate of the noise/signal ratio m/N_0 obtained from the amplitude of the photobleaching term (Fig. 2 *i*). This last parameter is reliably obtained only for the first few measurements in a series (*solid symbols* in Fig. 2 *i*), when the photobleaching decay is clearly visible in the ACF (later measurements can be fitted without the photobleaching term). Over these first few measurements, the relative importance of noise increases by several orders of magnitude. In this case, it appears to stabilize around a value of $m/N_0 = 1$, indicating that about half of the detected signal at this point comes from background fluorescence. This is consistent with the contrast observed in Fig. 2 *b*, where the studied nucleus is visible, but just barely.

For concentration measurements, however, the most important information contained in the ACFs is the amplitude of the combined diffusion terms, $G(0)$. For each series of FCS measurements that was performed (between 5 and 7 for each protein, performed in different embryos and on different days but using the same experimental conditions), the values obtained for $G(0)$ were found to depend on the count rate I as expected and as captured in Eq. 5 (Fig. 3, *a, d, and g*). For each series of successive measurements, fit of the data with Eq. 5 allowed a reliable in situ measurement of both molecular brightness B (Fig. 3, *b, e, and h*) and background I_B (Fig. 3, *c, f, and i*), just as for the fluorophore solutions described in the previous section. As long as the count rate was corrected for uneven illumination and detection across the field of view (as explained in the next section), the values of I_B were found to be similar for embryos expressing different proteins, ~10–15 kHz in the conditions of our experiments. The values of B were also reproducible but varied for different proteins and decreased from NLS-eGFP to Bcd-eGFP to Cic-sfGFP. As expected, the importance of noise was much larger in embryos than in solution, as demonstrated by the average values of m that were observed: $m = 2.5$ for NLS-eGFP, 4.3 for Bcd-eGFP, and 7.2 for Cic-sfGFP, indicating that measuring very low concentrations will be much more challenging in this case. Interestingly, for NLS-eGFP and Bcd-eGFP, the equilibrium concentration reached after several FCS measurements was still well above the point at which $G(0)$ starts noticeably decreasing because of background noise (Fig. 3, *a and d*), whereas for Cic-sfGFP, a strong decrease in $G(0)$ is observed for the later measurements in each series, and very low concentrations (for which $N < m$) are achieved at that point (Fig. 3 *g*). This difference can be traced back to the different behavior of these proteins in regard to nuclear import. -Whereas the nuclear concentration of fluorescent Bcd-eGFP and NLS-eGFP can completely recover in only a few minutes if a whole nucleus is photobleached (5), Cic-sfGFP only incompletely recovers (42), suggesting that the available pool of Cic in the cytoplasm of the embryo is very limited. Thus, the nuclear concentration reached at long times as an equilibrium between photobleaching and nuclear import is much lower for Cic than it is for the other two proteins. As a result, the estimate of B that can be made from each individual ACF by neglecting the effect of noise (i.e., using $B = G(0)I/\gamma$) only differs from the actual B by at most ~30% for NLS-eGFP and ~50% for Bcd-eGFP (Fig. 3, *b and e*, *small symbols*). But for Cic-sfGFP, the error made on the value of B when neglecting background noise can approach ~100% (Fig. 3 *h*).

Obtaining concentration maps

To obtain concentration maps of fluorescent proteins from confocal images, we followed the procedure illustrated in Fig. 4 for a *D. melanogaster* embryo expressing Cic-sfGFP.

First, a confocal image was acquired in the cortical region of the midsection of the embryo, just above the coverslip to

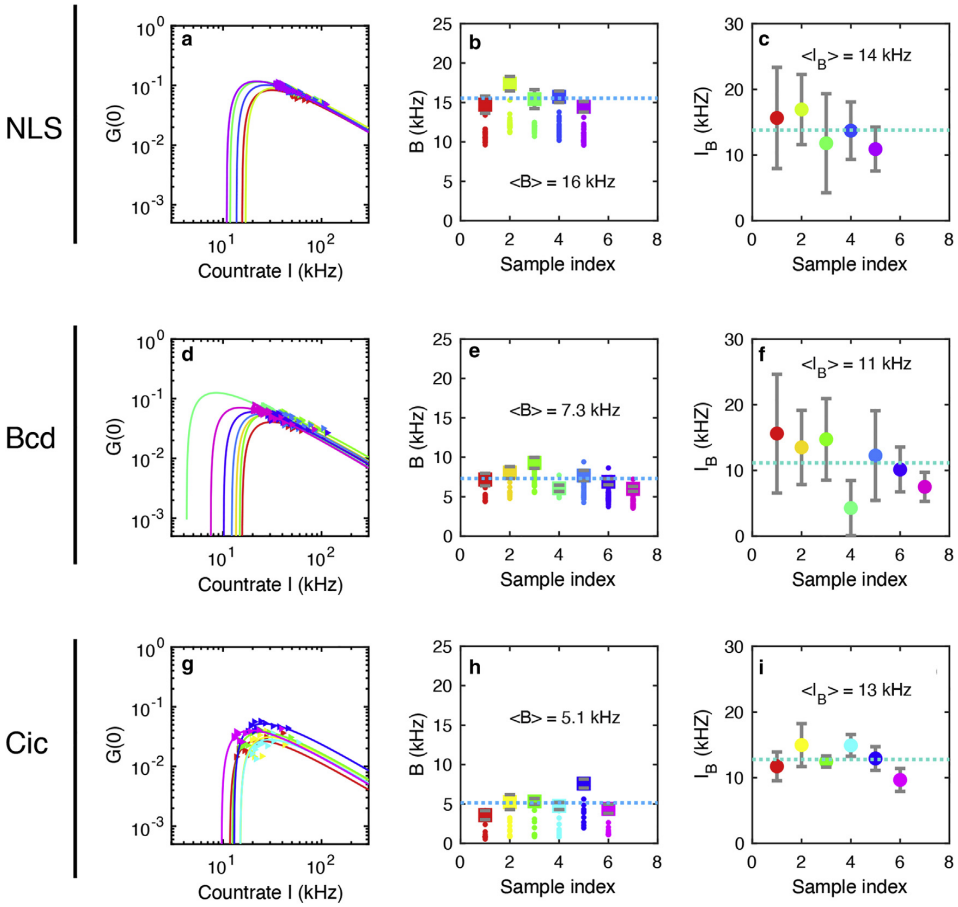

FIGURE 3 Relationship between $G(0)$, the amplitude of the diffusive part of the ACF, and I, the calibrated mean count rate, for embryos expressing NLS-eGFP (*a–c*, *top row*), Bcd-eGFP (*d–f*, *middle row*), and Cic-sfGFP (*g–i*, *bottom row*). (*a*, *d*, and *g*) For each type of embryo, at least five series of FCS measurements were performed, each resulting in a $G(0)$ vs. I sequence (represented by *symbols* of the *same color*), which was fitted with Eq. 5 (*lines*). (*b*, *e*, and *h*) Molecular brightness B (*large symbols*) and (*c*, *f*, and *i*) background noise I_B extracted from the fits of each of the $G(0)$ vs. I sequence. Error bars correspond to the 95% confidence interval obtained for these parameters. The mean values of B and I_B are indicated by a dashed line. In (*b*), (*e*), and (*h*), the values of B recovered from each individual FCS measurement, assuming $G(0) = \gamma B/I$ (no noise) is also shown (*small symbols*). To see this figure in color, go online.

reduce optical aberrations (Fig. 4 *a*). The original image was then corrected for the spatially uneven illumination and detection efficiency of our confocal instrument (Fig. 4 *b*). A previously acquired image of an AF488 solution for the exact same field of view was fitted to a broad two-dimensional Gaussian function, which was then normalized to 1, after which the pixel intensity at each point of the image was divided by the value of this normalized Gaussian function. The corrected intensity map then correctly displays nuclei with uniform fluorescence intensity across the field of view (Fig. 4 *b*), as expected here because Cic is known to have a uniform nuclear concentration in this region of the embryo (42).

The values of B and I_B measured in situ by performing a series of single-point FCS experiments (as explained in the previous section) were then used to convert the corrected intensity map into an absolute concentration map using Eq. 11 (Fig. 4 *c*). The single-point FCS measurements can be performed right after the acquisition of the image itself in the exact same field of view, or if the values of B and I_B can be shown to be reproducible for different regions of the sample, before imaging but in a different field of view (it is important that no FCS experiment is performed in the field of view before imaging to avoid photobleaching). It is also important that the intensity I used when fitting the dependence of $G(0)$ on I to obtain B and I_B is corrected for uneven illumination and detection in the same way as the pixels in the corrected image. A cross section through the concentration map shows that at this stage of development the nuclear concentration of Cic-sfGFP is ~200 nM (Fig. 4 *d*).

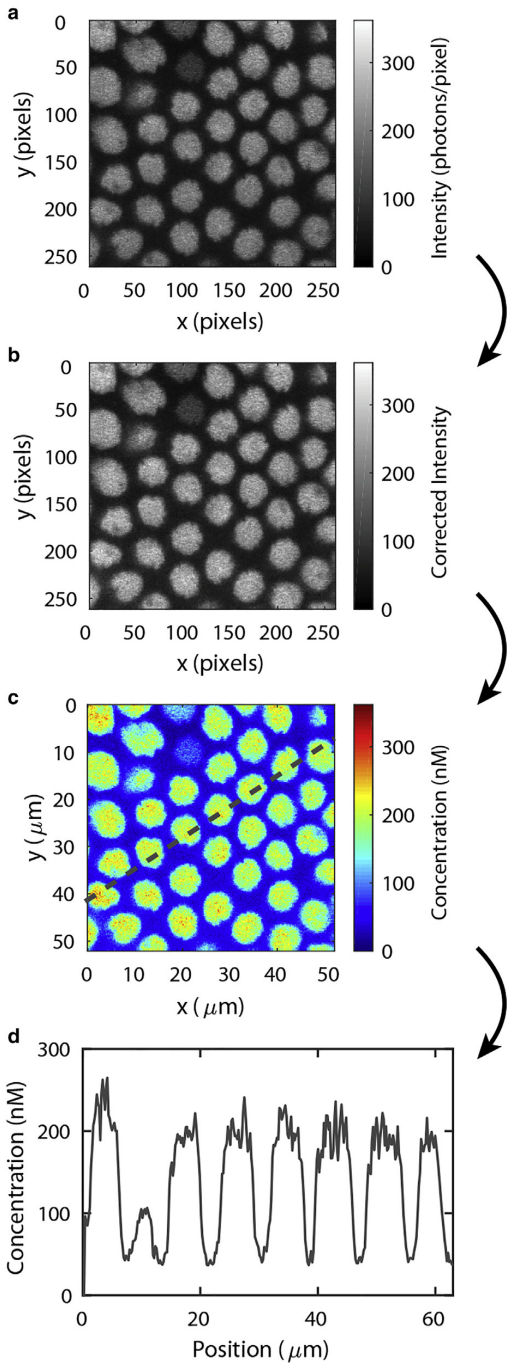

FIGURE 4 Conversion of an intensity image into a concentration map. (*a*) Confocal image of the cortical area in the midsection of a *D. melanogaster* embryo expressing Cic-sfGFP at nuclear cycle 14, ~1 h into the embryo development (261 × 261 pixels, pixel dwell time δ = 1 ms, pixel size 0.2 μm/pixel). (*b*) Intensity map after a correction taking into in account the uneven illumination and detection efficiency across the field of view. (*c*) Absolute concentration map calculated from the corrected pixel intensity map shown in (*b*), using Eq. 11 and values of B, I_B, and V obtained from single-point FCS data. (*d*) Concentration profile along the dashed line shown in (*c*). To see this figure in color, go online.

DISCUSSION

We propose here a method to obtain protein concentration maps that combines the strengths of confocal imaging (fluorescence signal measured in space and time and sensitivity to both immobile and mobile fluorophores) to those of FCS (sensitivity to absolute particle number) and, therefore, allows rapid measurements of absolute fluorophore concentration over large fields of view and with good temporal resolution. Obtaining a concentration map from a confocal image requires knowledge of both the molecular brightness B of the fluorophores present in the sample and the background intensity I_B. Our method is based on an accurate measurement of B that takes into account the effect of noise to avoid any systematic bias and that is performed in situ to recover the actual value of B in the cellular environment. We have demonstrated here through a series of in vitro and in vivo experiments that this could be achieved by establishing the dependence of $G(0)$ (a parameter directly accessible from FCS measurements) on I (average count rate), which we showed can be done by taking advantage of photobleaching, which causes a progressive decrease of I and provides the mean of performing a systematic titration of fluorophore concentration, just as one would do in vitro through serial dilutions. This strategy will be especially useful in systems in which photobleaching is prominent anyway or in which variations in concentration cannot be achieved by any other mean (e.g., uneven concentration across the sample or varying expression levels from cell to cell). Our approach should therefore be widely applicable to many types of eukaryotic systems.

It is important to note that some limitations are of course associated with the method proposed here. First, although the dependence between $G(0)$ and I captured in Eq. 5 is largely independent of the nature of the protein motions (it still holds for example if protein fractions with different mobilities are present), it relies on the assumption that the protein of interest has a mobile fraction with a single molecular brightness. Thus, Eq. 5 cannot be used in cases when the studied protein forms diffusing homo-oligomers. The method described here also relies on the assumption that the protein of interest is the only diffusing fluorescent species in the sample because the only type of noise considered in this work was uncorrelated noise (i.e., noise with no associated characteristic correlation time or with a correlation time outside of the 1-μs-to-10-s measurement window). Most types of noise expected in FCS experiments (detector noise, Rayleigh and/or Raman scattering, reflections at interfaces, and out-of-focus signal coming from fluorescent molecules) fulfill this condition. However, diffusing contaminants present in the detection

volume and with a concentration and molecular brightness comparable with that of the protein of interest would appear as an additional diffusion term in the ACF, changing its amplitude and throwing off concentration estimates. Contaminants present at very small concentrations (such as the sub-picomolar concentration of contaminant detected in buffer solutions in the presented in vitro experiments; indicated with a *dashed line* in Fig. 1, *g and h*) do not interfere with concentration estimates.

Second, only the long-term effect of the photobleaching occurring throughout the light cone and resulting in a slow exponential depletion of the fluorophore was considered here. The short-term effect of in-focus photobleaching on the ACF (as described, for example, in (43–45)) was not taken into account, although we expect that in most cases it would have a negligible effect on the retrieved value of $G(0)$—and therefore on the precision of concentration measurements. Third, even if $G(0)$ is correctly measured, the value of I_B can be obtained with good precision only if the regime in which $N < m$ (left of the inflection point on the $G(0)$ vs. I curve, when noise dominates) is reached. If only the regime in which $N > m$ is explored (as was the case here for NLS-eGFP or Bcd-eGFP that are continuously and robustly imported from the embryo's large cytoplasm and for which the $N < m$ regime was not reached; see Fig. 3, *a and d*), then B is precisely measured (Fig. 3, *b and e*) but not I_B (Fig. 3, *c and f*). Luckily, when it comes to concentration measurements, for these samples in which the signal/noise is always high (where $N \gg m$ and therefore $I - I_B \gg I_B$), uncertainties on I_B only result in small relative uncertainties on the measured absolute concentrations. For samples in which both the $N > m$ and $N < m$ regimes are spanned, as was the case here for Cic-sfGFP (for which nuclear import is limited, and the inflection point was reached on the $G(0)$ vs. I curve; see Fig. 3 *g*), a precise measurement of both molecular brightness and noise can be achieved (Fig. 3, *h and i*). In cultured cells, with a small cytoplasmic volume, nuclear import should always be limited by the finite available pool of cytoplasmic fluorescent proteins, and we therefore expect to always be in this favorable case. The method demonstrated here in fly embryos will thus be readily applicable in cultured cells, with the caveat that a double-exponential decay of the intensity might be expected.

Finally, a lot of possible artifacts (optical aberrations and fluorescence saturation) may affect the real or perceived size and shape of the confocal detection volume (46–49) and thus lead to systematic errors when using the value of the confocal volume V to calculate absolute protein concentrations from the estimated N (Eq. 11).

CONCLUSIONS

One tenet of our method is a systematic consideration of the effect of noise: when calculating absolute concentrations from pixel intensity (Eq. 11), when estimating B from $G(0)$ (Eq. 5), and when fitting the slow decay in the ACF due to photobleaching (Eq. 9). This proper accounting of noise means that we are able to accurately measure concentrations in cells, as intended. When applied to single-point FCS experiments, this relatively simple noise correction also means that we are able to measure picomolar concentrations in ideal fluorophore solutions (Fig. 1), lower than what is usually considered the lower limit for FCS measurements and entering instead the realm of what can be achieved using single particle detection (50,51).

Another tenet of our method is the fitting of ACFs that display a long-term decay due to photobleaching, for which we have developed a model (captured in Eq. 9) that incorporates both the effects of long-term photobleaching and background noise. This allows us to obtain accurate values for $G(0)$ because the count rate is continuously decreasing because of photobleaching, allowing us to explore the dependence of $G(0)$ on I over a wider range of count rates and to obtain as accurate an estimate of B as possible. Importantly, being able to fit ACFs that display photobleaching decays means that information about the dynamics of the fluorophores can be obtained from high signal/noise ratio data acquired at high excitation intensity and for long (10 s or more) measurement times and without having to wait until after the signal has stabilized to a low value. It resolves the conundrum of having to use low excitation intensities to avoid photobleaching at the cost of achieving only poor molecular brightness. We therefore expect it will change the way we think about performing single-point FCS experiments in cells by removing the obligation to avoid photobleaching at all costs.

APPENDIX: PHOTOBLEACHING TERM FOR DIFFERENT TYPES OF NORMALIZATION

The numerator of the ACF can be calculated as $\langle I \rangle^2_{0,t_M}$ (very simple normalization), $\langle I^2 \rangle_{0,t_M-\tau}$ (simple normalization), or $\langle I \rangle_{0,t_M-\tau}\langle I \rangle_{\tau,t_M}$ (symmetric normalization).

In the first case (very simple normalization), we obtain the following:

$$G_P(\tau) = e^{-t_M/\tau_P} \frac{\frac{\sinh[(t_M-\tau)/\tau_P]}{(t_M-\tau)/\tau_P}}{\left(\frac{1-e^{-t_M/\tau_P}}{t_M/\tau_P}\right)^2} - 1. \tag{12}$$

In the second case (simple normalization), we obtain the following (36):

$$G_P(\tau) = \frac{t_M - \tau}{2\tau_P} e^{-\tau/\tau_P} \frac{1 + e^{-(t_M-\tau)/\tau_P}}{1 - e^{-(t_M-\tau)/\tau_P}} - 1. \tag{13}$$

In the third case (symmetric normalization), we obtain the following (36):

$$G_P(\tau) = \frac{t_M - \tau}{2\tau_P} \frac{1 + e^{-(t_M-\tau)/\tau_P}}{1 - e^{-(t_M-\tau)/\tau_P}} - 1. \tag{14}$$

AUTHOR CONTRIBUTIONS

L.Z. designed the research, performed and analyzed all the experiments, and wrote the manuscript. C.P.-R. designed the research and performed preliminary experiments. N.D. designed the research. C.F. designed the research, performed analytical calculations, and wrote the manuscript.

ACKNOWLEDGMENTS

L.Z. and C.P.-R. were both recipients of Ontario Trillium Scholarships. The research was funded by the Natural Sciences and Engineering Research Council of Canada through a discovery grant to C.F. (RGPIN-2015-06362).

REFERENCES

1. Berg, H. C., and E. M. Purcell. 1977. Physics of chemoreception. *Biophys. J.* 20:193–219.
2. Bialek, W., and S. Setayeshgar. 2005. Physical limits to biochemical signaling. *Proc. Natl. Acad. Sci. USA.* 102:10040–10045.
3. Houchmandzadeh, B., E. Wieschaus, and S. Leibler. 2002. Establishment of developmental precision and proportions in the early Drosophila embryo. *Nature.* 415:798–802.
4. Ashe, H. L., and J. Briscoe. 2006. The interpretation of morphogen gradients. *Development.* 133:385–394.
5. Gregor, T., E. F. Wieschaus, …, D. W. Tank. 2007. Stability and nuclear dynamics of the bicoid morphogen gradient. *Cell.* 130:141–152.
6. Kicheva, A., P. Pantazis, …, M. González-Gaitán. 2007. Kinetics of morphogen gradient formation. *Science.* 315:521–525.
7. Harvey, S. A., and J. C. Smith. 2009. Visualisation and quantification of morphogen gradient formation in the zebrafish. *PLoS Biol.* 7:e1000101.
8. Sagner, A., and J. Briscoe. 2017. Morphogen interpretation: concentration, time, competence, and signaling dynamics. *Wiley Interdiscip. Rev. Dev. Biol.* 6:e271.
9. Shilo, B.-Z., and N. Barkai. 2017. Buffering global variability of morphogen gradients. *Dev. Cell.* 40:429–438.
10. Durrieu, L., D. Kirrmaier, …, M. Knop. 2018. Bicoid gradient formation mechanism and dynamics revealed by protein lifetime analysis. *Mol. Syst. Biol.* 14:e8355.
11. Elson, E. L., and D. Magde. 1974. Fluorescence correlation spectroscopy. I. Conceptual basis and theory. *Biopolymers.* 13:1–27.
12. Maiti, S., U. Haupts, and W. W. Webb. 1997. Fluorescence correlation spectroscopy: diagnostics for sparse molecules. *Proc. Natl. Acad. Sci. USA.* 94:11753–11757.
13. Weissman, M., H. Schindler, and G. Feher. 1976. Determination of molecular weights by fluctuation spectroscopy: application to DNA. *Proc. Natl. Acad. Sci. USA.* 73:2776–2780.
14. Koppel, D. E. 1974. Statistical accuracy in fluorescence correlation spectroscopy. *Phys. Rev. A.* 10:1938–1945.
15. Rigler, R., Ü. Mets, …, P. Kask. 1993. Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. *Eur. Biophys. J.* 22:169–175.
16. Wachsmuth, M., T. Weidemann, …, J. Langowski. 2003. Analyzing intracellular binding and diffusion with continuous fluorescence photobleaching. *Biophys. J.* 84:3353–3363.
17. Delon, A., Y. Usson, …, C. Souchier. 2006. Continuous photobleaching in vesicles and living cells: a measure of diffusion and compartmentation. *Biophys. J.* 90:2548–2562.
18. Abu-Arish, A., A. Porcher, …, C. Fradin. 2010. High mobility of bicoid captured by fluorescence correlation spectroscopy: implication for the rapid establishment of its gradient. *Biophys. J.* 99:L33–L35.
19. Hodges, C., R. P. Kafle, …, J.-C. Meiners. 2018. Fluorescence correlation spectroscopy with photobleaching correction in slowly diffusing systems. *J. Fluoresc.* 28:505–511.
20. Sezgin, E., F. Schneider, …, C. Eggeling. 2019. Measuring nanoscale diffusion dynamics in cellular membranes with super-resolution STED-FCS. *Nat. Protoc.* 14:1054–1083.
21. Widengren, J., and R. Rigler. 1998. Fluorescence correlation spectroscopy as a tool to investigate chemical reactions in solutions and on cell surfaces. *Cell. Mol. Biol.* 44:857–879.
22. Delon, A., Y. Usson, …, C. Souchier. 2004. Photobleaching, mobility, and compartmentalisation: inferences in fluorescence correlation spectroscopy. *J. Fluoresc.* 14:255–267.
23. Wachsmuth, M., C. Conrad, …, J. Ellenberg. 2015. High-throughput fluorescence correlation spectroscopy enables analysis of proteome dynamics in living cells. *Nat. Biotechnol.* 33:384–389.
24. Waithe, D., F. Schneider, …, C. Eggeling. 2018. Optimized processing and analysis of conventional confocal microscopy generated scanning FCS data. *Methods.* 140–141:62–73.
25. Ries, J., S. Chiantia, and P. Schwille. 2009. Accurate determination of membrane dynamics with line-scan FCS. *Biophys. J.* 96:1999–2008.
26. Fradin, C., A. Abu-Arish, …, M. Elbaum. 2003. Fluorescence correlation spectroscopy close to a fluctuating membrane. *Biophys. J.* 84:2005–2020.
27. Cluzel, P., M. Surette, and S. Leibler. 2000. An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science.* 287:1652–1655.
28. Weidemann, T., M. Wachsmuth, …, J. Langowski. 2003. Counting nucleosomes in living cells with a combination of fluorescence correlation spectroscopy and confocal imaging. *J. Mol. Biol.* 334:229–240.
29. Abu-Arish, A., P. Kalab, …, C. Fradin. 2009. Spatial distribution and mobility of the Ran GTPase in live interphase cells. *Biophys. J.* 97:2164–2178.
30. Politi, A. Z., Y. Cai, …, J. Ellenberg. 2018. Quantitative mapping of fluorescently tagged cellular proteins using FCS-calibrated four-dimensional imaging. *Nat. Protoc.* 13:1445–1464.
31. Heim, R., and R. Y. Tsien. 1996. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* 6:178–182.
32. Kremers, G.-J., J. Goedhart, …, T. W. Gadella, Jr. 2007. Improved green and blue fluorescent proteins for expression in bacteria and mammalian cells. *Biochemistry.* 46:3775–3783.
33. Nagy, A., J. Wu, and K. M. Berland. 2005. Observation volumes and γ-factors in two-photon fluorescence fluctuation spectroscopy. *Biophys. J.* 89:2077–2090.
34. Widengren, J., U. Mets, and R. Rigler. 1995. Fluorescence correlation spectroscopy of triplet states in solution: a theoretical and experimental study. *J. Phys. Chem.* 99:13368–13379.
35. Wachsmuth, M., W. Waldeck, and J. Langowski. 2000. Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *J. Mol. Biol.* 298:677–689.
36. Bacia, K. 2005. Dynamic processes in membranes studied by fluorescence correlation spectroscopy. Technische Universität Dresden, PhD thesis.
37. Petrásek, Z., and P. Schwille. 2008. Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. *Biophys. J.* 94:1437–1448.
38. Perez-Romero, C. A., H. Tran, …, N. Dostatni. 2018. Live imaging of mRNA transcription in Drosophila embryos. *In* Morphogen Gradients J. Dubrulle, ed.. Springer, pp. 165–182.
39. Patel, A. L., L. Zhang, …, S. Y. Shvartsman. 2021. A fast-acting transcriptional brake in the Drosophila embryo. *Curr. Biol.* Published online June 23, 2021. https://doi.org/10.1016/j.cub.2021.05.061.
40. Porcher, A., A. Abu-Arish, …, N. Dostatni. 2010. The time to measure positional information: maternal hunchback is required for the

synchrony of the Bicoid transcriptional response at the onset of zygotic transcription. *Development.* 137:2795–2804.

41. Vukojević, V., D. K. Papadopoulos, …, R. Rigler. 2010. Quantitative study of synthetic Hox transcription factor-DNA interactions in live cells. *Proc. Natl. Acad. Sci. USA.* 107:4093–4098.

42. Grimm, O., V. Sanchez Zini, …, E. Wieschaus. 2012. Torso RTK controls Capicua degradation by changing its subcellular localization. *Development.* 139:3962–3968.

43. Eggeling, C., J. Widengren, …, C. A. Seidel. 1998. Photobleaching of fluorescent dyes under conditions used for single-molecule detection: evidence of two-step photolysis. *Anal. Chem.* 70:2651–2659.

44. Dittrich, P., and P. Schwille. 2001. Photobleaching and stabilization of. fluorophores used for single-molecule analysis. with one-and two-photon excitation. *Appl. Phys. B.* 73:829–837.

45. Satsoura, D., B. Leber, …, C. Fradin. 2007. Circumvention of fluorophore photobleaching in fluorescence fluctuation experiments: a beam scanning approach. *Chemphyschem.* 8:834–848.

46. Hess, S. T., and W. W. Webb. 2002. Focal volume optics and experimental artifacts in confocal fluorescence correlation spectroscopy. *Biophys. J.* 83:2300–2317.

47. Nishimura, G., and M. Kinjo. 2004. Systematic error in fluorescence correlation measurements identified by a simple saturation model of fluorescence. *Anal. Chem.* 76:1963–1970.

48. Enderlein, J., I. Gregor, …, U. B. Kaupp. 2005. Performance of fluorescence correlation spectroscopy for measuring diffusion and concentration. *ChemPhysChem.* 6:2324–2336.

49. Nagy, A., J. Wu, and K. M. Berland. 2005. Characterizing observation volumes and the role of excitation saturation in one-photon fluorescence fluctuation spectroscopy. *J. Biomed. Opt.* 10:44015.

50. Berland, K. M., P. T. So, …, E. Gratton. 1996. Scanning two-photon fluctuation correlation spectroscopy: particle counting measurements for detection of molecular aggregation. *Biophys. J.* 71:410–420.

51. Friaa, O., M. Furukawa, …, C. Fradin. 2013. Optimizing the acquisition and analysis of confocal images for quantitative single-mobile-particle detection. *ChemPhysChem.* 14:2476–2490.

Chapter 4

Paper 2: Cic as a fast transcriptional brake

The following manuscript in this chapter is a published article in Current Biology:

Aleena L. Patel, Lili Zhang, Shannon E. Keenan, Christine A. Rushlow, Cécile Fradin, Stanislav Y. Shvartsman, Capicua is a fast-acting transcriptional brake, Current Biology, Volume 31, Issue 16, 2021, Pages 3639-3647.e5, ISSN 0960-9822, https://doi.org/10.1016/j.cub.2021.05.061.

Author contribution:

As a second author in this published work, L.Z. performed the FRAP and FCS experiments, analysed the FRAP and FCS data, helped with the sections of the paper related to FRAP and FCS, and prepared related figures: Fig. 1 in the main paper and Fig. 1 in the supplemental information. A.L.P., C.F., and S.Y.S. generated the ideas and designed the experiments with input from C.A.R. A.L.P. and L.Z. performed the optogenetics experiments. S.E.K. generated the hkb MS2 reporter line. A.L.P. wrote

the paper with input from C.A.R., C.F., and S.Y.S.

Research Background:

Capicua (Cic), a repressor morphogen, inhibits the transcription of a number of target genes in the early fly embryo. Cic only acts as a repressor in its dephosphorylated form. When phosphorylated, it is exported from nuclei and degraded in the cytoplasm, and stops acting as a repressor. Molecular mechanisms of gene repression are still poorly understood. In particular, it remains unclear how quickly a repressor such as Cic can stop gene transcription.

Research Purpose:

My goal for this second project was to use the methods described in the previous chapter to measure Cic concentration and diffusion in order to better understand the dynamic aspects of transcriptional repression by Cic.

Methods: My contribution was to perform Fluorescence Correlation Spectroscopy (FCS), and Fluorescence Recovery After Photobleaching (FRAP) experiments. Optogenetic perturbation experiments were performed by our collaborators.

Key points drawn from my data include:

- Cic concentration in the nuclei found in the middle of *D. melanogaster* embryos at NC 14 is $C \approx 250$ nM, corresponding to $\sim$ 20,000 Cic molecules per nucleus. At the poles, this parameter is reduced to 2,000 Cic molecules per nucleus.
- The effective diffusion coefficient of Cic is $D \approx 20$ $\mu m^2/s$.

- From the values of C and D, Cic average search time for its target DNA sequence in a target gene promoter region was estimated using Smoluchowski equation which assumes a simple diffusive search, and found to be 2.5 - 5.5 ms for C = 230 - 260 nM and D = 15 - 25 µm^2/s.

Significance: The optogenetic experiments carried out by our collaborators show that Cic is a fast transcriptional brake that can exert its repressive function on its target genes in only a few minutes. Our estimate of Cic nuclear concentration (at least 2,000 Cic molecules per nucleus) shows that there are enough Cic molecules to bind to all the Cic-binding loci in the *Drosophila* genome. Our estimate of Cic search time (a few ms) shows that many molecules of Cic approach these regions in the time necessary to apply the transcriptional brake (a few mins). Therefore, neither Cic concentration or mobility is a limiting factor for Cic repression.

Report

Current Biology

Capicua is a fast-acting transcriptional brake

Highlights

- Fluorescence correlation spectroscopy measurements estimate Capicua search time
- Optimized photoswitchable MEK provides direct and reversible Capicua control
- Transcription halts minutes after removing optogenetic signals antagonizing Capicua
- Gene repression by Capicua interferes with the emergence of transcriptional bursts

Authors

Aleena L. Patel, Lili Zhang, Shannon E. Keenan, Christine A. Rushlow, Cécile Fradin, Stanislav Y. Shvartsman

Correspondence

fradin@physics.mcmaster.ca (C.F.), stas@princeton.edu (S.Y.S.)

In brief

Patel et al. report the timescale of transcriptional repression by Capicua in the early *Drosophila* embryo. Biophysical properties of Capicua are measured in interphase nuclei. Optimized photoswitchable MEK toggles gene repression. Optogenetic perturbations and live reporters of nascent transcript production reveal fast-acting gene repression.

Patel et al., 2021, Current Biology *31*, 3639–3647
August 23, 2021
https://doi.org/10.1016/j.cub.2021.05.061

CellPress

Current Biology

CellPress

Report

Capicua is a fast-acting transcriptional brake

Aleena L. Patel,[1,2,3] Lili Zhang,[5] Shannon E. Keenan,[1,2,3] Christine A. Rushlow,[4] Cécile Fradin,[5,7,*] and Stanislav Y. Shvartsman[2,3,6,8,*]

[1]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08540, USA
[2]Department of Molecular Biology, Princeton University, Princeton, NJ 08540, USA
[3]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08540, USA
[4]Department of Biology, New York University, New York, NY 10003, USA
[5]Department of Physics and Astronomy, McMaster University, Hamilton, ON L8S 4K1, Canada
[6]Center for Computational Biology, Flatiron Institute, New York, NY 10010, USA
[7]Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, ON L8S 4K1, Canada
[8]Lead contact
*Correspondence: fradin@physics.mcmaster.ca (C.F.), stas@princeton.edu (S.Y.S.)
https://doi.org/10.1016/j.cub.2021.05.061

SUMMARY

Even though transcriptional repressors are studied with ever-increasing molecular resolution, the temporal aspects of gene repression remain poorly understood. Here, we address the dynamics of transcriptional repression by Capicua (Cic), which is essential for normal development and is commonly mutated in human cancers and neurodegenerative diseases.[1,2] We report the speed limit for Cic-dependent gene repression based on live imaging and optogenetic perturbations in the early *Drosophila* embryo, where Cic was originally discovered.[3] Our measurements of Cic concentration and intranuclear mobility, along with real-time monitoring of the activity of Cic target genes, reveal remarkably fast transcriptional repression within minutes of removing an optogenetic de-repressive signal. In parallel, quantitative analyses of transcriptional bursting of Cic target genes support a repression mechanism providing a fast-acting brake on burst generation. This work sets quantitative constraints on potential mechanisms for gene regulation by Cic.

RESULTS AND DISCUSSION

Transcriptional repressors provide molecular brakes on gene expression circuits at key moments in time and in precise spatial patterns during embryogenesis and homeostasis.[4–6] The high-mobility group (HMG)-box transcription factor and repressor Capicua (Cic) regulates cell fate decisions during development and acts as a tumor suppressor in adult tissues.[2,7,8] From fruit flies to humans, Cic mediates inductive receptor tyrosine kinase (RTK) signaling.[1,9–11] In the absence of RTK signals, Cic represses target genes, many of which are known oncogenes involved in cell proliferation.[12] Exposing cells to ligands that bind RTKs activates the extracellular signal-regulated kinase (ERK) cascade, which counteracts Cic repression, analogous to releasing the brake pedal of a car, to induce target gene transcription.[5,13]

The current quantitative models for Cic-dependent gene control rely on studies of the initial pulse of RTK activation in the *Drosophila* blastoderm, during 13 synchronous nuclear divisions spanning the first 2 h post-fertilization (hpf).[5,14] In this time window, Cic regulates expression of *tailless* (*tll*) and *huckebein* (*hkb*), genes required for distinguishing the head and tail from mid-body segmented structures of the emerging larva (Figure 1A).[15–17] A bipartite structure formed between the HMG-box and a C1 domain allows Cic to specifically recognize the conserved octameric DNA binding site "TGAATGAA" in the regulatory elements of *tll* and *hkb*.[9,18] Activation of ERK by locally produced ligands at the anterior and posterior poles phosphorylates Cic to relieve repression, causing Cic unbinding from DNA and export from the nucleus for eventual degradation.[19,20] Consequently, Cic de-repression is described as a two-step process: fast relief of repression upon Cic phosphorylation and unbinding from DNA, followed by slower changes in Cic subcellular localization and stability.

We currently lack such a detailed and quantitative understanding of the molecular mechanisms for establishing repression by Cic. It is particularly important to address this question in contexts where Cic levels are depleted by active and sustained ERK signaling. Signal-dependent control of Cic concentration might be a physiologically important mechanism for long-term memory of ERK activation; the slower steps of Cic de-repression may deplete enough repressor to sustain transcription after ERK signals are removed, a hypothesis that we are interested in testing. We manipulated Cic function with spatially uniform optogenetic signal perturbations in nuclei found in the middle of the *Drosophila* embryo after the 13th mitotic division (interphase of nuclear cycle "nc" 14). Prior studies of short (5-min) pulsed optogenetic signals suggest that this tissue has the potential to reveal the fastest timescales of transcriptional control, but these perturbations were insufficient to access the timescale of establishing a repressed state from scratch.[20] Here, we directly determine the speed limit of *de novo* Cic-responsive gene repression with hours-long sustained and step-like perturbations using a photoswitchable form of ERK's kinase, MEK.[21]

The optogenetic tool we used, optimized photoswitchable MEK (psMEK), activates ERK to at least the endogenous

CellPress

Current Biology
Report

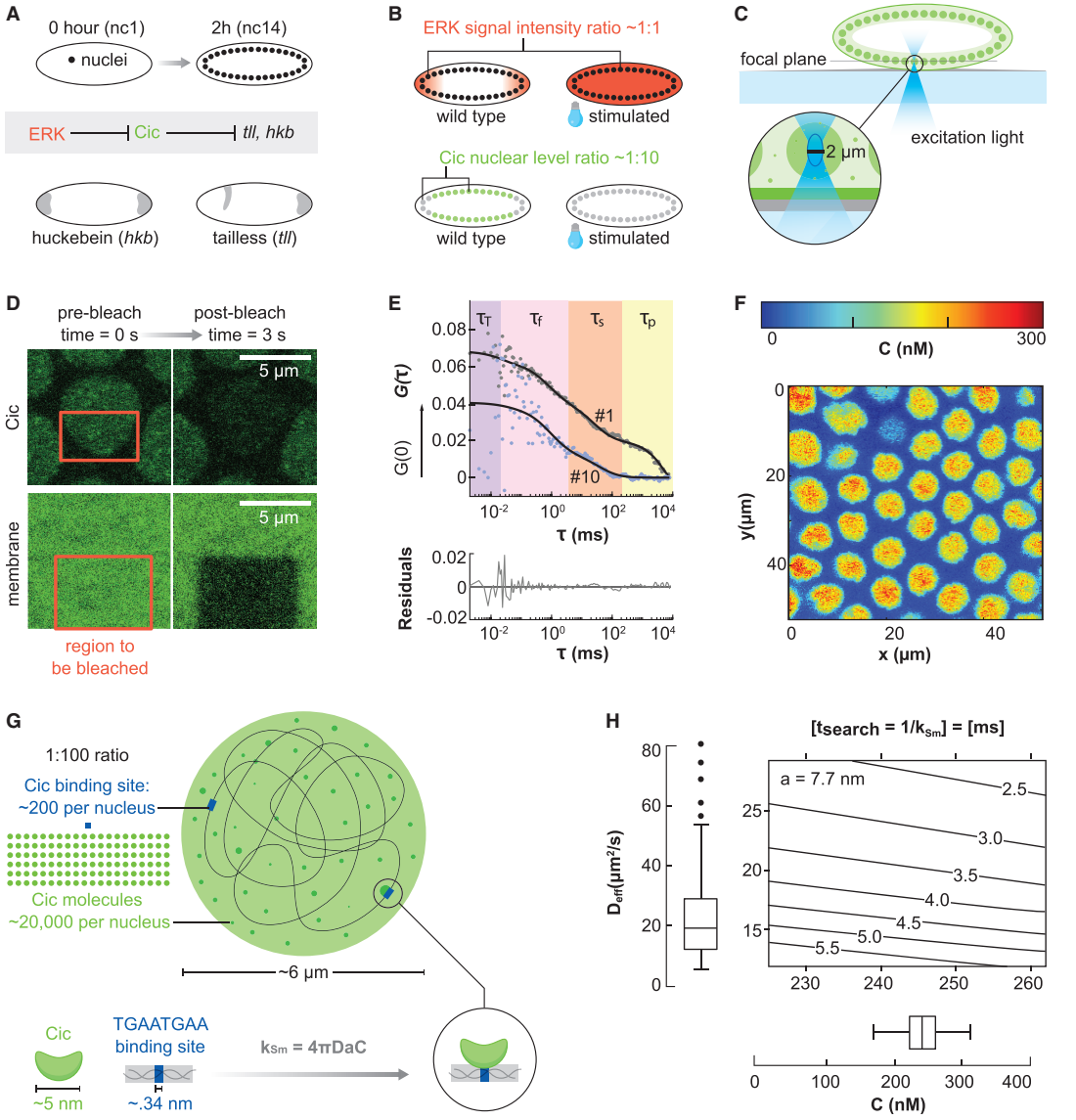

Figure 1. Measurements of Cic concentration and mobility in interphase nuclei

(A) 13 synchronous nuclear divisions (14 nuclear cycles, "nc") take place in a shared cytoplasm during the first 2 h of embryogenesis. The nuclei migrate to the periphery of the embryo, forming a uniform layer. In this time window, active ERK (red) signals antagonize Cic (green), a transcription factor that represses expression of the genes *tailless* (*tll*) and *huckebein* (*hkb*). Wild-type expression domains of *tll* and *hkb* are shown in gray.

(B) ERK is endogenously active at the poles (shown in red). Illuminating embryos expressing the optogenetic signaling tool optimized photoswitchable MEK activates ERK uniformly in the middle as well as at the poles. Previous quantifications suggest that levels of optogenetically activated ERK in the middle are at least equivalent to the levels at the anterior pole. Cic (green) is de-repressed in the poles of wild-type embryos, where ERK is active. Previous studies reported an ~10-fold reduction of Cic in ERK-activated nuclei at the poles compared to ERK-free nuclei in the middle. Optogenetic ERK signals that are at least as strong as endogenous ERK signals are expected to also de-repress Cic.

(C) Schematic of the confocal setup used to image Cic endogenously tagged with the fluorescent marker sfGFP in nuclei from the middle of an embryo. The confocal detection volume (dark blue ellipse in the inset) is smaller than a nucleus.

(legend continued on next page)

Current Biology
Report

active ERK levels at the poles, which have been shown to reduce nuclear Cic concentration about 10-fold (Figure 1B).[19,22] Knowing this, we measured Cic diffusivity and concentration in nuclei free of endogenous ERK signals in the middle of embryos to quantitatively gauge how rate limiting the mobility parameters might be during the early stages of embryogenesis. Cic endogenously tagged with superfolder GFP (sfGFP) was imaged via confocal microscopy (Figure 1C). Fluorescence recovery after photobleaching (FRAP) established that Cic-sfGFP molecules became uniformly distributed less than 3 s after photobleaching a portion of the nucleus, suggesting that there is no significant immobile fraction on the ~1-s timescale (Figure 1D). Fluorescence correlation spectroscopy (FCS) revealed two populations of Cic molecules: a fast-diffusing fraction with a residence time in the confocal detection volume of less than 1 ms (median diffusion coefficient 32 $\mu m^2/s$) and a slower moving population with a residence time of about 60 ms (Figure 1E). The slower population could consist of molecules that are transiently part of larger molecular complexes or phase-separated repressive droplets (moving with an apparent diffusion coefficient of 0.4 $\mu m^2/s$). It could also indicate transient DNA binding on the 60-ms timescale, which would be relevant for transcriptional repression.[23] No matter what the interpretation for this slowing down of a fraction of the Cic population is, the effective mobility of Cic (taking into account both the fast and slow population) can be given as D_{eff} = 20 $\mu m^2/s$ (median value for all measurements in nc 13 and 14).

Cic is not only quickly diffusing but also abundant compared to the number of binding loci in the genome. Concentration could be measured from the FCS experiments (Figure S1). In nc 14, the measured concentration was ~240 nM (Figure 1F). Because each nc 14 nucleus has a radius of ~3 μm, there are ~20,000 Cic molecules per nucleus, which is roughly two orders of magnitude larger than the number of Cic-binding loci in the *Drosophila* genome identified by chromatin immunoprecipitation sequencing (ChIP-seq) (Figure 1G).[20] Optogenetically ERK-activated nuclei would have around 2,000 molecules, putting Cic concentration in a range that may allow for sensitive and rapid switching of the transcriptional state.[19,24]

Our measurements of Cic diffusivity and concentration provide an estimate of the time for Cic to search for its binding site in the nucleus via the Smoluchowski equation (Figure 1G).[25] As a demonstration, we considered the target length "a" to be 7.7 nm (the sum of the size of a protein and an octameric binding site) and plotted the resulting search times using a range of plausible values for D_{eff} (12–29 $\mu m^2/s$) and *C* (225–262 nM), given the dispersion in our measurements (Figure 1H). The range of search times was ~3–6 ms and remains in the sub-second timescale, even when the length constraint is relaxed (3 ms for D = 23 $\mu m^2/s$, a = 5 nm, and C = 290 nM and 150 ms for D = 12 $\mu m^2/s$, a = 0.34 nm, and C = 130 nM). Reducing Cic levels by an order of magnitude with optogenetic ERK signals would still result in a search time on the sub-second timescale. These estimates provide a hypothesis that Cic represses transcription quickly if a strong optogenetic ERK signal is removed. However, mobility measurements of Cic do not address other features of transcription factor searching, such as local access to binding sites regulated by chromatin architecture or interactions with additional regulators like the co-repressor Groucho.[3,26] Thus, the biologically relevant response, transcription of Cic-responsive target genes, must be measured.

Here, we used the MS2-MCP (MS2 coat protein) system in *Drosophila* to report nascent mRNA production while optogenetically toggling ERK activation in nc 14 with optimized psMEK (Figure 2A).[27] As MS2 stem loops genetically engineered into the reporters are transcribed, MCP fluorescently tagged with mCherry binds to the loops. Concentrated mCherry is visible via confocal microscopy as a bright spot in nuclei. The optimized psMEK tool conveniently circumvents potential delays in signal transduction via upstream components of the pathway, as it directly phosphorylates ERK, and acts only one node away from Cic in the signal transduction cascade. 500 nm light activates optimized psMEK by dissociating domains that sterically hinder MEK's active site. 400 nm light inactivates it by closing the domains over the active site.[21,22,28]

(D) Dynamics of Cic molecules observed via fluorescence recovery after photobleaching (FRAP). Partial bleaching of a nucleus shows that all Cic molecules are mobile on the ~1-s timescale. Photobleaching of the embryo's vitelline membrane (lower panels) demonstrates that the bleach area is visible and well defined for immobile fluorophores. The bleach area is delineated in red. Right panels show the same samples 3 s after the beginning of the photobleaching step.

(E) Dynamics of diffusing Cic molecules as observed by fluorescence correlation spectroscopy (FCS). The first (no. 1) and the last (no. 10) autocorrelation functions in a representative series of 10 successive measurements obtained in a single nucleus are shown. Data (dots) were fit with a model (lines) that includes two diffusing Cic populations (characterized by times τ_f and τ_s), a photobleaching term (relaxation time τ_P), and a photophysics term (relaxation time τ_T). The lower panel shows the residuals of this fit for curve no. 1.

(F) Concentration heatmap calculated from the pixel intensities of a typical confocal image acquired in the middle of a Cic-sfGFP embryo, using the values of the background noise (I_B) and molecular brightness (*B*) extracted from FCS measurements. The nuclear Cic-sfGFP concentration is uniform among the nuclei in this field of view.

(G) From concentration, the number of Cic molecules per nucleus is estimated to be ~20,000 (assuming nuclei are spherical with a radius of 3 μm). There are ~200 Cic binding loci in the genome from a Cic ChIP-seq study. The measured concentration and diffusivity values are parameters that can be used to estimate the time for a single Cic molecule to search for and find its target region on the DNA. The Smoluchowski equation ($k_{Sm} = 4\pi DaC$) describes this rate constant in terms of diffusivity "D" and concentration "C" as well as a characteristic length scale "a." This length scale can be set by the average size of a protein (5 nm) and an estimate for the size of a base pair (0.34 nm).

(H) A range of search times ($1/k_{Sm}$) based on the Smoluchowski equation are shown (timescale is milliseconds). Box and whisker plots of the values measured for the effective diffusion coefficient (D_{eff}) and concentration (C) of nuclear Cic-sfGFP are shown on the y and x axis, respectively. For D_{eff}, each point in the dataset is the result of a single FCS measurement acquired during either nc 13 (30 measurements) or nc 14 (50 measurements), in 8 different nuclei. For C, each point in the dataset represents the average concentration in a single nucleus during early nc 14 (693 nuclei in total). The 25%–75% interval was used in the search time estimation shown in the contour plot. 7.7 nm was used as a fixed length scale "a," the sum of the average size of a protein and the Cic binding site "TGAATGAA" (0.34 nm/bp × 8 bp = 2.7 nm).

See also Figure S1.

CellPress

Current Biology
Report

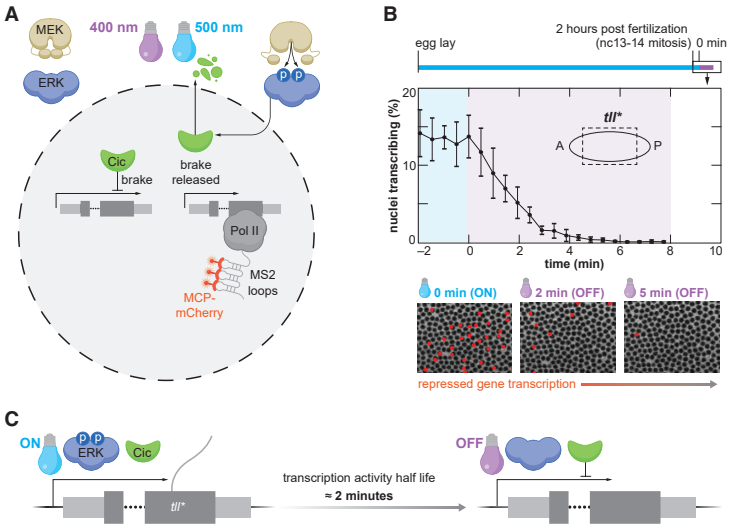

Figure 2. Transcriptional readouts of optogenetic perturbations reveal fast repression by Cic

(A) Optimized photoswitchable MEK (psMEK) reversibly controls phosphorylation (indicated by "P" in blue circles) and activation of ERK. 500-nm light dissociates photo-dimerizable Dronpa domains flanking the active site of MEK containing activating mutations, thereby allowing MEK to access its substrate, ERK. 400 nm illumination dimerizes the domains over the active site of MEK, blocking MEK-ERK interaction. This light-sensitive ERK-activating tool was genetically combined with a transgenic system for reporting live transcription with MS2-MCP reporters. Fluorescently tagged MCP-mCherry (red) binds to MS2 stem loops as RNA polymerase (Pol II) transcribes genes. Altogether, optimized psMEK and the MS2-MCP system enable real-time optogenetic control of the Cic transcriptional brake alongside measurements of the immediate transcriptional responses.

(B) Embryos were illuminated with activating light from the time of egg lay to nc 14, which spans approximately 2 h. Light was switched from activating (blue) to inactivating (purple) 2 min into nc 14, after the completion of the nc 13 to nc 14 mitosis. 0 min marks the time of photoswitching. Percentage of nuclei with a fluorescent dot indicating binding of MCP-mCherry to MS2 loops driven by fragments of the *tll* enhancer (*tll**) in the middle of the embryo are shown. Snapshots from the middle of the embryo at 0, 2, and 5 min after optimized psMEK inactivation are shown with transcriptionally active nuclei marked by red dots. Error bars represent one standard deviation (n = 5 embryos).

(C) The half-life of transcription activity when optimized psMEK is switched off, following long-term activation, is ~2 min.

See also Figure S2.

psMEK perturbations confirm that transcription repression is established very quickly, even after sustained illumination and ERK activation from the time of egg lay (spanning ~2 h). We combined the tool with an MS2 reporter for *tll* that contains fragments of the regulatory DNA (*tll**) used in previous studies.[20] Embryos were illuminated with psMEK-activating light from egg laying to nc 14. MS2 transcriptional activity was sustained with continued illumination for 2 min after the completion of the nc 13 to nc 14 mitosis. Signaling was then abruptly terminated by switching illumination to the psMEK-inactivating wavelength. *tll** transcription, quantified as the percent of the nuclei in the field of view with an MS2-MCP spot per time frame, declined rapidly upon inactivating optimized psMEK (Figure 2B). Transcriptional repression by Cic occurs within minutes of removing the sustained ERK signal (Figure 2C).

The MS2-MCP imaging reveals a highly regulated sequence of events, called "bursting," which reflects periods of active mRNA generation followed by transcriptional quiescence.[29,30] Fluctuating signals were characterized for transcriptional states established with two optogenetic perturbations: continuous and photoswitched MEK activation in nuclei from the middle of embryos in nc 14. MS2 loops inserted via CRISPR near the *hkb* gene body reported endogenous activity of an ERK target gene other than *tll* (Figure 3A). We illuminated embryos with psMEK-activating light from the time of egg lay to fully derepress transcription. For each time point, multiple bright foci of mCherry fluorescence, indicating MCP binding to MS2 loops, were detected. The maximum recorded spot intensity was plotted over time to display how transcriptional activity appeared throughout nc 14 (Figure 3Bi). Individually tracked spots revealed discontinuous transcription with several intensity peaks throughout nc 14, which collectively contributed to signal detection for at least 15 min (Figure 3Bii). In other embryos, we switched illumination to the psMEK-inactivating wavelength immediately after the nc 13 to nc 14 mitosis, thereby allowing Cic to repress transcription. In the movies of embryos subject to repression by Cic, because of the rapidly removed optogenetic ERK signal, the maximum spot intensity dropped dramatically after 5 min (Figure 3Ci). Individually tracked spots in this short time window of repression appeared to reach only one peak, indicative of single bursts (Figure 3Cii). Thus, Cic is a fast-acting brake on endogenous gene transcription that appears to limit bursting to an ~5-min time window.

To test whether the fast-acting brake is dependent on the presence of Cic binding sites, we performed similar perturbations and spot quantifications in embryos expressing a newly constructed RTK-sensitive reporter. We introduced 4 Cic binding sites (TGAATGAA) near the regulatory region of *bottleneck* (*bnk*), which does not contain Cic sites, driving MS2 loops (Figure 3D). In embryos expressing this reporter, MS2 activity was only apparent at the poles, reflecting endogenous ERK signals (Figure S2A). A similar reporter constructed with only one intact Cic site and three mutated Cic sites expressed uniformly (Figure S2B), suggesting that the 4 Cic binding sites were important for the observed restricted expression at the poles.

Activating optimized psMEK to continually lift the repressive Cic brake on transcription in embryos with the *bnk* + 4 Cic sites reporter led to sustained transcription in the middle of the embryo throughout nc 14 (Figure 3E). The apparent decline in maximum spot intensity could reflect a direct readout of falling

Current Biology
Report

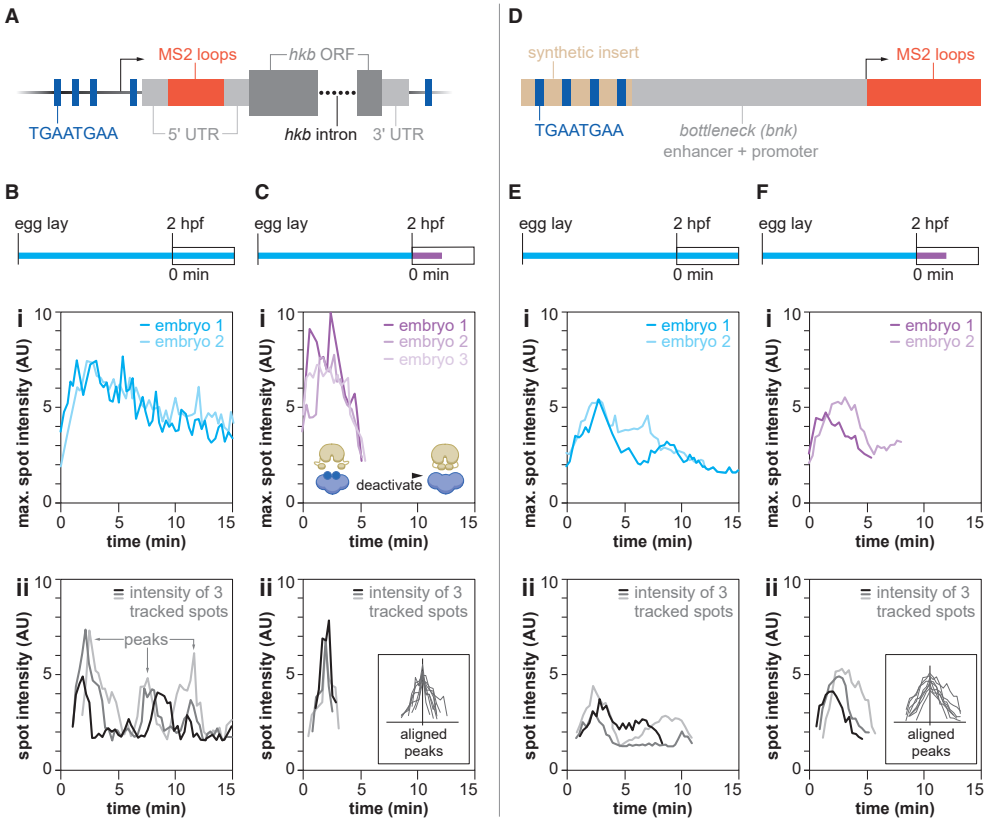

Figure 3. Photoswitching MEK limits the transcriptional bursting time window

(A) MS2 loops were inserted at the endogenous *hkb* locus via CRISPR. Thick dark gray bars designate the open reading frames (ORFs), thin dark gray bars are 3′ or 5′ UTRs, and dotted lines are introns. The *hkb* regulatory region has 5 strong Cic binding sites "TGAATGAA" (blue bars).

(B) Transcription responses in the middle of embryos expressing the endogenous *hkb* reporter and continuously illuminated with light that activates the optogenetic ERK signal. The colored bar represents the optogenetic illumination schedule, and the box represents the data collection time window. Developmental time is indicated above the light schedule. Activating light (blue) was provided from the time of egg lay throughout nc 14. Data collection starts at 0 min, which corresponds to 0 min in each of the plots below. (i) For each time frame, the maximum recorded spot intensity is plotted for 2 embryos. (ii) Examples of intensity time series from 3 individually tracked spots are shown and apparent burst "peaks" are denoted.

(C) Transcription responses in the middle of embryos expressing the endogenous *hkb* reporter and exposed to a change in optogenetic illumination wavelength after the nc 13 to nc 14 mitosis. Activating light (blue) was provided from the time of egg lay to the nc 13 to nc 14 mitosis (~2 hpf). Optogenetic illumination was then immediately switched to the inactivating wavelength (purple). The data collection time window starts at time of the optogenetic switch. (i) For each time frame, the maximum recorded spot intensity is plotted for 3 embryos. (ii) Examples of intensity time series from 3 individually tracked spots are shown. 10 tracked spots aligned by their maximum intensities are shown in the inset.

(D) A Cic-dependent reporter was constructed by inserting four strong Cic binding sites "TGAATGAA" (dark blue) spaced 21 bp apart directly upstream of the *bnk* enhancer and promoter driving MS2 loops.

(E) Transcription responses in the middle of embryos expressing the modified *bnk* reporter and continuously illuminated with the activating light (blue). The optogenetic light schedule was identical to the continuous light schedule for the *hkb* reporter described in (B). (i) Maximum recorded spot intensity per time frame for 2 embryos is shown. (ii) Example intensity time series from 3 individually tracked spots is shown.

(F) Transcription responses in the middle of embryos expressing the modified *bnk* reporter and exposed to a change in optogenetic illumination wavelength after the nc 13 to nc 14 mitosis. The optogenetic light schedule was identical to the photoswitched light schedule for the *hkb* reporter described in (C). (i) The maximum recorded spot intensity per time frame for 2 embryos is plotted. (ii) Example intensity time series from 3 individually tracked spots is shown, with 10 time series aligned by maximum intensity in the inset.

See also Figure S3.

levels of the transcriptional activator, Zelda, known to regulate *bnk*, as the embryo undergoes the maternal to zygotic transition (MZT).[31] A control reporter with no intact Cic sites added upstream of the *bnk* enhancer also showed a decline in transcription (Figure S3). Even for such a reporter that may be very sensitive to dynamics of the activator, removing the optogenetic ERK signal reapplied the Cic transcriptional brake quickly (Figure 3F).

CellPress
Current Biology
Report

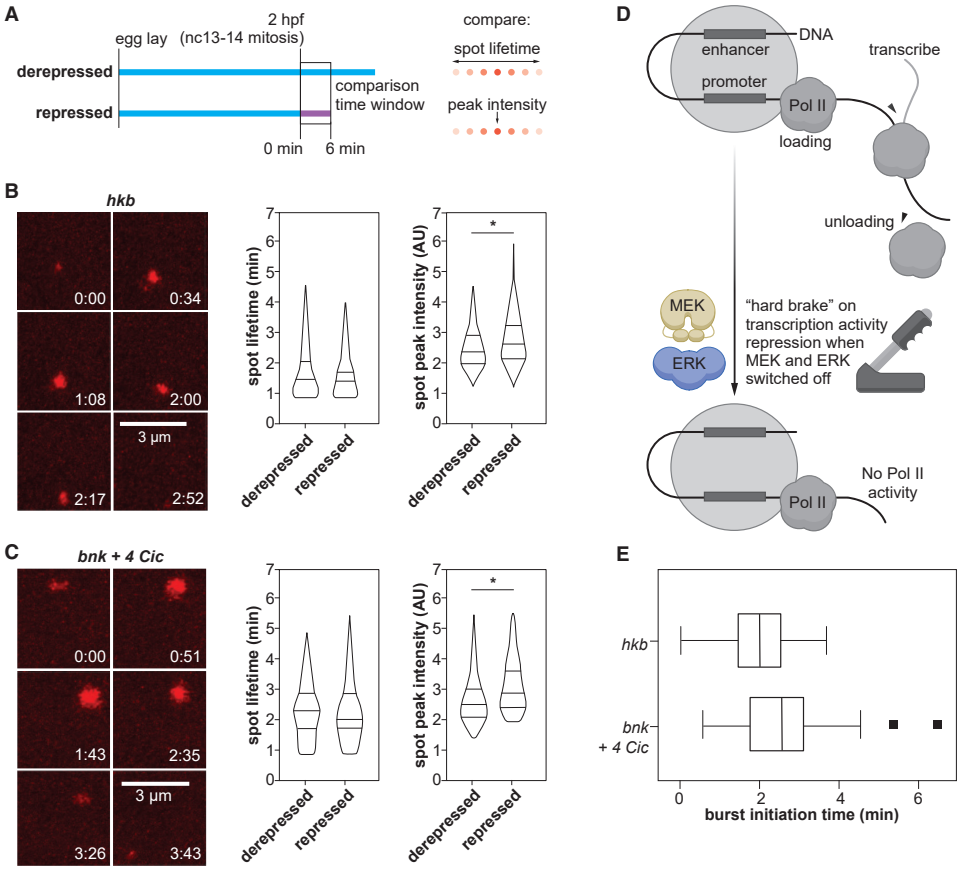

Figure 4. Repressed and de-repressed bursts are similar

(A) Tracked spots representing apparent transcriptional bursts were considered from the first 6 min of movies taken from embryos that were subject repression by Cic (optimized psMEK switched off at 0 min, purple) or de-repressed (continuous optimized psMEK activation, blue).

(B) Tracked spots were pooled for the de-repressed and repressed optogenetic conditions, respectively, from embryos expressing the *hkb* reporter. The lifetimes and peak intensities of the tracked spots for each condition are shown. The distributions of these values are displayed as truncated violin plots, with the median and quartiles represented (horizontal lines). *$p < 0.05$ for a 2-sample Kolmogorov-Smirnov test (KS test statistic for spot lifetimes = 0.094; KS test statistic for spot peak intensities = 0.161). Pooled numbers of spots used to generate the distributions were $n_{de\text{-}repressed} = 298$ and $n_{repressed} = 340$.

(C) Tracked spots were pooled for the de-repressed and repressed optogenetic conditions, respectively, from embryos expressing the modified *bnk* + 4 Cic sites reporter. The distributions of the spot lifetimes and peak intensities for each condition are shown. *$p < 0.05$ for a 2-sample Kolmogorov-Smirnov test (KS test statistic for spot lifetimes = 0.078; KS test statistic for spot peak intensities = 0.251). Pooled numbers of spots used to generate the distributions were $n_{de\text{-}repressed} = 220$ and $n_{repressed} = 106$.

(D) A "hard brake" on ERK signaling enabled by photoswitching MEK leads to severely interrupted transcription burst generation of Cic-responsive target genes. A hard brake on transcription may involve multiple factors that interact with enhancers and promoters and would prevent polymerase activity from being initiated.

(E) Box and whisker plots of the apparent burst initiation times, defined as the time it takes for a spot to be first detected after switching optimized psMEK off at the start of nc 14, for the *hkb* and *bnk* + 4 Cic sites reporters.

See also Figure S4.

We then compared quantitative features of the apparent bursts of transcription subjected to repression by Cic in nc 14 with bursts subjected to continued ERK activation (de-repressed transcription). Tracked fluorescent spots were extracted from the first 6 min of nc 14, to fairly compare the two conditions (Figure 4A). Here, spot tracking did not allow for periods of transcriptional quiescence, so sequential "bursts" in the same nucleus were recorded as multiple individually tracked spots. The lifetimes and maximum intensities (peaks) of spots from movies of embryos treated under the same optogenetic conditions (repressed or de-repressed transcription) were pooled to generate distributions for each parameter. For *hkb*, the spot lifetimes, corresponding to the apparent burst durations, of repressed and de-repressed bursts were found to belong to

Current Biology
Report

the same distribution by a Kolmogorov-Smirnov test. Spot peak intensity distributions were significantly different, but repression did not drastically dim bursts (Figure 4B).

Similar spot measurements performed for the modified *bnk* + 4 Cic sites reporter, which uses a different MS2 stem loop sequence, also suggest that Cic repression does not drastically alter individual bursts (Figure 4C).[32] To further ensure that our measurements were from a reporter that is sensitive enough to our optogenetic perturbation, we compared apparent bursts from the continually de-repressed nuclei to apparent bursts from the *bnk* reporter with no intact Cic binding sites (Figure S4). De-repressed bursts from the *bnk* + 4 Cic sites reporter were similar to bursts from the non Cic-dependent *bnk* reporter, indicating that the ERK signals are revealing the expected transcriptional activity. In sum, repression by Cic does not appear to change the nature of individual transcriptional bursts.

Bursts may remain unchanged upon repression by Cic if an immediate removal of the de-repressive ectopic ERK signal, or "hard brake" enabled by photoswitching MEK, only interferes with transcription burst generation (Figure 4D). We define a potential hard brake on transcription following the photoswitch of MEK and ERK as a mechanism that affects the onset of transcription but would not truncate nascent transcripts mid-elongation. This interpretation is consistent with previous work suggesting transcription initiation and RNA polymerase II (RNA Pol II) release are key points of gene control in this embryo and for mammalian genes.[33,34] If Cic primarily regulates burst generation, there could be a lag in gene silencing due to continued transcriptional elongation after the onset of repression. This scenario has been reported for the Zn-finger repressor Snail, where the time to complete transcriptional quiescence is a function of the size of the gene.[35] The burst initiation time at the start of nc 14 would be a lower bound on the time it takes for Cic to repress ERK target genes if Cic primarily acts to prevent a second burst (Figure 4E). This lower bound is slightly longer for the *bnk* + 4 Cic sites reporter, perhaps reflecting different dynamics of activator inputs to *bnk*.

These temporal bounds place quantitative limits on the molecular mechanisms of repression by Cic. Potential fast-acting mechanisms that would suppress burst generation include enhancer or promoter binding competition with activators or basal factors, disruption of the pre-initiation complex, altered interactions between RNA Pol II and elongation factors, or RNA Pol II pausing, implicated in repression by Cic's co-repressor Groucho.[31,36–40] The currently proposed mechanisms for gene repression by Cic in glioblastoma cells involve histone deacetylation.[41] It has not yet been shown that such epigenetic gene silencing mediates Cic repression in the early *Drosophila* embryo or whether histone deacetylation leads to rapid transcription shutdown. As an illustration, histone deacetylation occurs ~20 min after another mammalian repressor, Ikaros, binds to DNA. Interestingly, quicker changes to promoter accessibility, such as RNA polymerase eviction and altered nucleosome occupancy, silence Ikaros target gene transcription within 5–10 min, before histones are deacetylated.[42] The time window following the ERK photoswitch that we have captured encompasses the initial steps of Cic-mediated repression that rapidly deplete transcript, much like the immediate silencing of Ikaros targets. Here, we have not yet assayed potentially slower chromatin-level effects, such as how Cic-responsive gene loci might be repositioned into repressive micro-environments for more stable repression after loss of transcription.

Our work paints a picture of fast-acting mechanisms controlling nascent transcript generation in response to fluctuating signals in the early *Drosophila* embryo. Importantly, we also bring into question the function of two-step de-repression if Cic export and degradation do not limit the Cic binding search or transcriptional silencing rates. Quickly plunging a gene-regulatory system into a repressive state, as we have done using optogenetic manipulations of Cic, will be crucial for our emerging quantitative understanding of gene regulation relevant to both fundamental biology and disease.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - FRAP
  - FCS
  - Analysis of autocorrelation functions
  - Generation of transgenic flies
  - Combining optimized psMEK, MCP-mCherry, and MS2 reporters
  - Optogenetic illumination and live imaging
  - Quantification of percent nuclei transcribing
  - Spot detection of transcriptional bursts
  - Burst initiation time
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Quantification of transcription activity

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2021.05.061.

ACKNOWLEDGMENTS

We thank Shigehiro Yamada and Mengyu Li for the *bnk* MS2 reporter plasmid and Nareg Djabrayan for the synthetic Cic-responsive MS2 reporter sequence. We also thank Lucy Reading-Ikkanda for graphic design of the figures. We also thank Daniel Larson, Ido Golding, Mike Levine, Nicholas Treen, Gertrud Schüpbach, and Eric Wieschaus for helpful discussions. The research was supported by National Institutes of Health (NIH) research grants, DP2EB024247, F32GM119297, RO1HD0858870, and HD085870 to S.Y.S. and RO1GM63024 to C.A.R., and by Natural Sciences and Engineering Research Council of Canada (NSERC) grant RGPIN-2015-06362 to C.F.

AUTHOR CONTRIBUTIONS

A.L.P., C.F., and S.Y.S. generated the ideas and designed the experiments with input from C.A.R. A.L.P. and L.Z. performed the experiments. S.E.K. generated the *hkb* MS2 reporter line. A.L.P. wrote the paper with input from C.A.R., C.F., and S.Y.S.

CellPress

Current Biology
Report

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 27, 2021
Revised: March 18, 2021
Accepted: May 27, 2021
Published: June 23, 2021; corrected online: July 23, 2021

REFERENCES

1. Kim, J.W., Ponce, R.K., and Okimoto, R.A. (2021). Capicua in human cancer. Trends Cancer *7*, 77–86.
2. Lee, Y. (2020). Regulation and function of capicua in mammals. Exp. Mol. Med. *52*, 531–537.
3. Jiménez, G., Guichet, A., Ephrussi, A., and Casanova, J. (2000). Relief of gene repression by torso RTK signaling: role of capicua in Drosophila terminal and dorsoventral patterning. Genes Dev. *14*, 224–231.
4. Gaston, K., and Jayaraman, P.S. (2003). Transcriptional repression in eukaryotes: repressors and repression mechanisms. Cell. Mol. Life Sci. *60*, 721–741.
5. Patel, A.L., and Shvartsman, S.Y. (2018). Outstanding questions in developmental ERK signaling. Development *145*, dev143818.
6. Reynolds, N., O'Shaughnessy, A., and Hendrich, B. (2013). Transcriptional repressors: multifaceted regulators of gene expression. Development *140*, 505–512.
7. Bunda, S., Heir, P., Metcalf, J., Li, A.S.C., Agnihotri, S., Pusch, S., Yasin, M., Li, M., Burrell, K., Mansouri, S., et al. (2019). CIC protein instability contributes to tumorigenesis in glioblastoma. Nat. Commun. *10*, 661.
8. Lee, H., and Song, J.J. (2019). The crystal structure of Capicua HMG-box domain complexed with the ETV5-DNA and its implications for Capicua-mediated cancers. FEBS J. *286*, 4951–4963.
9. Ajuria, L., Nieva, C., Winkler, C., Kuo, D., Samper, N., Andreu, M.J., Helman, A., González-Crespo, S., Paroush, Z., Courey, A.J., and Jiménez, G. (2011). Capicua DNA-binding sites are general response elements for RTK signaling in Drosophila. Development *138*, 915–924.
10. Cinnamon, E., Helman, A., Ben-Haroush Schyr, R., Orian, A., Jiménez, G., and Paroush, Z. (2008). Multiple RTK pathways downregulate Groucho-mediated repression in Drosophila embryogenesis. Development *135*, 829–837.
11. Jiménez, G., Shvartsman, S.Y., and Paroush, Z. (2012). The Capicua repressor–a general sensor of RTK signaling in development and disease. J. Cell Sci. *125*, 1383–1391.
12. Lee, J.S., Kim, E., Lee, J., Kim, D., Kim, H., Kim, C.J., Kim, S., Jeong, D., and Lee, Y. (2020). Capicua suppresses colorectal cancer progression via repression of ETV4 expression. Cancer Cell Int. *20*, 42.
13. Lemmon, M.A., and Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. Cell *141*, 1117–1134.
14. Goyal, Y., Schüpbach, T., and Shvartsman, S.Y. (2018). A quantitative model of developmental RTK signaling. Dev. Biol. *442*, 80–86.
15. Pignoni, F., Baldarelli, R.M., Steingrímsson, E., Diaz, R.J., Patapoutian, A., Merriam, J.R., and Lengyel, J.A. (1990). The Drosophila gene tailless is expressed at the embryonic termini and is a member of the steroid receptor superfamily. Cell *62*, 151–163.
16. Brönner, G., and Jäckle, H. (1996). Regulation and function of the terminal gap gene huckebein in the Drosophila blastoderm. Int. J. Dev. Biol. *40*, 157–165.
17. Goldstein, R.E., Jiménez, G., Cook, O., Gur, D., and Paroush, Z. (1999). Huckebein repressor activity in Drosophila terminal patterning is mediated by Groucho. Development *126*, 3747–3755.
18. Forés, M., Simón-Carrasco, L., Ajuria, L., Samper, N., González-Crespo, S., Drosten, M., Barbacid, M., and Jiménez, G. (2017). A new mode of DNA binding distinguishes Capicua from other HMG-box factors and explains its mutation patterns in cancer. PLoS Genet. *13*, e1006622.
19. Grimm, O., Sanchez Zini, V., Kim, Y., Casanova, J., Shvartsman, S.Y., and Wieschaus, E. (2012). Torso RTK controls Capicua degradation by changing its subcellular localization. Development *139*, 3962–3968.
20. Keenan, S.E., Blythe, S.A., Marmion, R.A., Djabrayan, N.J., Wieschaus, E.F., and Shvartsman, S.Y. (2020). Rapid dynamics of signal-dependent transcriptional repression by Capicua. Dev. Cell *52*, 794–801.e4.
21. Zhou, X.X., Fan, L.Z., Li, P., Shen, K., and Lin, M.Z. (2017). Optical control of cell signaling by single-chain photoswitchable kinases. Science *355*, 836–842.
22. Patel, A.L., Yeung, E., McGuire, S.E., Wu, A.Y., Toettcher, J.E., Burdine, R.D., and Shvartsman, S.Y. (2019). Optimizing photoswitchable MEK. Proc. Natl. Acad. Sci. USA *116*, 25756–25763.
23. Clauß, K., Popp, A.P., Schulze, L., Hettich, J., Reisser, M., Escoter Torres, L., Uhlenhaut, N.H., and Gebhardt, J.C.M. (2017). DNA residence time is a regulatory factor of transcription repression. Nucleic Acids Res. *45*, 11121–11130.
24. Abu-Arish, A., Porcher, A., Czerwonka, A., Dostatni, N., and Fradin, C. (2010). High mobility of bicoid captured by fluorescence correlation spectroscopy: implication for the rapid establishment of its gradient. Biophys. J. *99*, L33–L35.
25. Atkins, P.W., and De, P.J. (2002). Atkins' Physical Chemistry (Oxford University).
26. Li, X.-Y., Thomas, S., Sabo, P.J., Eisen, M.B., Stamatoyannopoulos, J.A., and Biggin, M.D. (2011). The role of chromatin accessibility in directing the widespread, overlapping patterns of Drosophila transcription factor binding. Genome Biol. *12*, R34.
27. Garcia, H.G., Tikhonov, M., Lin, A., and Gregor, T. (2013). Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning. Curr. Biol. *23*, 2140–2145.
28. Zhou, X.X., Chung, H.K., Lam, A.J., and Lin, M.Z. (2012). Optical control of protein activity by fluorescent protein domains. Science *338*, 810–814.
29. Rodriguez, J., and Larson, D.R. (2020). Transcription in living cells: molecular mechanisms of bursting. Annu. Rev. Biochem. *89*, 189–212.
30. Tunnacliffe, E., and Chubb, J.R. (2020). What is a transcriptional burst? Trends Genet. *36*, 288–297.
31. Liang, H.L., Nien, C.Y., Liu, H.Y., Metzstein, M.M., Kirov, N., and Rushlow, C. (2008). The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. Nature *456*, 400–403.
32. Yamada, S., Whitney, P.H., Huang, S.K., Eck, E.C., Garcia, H.G., and Rushlow, C.A. (2019). The Drosophila pioneer factor Zelda modulates the nuclear microenvironment of a dorsal target enhancer to potentiate transcriptional output. Curr. Biol. *29*, 1387–1393.e5.
33. Bartman, C.R., Hamagami, N., Keller, C.A., Giardine, B., Hardison, R.C., Blobel, G.A., and Raj, A. (2019). Transcriptional burst initiation and polymerase pause release are key control points of transcriptional regulation. Mol. Cell *73*, 519–532.e4.
34. Fukaya, T., Lim, B., and Levine, M. (2017). Rapid rates of Pol II elongation in the Drosophila embryo. Curr. Biol. *27*, 1387–1391.
35. Bothma, J.P., Magliocco, J., and Levine, M. (2011). The snail repressor inhibits release, not elongation, of paused Pol II in the Drosophila embryo. Curr. Biol. *21*, 1571–1577.
36. Soutourina, J. (2018). Transcription regulation by the Mediator complex. Nat. Rev. Mol. Cell Biol. *19*, 262–274.
37. Fukaya, T., Lim, B., and Levine, M. (2016). Enhancer control of transcriptional bursting. Cell *166*, 358–368.
38. Zaborowska, J., Egloff, S., and Murphy, S. (2016). The pol II CTD: new twists in the tail. Nat. Struct. Mol. Biol. *23*, 771–777.
39. Shao, W., and Zeitlinger, J. (2017). Paused RNA polymerase II inhibits new transcriptional initiation. Nat. Genet. *49*, 1045–1051.
40. Chambers, M., Turki-Judeh, W., Kim, M.W., Chen, K., Gallaher, S.D., and Courey, A.J. (2017). Mechanisms of Groucho-mediated repression revealed by genome-wide analysis of Groucho binding and activity. BMC Genomics *18*, 215.

Current Biology
Report

41. Weissmann, S., Cloos, P.A., Sidoli, S., Jensen, O.N., Pollard, S., and Helin, K. (2018). The tumor suppressor CIC directly regulates MAPK pathway genes via histone deacetylation. Cancer Res. *78*, 4114–4125.

42. Liang, Z., Brown, K.E., Carroll, T., Taylor, B., Vidal, I.F., Hendrich, B., Rueda, D., Fisher, A.G., and Merkenschlager, M. (2017). A high-resolution map of transcriptional repression. eLife *6*, 1–24.

43. Hunter, C., and Wieschaus, E. (2000). Regulated expression of nullo is required for the formation of distinct apical and basal adherens junctions in the Drosophila blastoderm. J. Cell Biol. *150*, 391–401.

44. Gratz, S.J., Wildonger, J., Harrison, M.M., and O'Connor-Giles, K.M. (2013). CRISPR/Cas9-mediated genome engineering and the promise of designer flies on demand. Fly (Austin) *7*, 249–255.

45. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

46. Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2006). Engineering and characterization of a superfolder green fluorescent protein. Nat. Biotechnol. *24*, 79–88.

47. Petrásek, Z., and Schwille, P. (2008). Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. Biophys. J. *94*, 1437–1448.

48. Wachsmuth, M., Waldeck, W., and Langowski, J. (2000). Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. J. Mol. Biol. *298*, 677–689.

49. Porcher, A., Abu-Arish, A., Huart, S., Roelens, B., Fradin, C., and Dostatni, N. (2010). The time to measure positional information: maternal hunchback is required for the synchrony of the Bicoid transcriptional response at the onset of zygotic transcription. Development *137*, 2795–2804.

50. Vukojevic, V., Papadopoulos, D.K., Terenius, L., Gehring, W.J., and Rigler, R. (2010). Quantitative study of synthetic Hox transcription factor-DNA interactions in live cells. Proc. Natl. Acad. Sci. USA *107*, 4093–4098.

51. Widengren, J., Mets, U., and Rigler, R. (1995). Fluorescence correlation spectroscopy of triplet states in solution: A theoretical and experimental study. J. Phys. Chem. *99*, 13368–13379.

52. Bacia, K. (2005). Dynamic processes in membranes studied by fluorescence correlation spectroscopy. PhD thesis (Technische Universität Dresden).

53. Koppel, D.E. (1974). Statistical accuracy in fluorescence correlation spectroscopy. Phys. Rev. A *10*, 1938–1945.

54. Rigler, R., Mets, U., Widengren, J., and Kask, P. (1993). Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. Eur. Biophys. J. *22*, 169–175.

CellPress

Current Biology
Report

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: organisms/strains		
Drosophila melanogaster. UAS-psMEKE203K	Patel et al.[22]	N/A
Drosophila melanogaster. MCP-mCherry	Levine lab	N/A
Drosophila melanogaster. tll-MS2	Keenan et al.[20]	N/A
Drosophila melanogaster. Hkb-MS2	This study	N/A
Drosophila melanogaster. Bnk+4Cic-MS2	This Study	N/A
Drosophila melanogaster. P(matα-GAL-VP16)mat67; P(matα-GAL-VP16)mat15	Hunter and Wieschaus[43]	N/A
Drosophila melanogaster. Sp/Cyo ; Dr/Tm3	N/A	N/A
Drosophila melanogaster. Crey;+;D/Tm3	Bloomington Stock 851	RRID: BDSC_851
Drosophila melanogaster. CicsfGFP	Keenan et al.[20]	N/A
Oligonucleotides		
4 Cic sites insert to *bnk*-MS2, 5′-ttcgtttaaacggccgtgaatgaaTATCTATGATCACTAGTCTCG tgaatgaaATGTCAGGAGATCTCCAGTTTAtgaatgaaTTTAC TAAATGAGCTCAGTCGtgaatgaacggccggccagatcca-3′	This paper	N/A
1 Cic site insert to *bnk*-MS2, 5′-ttcgtttaaacggccgTGAAG CTATATCTATGATCACTAGTCTCGTGAAGCTAATGTCA GGAGATCTCCAGTTTATGAAGCTATTTACTAAATGAGC TCAGTCGtgaatgaacggccggccagatcca-3′	This paper	N/A
0 Cic site insert to *bnk*-MS2, 5′-ttcgtttaaacggccgTGAAG CTATATCTATGATCACTAGTCTCGTGAAGCTAATGTCA GGAGATCTCCAGTTTATGAAGCTATTTACTAAATGAG CTCAGTCGTGAAGCTAcggccggccagatcca-3′	This paper	N/A
Guide for *hkb* reporter CRISPR, 5′ CTTCGCGACACTAA ATCACTTGGA 3′	This paper	N/A
Guide for *hkb* reporter CRISPR, 5′ AAACTCCAAGTG ATTTAGTGTCGC 3′	This paper	N/A
5′ homology arm of *hkb* amplification, agtgcatatgt ccgcggccgGGATGGAACACTTGTGATTATGATTTTG	This paper	N/A
5′ homology arm of *hkb* amplification, cccttcgct gaagcaggtggGCCAGTAAAGTTTTTCTCAAGCACC	This paper	N/A
3′ homology arm of *hkb* amplification, tacgaagttata gaagagcaAAGTGATTTAGTGTCGCGAGAGAGC	This paper	N/A
3′ homology arm of *hkb* amplification, gagcctcgag ctgcagaaggGTATGAGTACATGGGCACGAAGATG	This paper	N/A
Recombinant DNA		
pb-phi-bnkMS2	Rushlow lab	N/A
pU6-BbsI-chiRNA	Gratz et al.[44]	RRID: Addgene_45946
pHD-dsRed-24xMS2	Levine lab	N/A
Software and algorithms		
Imaris	https://imaris.oxinst.com/packages	RRID: SCR_007370
FIJI	Schindelin et al.[45]	RRID: SCR_002285
Python	https://www.python.org/	RRID: SCR_008394
MATLAB	https://www.mathworks.com/products/matlab/	RRID: SCR_001622
GraphPad Prism 9	https://www.graphpad.com:443/	RRID: SCR_002798
Ilastik	https://ilastik.org/	RRID: SCR_015246

Current Biology
Report

CellPress

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Stas Shvartsman (stas@princeton.edu).

Materials availability

Fly lines generated in this study are listed in the Key resources table and are available at the laboratory upon request.

Data and code availability

Raw image data is available upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cic-sfGFP/+,[20] UAS-psMEKE203K[22] *tll*-MS2, MCP-mCherry (provided by the Levine lab), *hkb*-MS2 (this study), *bnk+4Cic*-MS2 (this study) and P(matα-GAL-VP16)mat67; P(matα-GAL-VP16)mat15 stocks[43] were used in this study. Cic-sfGFP represents Cic endogenously tagged with superfolder GFP.[46] Flies were kept at room temperature in vials containing a standard mixture of agar, cornmeal, and yeast, provided by the *Drosophila* Media Core Facility within the Princeton Molecular Biology Department. To collect embryos, flies were placed in cages with an agar plate made with apple juice and supplemented with a yeast paste.

METHOD DETAILS

FRAP

FRAP measurements were conducted for embryos endogenously expressing Cic-sfGFP. Embryos laid on agar plates were collected, dechorionated by hand, placed on their lateral side on a microscope coverslip covered with a very thin layer of heptane glue, then covered in halocarbon oil 700. Imaging and photobleaching were performed using a Nikon Eclipse Ti inverted confocal microscope. A rectangular region covering about half of a nucleus in the mid section of the embryo was photobleached at high laser power for 3 s, then the whole nucleus was immediately imaged at low laser power. As a control, the vitelline membrane of the embryo (which contains immobile fluorescence molecules) was also photobleached and imaged in the same conditions as the nucleus.

FCS

Embryos were collected and prepared as explained in the FRAP section. FCS measurements were performed using an Insight confocal instrument (Evotec Technologies, Hamburg). The radius of the detection volume, w_0, half-height Sw_0, and volume $V = \pi^{3/2} S w_0^3$, were obtained from measurements of the diffusion of a well characterized fluorophore (Alexa 488, diffusion coefficient $D = 435\ \mu m^2/s$).[47] All measurements were performed in the middle part of the embryo, only $\sim$10 μm above the glass coverslip to avoid optical aberrations. In each studied nucleus, 10 successive FCS measurements (each lasting between 5 and 20 s) were performed in the middle of the nucleus. Noticeable photobleaching was systematically observed, even though a low 20 μW excitation intensity was used. Consequently, the resulting autocorrelation functions were analyzed using a model taking into account photobleaching, as well as two diffusing components and the presence of a photophysics term (see Analysis of autocorrelation functions) and returning in particular the amplitude of the diffusive term of the autocorrelation function ($G(0)$), the characteristic times associated with the transit of fast and slow molecules through the detection volume (τ_f and τ_s), and the fraction of slow molecules (f). For each measurement, characteristic times were turned into diffusion coefficients using $D = w_0^2/(4\tau)$, after which an effective diffusion coefficient was calculated as $D_{eff} = (1\text{-}f)\ D_f + f\ D_s$.

For each series of FCS measurements, the amplitude of the autocorrelation function, $G(0)$, was plotted as a function of the average fluorescence signal (I) and was fit to obtain the molecular brightness (B) of the fluorescent Cic molecules, and the average background noise (I_B) at this position in the embryo, using the expected dependence:

$$G(0) = (B / \gamma)(I - I_B)/I^2,$$

where $\gamma = 2^{3/2}$ *is a geometrical factor.* Multiple measurements over several days showed that $B = 5.1 \pm 1.4$ kHz and $I_B = 13 \pm 2$ kHz in the conditions of our experiments. Confocal images (typically acquired with pixel size of $d = 0.2$ μm and pixel dwell time of $\delta = 1$ ms) were first adjusted for uneven illumination using a reference image acquired in a fluorophore solution, then each pixel intensity (i) was turned into a concentration (c) using $c = (i / \delta - I_B)/((B/\gamma)V)$. The software ilastik was then used to segment all nuclei in the field of view and obtain their average Cic concentration.

Analysis of autocorrelation functions

General form of the autocorrelation functions

The autocorrelation functions, $G(\tau)$, obtained as a result of FCS experiments in embryos expressing Cic-sfGFP were fitted with a model accounting for two mobile components. For commodity these components were assumed to be both diffusive (an assumption

CellPress Current Biology Report

commonly made when analyzing FCS data of nuclear proteins).[48–50] An additional term, $G_P(\tau)$, was included in the model, in order to take into account the severe photobleaching that was observed during experiments (even though a low 20 μW excitation intensity, and experiment times as short as 5 s, were used). The function used to fit the data was:

$$G(\tau) = G(0)\left(1 + \frac{T}{1-T}e^{-\tau/\tau_T}\right)\left(\frac{f}{(1+\tau/\tau_f)\left(1+\tau\Big/\left(S^2\tau_f\right)\right)^{1/2}} + \frac{1-f}{(1+\tau/\tau_s)\left(1+\tau\Big/\left(S^2\tau_s\right)\right)^{1/2}}\right) + G_P(\tau) \quad \text{(Equation 1)}$$

When fitting the data, the value of the aspect ratio of the confocal detection volume was fixed to the value determined during calibration experiments, $S = 7$. All other parameters were left free to vary. The mobility of the proteins is characterized by the characteristic times associated with fast proteins (τ_f) and slow proteins (τ_s), and the fraction of fast proteins (f). Diffusion coefficients can be calculated from these characteristic times using $D_{s,f} = w^2/(4\tau_{s,f})$. The value of the $1/e^2$ radius of the confocal detection volume, $w = 301 \pm 7$nm, was determined in calibration experiments involving the diffusion of the fluorophore Alexa 488, which has a known diffusion coefficient $D = 435 \sim \mu m^2/s$.[47] The diffusive part of the correlation function also accounts for the presence of a small but noticeable photophysics term for the sfGFP fluorophore, with a fraction T of dark molecules, and a relaxation time τ_T for the dark state.[51]

Photobleaching

Although photobleaching occurred too slowly to interfere with the measurement of the diffusion characteristic time, it caused a regular decrease in the number of observed fluorescent molecules in the nucleus under study. This translated in a slow change in the average fluorescence signal, I, that could be well approximated by a decaying exponential with characteristic decay time τ_P on the order of the duration of the experiments $(t_M \approx 10\text{s})$. Because this characteristic time is well-separated from the other characteristic times in the system (τ_T, $\tau_f$and τ_s), this slow decay in the fluorescence results in a separate term in the correlation which can be well-approximated by:[52]

$$G_P(\tau) = \frac{t_M - \tau}{2\tau_P}\coth\left[\frac{t_M - \tau}{2\tau_P}\right] - 1 \quad \text{(Equation 2)}$$

Amplitude of the autocorrelation functions

In the absence of background noise, the amplitude of the diffusive part of the correlation function takes the simple form $G(0) = 1/N$, where N is the average number of fluorescent molecules present in the confocal detection volume, $V = \pi^{3/2}Sw^3$. However, experiments in embryos are characterized by a fair amount of background fluorescence, especially after several measurements have already been taken in a particular nucleus and photobleaching has reduced the number of fluorescent proteins. In the presence of background noise with average value I_B, the amplitude of the diffusive part of the correlation function takes the modified form[53,54]

$$G(0) = \frac{1}{N}\frac{1}{(1 + I_B/[I - I_B])^2} \quad \text{(Equation 3)}$$

Concentration measurements

In the absence of background noise and photobleaching, measuring $G(0)$ leads to a straightforward measurement of the absolute concentration of the fluorophore under study, $c = N/V = 1/(VG(0))$. However, since the value of $G(0)$ is affected by noise (Equation 3) when working in embryos determining I_B is important. In addition, $G(0)$ reflects only the concentration of visible fluorophores, which in our case was severely affected by photobleaching even after a single short FCS measurement. We thus decided to calculate concentration instead from the pixel intensity measured from images acquired before performing any FCS experiment.

The signal intensity (whether it is the pixel intensity in the image, or the average signal intensity of an FCS measurement, since both were acquired with the same instrument), I, is directly related to the average number of observed fluorophores, N, through:

$$I = \frac{B}{\gamma}N + I_B \quad \text{(Equation 4)}$$

where B is the effective molecular brightness of the fluorescent protein.

Combining Equations 3 and 4 shows that there is a direct relationship between $G(0)$ and I:

$$G(0) = \frac{B}{\gamma} \times \frac{I - I_B}{I^2} \quad \text{(Equation 5)}$$

If the number of fluorescent proteins, and therefore I, can be made to vary, as is the case in the presence of photobleaching, successive FCS measurements can be used to measure $G(0)$ as a function of I, and fitting this data with the above equation then allows retrieving both B and I_B. Equation 4 can then be used to measure N (either at the pixel of a confocal image or the location of an FCS experiment) and subsequently the absolute fluorophore concentration, c, at that location.

Generation of transgenic flies

Cic-sfGFP generation is described in Keenan et al.[20]

Current Biology
Report

Optimized psMEK

The psMEK1tight construct is available at Addgene plasmid #89361. The optimizing E203K substitution was made by changing the GGA codon to AAG. The construct was assembled into pTIGER as described previously[22] and integrated into the second chromosome using the phiC31 integration system at the attP site and balanced with CyO by The BestGene.

*tll** MS2. Described in Keenan et al.[20]

Endogenous hkb MS2

For insertion of MS2 stem loops into the 5′ UTR of the *hkb* locus, pU6-BbsI-chiRNA expression plasmid[44] and pHD-dsRed-24xMS2 donor plasmid (gifted by the Levine Lab) were coinjected to yw;nos-Cas9(II-attP40) embryos. Microinjection was performed by BestGene. dsRed was used for subsequent screening.

To generate the guide for a cut in the 5′ UTR of *hkb*, two DNA oligos, 5′ CTTCGCGACACTAAATCACTTGGA 3′ and 5′ AAACTC CAAGTGATTTAGTGTCGC 3′, were annealed and inserted into pU6-BbsI-chiRNA plasmid using BbsI sites.

To insert homology arms into the pHD_dsRed_24MS2 plasmid (for homology directed repair), we first amplified the 1000 bp 5′ homology arm of *hkb* from genomic DNA of OreR flies using two primers, cccttcgctgaagcaggtggGCCAGTAAAGTTTTTCTCAAGCACC and agtgcatatgtccgcggccgGGATGGAACACTTGTGATTATGATTTTG. These primers contain overhangs that over lap with the pHD_dsRED_24xMS2 plasmid. The pHD_dsRED_24xMS2 plasmid was linearized by cutting with EcoRI-HF and NheI-HF restriction enzymes (Upstream of the MS2 loops). The 5′ homology arm of *hkb* was then inserted into the plasmid using NEB HiFi assembly master mix. Subsequently, the 1000 bp 3′ homology arm of *hkb* was amplified from genomic DNA using two primers, tacgaagttata gaagagcaAAGTGATTTAGTGTCGCGAGAGAGC and gagcctcgagctgcagaaggGTATGAGTACATGGGCACGAAGATG. The pHD_dsRED_5′Arm_24xMS2 was linearized by cutting with SpeI and StuI restriction enzymes (downstream of the MS2 loops and the dsRED). The 3′ homology arm was then inserted into the plasmid using NEB HiFi assembly master mix.

The dsRED marker is flanked by loxP sites. To remove the dsRED marker from the locus, *hkb*-24xMS2-dsRED flies were crossed to Crey;+;D/Tm3 flies (Bloomington Stock 851). After screening for removal of dsRED, the Crey allele was crossed-out of the final stock.

Modified bnk reporters

All modifications were made from a *bnk* MS2 reporter gerenated in the Rushlow lab. A 130 bp fragment was inserted at the EagI site via in-fusion cloning upstream of the bnk enhancer and promoter driving MS2v7 loops from.[32] The fragment contained "TGAATGAA" for the Cic binding sites. "TGAAGCTA" was used for the mutated Cic binding sites. The fragment inserted at the EagI site to generate the modified reporter with 0 Cic sites was 5′-ttcgtttaaacggccgTGAAGCTATATCTATGATCACTAGTCTCGTGAAGCTAATGTCAGGA GATCTCCAGTTTATGAAGCTATTTACTAAATGAGCTCAGTCGTGAAGCTAcggccggccagatcca-3′. The fragment inserted at the EagI site to generate the modified reporter with 1 Cic site was 5′-ttcgtttaaacggccgTGAAGCTATATCTATGATCACTAGTCTCGTGAAGC TAATGTCAGGAGATCTCCAGTTTATGAAGCTATTTACTAAATGAGCTCAGTCGtgaatgaacggccggccagatcca-3′. The fragment inserted at the EagI site to generate the modified reporter with 4 Cic sites was 5′-ttcgtttaaacggccgtgaatgaaTATCTATGATCACT AGTCTCGtgaatgaaATGTCAGGAGATCTCCAGTTTAtgaatgaaTTTACTAAATGAGCTCAGTCGtgaatgaacggccggccagatcca-3′.
These fragments were designed from a region of a synthetic Cic-responsive construct courtesy of Nareg Djabrayan and the Jimenez lab. The construct was integrated into the 3rd chromosome using the phiC31 system by BestGene Inc (line #9750).

Combining optimized psMEK, MCP-mCherry, and MS2 reporters

The optimized psMEK transgenic flies were double balanced with Sp/Cyo ; Dr/Tm3 and crossed with MCP-mCherry on the third chromosome to generate optimized psMEK ; MCP-mCherry. P(matα-GAL-VP16)mat67; P(matα-GAL-VP16)mat15 was crossed with Sp/Cyo ; Dr/Tm3 to generate P(matα-GAL-VP16)mat67; Dr/Tm3, which was crossed with miFP-Histone on the third chromosome. Virgin females with P(matα-GAL-VP16)mat67 driving optimized psMEK expression and expressing MCP-mCherry were placed in a cage with MS2 reporter males for the experimental cross generating embryos to be imaged. The miFP-Histone was not used in this study.

Optogenetic illumination and live imaging

Embryos were collected on a yeasted apple juice plate in a cage placed in an aluminum foil-lined box under a 505 nm LED panel made in-house, described in Patel et al.[22] The T4 3/4 LEDs were purchased from https://www.superbrightleds.com/. Voltage was supplied by the KORAD KA3005D power supply from amazon.com. Embryos were collected for 1-2 hours under light. Embryos that did not have visible pole cells, which were visualized in halocarbonoil on the collection plate, were manually dechorionated, and mounted on their lateral side in a live imaging chamber consisting of a gas permeable membrane with halocarbon oil and a coverslip. All imaging was performed on a Leica SP5 point scanning confocal microscope. Embryos were illuminated with 10% 488 nm laser to activate optimized photoswitchable MEK and to image the cytoplasmic Dronpa fluorescent signal. 10% 405 nm laser was used to inactivate embryos. Embryos were staged by the density of apparent nuclei at the surface. The 561 nm laser was used to image the MCP-mCherry fluorescent signal. For *tll**MS2, the 63x objective was used and pinhole was also opened to 1.6 AU. 12.65 μm stacks with 0.55 μm steps were taken every 29 s with an imaging frequency of 400 Hz. For *hkb* and *bnk* + 4cic, the 63x objective was used and zoomed to 2.5x. Imaging frequency was 700 Hz. 14.77μm stacks with 0.67 μm steps were taken every 17 s, and 15% 515nm laser was used to activate the optogenetic tool.

CellPress

Current Biology
Report

Quantification of percent nuclei transcribing

Max projections of the mCherry channel were made in FIJI[45] for all movies. A Gaussian blur with a radius of 4 was applied to the max projection and then subtracted from the original max projection. The image sequences of background subtracted max projections were processed in MATLAB to count the number of MS2 dots. The open source FastPeakFind MATLAB function with a user supplied threshold was used to count the number of dots at each time point. The total number of nuclei was counted by inverting the image from the Dronpa channel (cytoplasmic Dronpa outlines nuclei) and using imfindcircles in MATLAB. The % nuclei transcribing was calculated by dividing the number of dots detected per frame by the number of nuclei for the embryo counted from the middle slice of the z stack immediately before the 488 nm laser was switched off.

Spot detection of transcriptional bursts

The spot detection algorithm in the imaging software Imaris was used. Spots detection parameters were as follows: 1 μm for hkb and 1.5 μm for *bnk* + 4 cic, user specified intensity standard deviation threshold above the noise, maximum distance 3 μm and no allowed gap size. Autoregressive motion was selected for spot tracking. For the intensity traces of example tracked spots in Figure 3, a maximum gap distance of the number of frames was allowed, and gaps were filled with all detected objects. Spot intensity means were normalized by the intensity of the cytoplasmic mCherry signal from the time frame at the start of nc 14. Intensity was measured in FIJI, from a reconstructed image exported from Imaris. All images exported had consistent parameters in the Display Adjustment window before measuring intensity in FIJI.

Burst initiation time

The burst initiation time was defined as the time after the photoswitch from activating to inactivating light at which the individually tracked spot was first detected.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of transcription activity

The mean and standard deviations of the percent nuclei transcribing were calculated in MATLAB. 2-sample Kolmogorov-Smirnov tests (two-tailed) from the scipy.stats package in Python were used to calculate the KS-test statistics and p values of the spot life-times and peak intensities. For the *hkb* reporter, the numbers of de-repressed and repressed bursts were 298 (from 2 embryos) and 340 (from 3 embryos) respectively. For the *bnk* + 4cic reporter, the numbers of de-repressed and repressed bursts were 220 (from 2 embryos) and 106 (from 2 embryos) respectively. Truncated violin plots show the median and quartiles with thick and thin lines respectively. Box and whisker plots with Tukey whiskers generated in GraphPad Prism 9 are shown to represent the burst initiation times.

Current Biology, Volume *31*

Supplemental Information

Capicua is a fast-acting transcriptional brake

Aleena L. Patel, Lili Zhang, Shannon E. Keenan, Christine A. Rushlow, Cécile Fradin, and Stanislav Y. Shvartsman

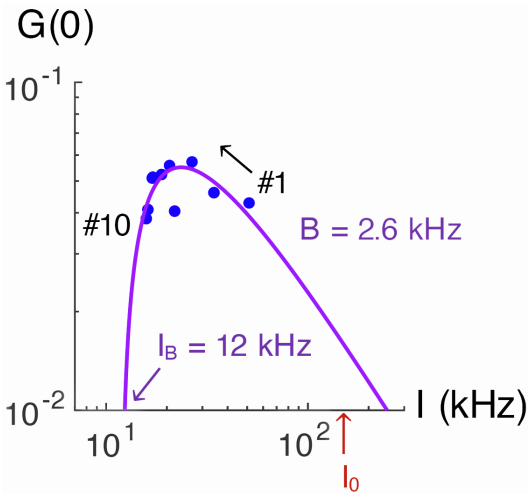

Figure S1. The relationship between the amplitude of the diffusive part of the autocorrelation function and the average measured intensity. Related to Figure 1. The relationship between the amplitude of the diffusive part of the autocorrelation function, G(0), and the average measured intensity, I, shown for a series of 10 successive FCS measurements during which I monotonously decreased due to slow continuous photobleaching. The line is a fit of the data, which allows retrieving the value of the effective molecular brightness, B, and background fluorescence, I_B, in the embryo. The value of the signal at the position of the FCS measurement, measured from the image acquired immediately before the start of the first FCS measurement, I_0, is shown on the figure.

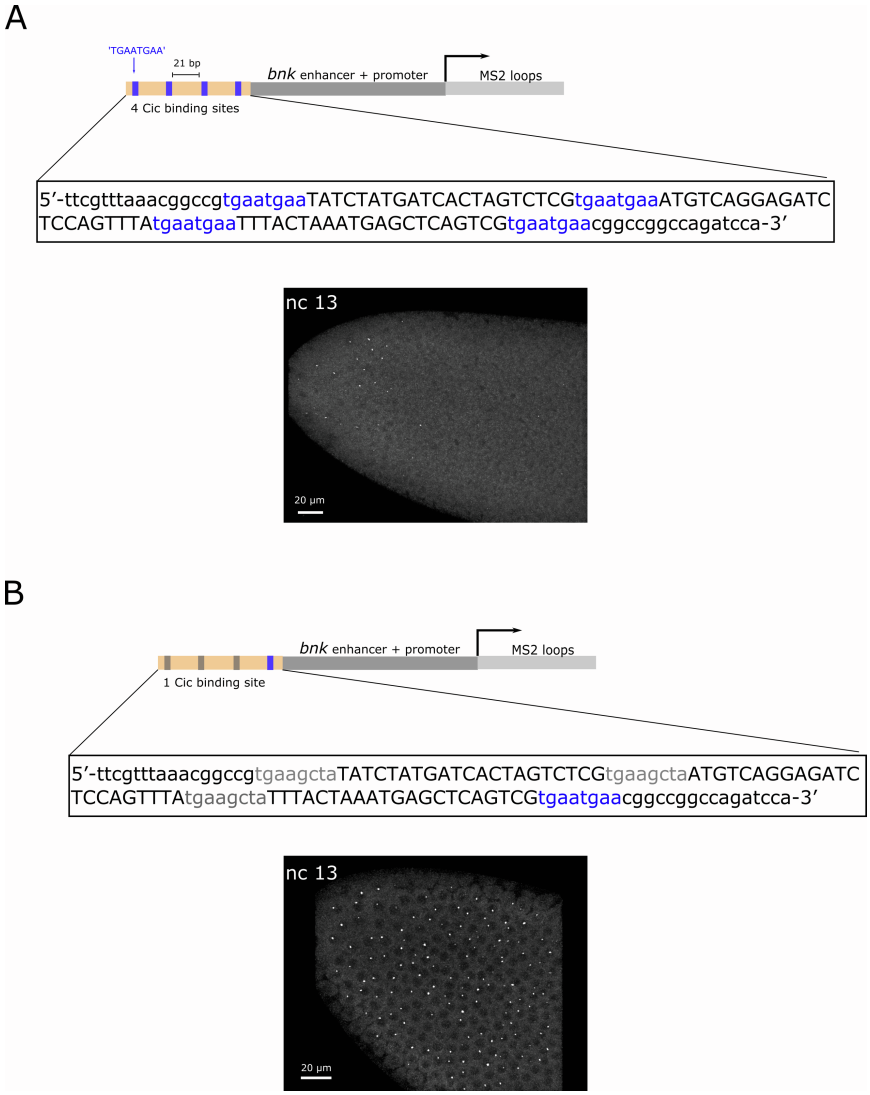

Figure S2. Modified *bnk* reporter design. Related to Figure 3. (A) Design for the modified *bnk* reporter. The Cic binding site 'TGAATGAA' (blue) was inserted 4 times upstream of the *bnk* enhancer and promoter driving MS2 loops. Image shows the anterior pole of an embryo in nc 13. MS2 dots are localized to the pole, reflecting endogenous ERK signaling. (B) The same reporter construct with the first 3 sites mutated ('TGAAGCTA', gray). Images shows the anterior pole of an embryo in nc 13. MS2 spots are not localized to the pole, suggesting this reporter is not sensitive to endogenous ERK signaling

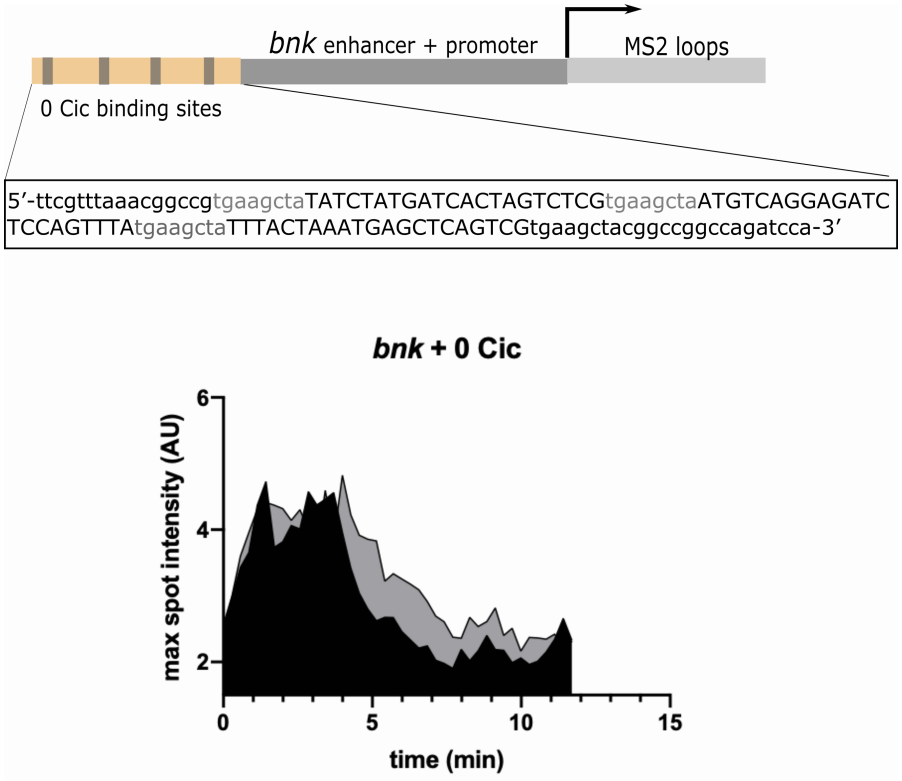

Figure S3. Control modified *bnk* reporter with no Cic binding sites. Related to Figure 3. A control construct with all four Cic sites mutated (*bnk* + 0 Cic) was generated. The maximum spot intensity of all spots recorded per time frame is shown for 2 embryos. Signal appears to drop off significantly throughout nc 14.

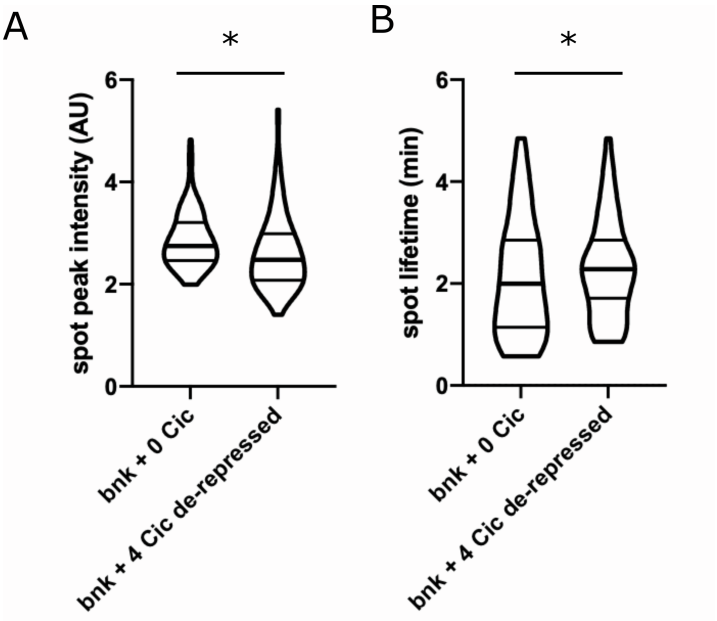

Figure S4. Spot parameter comparison between *bnk* reporters with 0 and 4 Cic sites. Related to Figure 4. The lifetimes and peak intensities of tracked spots from embryos expressing the bnk + 0 Cic sites reporter were compared to the same parameters from the *bnk* + 4 Cic sites reporter de-repressed with continuously activated optimized psMEK. Tracked spots were only considered from the first 6 minutes of nc 14. Nbnk0Cic = 228, Nbnk4cic,derepressed = 220. * $p < .05$

Chapter 5

Paper 3: Comparing the dynamics of two morphogens with opposite functions

The following draft manuscript in this chapter is in the preparation process for submission to a research journal:

Lili Zhang, Ahmad Mahmood, Carmina Perez-Romero, and Cécile Fradin. "Activator vs. repressor: differences and similarities in the mobility of two early fly embryo transcription factors with opposite functions." (2021).

Author contribution:

L.Z. designed the research, performed and analysed all the experiments, and wrote the manuscript. A.A. helped analyze the slow FRAP data. C.P.-R. designed the research and performed preliminary experiments. C.F. designed the research, performed analytical calculations, and wrote the manuscript.

Research Background:

Morphogens that are transcription factors come in two varieties: they are either transcriptional activators or transcriptional repressors. In the early fly embryos, these two types of morphogens work together to define regions in which different sets of genes are expressed. Little is known about the differences or similarities in the mechanism by which these two types of morphogens influence transcription.

Research Purpose:

The goal of this third project was to employ the same methods that were used in the previous chapter for our preliminary study of Cic dynamics in order to directly compare the dynamics of two morphogens, an activator (Bcd) and a repressor (Cic), to see if their opposite functions were driven by difference in their dynamics.

Methods: Fluorescence Correlation Spectroscopy (FCS), Fluorescence Recovery After Photobleaching (FRAP), Monte Carlo simulations.

Highlights of this paper:

- The nucleo-cytoplasmic transport properties of Bcd and Cic are strikingly different: whereas Bcd rapidly shuttles between nucleus and cytoplasm, unphosphorylated Cic is trapped inside nuclei.
- In contrast, FRAP and FCS show that the intranuclear dynamics of Bcd and Cic are very similar: both are very mobile inside nuclei, with effective diffusion coefficients in the 20 $\mu m^2/s$ range, and no immobile fraction.

- Interestingly, FCS experiments show that both Bcd and Cic have a fast freely diffusing population and a less abundant slow or transiently associated to DNA population.

Significance: Despite having opposite transcriptional functions, Bcd and Cic have similar intranuclear mobilities. This finding, along with similarities in the structures of these two morphogens (a well-defined DNA-binding region flawed by disordered regions) suggest they might use similar mechanisms to influence transcription, in spite of the opposite outcome.

Activator vs. repressor: The differences and similarities in the mobilities of two early fly embryo transcription factors with opposite functions

Lili Zhang[a], Ahmad Mahmood[a], Carmina Perez-Romero[a], Cécile Fradin[a,b]

[a]*Department of Physics and Astronomy, McMaster University, 1280 Main St. W, Hamilton, ON, L8S 4M1, Canada*
[b]*Department of Biochemistry and Biomedical Sciences, McMaster University, 1280 Main St. W, Hamilton, ON, L8S 4K1, Canada*

Abstract

Transcription factors, including activators and repressors, play an essential role in pattern formation during early embryo development. Understanding the mechanisms by which these life-sustaining proteins exert their function necessitates understanding their dynamics. Here we compare the dynamics of two morphogens, a transcription activator, Bicoid (Bcd), and a transcription repressor, Capicua (Cic), both active in the early fly embryo, which we studied using Fluorescence Correlation Spectroscopy (FCS) and Fluorescence Recovery After Photobleaching (FRAP). We find that in spite of their opposite functions, Bcd and Cic have very similar nuclear dynamics, characterized by the existence of both a fast freely diffusing fraction and a much slower and less abundant fraction, a feature fully consistent with inclusion of these proteins into transcriptional condensates. This may explain how the transcriptional response they elicit can be at the same time so fast and so precise. In contrast, Bcd and Cic have strikingly different nucleo-cytoplasmic transport properties, where Bcd rapidly shuttles in and out of nuclei, whereas in the mid-embryo ventral region in which our experiments were performed and where Cic concentration is the highest, Cic nuclear export is severely restricted. This difference can be linked to the very different mechanisms by which the concentration gradients of these two morphogens are formed. Although our study is focused on the early fly embryo, Cic has a homolog gene in human (proposed to be a major tumour suppressor) and our results fits in with the recent proposal that many if not all transcription factors might share a common target search strategy that make use of their large unstructured regions.

Keywords: Morphogen, transcription factors, activation, repression, Bicoid, Capicua, *Drosophila melanogaster*, FCS, FRAP, diffusion, nucleo-cytoplasmic transport, DNA binding, transcriptional condensates, disordered regions

1. Introduction

Transcription factors, by regulating gene expression in space and time, play a crucial role in all cells. They come in two opposite flavours, activator or repressor. Both types of transcription factors, however, act on transcription via tight binding to specific target sequences in the regulatory regions of the genes they regulate. Many questions remain, however, regarding how they can quickly find their target sequences in the crowded nuclear environment, and by which mechanisms exactly they affect transcription.

In the early embryo, the role of morphogen transcription factors (which form concentration or activity gradients and affect transcription in a concentration or activity-dependent manner) and the interplay between gene activation and repression, is especially important, as it sets the boundaries between different gene expression domains destined to lead later on to tissue differentiation. Bicoid (Bcd) and Capicua (Cic) are two morphogens involved in the control of body patterning along the anterior-posterior axis in the early fly embryo. While Bcd is a transcriptional activator, Cic is a transcriptional repressor. Bcd is one of the first discovered and the most studied morphogens since its discovery in *Drosophila* in the 1980s [1, 2, 3]. It is one of the many important maternal genes that is deposited at the anterior pole of the egg by its mother [4]. Bcd is a homeobox transcriptional activator having at least 66 target genes including *hunchback* (*hb*) [5]. Cic, on the other hand, is a high-mobility group (HMG)-box transcription repressor and an evolutionarily conserved gene that was discovered more recently, in 2000 [6]. Cic is down-regulated by Torso from the receptor tyrosine kinase (RTK) pathway at both the anterior and the posterior pole of the embryo and thus represses its target genes, including *tailless* (*tll*) and *huckebein* (*hkb*), only in the central region of the embryo. Abnormal behaviours of Cic in human have been linked to several health conditions such as lung cancer and brain degenerative diseases [7, 8].

Bcd and Cic both have a primarily nuclear localization, but otherwise form very different types of concentration gradients: exponential for Bcd with the highest concentration at the anterior pole [9, 10] and flat across most of the embryo for Cic with low concentration at both poles [11]. These gradients are formed and maintained through very different mechanisms: spatially restricted synthesis, diffusion and uniform degradation for Bicoid [12, 13, 14] and downregulation by maternal Torso at both poles for Cic, more precisely nuclear export enahnced by Torso RTK activation followed by cytoplasmic degradation [11]. Bcd and Cic are thought to act in an antagonistic manner in the early fly embryo ($\sim$ 1 to 2 hr after egg laying, nc 8 to 14) to establish the boundaries of the expression domains of head genes [15], with the help of other morphogens, such as Zelda

[16]. One example of the entanglement between gene regulation by Bcd and Cic is the competition between these two proteins for phosphorylation by the mitogen-activated protein kinase (MAPK), which is most intense at the anterior pole where Bcd is most abundant, and results in the asymmetric concentration profile of Cic [17].

A common mechanism for controlling gene transcription is Histone acetylation, which results in the unwinding of DNA and opening up the chromatin structure, and is generally associated with active transcription, or deacetylation, usually associated with repression of transcription [18]. Histone deacetylation is thought to play a role in Cic repression in mammalian cells where Cic was shown to recruit Histone deacetylases [19, 20, 21], and in a recent study the question was raised as to whether DNA accessibility might be catalyzed by Bcd through histone acetylation [22]. However, these processes are thought to occur on relatively long timescales, of the order of tens of minutes. In contrast, in early embryos where nuclear divisions occur in rapid successions and interphase may last only a few minutes, the transcriptional response needs to be established quickly. Indeed, the time necessary for the transcriptional response to Bcd and Cic take effect is very similar for both proteins, and very short, on the order of a couple of minutes [23, 24]. This suggests that activation (for Bcd) or repression (for Cic) occurs instead *via* relatively direct mechanisms such as physical chromatin remodelling, competitive binding or physical exclusion of other transcription factors.

For both Bcd and Cic, small changes in concentration along the embryo anterior-posterior axis give rise to a very sharp transcriptional response of their target genes, and thus to well-defined expression domains [14, 25]. It is clear that cooperative binding to gene regulatory regions, who often possess multiple binding sites for the same TF, must play a role in this response. At least in the case of Bcd, there is experimental evidence that cooperativity is involved in the binding of this protein to DNA target sites [26, 5], and this cooperative binding has been proposed to explain the steepness of the *hb* transcriptional response [27, 28, 29]. Cooperative binding is also thought to play a role in Cic repression [24]. Also the transcription of *hb* from Bcd is found to be bursty [30], which is attributed to the switching of *hb* promoter between on and off states.

How such a precise response is possible within the short time frame in which the transcriptional response of these genes is established is still the object of debate, and raise the question of the mechanism (and the speed) with which these TFs search and find their target binding sites [28]. Different models have been proposed to explain the great speed at which target search seem to occur. Facilitated diffusion, for example, *i.e.* the idea that DNA-binding proteins might unspecifically bind to DNA and diffuse along it in search of their target site, has been shown to speed up the target search for some transcription factors [31, 32], and proposed as a possible mechanism to explain how Bcd could elicit such a fast and precise response [33]. A different and emerging model of transcription regulation suggests the existence of condensates (or "hubs") creating high local concentrations of TFs at the active transcription foci [34, 35, 36, 37]. Hubs of Bcd and Zelda have in fact been observed in live *Drosophila* embryos and their activity are thought to coordinate active transcription [38, 39], even though the mechanism leading to the clustering of the proteins inside these hubs is not entirely clear.

Many questions thus remain about the mechanisms of transcriptional activation and repression, starting with how the TFs reach their target genes and how they then promote either activation or repression. Could it be that they physically exclude the transcription machinery by forming domains [40]? Do they sit tight on the gene? Or do they kiss and run, modifying the chromatin and then leaving to go and activate or repress other genes? Are activators and repressors acting through identical or different processes? We hypothesize that different mechanisms for activation and repression would lead Bcd and Cic to have different dynamics inside nuclei, and look at the transcriptional function of these two morphogens through the lens of their mobility, studied here by fluorescence methods.

2. Results

To compare and contrast the dynamics of nuclear Bcd and Cic, we performed three types of experiments, namely: i) slow FRAP, ii) fast FRAP and iii) FCS, covering a wide range of dynamical time scales (0.01 ms $\sim$ 100 s) and length scales (300 nm $\sim$ 10 μm). These experiments were performed in live *Drosophila melanogaster* embryos expressing the protein of interest fused to a fluorescent protein: either Cic-sfGFP or Bcd-eGFP. In addition, control experiments were performed in embryos expressing a control fluorescent protein fused to a nuclear localization signal, NLS-eGFP, that would localize to nuclei but have no specific interaction with DNA. All experiments were performed at nuclear cycle (nc) 13 or 14, and in nuclei found in the middle of the embryo (mid-way between the anterior and posterior poles). Experimental results are presented in order of decreasing length scale, starting with slow FRAP experiments which test the exchange of protein between cytoplasm and nucleus and finishing with FCS experiments which probe micron-scale dynamics.

2.1. Slow FRAP

For slow FRAP experiments, a majority of the fluorescent molecules inside a single target nucleus were photobleached by focusing a high power laser beam at the centre of that nucleus for 3 s. The slow fluorescence recovery due to the progressive replacement of photobleached proteins by fluorescent proteins from the cytoplasmic pool was then monitored over a 20 min period. The goal of these experiments was to obtain information about the kinetics of the transport of the transcription factors through the nuclear pore complexes found in the nuclear envelop (NE). Fig. 1A shows representative confocal images of nuclei pre-photobleaching, immediately post-photobleching and 20 min after photobleaching, for embryos expressing Cic-sfGFP, Bcd-eGFP or NLS-eGFP. It immediately appears that nuclear Cic-sfGFP fluorescence does not recover even after 20 mins while that of Bcd-eGFP and NLS-eGFP mainly does. To

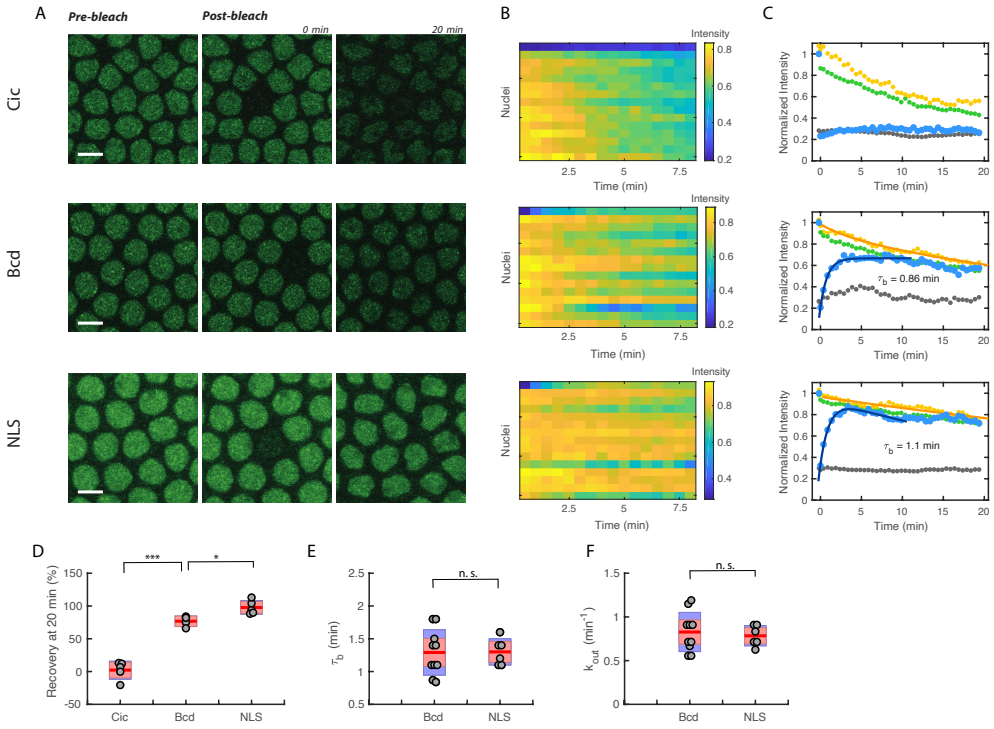

Figure 1: **Results of slow FRAP experiments.** (**A**) Confocal images of nuclei in *Drosophila melanogaster* embryos expressing Cic-sfGFP (top), Bcd-eGFP (middle), and NLS-eGFP (bottom) pre-bleach, immediately post-bleach (t = 0 min) and t = 20 min post-bleach. Scale bar is 5 μm. Imaging conditions were the same for all three proteins. (**B**) Carpet plot showing the time evolution of the mean fluorescence intensity (normalized to the pre-photobleaching intensity) of the photobleached nucleus and of the 15 closest nuclei, in order of proximity from top to bottom. (**C**) Mean intensity of the photobleached nucleus (blue circles), the 5 nearest-neighbour nuclei (green circles), the 5 furthest away nuclei (yellow circles) and the cytoplasm (grey circles). All intensities are normalized to the average pre-photobleaching nuclear intensity. Solid lines show fits with either a single exponential (far away nuclei, yellow line) or a double exponential (photobleached nucleus, blue line), and the corresponding recovery time (τ_b) is indicated. (**D**) Percent recovery at 20 min post-bleach measured for the three proteins. The measured values of τ_b, and the values of $k_{\text{out}} = 1/\tau_b$ are shown in (**E**) and (**F**), respectively. In these three plots, the mean is indicated by the tick red line, the 1.96 SEM (95% confidence interval) is represented by the red box, and the standard deviation by the purple box. Three stars ($***$) indicate that the p value from the t-test is less than 0.001, one star ($*$) that it is less than 0.05, and n.s. that it is larger than 0.5, meaning there is no significant difference between the two sets of data.

examine this phenomenon quantitatively, an automatic segmentation procedure was performed, to retrieve the total fluorescence intensity recorded from each nucleus in the field of view (Fig. 1B,C). The first 10 min of the recovery curves obtained for the photobleached nucleus were fitted with a double exponential, the first accounting for the fluorescence recovery due to nucleo-cytoplasmic transport, and the second for continuous photobleaching while imaging (Eq. 1). The average recovery times, τ_b, and recovery percentages at 20 min, obtained from repeated experiments, are shown in Fig. 1D,E. Whereas the recovery observed for Bcd ($77 \pm 8\%$) approaches the 100 % recovery obtained for the control NLS-eGFP construct, there was no detectable recovery for Cic-sfGFP in the 20 min time window of the experiment. The recovery time obtained for Bcd-eGFP (1.29 ± 0.35 min) was similar to that reported previously [41], and undistinguishable from that of the control NLS-eGFP (1.30 ± 0.20 min). Assuming that the cytoplasmic concentration of Bcd-eGFP and NLS-eGFP is constant (a reasonable assumption given that at this stage of development the fly embryo is a syncytium with a very large cytoplasmic volume, that the cytoplasmic fluorescence seems to be constant and that the recovery is almost complete), the recovery time should depend only on the proteins nuclear export rate (k_{out}), where $\tau_b = 1/k_{\text{out}}$ (the nuclear export rate measured for Bcd-eEGFP and NLS-eEGFP is shown in Fig. 1F). In effect, the values of τ_b measured for Bcd-eGFP and NLS-eEGFP thus corresponds to the average time spent by these proteins in the nucleus. In contrast, the fact that no recovery is observed for Cic-sfGFP indicates that there is little to no nuclear export for this protein at this stage of development and at this position in the embryo - the slow FRAP experiments, however, cannot distinguish between a scenario where Cic molecules are not exported because they are immobilized on the DNA and a scenario where they are mobile within the nucleus but unable to cross the NE.

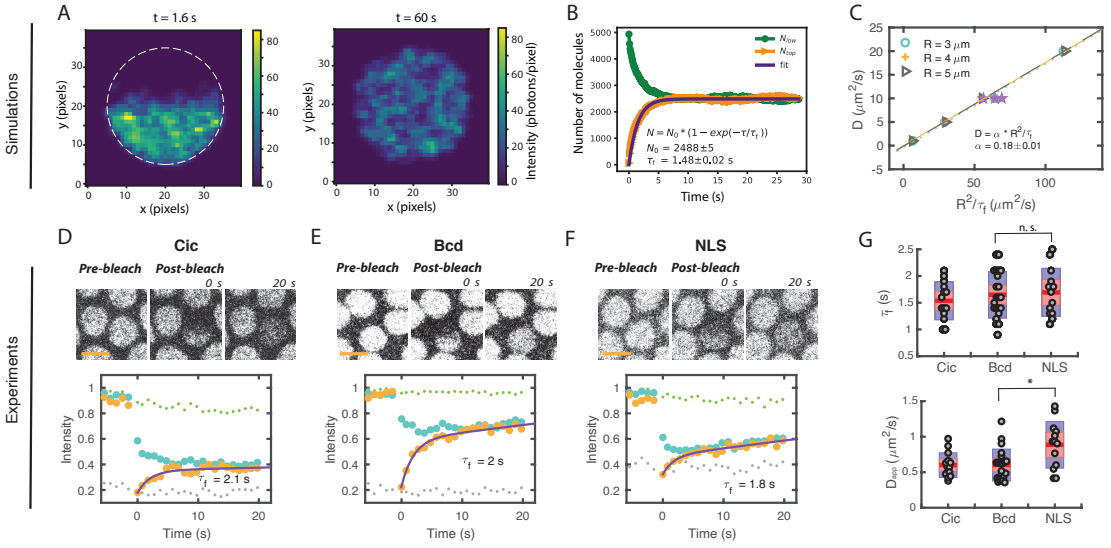

Figure 2: **Results of fast FRAP experiments.** (**A**) Simulation of a fast FRAP experiment, showing the simulated confocal image of a nucleus with fluorescent proteins initially placed only in the lower half of the nucleus (t = 1.6 s) and after the proteins have been allowed to diffuse for t = 60 s. The nucleus is represented by a 6 μm-diameter sphere and the size of the cubic simulation box is 8 μm. The number of molecules in this simulation was 1000, all with a 0.1 $\mu m^2/s$ diffusion coefficient. (**B**) Time evolution of the number of molecules for the lower and upper halves of the nucleus for the simulation shown in (A). The solid line shows the fit of the recovery curve for the photobleached half of the nucleus with a single exponential function. The number of molecules at equilibrium (N_0) and recovery time (τ_f) extracted from this fit are indicated. (**C**) Relationship between D and R^2/τ_f obtained for a number of simulations performed for different values of R and D. The line is a fit of the data assuming a linear relationship between D and R^2/τ_f, and returning $\alpha = 0.18$. The stars are for different sizes of the photobleached area. (**D, E, F**) Confocal images of the nuclei pre-photobleaching, and 0 or 20 s post-photobleaching (upper panels), for embryos expressing Cic-sfGFP (**D**), Bcd-eGFP (**E**), or NLS-eGFP (**F**) at nc 14. Scale bar is 5 μm. The corresponding recovery curves are shown in the lower panels. The solid line is a fit (with Eq. 3) of the recovery curve for the photobleached half of the nucleus. The recovery time (τ_f) obtained from this fit is also shown. (**G**) Values measured for τ_f and for the apparent diffusion coefficient D_{app} for the three different proteins. The mean is indicated by the tick red line, the 1.96 SEM (95% confidence interval) is represented by the red box, and the standard deviation by the purple box. One star ($*$) indicates that the p value from the t-test is less than 0.05, and n.s. that it is larger than 0.5, meaning there is no significant difference between the two sets of data.

2.2. *Fast FRAP*

For fast FRAP experiments, photobleaching was restricted to one half of a nucleus, the photobleaching step was performed as quickly as possible (1 s), and recovery was monitored for a time (20 s) shorter than the average residence time of a protein in the nucleus (i.e. shorter than τ_b), in order to concentrate on the redistribution of the proteins inside the nucleus. The primary goal of these experiments was to check for the eventual presence of an immobile fraction of the proteins, but they were also used to place a lower limit on their intranuclear diffusion coefficients. Generally, recovery times are extracted from a FRAP curves using an appropriate recovery model (e.g., single exponential, double exponential, or more complex ones [42]). However, obtaining accurate diffusion coefficients from the recovery time is not trivial and depend on the geometry of the photobleached region [41, 43, 44]. The approach we took here was to perform Monte Carlo simulations reproducing the geometry of our experiments to obtain an empirical relationship between the protein diffusion coefficient (D), the nucleus radius (R) and the recovery time obtained from a single exponential fit of the FRAP data (τ_f). From dimensional analysis, we expect $D = \alpha \times R^2/\tau_f$, and we used the simulations to estimate the value of α. All simulated proteins were made to follow a random walk corresponding to diffusion with D in the 1 to 20 $\mu m^2/s$ range (i.e. there was no immobile protein fraction), but they were not allowed to cross the NE. Typical results from a fast FRAP simulation are presented in Fig. 2A,B, showing simulated confocal images at t= 1.6 s (where all fluorescent molecules are placed in the bottom half of a 6 μm-diameter sphere representing a nucleus, to mimic the photobleaching of the upper half of the nucleus) and t = 60 s (after molecules have redistributed homogeneously throughout the nucleus). The recovery curve for the photobleached half of the nucleus can be fitted well with a single exponential from which the fitted recovery time, τ_f, can be obtained (Fig. 2B). The empirical relationship between D and the ratio R^2/τ_f obtained from a series of simulations performed for different nucleus radii and different diffusion constants, as well as slightly different sizes of the photobleached region, is shown in Fig. 2C. It is, as expected, a linear relationship and the coefficient, α, was found to be 0.18 ± 0.01.

Examples of representative fast FRAP experiments for Cic-sfGFP, Bcd-eEGFP and NLS-eGFP are shown in Fig. 2D,E,F, respectively, where confocal images of the nucleus pre- and post-photobleaching are shown (top), as well as recovery curves from both halves of the nucleus (bottom). It is immediately evident from these curves that none of these protein populations

have a significant immobile fraction on the timescale of the experiments, since after a few seconds they are homogeneously redistributed across the half-photobleached nuclei. Quantification of the difference between the fluorescence signals coming from either half of the nuclei show that both Bcd and Cic have an immobile fraction of less than 5% (Fig. S1E in the Supplementary Information). It is also evident that the fast intranuclear redistribution that quickly leads to the equalization of the average fluorescence in both halves of the nucleus is, in the case of Bcd-eGFP and NLS-eGFP, followed as can be expected by a slower overall increase in fluorescence due to nuclear import from the cytoplasm. Also as expected, this slow increase is not observed for Cic-sfGFP. The recovery curves obtained for the photobleached half of the nuclei were therefore fitted with a single exponential function with characteristic time τ_f accounting for the intranuclear redistribution to which a linear term was added accounting for the much slower nucleo-cytoplasmic exchange (Eq. 3). The values of τ_f obtained from repeated fast FRAP experiments for the three proteins are shown in Fig. 2G, as well as $D_{\text{app}} = \alpha \times R^2/\tau_f$ calculated from these values, taking into account the size of each individual studied nucleus, and using the value $\alpha = 0.18$ obtained from the simulations. The average recovery time for all three proteins is ~ 1.5 s, only just slightly longer than the duration of the photobleaching step (1 s). This indicates that molecules already significantly redistribute during the photobleaching step, a known issue when trying to capture the motion of fast molecules with imaging FRAP experiments, which leads to an underestimate of the diffusion coefficient [45, 46]. The calculated values of D_{app} thus only represent a lower limit for the average diffusion coefficient of the proteins. From these experiments, we can then conclude that all three of the studied proteins are fully mobile inside nuclei at this position in the embryo and at this stage of development, and that their mobility is consistent with a diffusion process with $D > 0.6\ \mu\text{m}^2/\text{s}$ for Cic-sfGFP and Bcd-eGFP, and $D > 0.9\ \mu\text{m}^2/\text{s}$ for NLS-eGFP. Of note, despite the limitations of these experiments, and whereas there is no significant difference between the values of τ_f measured for NLS-eGFP and for the other two proteins, once the size of the nuclei is taken into account (see also Fig. S1E in the Supplementary Information for more information on the size variation of nuclei between nc 13 and nc 14), we find that NLS-eGFP is significantly more mobile than Cic-sfGFP and Bcd-eGFP, as expected since it is a smaller protein not expected to interact with other molecules except for transport factors.

2.3. FCS

To fully capture the fast motions of Bcd and Cic, and also to explore these motions on a sub-μm-scale (as opposed to the $\approx 10\ \mu$m scale explored with FRAP experiments), we performed FCS experiments. Series of 10 to 20 single-point FCS measurements (each lasting 5, 10, or 20 s) were performed in embryos expressing either of the recombinant proteins, Cic-sfGFP, Bcd-eGFP or NLS-eGFP, at nc 13 or 14, and always at the centre of a nucleus in the mid-embryo. For all studied proteins, the autocorrelation function (ACF) obtained from the first measurement in each series systematically showed a pronounced effect due to the continuous photobleaching of fluorescent molecules in the small nuclear compartment, in the form of a decay at 10 s lag times (orange curves in Fig. 3A,B,C). This effect decreased then disappeared in subsequent measurements (pink and blue curves in Fig. 3A,B,C) as the fluorescence signal stabilized. The original ACFs were fitted first assuming two diffusing populations (Two-Component Model or TCM), and taking into account both the background noise and the photobleaching [47]. From this fit, four types of characteristic times with very different orders of magnitude are obtained: the fluorophore triplet relaxation time time ($\approx 20\ \mu$s), τ_T, the characteristic diffusion time of the fast diffusing fraction ($\approx 1\ m$s), τ_{Df}, the characteristic diffusion time of the slow diffusion fraction (≈ 10 to $100\ m$s), τ_{Ds}, and the photobleaching time (≈ 40 s), τ_P. The measured values of these four parameters are shown for all three proteins in Fig. 3D,E,F. The relative amplitude of the fast diffusion term, that can be interpreted as the fraction of fast molecule, p, as long as both fast and slow molecules have the same brightness, is shown in Fig. 3G,H,I. Four series of measurements with different measurement times (represented by symbols of different colors) are shown for each protein to demonstrate the consistency of the results. Remarkably, Bcd and Cic have very similar dynamics, dominated by a fast diffusing fraction ($\tau_{\text{Df}} \simeq 0.8$ ms corresponding to a diffusion coefficient $D_f \simeq 30\ \mu m^2/s$), but with a significant fraction of slower molecules ($\tau_{\text{Ds}} \simeq 50$ to 60 ms, corresponding to $D_s \simeq 0.4\ \mu m^2/s$). For both proteins, the contribution of the slow molecules decreases over time from 50 to 25 %, with the same characteristic time (≈ 40 s) as the continuous photobleaching of the overall signal, an indication that this decrease must be due to the photobleaching or to a decrease of the brightness of the slower particles. The smaller control nuclear protein NLS-eGFP has a freely diffusing population which is slightly faster ($D_f \simeq 30\ \mu m^2/s$) than that of Cic-sfGFP and Bcd-eGFP, and in stark contrast to what is observed for the transcription factors it has only a very small (and constant) fraction of slowly diffusing molecules ($\simeq 10$ %) - and these "slow" NLS-eGFP molecules are also significantly faster (with $D_f \simeq 1.5 \mu m^2/s$) than the slow Cic-sfGFP and Bcd-eGFP molecules.

A different assumption can be made when analyzing the FCS data, namely that proteins can transiently associate with an immobile or very slow structure (for example DNA) and become temporarily immobilized. The data should then be analyzed with a Stick-and-Diffuse Model (SDM) [48, 46] instead of the TCM. The goodness of fit obtained when using the SDM, though good enough, is not as good as that obtained with the TCM (see Fig. S2 in Supplementary). The parameters characterizing a potential binding to and unbinding from DNA obtained from fitting the ACFs with the SDM, namely the effective binding rate, k_{on}, the unbinding rate, k_{off}, the fraction of unbound diffusing molecules, $p = k_{\text{off}}/(k_{\text{on}} + k_{\text{off}})$, and the characteristic diffusion time, τ_D, are shown in Fig. 4.

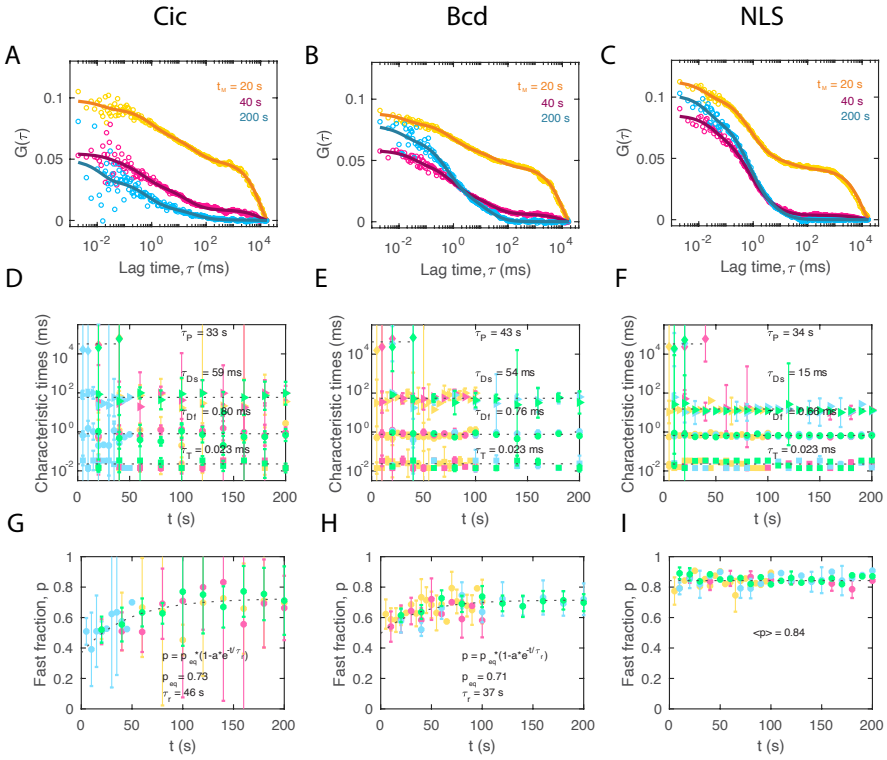

Figure 3: **Results of FCS experiments analyzed with the Two-Component Model (TCM).** (**A-C**) ACFs obtained from the first (orange), second (pink) and tenth (blue) measurements in a series of 20 s single-point FCS measurements performed in the center of a nucleus in *Drosophila* embryos expressing Cic-sfGFP (**A**), Bcd-eGFP (**B**), or NLS-sfGFP (**C**). Lines are fit of the ACFs with a TCM taking into account noise and photobleaching (**D-F**) Characteristic times obtained from the fit of the ACFs with the TCM: triplet relaxation time, τ_T, fast diffusion characteristic diffusion time, τ_{Df}, slow diffusion characteristic diffusion time, τ_{Ds}, and photobleaching time, τ_p, shown in **D** for Cic-sfGFP, **E** for Bcd-eGFP and **F** NLS-eGFP. The value of τ_p is shown only for the first two or three measurements in each series, since the photobleaching contribution becomes negligible after that, and the values of τ_p obtained from the fit of the ACF become unreliable. (**G-I**) Relative amplitude of the term corresponding to fast-diffusing proteins, p, shown in **G** for Cic-sfGFP, **H** for Bcd-eGFP and **I** NLS-eGFP. In (**D-I**), symbols of diifferent colors represent different series of FCS measurements.

3. Discussion

The purpose of this study was to gain a better understanding of the dynamics of two morphogens acting as transcription factors in the early *Drosophila* embryos, an activator (Bcd) and a repressor (Cic), with a particular focus on their regulation of gene transcription. A lot of attention has been given already to the mobility of Bcd (often in the context of gradient formation), which has been examined using slow FRAP [41, 49, 50], FCS [25, 50] and single particle tracking [38, 39]. Some *Drosophila* studies also touched on the dynamics of Cic, initially focusing on nucleo-cytoplasmic transport, using either FRAP or FCS [11, 24]. The premise of this current study was to study the nuclear dynamics of these two proteins with opposite functions side by side, using the exact same experimental protocol, with the idea that if their opposite transcriptional functions were executed using different mechanisms, this should be reflected in their nuclear dynamics. We used both FRAP and FCS in order to cover as large a range of time and length scales as possible. The different parameters measured for both proteins (as wall as for the control NLS-eGFP protein) are shown in Table 1.

The one obvious difference in the behaviour of Cic and Bcd is their nucleo-cytoplasmic shuttling. With the slow FRAP experimental conditions used in this study, nuclear Cic had a recovery percentage of $\sim 0\%$ whereas Bcd recovered almost entirely (Bcd recovery was $77\pm8\%$, approaching the control NLS recovery of $\sim 98\pm11\%$). This pattern of results is consistent with previous literature, which demonstrated highly efficient shuttling between nucleus and cytoplasm for Bicoid with a characteristic nucleo-cytoplasmic shuttling time of $\sim$ 60 s [41], whereas for Cic nuclear fluorescence recovery after photobleaching was found to be incomplete at best and shown to be under tight control of the TOR kinase, where the phosphorylation of Cic which naturally occurs at the poles of the embryo significantly increasing nuclear export rate and decreasing nuclear import (leading to Cic degradation in the cytoplasm and the formation of the Cic concentration gradient) [11]. Our results put

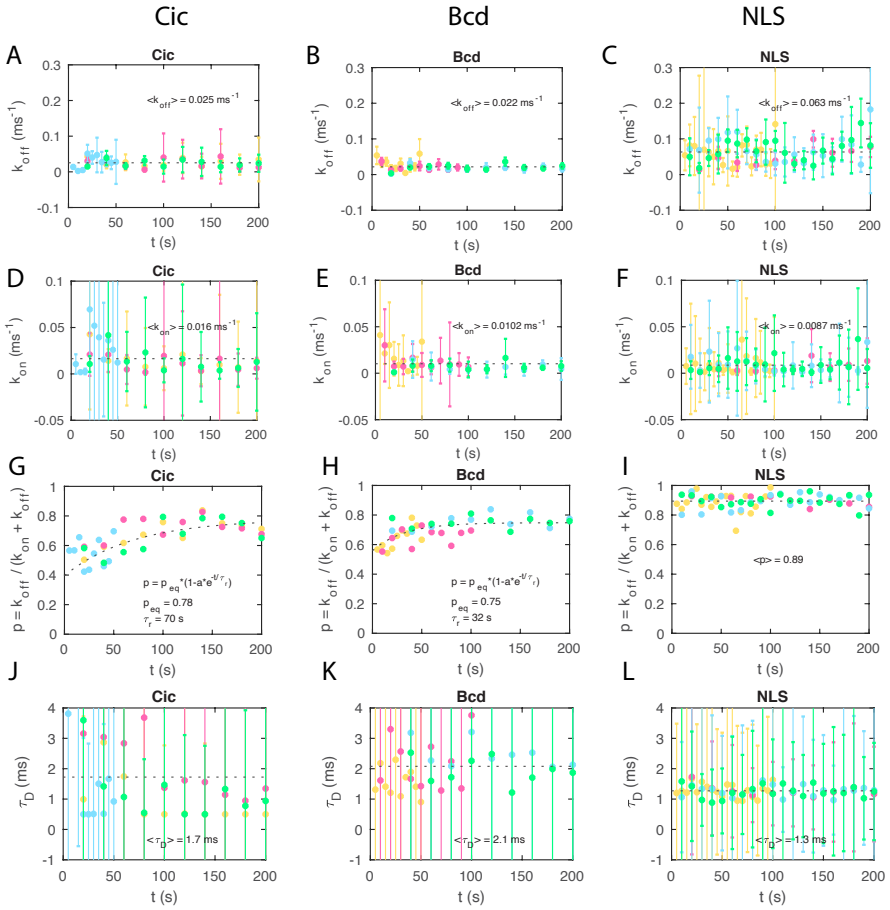

Figure 4: **Results of FCS experiments analyzed with the Stick-and-Diffuse Model (SDM).** The DNA unbinding rate, k_{off}, binding rate, k_{on}, the effective fast fraction, $p = k_{\text{off}}/(k_{\text{on}} + k_{\text{off}})$, and the characteristic diffusion time, τ_D, from fitting the ACFs with Stick-and-Diffuse Model, as well as their averaged values represented by dotted grey lines are shown. Different colors of symbols represents different series of measurement with measurement time of 5, 10 or 20 seconds. The errorbars represent 50% confidence interval.

this contrast in focus, by showing that in the ventral side of the mid-embryo, where all our experiments were performed, the nucleo-cytoplasmic shuttling of Cic is negligible. What we actually observe is that there is no recovery after photobleaching of the nuclear Cic population. Two conditions must be met for nuclear recovery: there must be a pool of available molecules to be transported in the cytoplasm, and these molecules must be recognized as legitimate nuclear import cargos. Therefore there could be two possible reasons for the absence of Cic nucleo-cytoplasmic shuttling: either there is no Cic in the cytoplasm available to be imported, or Cic is present in the cytoplasm but its import rate is very low. Note that Cic is found to be mainly and efficiently degraded in the cytoplasm of *Drosophihla* embryos [51] (whereas it is unknown if Bcd is degraded mainly in the cytoplasm or nucleus, however the two TFs seem to have a similar lifetime in the mid-embryo, 25 min for Bcd [52] and 30 min for Cic in the absence of Torso signalling [11]). In the ventral mid-embryo, Cic is acting to repress both terminal genes such as *tailless (tll)* and *huckebein(hkb)* and dorsal genes such as *zerknullt (zen)*. It thus seems likely that the irreversibility of Cic nucleo-cytoplasmic shuttling in this area of the embryo is due to the combination of an extremely low rate of nuclear export (thus maintaining high Cic concentrations by avoiding its degradation in the cytoplasm) and of a cytoplasmic degradation process which further decreases the concentration of any remaining cytoplasmic pool of Cic. What we observe here is the tight control of the nuclear localization of Cic by the embryo, in relation to its concentration and activity gradient. In contrast, the Bcd gradient is mainly formed by diffusion along the AP axis of the embryo [50], and although already stabilized by nc 13, it might require that Bcd is free to shuttle from one nucleus to the next.

In sharp contrast with what is observed for nucleo-cytoplasmic transport, there is no discernible difference be-

Table 1: Parameters obtained from slow FRAP, fast FRAP and FCS using Two Component Model and Stick-and-Diffuse Model. Value in parentheses is one standard deviation.

Experiments	Parameters	Cic	Bcd	NLS
Slow FRAP	Recovery at 20 min	2 ± 14 %	77 ± 8 %	98 ± 11 %
	τ_b (min)	-	1.29 (0.35)	1.30 (0.20)
	k_{out} (min^{-1})	-	0.83 (0.22)	0.78 (0.12)
	k_{in} (min^{-1})	-	6.6 (1.8)	6.2 (1.0)
Fast FRAP	Immobile fraction	2 ± 4 %	4 ± 6 %	3 ± 6 %
	D ($\mu m^2/s$)	>0.6 (0.17)	>0.6 (0.22)	>0.9 (0.33)
FCS (Two-Component Model)	τ_p (s)	33 (19)	43 (24)	34 (17)
	D_f ($\mu m^2/s$)	28 (18)	30 (8)	34 (5)
	D_s ($\mu m^2/s$)	0.38 (0.20)	0.42 (0.19)	1.5 (0.8)
	D_{eff} ($\mu m^2/s$)	20 (14)	21 (6)	28 (5)
	$p_{t=0}$	37 ± 4 %	50 ± 7 %	84 ± 3 %
	$p_{t\to\infty}$	73 ± 3 %	71 ± 1 %	84 ± 3 %
FCS (Stick-and-Diffuse Model)	k_{off} (ms^{-1})	0.025 (0.012)	0.022 (0.012)	0.063 (0.032)
	k_{on} (ms^{-1})	0.016 (0.015)	0.010 (0.009)	0.009 (0.008)
	K_D	1.6 (2.2)	2.2 (3.2)	7.2 (10)
	D_f ($\mu m^2/s$)	13 (15)	11 (5)	18 (3)
	$p_{t=0}$	38 ± 6 %	56 ± 7 %	89 ± 5 %
	$p_{t\to\infty}$	77 ± 4 %	75 ± 4 %	89 ± 5 %

tween the intranuclear mobility of activator Bcd and that of repressor Cic: *(i)* neither protein has a significant immobile nuclear fraction (as shown by our fast FRAP experiments, which demonstrate that no more than 4% of the proteins is immobilized for more than 1 s), *(ii)* both proteins have an abundant freely diffusing nuclear population, with a diffusion coefficient $D_f \sim 30\,\mu m^2/s$ (measured by FCS), and *(iii)* for both proteins a second nuclear fraction with much slower dynamics is also detected, which must correspond to binding with other proteins or structures inside the nucleus which restricts the motion of these molecules in some way.

The fact that only a very small fraction of the proteins is immobilized for more than 1 s, combined with the high diffusion coefficient of the freely diffusing form of the proteins show that the TFs can quickly redistribute across the nucleus. Transcription factors specifically bound to their target DNA sequence can be expected to remain bound for 1 s (as measured for Bcd [38]) or longer. But although each of the TFs studied here has tens of target genes and hundreds of specific binding sites on the *Drosophila* genome [53], in the mid-embryo the concentration of both proteins is in the 100 nM range [41, 50, 24], meaning that there are tens of thousand of molecules of Cic and Bcd in each nucleus. Thus the specifically bound fraction of Cic or Bcd will always only be a very small fraction, and in regions where these morphogens are active, a great many of them are present in nuclei ready to act on their target genes. This pool of available TF has a rermarkably large mobility, where despite being quite a bit larger ($\sim$ 200 kDa for Cic-sfGFP and $\sim$ 80 kDa for Bcd-eGFP) than the control NLS-eGFP protein ($\sim$ 30 kDa), the unbound fractions of Bcd and Cic have a diffusion coefficient that is only very slightly smaller than that of the control. How is that possible? The answer might lay in the particular structure of Cic and Bcd, which (like many transcription factors [54, 55, 56]) are both predicted to possess very large intrinsically disordered regions on either side of their well-structured DNA-binding domain. Disordered proteins have been shown to diffuse faster than smaller globular proteins in crowded conditions [57], a counterintuitive feature thought to be due at least in part to the greater conformational flexibility of the disordered proteins [58]. In the context of the crowded conditions inside nuclei, and the necessity for fast target search by TFs, the overall lack of structure of Bcd and Cic could be an advantage by allowing them to diffuse faster.

The presence of a second (slower) nuclear fraction of Cic and Bcd, unmistakable in our FCS experiments (and which has been reported for a variety of transcription factors [59, 60, 61, 62]), can be interpreted in at least two ways, corresponding to two possible target search mechanisms, as illustrated in Fig. 5. The first possible search mechanism (left panel) is a facilitated diffusion mechanism, where each TF alternates between periods of three-dimensional diffusion in the nucleoplasm and periods of one-dimensional diffusion while non-specifically and transiently interacting with DNA [63, 31]. As long as the one-dimensional translational motion of the TF on the DNA is slow enough, the protein can be considered immobile from the point of view of an FCS experiment, and ACFs obtained in such ex-

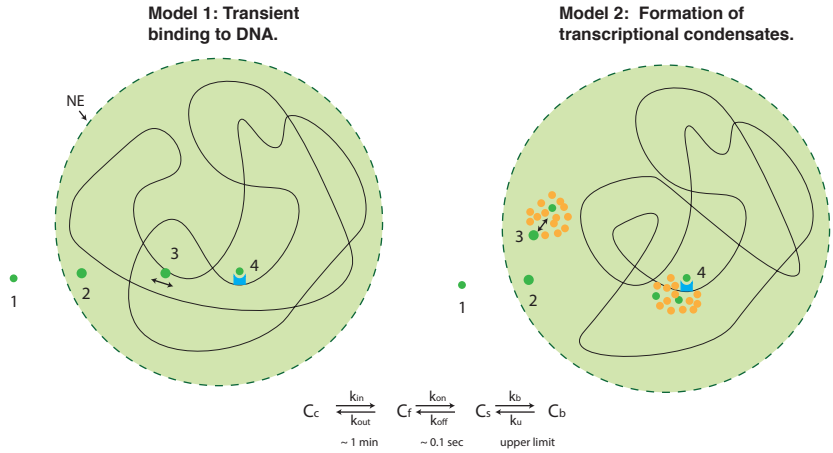

Figure 5: **Two possible models for transcription factors activity.** Two possible models are considered to explain the dynamics of the TFs observed in this study. In both models, TFs are divided into four major populations: (1) a freely diffusing cytoplasmic population (concentration C_c), (2) an abundant monomeric and freely diffusing nuclear population (concentration C_f), (3) a second slightly less abundant nuclear population interacting with nuclear structures (concentration C_s) and (4) a third nuclear population specifically bound to a gene promoter. The two proposed models differ only with regard to the structures with which the TFs interact in the nucleus. In the first model (left panel), the TFs transiently bind to DNA *via* non-specific electrostatic interactions involving positively charged patches on the DNA-binding domain of the TFs. In the second model (right panel), the TFs participate in the formation of phase-separated transcriptional condensates or hubs *via* low-specificity protein-protein interactions mediated by their large unstructured regions. In both models, the freely diffusing cytoplasmic and nuclear populations can exchange through import (import rate k_{in}) or export (export rate k_{out}) through the NE. In the nucleus, the exchange between the freely diffusing (unbound) population) and slow or immobile (non-specifically bound to DNA or part of a condensate) population is characterized by an on-rate (k_{on}) and an off-rate (k_{off}). Non specifically bound molecules on the DNA or in condensates can then specifically bind (binding rate k_b) or unbind (unbinding rate k_u) to target sequences in the promoter regions of their target genes.

periments should be well-fitted by a stick-and-diffuse model, which will then return the unspecific binding rate k_{on} and unbinding rate k_{off} of the TF to the DNA. Analyzing our ACFs with the stick-and-diffuse model (Fig. 4) indeed results in credible values of k_{on} and k_{off} (see Table 1), consistent with the timescale for interactions between Bcd and DNA that have been reported previously from single particle tracking experiments [38]. However, it is hard to reconcile this model with the fact that the bound fraction of TFs ($1 - p$) is found to decrease over time (Fig. 4G,H), since each TF is supposed to alternate quickly between a free and a bound state, and therefore $1 - p = k_{\text{on}}/(k_{\text{on}} + k_{\text{off}})$ should be constant.

The second possible search mechanism (Fig. 5, right panel) involves the formation of small condensates containing several of the fluorescent TFs, that would diffuse much slower than a free TF and therefore account for the slow component observed in FCS experiments. Phase separation is increasingly being recognized as playing an important role in nuclear compartmentalization [64]. In this context, the formation of small transcriptional condensates or hubs with high concentration of transcription factors, has been proposed to participate in transcription regulation in general (a suggestion relying in part on the observation that most TFs contain intrinsically disordered domains, a feature associated with liquid-liquid phase separation) [65, 66, 67, 36, 68, 37], and in the case of Bcd in particular [38, 39]. By locally increasing the concentration of TFs target gene regulatory regions, such condensates could accelerate target search and therefore transcription regulation as well as make it more precise. From the point of view of an FCS experiment, a TF which would become associated with a small condensate or pre-condensate would appear as a slowly diffusing molecule, and the ACFs would be well fit by a two-component model. Further, because the motion of a condensate containing n fluorescent proteins would disproportionately contribute to the ACF (with an amplitude proportional to n^2), a small amount of diffusing condensates would be enough to result in what may look like a large amount of slowly diffusing proteins. This feature can also explain why the apparent fraction of slowly diffusing fluorescent TFs captured from TCM analysis seem to decrease over time (Fig. 3G,H), since continuous photobleaching would reduce the value of n, and therefore the contribution of the condensates to the ACF, even if the concentration and size of condensates remain the same. In conclusion, our FCS data is fully compatible with the idea that Cic and Bcd participate in transcriptional condensates. Transcriptional condensate formation would also explain the observed cooperativity of the binding of Bcd [26] and Cic [24] to DNA (a cooperativity shown to be mediated, in the case of Bcd, by its intrinsically disordered region [26]), and the fact that both proteins can elicit a precise transcriptional response in only a few minutes [23, 24].

In conclusion, both models (facilitated diffusion and transcriptional condensate formation) can help explain the speed and precision of the transcriptional response elicited by Cic and Bcd and both might in fact act in parallel [69], where each is driven by a separate structural feature shared by these two TFs (positively charged DNA binding domain allowing for non-

specific DNA binding interactions for facilitated diffusion, and large disordered regions for condensate formation). However, our FCS data favours the condensate formation model. What is very interesting is that both Cic and Bcd have the same dynamical signature, thus the mechanism by which they accelerate their target search is almost certainly similar, and therefore shared by both activators and repressors. A slight difference between the two proteins is that the apparent fraction of bound/interacting TF at $t = 0$, $1 - p_{t=0}$, is larger for Cic than for Bcd, which could indicate that Cic represses genes by physically blocking access to regulatory regions.

The values we report here for the diffusion coefficients of the fast and slow fractions of Bcd are higher than previously reported, where a prior FCS studies led to $D_f = 7.7 \pm 0.3\ \mu m^2/s$, $p = 0.57 \pm 0.01$, $D_s = 0.22 \pm 0.01\ \mu m^2/s$ (our values of the diffusion coefficient of NLS-eGFP are also higher than previously reported) [25]. This difference could be due to a number of reasons: *(i)* our use in this study of a modified TCM which takes into account photobleaching effect and allows analyzing the first few curves in a series of measurements, which would usually be rejected, but are in fact less noisy and more reliable, *(ii)* fit of individual ACFs instead of an averaged ACF from multiple measurements, which also allowed to check for variation in parameters over time, *(iii)* measurements made in a different part of the embryo (ventral vs. anterior pole) where the optical path is shorter and Bcd is present at a lower concentration, *(iv)* possibly different temperatures (as the effect of temperature on TF mobility is not well characterized but could be important), and *(v)* use of different dye for calibration of the confocal detection volume. With higher values of diffusion coefficients, our results continue to support the synthesis-diffusion-degradation model for gradient formation and makes it easier to reconcile the very high precision of the Bcd gradient readout with the speed of the transcriptional response [50, 24, 33].

Certain limitations of this study could be addressed in future research. For example, fast FRAP experiments with a higher time resolution and shorter photobleaching step could be preformed to obtain an independent and accurate measurement of the TFs effective intranuclear diffusion. The cytoplasmic concentration of Cic and Bcd could be measured accurately (although this is a difficult endeavour because of the very low S/N ratio in the cytoplasm) to accurately determine the nuclear-to-cytoplasmic concentration ratio and firmly establish the relationship between nuclear import and export rates.

4. Conclusion

Taken together, our findings indicate that both the activator Bcd and the repressor Cic are very mobile inside nuclei, with a mix of fast and slow dynamics that hints towards the formation of transcriptional condensates (a mechanism that can help explain the high speed and precision of the transcriptional response to the presence of Cic and Bcd), and at their importance for both activation and repression. In contrast, the nucleocytoplasmic transport properties of these two proteins are quite different as Bcd can quickly shuttle back and forth between nucleus and cytoplasm, while Cic (at least in the ventral mid-embryo where our experiments were carried out) is strictly confined to the nucleus during interphase.

5. Materials and Methods

5.1. Experiments

5.1.1. Drosophila embryo preparation

Drosophila embryos were prepared for imaging following the protocol detailed in ref. [70]. Briefly, *D. melanogaster* fly strains expressing NLS-eGFP, Bcd-eGFP (a kind gift of Dr. Wieschaus) [41] or Cic-sfGP (a kind gift of Dr. Shvartsman) [24] were maintained in a 25°C incubator with alternating day-night lighting. To collect embryos for experiment, the plastic tube containing flies were inverted on an embryo collection plate with yeast paste in the centre. After about 3 hours, the embryos on the collection plate were transferred with a tweezer to a double-sided tape to remove the chorion. The dechorionated embryos were transferred to a thin line of glue on a 0.17 m coverslip. The ventral side of the embryo was placed in contact with the coverslip such that as many nuclei as possible in the embryo could be observed in a single field of view just above the surface of the coverslip. Lastly, a small drop of Halocarbon oil 700 (Sigma) was added on top of the embryos to prevent water evaporating from the dechorionated embryo while allowing oxygen permeation for necessary metabolism.

5.1.2. Imaging Fluorescence Recovery After Photobleaching

All imaging FRAP experiments were performed on an eclipse Ti inverted microscope equipped with NIS-Element software (Nikon, Mississauga, ON, Canada). The objective used was either an air-immersion Nikon Plan Apo λ (60× magnification, NA 0.95) in conjunction with a 46 μm pinhole (slow FRAP), or an oil-immersion Nikon Plan Apo λ (60× magnification, NA 1.4), in conjunction with a 38 μm pinhole (fast FRAP). Fluorescence excitation and photobleaching were achieved with a 488 nm laser. All FRAP experiments were performed at room temperature.

Slow FRAP. For slow FRAP experiments, a small circular region (diameter ~2 μm) at the centre of a nucleus was photobleached for 3 s using a laser power of 100% (0.5 mW). The region around that nucleus was imaged both before and after the photobleaching (for 10 to 20 min at 30 s intervals) using a smaller 6.2% (30 μW) laser power, a pixel dwell time 4.6 μs, and a 1024×1024 pixels image size. The software Ilastik was used to extract nuclei position in each image with a machine learning algorithm [71]. Fitting of the recovery curves was done using MATLAB.

Fast FRAP. For fast FRAP experiments, a rectangular region covering one half of a nucleus was photobleached for 1 s using a laser power of 80% (0.4 mW). The region around that nuclei was imaged imaged both before and after the photobleaching step (for 20 s at 1 s intervals) using a 5.1 % (25 μW) laser power, a 2.7 μs pixel dwell time, a pixel size of 0.1 μm/pixel, and an image size of 128×128 pixels. The software Fiji was used to extract the intensity of the bleached and unbleached halves of the nucleus [72]. Fitting of the recovery curves was done using MATLAB.

5.1.3. Fluorescence correlation spectroscopy

Single-point FCS data were acquired on an Insight Cell confocal microscope (Evotec Technologies, Hamburg, Germany, now PerkinElmer, Waltham, MA, USA). Excitation was performed using a 488 nm continuous wave solid state diode-pumped laser (Sapphire 488-20/460-10, Coherent, Santa Clara, CA, USA). The excitation beam was set to underfill the back-aperture of the water-immersion objective (UApoN, 40×, 1.15 NA, Olympus Canada, Richmond Hill, ON, Canada) and used in conjunction with a 40 μm pinhole. Calibration measurements were performed with solution of Alexa Fluor 488 (Invitrogen, now Life Technologies, Carlsbad, CA, USA) which has a known diffusion coefficient $D = 435\,\mu m^2/s$ at 22.5 $^\circ C$ [73]. The radius of the detection volume ω_0 was then calculated via $\omega_0 = \sqrt{4D\tau}$. All FCS experiments were performed at room temperature. ACFs obtained as a result of FCS experiments were analyzed using either FCS+plus Analyze (Evotec) or Matlab.

5.2. Data analysis

5.2.1. Slow FRAP

The recovery curves for the mean intensity of the photobleached nucleus in slow FRAP experiments were assumed to take the form:

$$I_b(t) = [I_\infty - be^{-t/\tau_b}] \times e^{-t/\tau_p}, \tag{1}$$

where τ_b is the photobleaching recovery time and τ_p is the characteristic decay time of the overall fluorescence due to continuous photobleaching during imaging. I_∞ represents the equilibrium intensity that would be reached in the absence of continuous photobleaching. b is a constant whose value depends on the parameters of the photobleaching step.

The amount of recovery was obtained for $t \gg \tau_b$ (at t = 20 min) by comparing the mean fluorescence in the photobleached nucleus, $I_b(t)$, to the mean fluorescence in nearby nuclei, $I_u(t)$. Taking into account an eventual fluorescence background, I_{bkg}, the fraction of recovered signal was calculated as:

$$\%\ \text{recovery} = \frac{I_b(t) - I_{bkg}}{I_u(t) - I_{bkg}}. \tag{2}$$

Assuming this recovery is due to nucleo-cytoplasmic transport, and further assuming that the concentration of fluorescent proteins in the cytoplasm (C_c) is constant, then the recovery time only depend on the nuclear export rate, $\tau_b = 1/k_{\text{out}}$. The value of k_{out} can then be straightforwardly calculated. At equilibrium, the import rate, k_{in}, has to satisfy $k_{\text{in}} = C_c/C_n \times k_{\text{out}}$ (where C_n is the equilibrium nuclear concentration of the proteins) such that the concentration in the nucleus is much higher than that in the nucleus, as observed in the experiments. The ratio of nucleo-cytoplasmic concentration for Bcd is about 8 [49], and we assumed it to be the same for NLS.

5.2.2. Fast FRAP

We assumed that there were two separate processes contributing to fluorescence recovery after fast photobleaching of half a nucleus: the fast exchange of fluorescent molecules between the two halves of the nucleus due to diffusion (characteristic time τ_f), and the slower exchange of molecules between nucleus and cytoplasm (characteristic time τ_s). Thus, the evolution of intensity in the bleached and unbleached part of the nucleus were described as:

$$I_b(t) = I_{\text{eq}} - I_{im} + ae^{-t/\tau_f} + bt/\tau_b, \tag{3}$$

$$I_u(t) = I_{\text{eq}} + ae^{-t/\tau_f} + bt/\tau_b, \tag{4}$$

In the equations above, a single exponential is used to describe the fast redistribution of fluorescent proteins inside the nucleus, where we assumed that the exchange between fast and slow diffusing populations was fast enough that the motions of the proteins on this scale would appear as a single population with an effective diffusion coefficient $D_{\text{app}} \propto R^2/\tau_f$. These equations are valid for $t \ll \tau_b$, in which case $e^{-t/\tau_b} = 1 - t/\tau_b$

The fraction of immobile molecules, p_{im}, can be calculated from I_u, the mean intensity of the unbleached half of nucleus, I_b, the mean intensity of the bleached half nucleus, and I_0, the mean intensity of a control nucleus that is not being bleached, $t \ll \tau_f$:

$$p_{\text{im}} = (I_u - I_b)/I_0 = (c_u^{eq} - c_b^{eq})/c_0^{eq}. \tag{5}$$

In practice, we used the mean from the last five data points when calculating I_u, I_b and I_0.

5.2.3. FCS

Two-Component Model

A Two-Component Model (TCM) was applied to the analysis of ACFs obtained from FCS measurements. This model was modified to take into account explicitly both the effects of noise and photobleaching, as detailed in [47]. According to this model, the ACF given by $G(\tau) = G_D(\tau) + G_P(\tau)$, where $G_D(\tau)$ accounts for the fluorophore dynamics and $G_P(\tau)$ global continuous photobleaching in the small compartment where the FCS experiment takes place (in our case a nucleus).

If we assume that there are two species of diffusing fluorescent molecules in the detection volume, the part of the ACF reflecting fluorophore dynamics should be [74]:

$$G_D(\tau) = G(0)\left(1 + T/(1-T)e^{-\tau/\tau_T}\right) \left[\frac{p}{(1+\tau/\tau_{\text{Df}})(1+\tau/(S^2\tau_{\text{Df}}))^{1/2}} + \frac{1-p}{(1+\tau/\tau_{\text{Ds}})(1+\tau/(S^2\tau_{\text{Ds}}))^{1/2}}\right]. \tag{6}$$

p is the fraction of the fast diffusing particles (assuming all particles, fast and slow, have the same brightness), with a characteristic diffusion time τ_{Df}. The second (slower) species diffuse with a characteristic time of τ_{Ds}. $G(0)$ is the amplitude of the diffusive part of the ACF. T is the fraction of molecules in the dark state with a relaxation time of τ_T. S is the aspect ratio of the detection volume.

The photobleaching contribution is given by [47]:

$$G_P(\tau) = \left(\frac{t_M-\tau}{2\tau_P}\coth\left[\frac{t_M-\tau}{2\tau_P}\right]-1\right) / \left(1+\frac{m}{N_0}\frac{t_M-\tau}{\tau_P}\frac{1+e^{\tau/\tau_P}}{1-e^{-\frac{t_M-\tau}{\tau_P}}}+\left(\frac{m}{N_0}\right)^2\left(\frac{t_M-\tau}{\tau_P}\right)^2\frac{e^{\tau/\tau_P}}{1-e^{-\frac{t_M-\tau}{\tau_P}}}\right). \quad (7)$$

τ_P is the photobleaching characteristic time and t_M is the FCS measurement time. N_0 represents the number of fluorescent molecules in the detection volume at $t = 0$. m is a constant related to the background noise and molecular brightness where N/m is a measure of the signal-to-noise [47].

The diffusion coefficient, D, is related to the characteristic diffusion time, τ, via $D = \omega_0^2/(4\tau)$, where ω_0 is the $1/e^2$ radius of the Gaussian-shaped detection volume. Thus $D_{\mathrm{f,s}} = \omega_0^2/(4\tau_{\mathrm{f,s}})$. The effective diffusion constant, D_{eff} can be used to describe the collective diffusion property of the molecules with two diffusing populations: $D_{\mathrm{eff}} = p \times D_f + (1-p) \times D_s$. $D_{\mathrm{eff}} = p \times D_f$ if $D_s \ll D_f$.

Stick-and-Diffuse Model

Unlike for the Two-Component Model where fluorophore are assumed to remain in a given state (slow or fast) as they cross the detection volume, according to the Stick-and-Diffuse Model (SDM) there is only one diffusing population but it alternates between a freely diffusing state (i.e. and unbound state) and a transiently bound state with $D \approx 0$. The ACF derived from the SDM is shown in Eq. 8 [48, 46].

k_{on} and k_{off} are the binding and unbinding rate, respectively. N, T, τ_T, S were defined above. τ_D is he characteristic diffusion time of molecules in the unbound state. It has been tested previously that n can be truncated at 7 in the Taylor series [46] thus we set $n = 1, 2, ..., 7$ when fitting data with Eq. 8.

The dissociation constant K_D represents the equilibrium between bound and unbound species, and can be calculated through $K_D = k_{\mathrm{off}}/k_{\mathrm{on}}$.

5.3. Simulations

5.3.1. TF motion

Dynamic modeling was carried out via Monte Carlo simulation written in Python (the code is given in the supplementary information). A simulation box of size $12 \times 12 \times 12\,\mu m^3$ was used. A sphere (with a radius of 3 to 5 μm) representing the nucleus was placed at the centre of the cubic box. Particles were placed either randomly in the simulation box, or to reflect a particular initial configuration, as created for example by nuclear import or by photobleaching. At each simulation step (each corresponding to a duration $\delta t = 1$ ms), the coordinates of each particle were updated by a step size drawn from a Gaussian distribution with a mean of 0 and a variance of $\sqrt{2D\delta t}$. Moves bringing particles outside of the simulation box were refused. Moves causing particles to cross the NE were accepted with a probability chosen to reflect the probability of import in or export from the nucleus.

5.3.2. Confocal Imaging

Simulated confocal images were produced at different times during a simulation. In confocal imaging, the laser focus is scanned point by point across the chosen field of view and at each point the signal is recorded in a single photon detector. In the simulation, the signal emitted by each individual particle at each pixel in the image was first calculated according to:

$$g(x,y,x) = B\tau\exp(-\frac{2((x-x_0)^2+(y-y_0)^2)}{\omega_0^2}-\frac{2(z-z_0)^2}{\omega_z^2}), \quad (9)$$

where B is the molecular brightness of the fluorescent particles and x_0, y_0, z_0 is its position. τ is the pixel dwell time, x, y, z is the position of the pixel, and ω_0, ω_z are the short and long axis of the Gaussian confocal detection volume, respectively. The final confocal image is a superposition of pixel intensity from all particles [75]. Most images were generated using a pixel size (distance between two consecutive pixels) of 0.1 μ m/pixel.

References

[1] C. Nüsslein-Volhard, E. Wieschaus, Mutations affecting segment number and polarity in drosophila, Nature 287 (5785) (1980) 795–801.

[2] W. Driever, J. Ma, C. Nüsslein-Volhard, M. Ptashne, Rescue of bicoid mutant drosophila embryos by bicoid fusion proteins containing heterologous activating sequences, Nature 342 (6246) (1989) 149–154.

[3] W. Driever, C. Nüsslein-Volhard, The bicoid protein is a positive regulator of hunchback transcription in the early drosophila embryo, Nature 337 (6203) (1989) 138–143.

[4] D. St Johnston, C. Nüsslein-Volhard, The origin of pattern and polarity in the drosophila embryo, Cell 68 (2) (1992) 201–219.

[5] H. Chen, Z. Xu, C. Mei, D. Yu, S. Small, A system of repressor gradients spatially organizes the boundaries of bicoid-dependent target genes, Cell 149 (3) (2012) 618–629.

[6] G. Jiménez, A. Guichet, A. Ephrussi, J. Casanova, Relief of gene repression by torso rtk signaling: role of capicua in drosophila terminal and dorsoventral patterning, Genes & development 14 (2) (2000) 224–231.

[7] M. Tanaka, T. Yoshimoto, T. Nakamura, A double-edged sword: The world according to capicua in cancer, Cancer science 108 (12) (2017) 2319–2325.

[8] R. A. Okimoto, F. Breitenbuecher, V. R. Olivas, W. Wu, B. Gini, M. Hofree, S. Asthana, G. Hrustanovic, J. Flanagan, A. Tulpule, et al., Inactivation of capicua drives cancer metastasis, Nature genetics 49 (1) (2017) 87–96.

[9] W. Driever, C. Nüsslein-Volhard, A gradient of bicoid protein in drosophila embryos, Cell 54 (1) (1988) 83–93.

[10] W. Driever, C. Nüsslein-Volhard, The bicoid protein determines position in the drosophila embryo in a concentration-dependent manner, Cell 54 (1) (1988) 95–104.

[11] O. Grimm, V. S. Zini, Y. Kim, J. Casanova, S. Y. Shvartsman, E. Wieschaus, Torso rtk controls capicua degradation by changing its subcellular localization, Development 139 (21) (2012) 3962–3968.

[12] L. Wolpert, Positional information and the spatial pattern of cellular differentiation, Journal of theoretical biology 25 (1) (1969) 1–47.

[13] T. Gregor, W. Bialek, R. R. D. R. Van Steveninck, D. W. Tank, E. F. Wieschaus, Diffusion and scaling during early embryonic pattern formation, Proceedings of the National Academy of Sciences 102 (51) (2005) 18403–18407.

[14] B. Houchmandzadeh, E. Wieschaus, S. Leibler, Precise domain specification in the developing drosophila embryo, Physical Review E 72 (6) (2005) 061920.

[15] U. Löhr, H.-R. Chung, M. Beller, H. Jäckle, Antagonistic action of bicoid and the repressor capicua determines the spatial limits of drosophila head gene expression domains, Proceedings of the National Academy of Sciences 106 (51) (2009) 21695–21700.

$$G(\tau) = \frac{1}{N}(1 + \frac{T}{1-T}e^{-\frac{\tau}{\tau_T}})[\frac{e^{-k_{\text{off}}\tau}}{1+\frac{k_{\text{off}}}{k_{\text{on}}}} + \frac{1}{1+\frac{k_{\text{on}}}{k_{\text{off}}}}\frac{e^{-k_{\text{on}}\tau}}{(1+\tau/\tau_D)\sqrt{1+\tau/(S^2\tau_D)}} + \frac{k_{\text{on}}k_{\text{off}}}{k_{\text{on}}+k_{\text{off}}}\sum_{n=1}^{\infty}\frac{1}{(n-1)!n!}\int_0^{\tau} ds \frac{e^{-k_{\text{off}}(\tau-s)-k_{\text{on}}s}}{(1+s/\tau_D)\sqrt{1+s/(S^2\tau_D)}}(2n + k_{\text{off}}s + k_{\text{on}}(\tau-s))(k_{\text{on}}k_{\text{off}}s(\tau-s))^{n-1}] + C. \quad (8)$$

[16] U. Löhr, H.-R. Chung, M. Beller, H. Jäckle, Bicoid-morphogen function revisited: Morphogen function revisited, Fly 4 (3) (2010) 236–240.

[17] Y. Kim, M. Coppey, R. Grossman, L. Ajuria, G. Jiménez, Z. Paroush, S. Y. Shvartsman, Mapk substrate competition integrates patterning signals in the drosophila embryo, Current biology 20 (5) (2010) 446–451.

[18] I. M. Adcock, G. Caramori, Chapter 31 - transcription factors, in: P. J. Barnes, J. M. Drazen, S. I. Rennard, N. C. Thomson (Eds.), Asthma and COPD (Second Edition), second edition Edition, Academic Press, Oxford, 2009, pp. 373 – 380.

[19] K. Gaston, P.-S. Jayaraman, Transcriptional repression in eukaryotes: repressors and repression mechanisms, Cellular and Molecular Life Sciences CMLS 60 (4) (2003) 721–741.

[20] N. Reynolds, A. O'Shaughnessy, B. Hendrich, Transcriptional repressors: multifaceted regulators of gene expression, Development 140 (3) (2013) 505–512.

[21] S. Weissmann, P. A. Cloos, S. Sidoli, O. N. Jensen, S. Pollard, K. Helin, The tumor suppressor cic directly regulates mapk pathway genes via histone deacetylation, Cancer research 78 (15) (2018) 4114–4125.

[22] E. Eck, J. Liu, M. Kazemzadeh-Atoufi, S. Ghoreishi, S. A. Blythe, H. G. Garcia, Quantitative dissection of transcription in development yields evidence for transcription-factor-driven chromatin accessibility, Elife 9 (2020) e56429.

[23] T. Lucas, H. Tran, C. A. Perez Romero, A. Guillou, C. Fradin, M. Coppey, A. M. Walczak, N. Dostatni, 3 minutes to precisely measure morphogen concentration, PLoS genetics 14 (10) (2018) e1007676.

[24] A. L. Patel, L. Zhang, S. E. Keenan, C. A. Rushlow, C. Fradin, S. Y. Shvartsman, Capicua is a fast-acting transcriptional brake, Current Biology 31 (16) (2021) 3639–3647.

[25] A. Porcher, A. Abu-Arish, S. Huart, B. Roelens, C. Fradin, N. Dostatni, The time to measure positional information: maternal hunchback is required for the synchrony of the bicoid transcriptional response at the onset of zygotic transcription, Development 137 (16) (2010) 2795–2804.

[26] X. Ma, D. Yuan, K. Diepold, T. Scarborough, J. Ma, The drosophila morphogenetic protein bicoid binds dna cooperatively, Development 122 (4) (1996) 1195–1206.

[27] F. J. Lopes, A. V. Spirov, P. M. Bisch, The role of bicoid cooperative binding in the patterning of sharp borders in drosophila melanogaster, Developmental biology 370 (2) (2012) 165–172.

[28] H. Tran, J. Desponds, C. A. Perez Romero, M. Coppey, C. Fradin, N. Dostatni, A. M. Walczak, Precision in a rush: Trade-offs between reproducibility and steepness of the hunchback expression pattern, PLoS computational biology 14 (10) (2018) e1006513.

[29] H. Tran, A. M. Walczak, N. Dostatni, Constraints and limitations on the transcriptional response downstream of the bicoid morphogen gradient, Current topics in developmental biology 137 (2020) 119–142.

[30] J. Desponds, H. Tran, T. Ferraro, T. Lucas, C. Perez Romero, A. Guillou, C. Fradin, M. Coppey, N. Dostatni, A. M. Walczak, Precision of readout at the hunchback gene: analyzing short transcription time traces in living fly embryos, PLoS computational biology 12 (12) (2016) e1005256.

[31] L. Mirny, M. Slutsky, Z. Wunderlich, A. Tafvizi, J. Leith, A. Kosmrlj, How a protein searches for its site on dna: the mechanism of facilitated diffusion, Journal of Physics A: Mathematical and Theoretical 42 (43) (2009) 434013.

[32] O. Bénichou, C. Chevalier, B. Meyer, R. Voituriez, Facilitated diffusion of proteins on chromatin, Physical review letters 106 (3) (2011) 038102.

[33] C. Fradin, On the importance of protein diffusion in biological systems: The example of the bicoid morphogen gradient, Biochimica Et Biophysica Acta (BBA)-Proteins and Proteomics 1865 (11) (2017) 1676–1686.

[34] M. Kato, T. W. Han, S. Xie, K. Shi, X. Du, L. C. Wu, H. Mirzaei, E. J. Goldsmith, J. Longgood, J. Pei, et al., Cell-free formation of rna granules: low complexity sequence domains form dynamic fibers within hydrogels, Cell 149 (4) (2012) 753–767.

[35] S. Chong, C. Dugast-Darzacq, Z. Liu, P. Dong, G. M. Dailey, C. Cattoglio, A. Heckert, S. Banala, L. Lavis, X. Darzacq, et al., Imaging dynamic and selective low-complexity domain interactions that control gene transcription, Science 361 (6400) (2018).

[36] A. Tsai, R. Galupa, J. Crocker, Robust and efficient gene regulation through localized nuclear microenvironments, Development 147 (19) (2020) dev161430.

[37] T. Jana, S. Brodsky, N. Barkai, Speed–specificity trade-offs in the transcription factors search for their genomic binding sites, Trends in Genetics 37 (5) (2021) 421–432.

[38] M. Mir, A. Reimer, J. E. Haines, X.-Y. Li, M. Stadler, H. Garcia, M. B. Eisen, X. Darzacq, Dense bicoid hubs accentuate binding along the morphogen gradient, Genes & development 31 (17) (2017) 1784–1794.

[39] M. Mir, M. R. Stadler, S. A. Ortiz, C. E. Hannon, M. M. Harrison, X. Darzacq, M. B. Eisen, Dynamic multifactor hubs interact transiently with sites of active transcription in drosophila embryos, Elife 7 (2018) e40497.

[40] B. Lim, M. S. Levine, Enhancer-promoter communication: hubs or loops?, Current Opinion in Genetics & Development 67 (2021) 5–9.

[41] T. Gregor, E. F. Wieschaus, A. P. McGregor, W. Bialek, D. W. Tank, Stability and nuclear dynamics of the bicoid morphogen gradient, Cell 130 (1) (2007) 141–152.

[42] F. Mueller, P. Wach, J. G. McNally, Evidence for a common mode of transcription factor interaction with chromatin as revealed by improved quantitative fluorescence recovery after photobleaching, Biophysical journal 94 (8) (2008) 3323–3339.

[43] M. Kang, C. A. Day, A. K. Kenworthy, E. DiBenedetto, Simplified equation to extract diffusion coefficients from confocal frap data, Traffic 13 (12) (2012) 1589–1600.

[44] D. Dey, S. Marciano, A. Nunes-Alves, V. Kiss, R. C. Wade, G. Schreiber, Line-frap, a versatile method to measure diffusion rates in vitro and in vivo, Journal of molecular biology 433 (9) (2021) 166898.

[45] F. Waharte, C. M. Brown, S. Coscoy, E. Coudrier, F. Amblard, A two-photon frap analysis of the cytoskeleton dynamics in the microvilli of intestinal cells, Biophysical journal 88 (2) (2005) 1467–1478.

[46] A. Abu-Arish, P. Kalab, J. Ng-Kamstra, K. Weis, C. Fradin, Spatial distribution and mobility of the ran gtpase in live interphase cells, Biophysical journal 97 (8) (2009) 2164–2178.

[47] L. Zhang, C. Perez-Romero, N. Dostatni, C. Fradin, Using fcs to accurately measure protein concentration in the presence of noise and photobleaching, Biophysical Journal (2021).

[48] C. Yeung, M. Shtrahman, X.-l. Wu, Stick-and-diffuse and caged diffusion: a comparison of two models of synaptic vesicle dynamics, Biophysical journal 92 (7) (2007) 2271–2280.

[49] O. Grimm, E. Wieschaus, The bicoid gradient is shaped independently of nuclei, Development 137 (17) (2010) 2857–2862.

[50] A. Abu-Arish, A. Porcher, A. Czerwonka, N. Dostatni, C. Fradin, High mobility of bicoid captured by fluorescence correlation spectroscopy: implication for the rapid establishment of its gradient, Biophysical Journal 99 (4) (2010) L33–L35.

[51] G. Jiménez, S. Y. Shvartsman, Z. Paroush, The capicua repressor–a general sensor of rtk signaling in development and disease, Journal of cell science 125 (6) (2012) 1383–1391.

[52] L. Durrieu, D. Kirrmaier, T. Schneidt, I. Kats, S. Raghavan, L. Hufnagel, T. E. Saunders, M. Knop, Bicoid gradient formation mechanism and dynamics revealed by protein lifetime analysis, Molecular systems biology

14 (9) (2018) e8355.
[53] C. E. Hannon, S. A. Blythe, E. F. Wieschaus, Concentration dependent chromatin states induced by the bicoid morphogen gradient, Elife 6 (2017) e28275.
[54] J. Liu, N. B. Perumal, C. J. Oldfield, E. W. Su, V. N. Uversky, A. K. Dunker, Intrinsic disorder in transcription factors, Biochemistry 45 (22) (2006) 6873–6888.
[55] Y. Minezaki, K. Homma, A. R. Kinjo, K. Nishikawa, Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation, Journal of molecular biology 359 (4) (2006) 1137–1149.
[56] M. Fuxreiter, I. Simon, S. Bondos, Dynamic protein–dna recognition: beyond what can be seen, Trends in biochemical sciences 36 (8) (2011) 415–423.
[57] Y. Wang, L. A. Benton, V. Singh, G. J. Pielak, Disordered protein diffusion under crowded conditions, The journal of physical chemistry letters 3 (18) (2012) 2703–2706.
[58] J. Wang, Y. Bian, X. Cao, N. Zhao, Understanding diffusion of intrinsically disordered proteins in polymer solutions: A disorder plus collapse model, AIP Advances 7 (11) (2017) 115120.
[59] A. Michelman-Ribeiro, D. Mazza, T. Rosales, T. J. Stasevich, H. Boukari, V. Rishi, C. Vinson, J. R. Knutson, J. G. McNally, Direct measurement of association and dissociation rates of dna binding in live cells by fluorescence correlation spectroscopy, Biophysical journal 97 (1) (2009) 337–346.
[60] V. Vukojević, D. K. Papadopoulos, L. Terenius, W. J. Gehring, R. Rigler, Quantitative study of synthetic hox transcription factor–dna interactions in live cells, Proceedings of the National Academy of Sciences 107 (9) (2010) 4093–4098.
[61] P. Brazda, T. Szekeres, B. Bravics, K. Toth, G. Vámosi, L. Nagy, Live-cell fluorescence correlation spectroscopy dissects the role of coregulator exchange and chromatin binding in retinoic acid receptor mobility, Journal of cell science 124 (21) (2011) 3631–3642.
[62] G. Kaur, M. W. Costa, C. M. Nefzger, J. Silva, J. C. Fierro-González, J. M. Polo, T. D. Bell, N. Plachta, Probing transcription factor diffusion dynamics in the living mammalian embryo with photoactivatable fluorescence correlation spectroscopy, Nature communications 4 (1) (2013) 1–13.
[63] P. H. von Hippel, O. G. Berg, Facilitated target location in biological systems, Journal of Biological Chemistry 264 (2) (1989) 675–678.
[64] A. Zidovska, The rich inner life of the cell nucleus: dynamic organization, active flows, and emergent rheology, Biophysical Reviews (2020) 1–14.
[65] B. R. Sabari, A. Dall?Agnese, A. Boija, I. A. Klein, E. L. Coffey, K. Shrinivas, B. J. Abraham, N. M. Hannett, A. V. Zamudio, J. C. Manteiga, et al., Coactivator condensation at super-enhancers links phase separation and gene control, Science 361 (6400) (2018).
[66] S. Brodsky, T. Jana, K. Mittelman, M. Chapal, D. K. Kumar, M. Carmi, N. Barkai, Intrinsically disordered regions direct transcription factor in vivo binding specificity, Molecular cell 79 (3) (2020) 459–471.
[67] S. Brodsky, T. Jana, N. Barkai, Order through disorder: The role of intrinsically disordered regions in transcription factor binding specificity, Current Opinion in Structural Biology 71 (2021) 110–115.
[68] D. A. Garcia, T. A. Johnson, D. M. Presman, G. Fettweis, K. Wagh, L. Rinaldi, D. A. Stavreva, V. Paakinaho, R. A. Jensen, S. Mandrup, et al., An intrinsically disordered region-mediated confinement state contributes to the dynamics and function of transcription factors, Molecular Cell 81 (7) (2021) 1484–1498.
[69] F. Lu, T. Lionnet, Transcription factor dynamics, Cold Spring Harbor Perspectives in Biology (2021) a040949.
[70] C. A. Perez-Romero, H. Tran, M. Coppey, A. M. Walczak, C. Fradin, N. Dostatni, Live imaging of mrna transcription in drosophila embryos, in: Morphogen Gradients, Springer, 2018, pp. 165–182.
[71] S. Berg, D. Kutra, T. Kroeger, C. N. Straehle, B. X. Kausler, C. Haubold, M. Schiegg, J. Ales, T. Beier, M. Rudy, et al., Ilastik: interactive machine learning for (bio) image analysis, Nature Methods 16 (12) (2019) 1226–1232.
[72] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al., Fiji: an open-source platform for biological-image analysis, Nature methods 9 (7) (2012) 676–682.
[73] Z. Petrášek, P. Schwille, Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy, Biophysical journal 94 (4) (2008) 1437–1448.
[74] M. Wachsmuth, W. Waldeck, J. Langowski, Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy, Journal of molecular biology 298 (4) (2000) 677–689.
[75] M. Rose, M. Kurylowicz, M. Mahmood, S. Winkel, J. M. Moran-Mirabal, C. Fradin, Direct measurement of the affinity between tbid and bax in a mitochondria-like membrane, International journal of molecular sciences 22 (15) (2021) 8240.

Supplemental Material
Activator vs. repressor: The differences and similarities in the mobilities of two early fly embryo transcription factors with opposite functions

Lili Zhang, Ahmad Mahmood, Carmina Perez-Romero, Cécile Fradin

1 Influence of nuclear size on the estimate of diffusion coefficients from fast FRAP data

Since D depends not only on the recovery time measured by fast FRAP but also on the radius of the nucleus in which the experiment was performed, and since nuclear size changes with nuclear cycle, we examined further the nuclear cycle and nuclear radius dependence of the measured recovery time and estimated diffusion coefficient. The radius of nuclei is around 3 μm at nc13 and around 2 μm at nc14 for all three types of proteins (Fig. S1A). However, the recovery time does not visibly change between nc13 and nc14 (Fig. S1C). Consequently, the diffusion coefficient for nc13 seems slightly higher than for nc 14 (Fig. S1B) - however one has to remember that the estimated value is only a lower limit, and the difference we see can simply be explained by the fact that the larger nuclei in nc13 allow a more stringent test of the diffusion coefficient of the fluorescence proteins using fast FRAP, allowing us to push the limit for D a bit hhigher. The relationship between recovery time and $Diameter^2$ is also shown in Fig. S1D.

2 Comparing the goodness of fit for the two-component model and the stick-and-diffuse model

A comparison of the fit of autocorrelation functions (ACFs) with either the Two-Component Model (TCM) or the Stick-and-Diffuse Model (SDM), for the three types of proteins studied, is shown in Fig. S2. The goodness of fit is good for both the TCM and the SDM, with the TCM having a slightly lower mean-squared difference between the fit and the data.

3 Simulations of protein diffusion

The code used to simulate fluorescent protein diffusion and confocal imaging can be found in the Appendix for simulation.

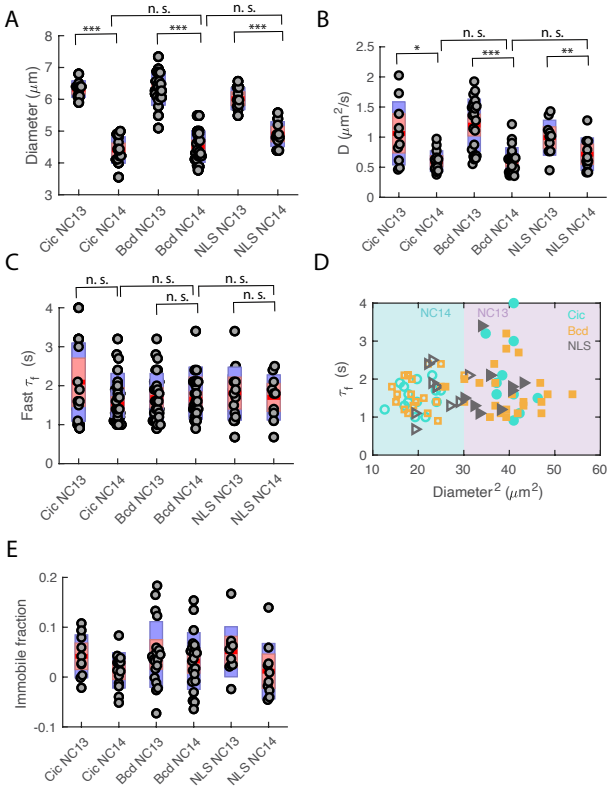

Figure S1: **Effect of nuclear radius on results from Fast FRAP experiments.** (**A**) Diameter of the bleached nuclei in fast FRAP experiments for Cic, Bcd and NLS at nc13 and nc14. (**B**) Diffusion coefficient, D, and (**C**) recovery time,τ_f, and (**E**) Immobile fraction for the bleached nuclei in Fast FRAP experiments for Cic, Bcd, NLS at NC 13 and NC 14. (**D**) Scatter plot of recovery time,τ_f, as a function of diameter squared, of the bleached nuclei in Fast FRAP experiments for Cic, Bcd, NLS at nc13 and nc14. Filled symbols represent data from NC13, while open symbols represent data from NC14. Three stars ($*$) indicate that p value from t test is less or equal than 0.001, two stars indicate p value is less or equal than 0.01, one star indicates p value is less or equal than 0.05, and n.s. indicates p is larger than 0.5, meaning there is no significant difference between the two sets of data.

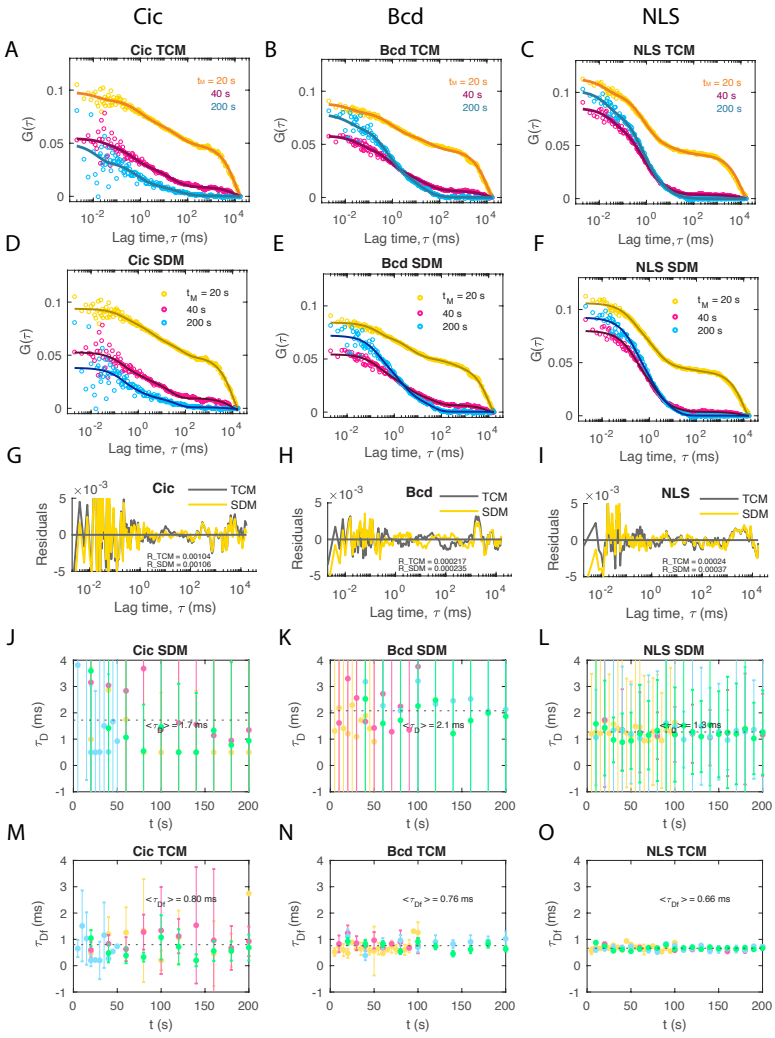

Figure S2: **Comparison between the TCM and the SDM.** (**A,B,C**) Representative data (open circles) and fit with the TCM including a photobleaching term (solid line) of the first (orange), second (pink) and tenth (blue) ACFs from a series of single-point FCS measurements in the center of a nucleus in Drosophila embryos expressing Cic-sfGFP (**A**), Bcd-eGFP (**B**), or NLS-eGFP (**C**). (**D,E,F**) Same ACFs as in (**A,B,C**), but fit with the SDM plus a constant term, and cutting off the data at longer lag times to account for photobleaching in a simple way. Each measurement in the series lasted 20 s in this case, thus the first, second and tenth ACFs corresponds to $t_M = 20$ s, 40 s, and 200 s, respectively. (**G,H,I**) Comparison of the residuals obtained from fitting the first ACF with the TCM (grey line) and the SDM (yellow line) are shown in the third row, with the mean-squared residuals shown at the bottom. To further compare the goodness of fit from the two models, the characteristic time τ_D obtained from the SDM (**J,K,L**) and the fast diffusion characteristic time τ_{Df} from the TCM (**M,N,O**) are also shown, where symbols of different color represent different series of data with different measurement time (5, 10, 20s). The error bars for τ_D from SDM are 50 % confidence interval (CI) while the error bars for τ_{Df} from TCM are 50 % CI.

Chapter 6

Conclusion and Future Outlook

In this thesis, I reported on my study of the concentration and dynamics of two morphogens, an activator Bicoid and a repressor Capicua, in early *Drosophila* embryos, using confocal fluorescence microscopy techniques (FCS and FRAP). The main result of this thesis is that in spite of their opposite function, Bcd and Cic have similar nuclear concentrations (on the order of a few 100 nM, or $\sim$ 10,000 molecules per nucleus) and similar intranuclear dynamics (no immobile fractions, the capability to redistribute within a nucleus in about 1 s, and the presence of a small fraction of molecules seemingly transiently slower). This suggests that there are strong similarities in the mechanisms by which these proteins act on transcription. The presence of disordered regions on either side of their DNA binding regions further suggests they might both participate in the formation of transcriptional condensates, i.e. phase separated regions with slower mobility but higher concentration of transcription factors, that sit at transcription sites. In contrast, in terms of nucleo-cytoplasmic dynamics, Bcd and Cic behave quite differently where Bcd can easily pass though the nuclear envelop from the cytoplasm while Cic cannot even after 20 min. This is probably closely

related to the different ways in which these proteins form their gradient. While Bcd gradient is formed by diffusion (thus requiring Bcd to be able to move from nucleus to nucleus), Cic gradient is formed by local degradation of the protein (thus requiring Cic staying in place in order for the gradient to be maintained). The details of the new comparison of dynamics of Bcd and Cic are described in Chapters 4 and 5. Chapter 4 contains results of concentration and dynamics of Cic. All of these results regarding concentration and dynamics of Bcd and Cic provide information for a better understanding of the mechanism of transcriptional activation and repression, which are essential for proper development of eukaryotic organisms.

An improved method to accurately acquire absolute concentration of morphorgens *in vivo* from FCS is described in Chapter 3. Before the measurements applied in Chapters 4 and 5 were possible, I had to find ways to exploit FCS data acquired in the difficult conditions imposed by the developing embryo. The measurements in fly embryo are challenging for a number of reasons. They have to be done at a precise location (mid-embryo) and at a precise time (NC 13 to 14). More importantly, they have to be done in the presence of background autofluorescence (very prevalent in the young fly embryo) and in spite of strong photobleaching. The strategy I proposed in Chapter 3 to exploit FCS data in spite of these issues should be useful in the future to many performing FCS experiments *in vivo*.

As for future work, at least three routes based off this thesis exist that could be explored further. 1) The proposed method of obtaining accurate protein concentration in the presence of noise and photobleaching (Chapter 3) could be applied to many other types of transcription factors, morphogens, and proteins of interest in *Drosophila* embryos or other organisms. With more accurate concentration numbers, their modes

of function could be made clearer to us. With the advancement of instrumentation, four dimensional (XYZT) concentration maps of those proteins of importance could be a valuable resource for researchers in the fields of biophysics, biochemistry and biology. 2) The existence of condensates of Bcd in the nuclei has been deduced from the observation of single Bcd trajectories (Mir *et al.*, 2017, 2018). It is thought that these condensates could help buffering the concentration of TFs in dilute phases. Selective plane illumination microscopy (SPIM) or STED microscopy could be used to image nuclei with better time resolution (SPIM) or spatial resolution (STED). This could definitely prove the existence of these elusive structures. 3) Dynamical simulations performed for various dynamical models (several diffusing populations, transient interaction with DNA, and condensate formation) should be performed to select out a model that best fits all of the experimental observations for Bcd and Cic. One should then see if that model can be generalized to explain the motions of other transcription factors and morphogens in *Drosophila* and beyond.

Appendix A

Appendix for fluorescence experiments

A.1 Laser power stability during FCS experiments

A stable and reliable source of laser in both FCS and FRAP experiments is a key factor for acquiring decent data. It is best to allow the laser to warm up for at least half an hour before each experiment. It is also recommended to record the laser power before, during and after an experiment, as many times as needed. In my thesis projects, the laser power for the FCS instrument is set by turning a knob and measured using a power meter, whereas the laser power for FRAP instrument is set digitally on its software. Below is an example of the laser power monitored for about 50 min before an FCS experiment (Fig. A.1). It can be seen that the laser power reaches a plateau after about 20 min and the variation in laser power afterwards is ~ 2 µW.

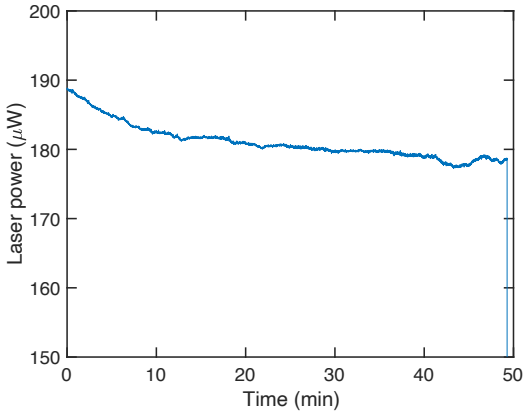

Figure A.1: An example of the laser power monitored during the warm-up of an FCS experiment.

A.2 Accuracy of the set bleaching area

During a FRAP experiment, the bleaching area is manually selected by the experimentalist by dragging a circle or a square or even drawing freehand. As the radius of a nucleus is about 3 µm in an early developing embryo in the experiment, it is of great importance that the set bleaching area is as accurate as possible, especially during the fast FRAP experiment where a half nucleus is bleached and the other half is unbleached, in order to study the intranulcear dynamics of morphogens. Indeed, the accuracy of the set bleaching area is quite accurate as shown in the example in Fig. A.2, where bleaching area of various shapes and sizes including a small dot, a square, a half circle, a star, and a letter 'Z' were bleached on the vitelline membrane of a *D. melanogaster* embryo expressing Bcd-eGFP. The scale bar is 10 µm.

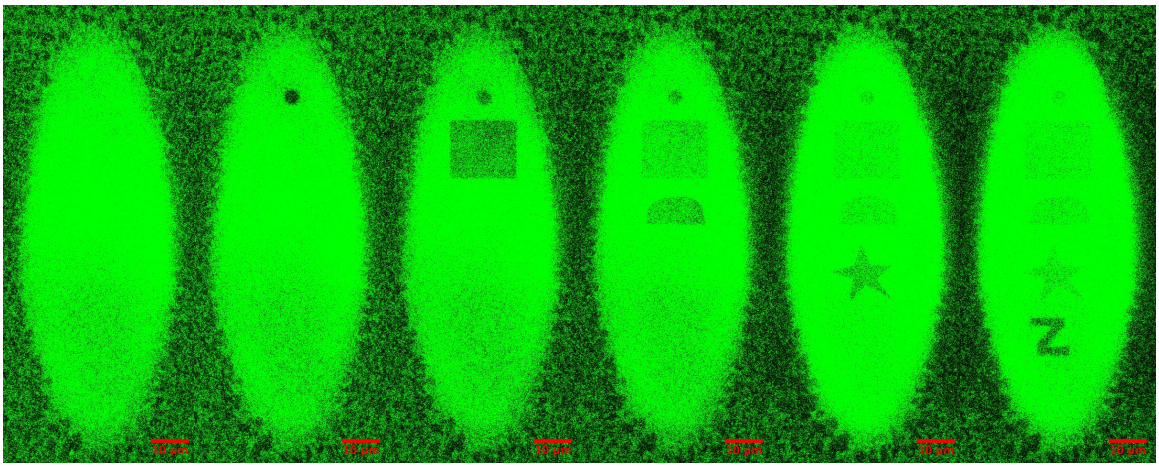

Figure A.2: A test of the accuracy of the set bleaching area for FRAP experiments. Bleaching area of various shapes and sizes including a small dot, a square, a half circle, a star, and a letter 'Z' were bleached on the vitelline membrane of a *D. melanogaster* embryo expressing Bcd-eGFP. The scale bar is 10 µm.

A.3 Solving for particle number from its quadratic equation

When the fluorescence signal is equavalent to that from the background noise, the amplitude of the Autocorrelation function, $G(0)$ is no longer simply the inverse of the particle number, N, but should be expressed as in Eq. 2.3.9. By rearranging it into a quadratic function of N, one obtains:

$$N^2 + (2m - \frac{1}{G(0)})N + m^2 = 0. \tag{A.3.1}$$

where $m = I_B/(\gamma B)$ as defined earlier in the Method Chapter. during the fitting of ACF with appropriate model, including the one-component model, values of I_B, B, and $G(0)$ can be obtained and γ is simply a constant of $2^{3/2}$. Thus Eq. A.3.1 can be solved and an example of the two solution from solving this quadratic equation

is shown in Fig. A.3, as well as the comparison between calculated N using $G(0)$ with and without the background noise. The samples are serial dilution solutions of Alexa Fluor 488 and the star symbol represents the deionised water used to dilute the solution.

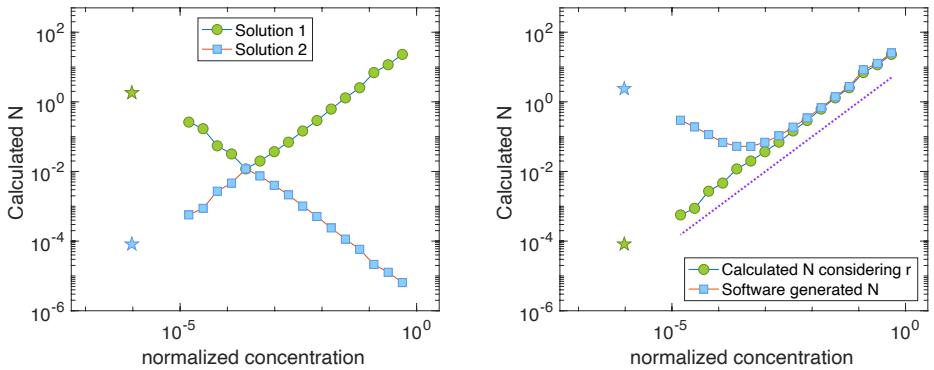

Figure A.3: An example of solving for particle number N from its quadratic equation, as well as a comparison between calculated N using $G(0)$ with and without the background noise. The samples are serial dilution solutions of Alexa Fluor 488 and the star symbol represents the deionised water used to dilute the solution.

A.4 Correction for inhomogeneous illumination

It is almost inevitable that the illumination in the widest field of view is not completely evenly distributed and proper correction for the inhomogeneous illumination should be carried out. The procedure that we take for this correction is as follows. First a calibration image of a dye solution, in this case, Alexa Fluor 488, is obtained, as shown in Fig. A.4 a, and this 2D image is fitted with a 2D Gaussian (Fig. A.4 b) model, as in Eq. A.4.1, which is a combination of the unrotated 2D Gaussian (Eq.

A.4.2), with the effect of a rotation matrix (Eq.A.4.3) (Diaz, 2021).

$$f(x,y) = Aexp(-(\frac{((x-x_0)cos\theta-(y-y_0)sin\theta)^2}{2{\sigma_X}^2}+\frac{((x-x_0)sin\theta-(y-y_0)cos\theta)^2}{2{\sigma_Y}^2})) \tag{A.4.1}$$

$$f(x,y) = Aexp(-(\frac{(x-x_0)^2}{2{\sigma_X}^2}+\frac{(y-y_0)^2}{2{\sigma_Y}^2})) \tag{A.4.2}$$

$$R = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \tag{A.4.3}$$

The coefficient A is the amplitude. x_0, y_0 are the coordinates of centre of the beam. σ_x, σ_y are the spreads along x and y axis, respectively. The angle, θ, is the rotation of the 2D Gaussian with respect to the positive x axis. The corrected filed of view is then the resultant image of the original dye image divided by the fitted image, as shown in Fig. A.4 c. Indeed, the histogram of the corrected image only shows one peak whereas that of the original image shows two peaks (Fig. A.4 d), demonstrating that the correction makes the illumination field of view more homogeneous.

A.5 Time lapse confocal fluorescent images of embryos expressing Bcd-eGFP and Cic-sfGFP

The work of this thesis focuses on the early development ($0 \sim 200$ min after egg deposition) of the *Drosophila melanogaster* embryo and how those morphogenic protein gradients form during this period. There are 14 mitotic cycles (i.e. nuclear cycles)

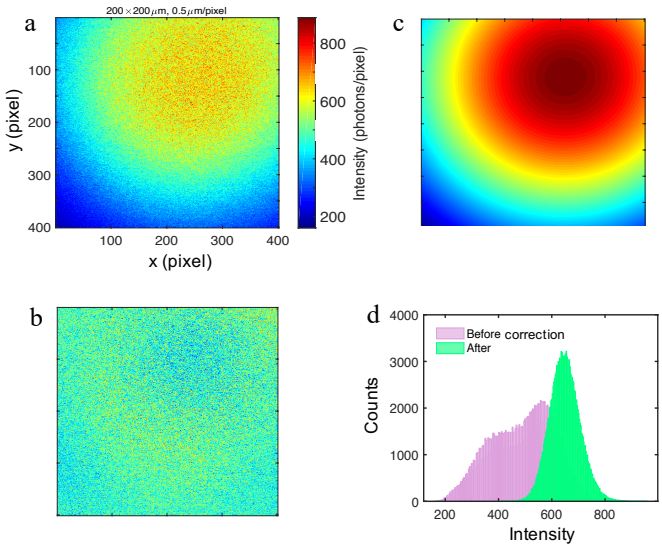

Figure A.4: Procedures for correcting the inhomogeneity of the filed of view during confocal microscopy imaging. First a calibration image of a dye solution, in this case, Alexa Fluor 488, is obtained (a), and this 2D image is then fitted with a 2D Gaussian (b). The corrected filed of view is the resultant image of the original dye image divided by the fitted image (c). Indeed, the histogram of the corrected image only shows one peak whereas that of the original image shows two peaks (d), demonstrating that the correction makes the illumination field of view more homogeneous.

from the original fertilized egg to a gastrulated embryo where anterior-posterior features start appearing Campos-Ortega and Hartenstein (2013). Mitotic cycle 1 - 13 last less than 30 min individually, while cycle 14 lasts more than 30 min and nuclei can be clearly observed. FCS on embryos are usually usually performed during cycle 12 - 14.

In order to pinpoint each cycle, especially cycle 13 and cycle 14, it is necessary to obtain time lapse images of a developing embryo serving as roadmap. Fig. A.5 and Fig. A.6 show the time lapse confocal fluorescent images of embryos expressing Bcd-eGFP and Cic-sfGFP, respectively.

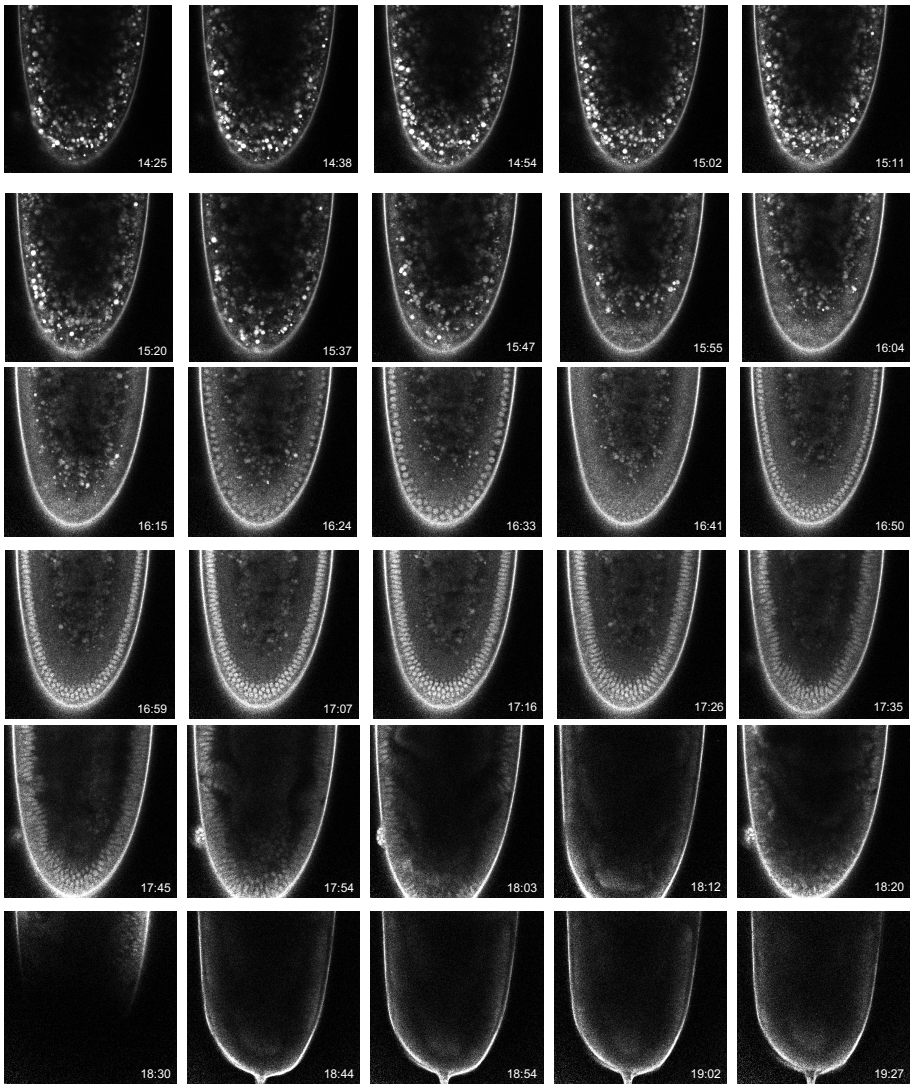

Figure A.5: Time lapse confocal fluorescent images of the *Drosophila melanogaster* embryo expressing Bcd-eGFP. The images contain the anterior end of the embryo, which is immersed in carbonhalo oil on a 0.17 mm thick coverslip. These images have been calibrated with calibration image from dye solution. The size of the image is 200 µm × 200 µm. The pixel size is 0.5 µm/pixel. The time stamp at the bottom right is in format of hh:mm. The half dark image at 18:30 was a result of insufficient immersion of water on the objective due to long hours of evaporation and was refilled before the next image was taken.

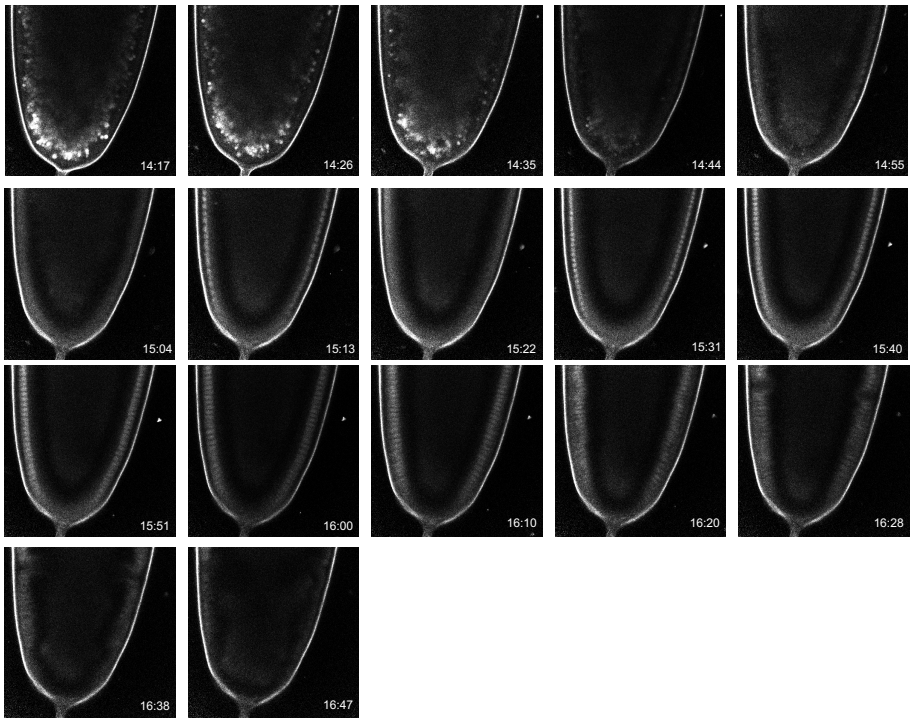

Figure A.6: Time lapse confocal fluorescent images of the *Drosophila melanogaster* embryo expressing Cic-sfGFP. The images contain the anterior end of the embryo, which is immersed in carbonhalo oil on a 0.17 mm thick coverslip. These images have been calibrated with calibration image from dye solution. The size of the image is 200 µm × 200 µm. The pixel size is 0.5 µm/pixel. The time stamp at the bottom right is in format of hh:mm.

Appendix B

Appendix for simulations

B.1 Flowchart of a dynamic simulation

Fig. B.1 is a flowchart of a dynamic simulation where there are three types of molecules, i.e. slow diffusing molecules in the nucleus (state = 0), fast diffusing molecules in the nucleus (state = 1) and fast diffusing molecules in the cytoplasm (state = 2). Molecules 1 and 2 can exchange between each other when they pass through the nucleus envelop, and the import and export of molecules happen with a probability of p_{in} and p_{out}, respectively. Molecules 0 and 1 can exchange between each other when a molecule binds or unbinds from a binding site, and the binding and unbinding happen with a probability of p_{on} and p_{off}, respectively. During each simulation step, each particle undergoes random walk along each coordinate with a step size drawn from a Gaussian distribution with mean at 0 and variance $\sqrt{2D\delta t}$ where D is the diffusion coefficient and δt is the step time, if it is allowed to move, i.e. passing those probability test if it is about to change state. Otherwise, it remains at its location. In the example shown here, slow diffusing molecules, i.e. molecules

tagged with state = 0, are not allowed to move during a simulation step. Ellipse element represents an iterative `for` loop. Diamond element represents an `if` expression. Rectangular element represents an execution statement.

B.2 Simulation code example

A simulation code example written in Python is shown here. More resources on the Python code used to simulate confocal images are available in the publicly accessible repository: `https://github.com/cecilefradin/BidBax_Simulation_and_Analysis` (accessed on 30 July 2021) (Rose *et al.*, 2021).

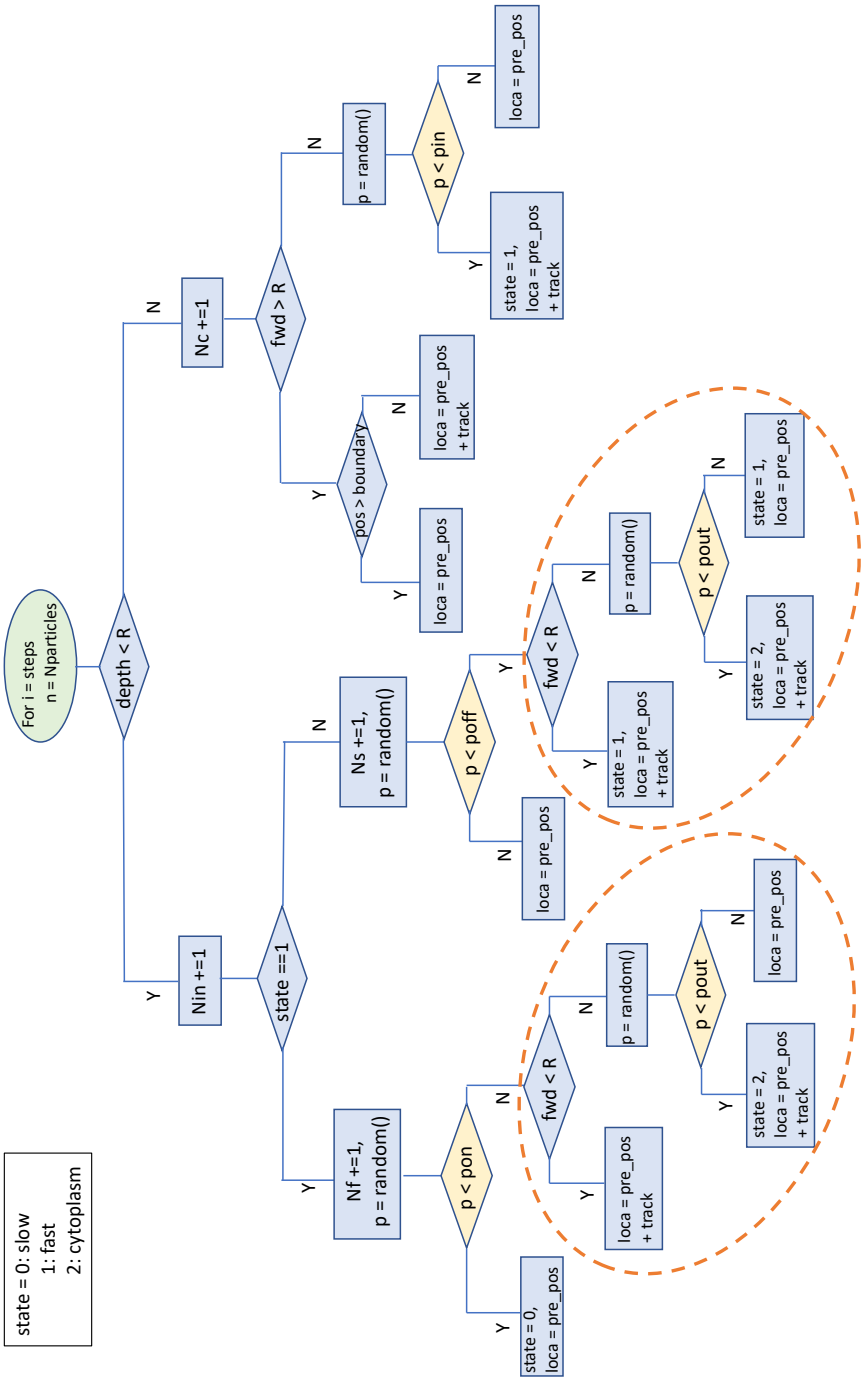

Figure B.1: A flowchart of a dynamic simulation with three types of molecules, i.e. slow diffusing molecules in the nucleus (state =0), fast diffusing molecules in the nucleus (state = 1) and fast diffusing molecules in the cytoplasm (state =2).

```
#!/usr/bin/env python
# coding: utf-8

# The principle behind the confocal image simulation - FRAP simulation
# This program intends to simulate the scenario immediately after the
bleaching event, where a whole nucleus is photobleached.

#_____________________________________________________________________
# Importing Packages.

import numpy as np

#_____________________________________________________________________
# Image properties.

image_size = 40 # px (The image will be square 40 x 40 pixels)
pixel_size = 0.2 # µm/pixel
boundary = image_size * pixel_size
radius = 3 ; # µm
dwell_time = 0.001 # s
psf_width = 0.3 # µm (Width of the point spread function in focus)
psf_height = 1.5 #
diff_const = 0.1 # µm^2/s (diffusion coefficient of mobile particles)
step_time = 0.001 # s
B = 1e4 # Brightness, Hz

Nparticles = 2000

center_pos = [4, 4,4] # the centre of the sphere

#_____________________________________________________________________
# Generate initial positions of particles, which are outside of
nucleus.

start_pos = np.zeros((Nparticles,3))
for n in range(Nparticles):
    temp = start_pos[n,:]
    while temp[0]**2 + temp[1]**2 + temp[2]**2 == 0:
        x = np.random.rand(3) * boundary
        if ((x[0] - 4)**2 + (x[1] - 4)**2 + (x[2] - 4)**2) > radius**2
and (x[0]**2 + x[1]**2 + x[2]**2) > 0:
            start_pos[n,:] = x

#_____________________________________________________________________
# Calculating the pixel intensity.
# The pixel intensity is dependent on the distance from the optical
axis.

def GaussianBeam( start_pos, beam_pos, psf_width, psf_height):
    if start_pos.shape[0] == 2:
```

```
        GB = B*step_time*np.exp(- 2* ((start_pos -
beam_pos)**2).sum()/ psf_width**2)
    else:
        GB = B*step_time*np.exp(- 2* ((start_pos[0:2] -
beam_pos[0:2])**2).sum()/ psf_width**2) * np.exp(-2*((start_pos[2]-
beam_pos[2])**2/psf_height**2))

    return GB

#_____________________________________________________________________
# More parameters for the movement of particles.

pout = 0.01 #  flow rate from nucleus to cytoplasm

pin = 0.04 # flow rate from cytoplasm to nucleus

steps = 60000 # number of steps in the simulation

pre_pos = np.zeros((steps+1,3,Nparticles)) # a 3D matrix storing the
previous position of particles
pre_pos[0,:,:] = np.transpose(start_pos)
depth = np.zeros((steps,Nparticles)) # the distance form the particle
to the center

# the size of step along x,y,z coordinate
track =
np.random.normal(loc=0,scale=np.sqrt(2*diff_const*step_time),size=(ste
ps,3,Nparticles))

loca = np.zeros((steps,3,Nparticles))

#_____________________________________________________________________
# Movement for each particle during each step.

for n in range(Nparticles):
    for i in range(steps):

        depth[i,n] = np.sqrt(((pre_pos[i,:,n] - center_pos)**2).sum())
        forwd = np.sqrt(((pre_pos[i,:,n] + track[i,:,n] -
center_pos)**2).sum())

        if depth[i,n]  <= radius: # radius = image_size * pixel_size /
4

            if forwd <= radius:
                loca[i,:,n] = pre_pos[i,:,n] + track[i,:,n]

            else:

                proba = np.random.rand()
```

```
                if proba >= 0 and proba <= pout :
                    loca[i,:,n] = pre_pos[i,:,n] + track[i,:,n]

                else:
                    loca[i,:,n] = pre_pos[i,:,n]
            else:
                if forwd >= radius:
                    x  = pre_pos[i,0,n] + track[i,0,n]
                    y  = pre_pos[i,1,n] + track[i,1,n]
                    z  = pre_pos[i,2,n] + track[i,2,n]

                    if  x > boundary or x < 0:
                         loca[i,0,n] = pre_pos[i,0,n]
                    else:
                        loca[i,0,n] = pre_pos[i,0,n] + track[i,0,n]

                    if y > boundary or y < 0:
                        loca[i,1,n] = pre_pos[i,1,n]
                    else:
                        loca[i,1,n] = pre_pos[i,1,n]+ track[i,1,n]

                    if z > boundary or z < 0:
                        loca[i,2,n] = pre_pos[i,2,n]
                    else:
                        loca[i,2,n] = pre_pos[i,2,n] + track[i,2,n]

                else:

                    proba = np.random.rand()
                    if proba >= 0 and proba <= pin :
                        loca[i,:,n] = pre_pos[i,:,n] + track[i,:,n]

                    else:
                        loca[i,:,n] = pre_pos[i,:,n] # - track[i,:,n]

            pre_pos[i+1,:,n] = loca[i,:,n]

    #_____________________________________________________________________
    # Calculate the intensity array at t = start and t = end for the
    centre z slice.
    z_slice = [19, 19]
    kk = [0, steps - image_size*image_size]  # the index for the start of
    the scanning

    image_array = np.zeros((image_size,image_size,len(z_slice)))
    image_array_mobile = np.zeros((image_size,image_size,len(z_slice)))


    for n in range(Nparticles):
```

```
    for k in range(0,2): # two images, one at the begninning, one at
the end
        for j in range(image_array.shape[1]): # x
            for i in range(image_array.shape[0]): # y
                beam_pos = np.array([i,j,19]) * pixel_size # only scan
the Z/2

                particle_pos = loca[ i + image_size * j + kk[k] ,:,n]
                image_array[i,j,k] +=
GaussianBeam(particle_pos,beam_pos,psf_width,psf_height)
image_array_mobile = np.array(image_array)
```

Bibliography

Abolaji, A., Kamdem, J., Farombi, E., and Rocha, J. (2013). Drosophila melanogaster as a promising model organism in toxicological studies. *Arch Bas App Med*, **1**, 33–38.

Abu-Arish, A., Kalab, P., Ng-Kamstra, J., Weis, K., and Fradin, C. (2009). Spatial distribution and mobility of the ran gtpase in live interphase cells. *Biophysical journal*, **97**(8), 2164–2178.

Abu-Arish, A., Porcher, A., Czerwonka, A., Dostatni, N., and Fradin, C. (2010). High mobility of bicoid captured by fluorescence correlation spectroscopy: implication for the rapid establishment of its gradient. *Biophysical Journal*, **99**(4), L33–L35.

Adcock, I. M. and Caramori, G. (2009). Chapter 31 - transcription factors. In P. J. Barnes, J. M. Drazen, S. I. Rennard, and N. C. Thomson, editors, *Asthma and COPD (Second Edition)*, pages 373 – 380. Academic Press, Oxford, second edition edition.

Allen, G. E. (1975). The introduction of drosophila into the study of heredity and evolution: 1900-1910. *Isis*, **66**(3), 322–333.

Auer, J. M., Stoddart, J. J., Christodoulou, I., Lima, A., Skouloudaki, K., Hall,

H. N., Vukojević, V., and Papadopoulos, D. K. (2020). Of numbers and movement–understanding transcription factor pathogenesis by advanced microscopy. *Disease Models & Mechanisms*, **13**(12), dmm046516.

Baumgartner, S. (2018). Seeing is believing: the bicoid protein reveals its path. *Hereditas*, **155**(1), 1–6.

Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., *et al.* (2019). Ilastik: interactive machine learning for (bio) image analysis. *Nature Methods*, **16**(12), 1226–1232.

Bushweller, J. H. (2019). Targeting transcription factors in cancer—from undruggable to reality. *Nature Reviews Cancer*, **19**(11), 611–624.

Campos-Ortega, J. A. and Hartenstein, V. (2013). *The embryonic development of Drosophila melanogaster*. Springer Science & Business Media.

Chen, H., Xu, Z., Mei, C., Yu, D., and Small, S. (2012). A system of repressor gradients spatially organizes the boundaries of bicoid-dependent target genes. *Cell*, **149**(3), 618–629.

Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996). Facs-optimized mutants of the green fluorescent protein (gfp). *Gene*, **173**(1), 33–38.

Cranfill, P. J., Sell, B. R., Baird, M. A., Allen, J. R., Lavagnino, Z., De Gruiter, H. M., Kremers, G.-J., Davidson, M. W., Ustione, A., and Piston, D. W. (2016). Quantitative assessment of fluorescent proteins. *Nature methods*, **13**(7), 557–562.

Dalcq, A. (1938). *Form and causality in early development*. CUP Archive.

Diaz, M. A. (2021). Fit 1d and 2d gaussian to noisy data. https://www.mathworks.com/matlabcentral/fileexchange/55033-fit-1d-and-2d-gaussian-to-noisy-data/. [MATLAB Central File Exchange; Retrieved August 16, 2021].

Driever, W. and Nüsslein-Volhard, C. (1989). The bicoid protein is a positive regulator of hunchback transcription in the early drosophila embryo. *Nature*, **337**(6203), 138–143.

Dubuis, J. O., Tkačik, G., Wieschaus, E. F., Gregor, T., and Bialek, W. (2013). Positional information, in bits. *Proceedings of the National Academy of Sciences*, **110**(41), 16301–16308.

Durrieu, L., Kirrmaier, D., Schneidt, T., Kats, I., Raghavan, S., Hufnagel, L., Saunders, T. E., and Knop, M. (2018). Bicoid gradient formation mechanism and dynamics revealed by protein lifetime analysis. *Molecular systems biology*, **14**(9), e8355.

Elson, E. L. and Magde, D. (1974). Fluorescence correlation spectroscopy. i. conceptual basis and theory. *Biopolymers: Original Research on Biomolecules*, **13**(1), 1–27.

Ferguson, E. L. and Anderson, K. V. (1992). Decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the drosophila embryo. *Cell*, **71**(3), 451–461.

Fradin, C. (2017). On the importance of protein diffusion in biological systems: The example of the bicoid morphogen gradient. *Biochimica Et Biophysica Acta (BBA)-Proteins and Proteomics*, **1865**(11), 1676–1686.

Gallet, A. (2011). Hedgehog morphogen: from secretion to reception. *Trends in cell biology*, **21**(4), 238–246.

Gaston, K. and Jayaraman, P.-S. (2003). Transcriptional repression in eukaryotes: repressors and repression mechanisms. *Cellular and Molecular Life Sciences CMLS*, **60**(4), 721–741.

Gregor, T., Bialek, W., Van Steveninck, R. R. D. R., Tank, D. W., and Wieschaus, E. F. (2005). Diffusion and scaling during early embryonic pattern formation. *Proceedings of the National Academy of Sciences*, **102**(51), 18403–18407.

Gregor, T., Tank, D. W., Wieschaus, E. F., and Bialek, W. (2007a). Probing the limits to positional information. *Cell*, **130**(1), 153–164.

Gregor, T., Wieschaus, E. F., McGregor, A. P., Bialek, W., and Tank, D. W. (2007b). Stability and nuclear dynamics of the bicoid morphogen gradient. *Cell*, **130**(1), 141–152.

Grimm, O. and Wieschaus, E. (2010). The bicoid gradient is shaped independently of nuclei. *Development*, **137**(17), 2857–2862.

Hammar, P., Leroy, P., Mahmutovic, A., Marklund, E. G., Berg, O. G., and Elf, J. (2012). The lac repressor displays facilitated diffusion in living cells. *Science*, **336**(6088), 1595–1598.

Holloway, D. M. and Spirov, A. V. (2015). Mid-embryo patterning and precision in drosophila segmentation: Krüppel dual regulation of hunchback. *PLoS One*, **10**(3), e0118450.

Houchmandzadeh, B., Wieschaus, E., and Leibler, S. (2005). Precise domain specification in the developing drosophila embryo. *Physical Review E*, **72**(6), 061920.

Houtsmuller, A. B. (2005). Fluorescence recovery after photobleaching: application to nuclear proteins. *Microscopy Techniques*, pages 177–199.

Hülskamp, M., Pfeifle, C., and Tautz, D. (1990). A morphogenetic gradient of hunchback protein organizes the expression of the gap genes krüppel and knirps in the early drosophila embryo. *Nature*, **346**(6284), 577–580.

Jayavelu, N. D., Jajodia, A., Mishra, A., and Hawkins, R. D. (2020). Candidate silencer elements for the human and mouse genomes. *Nature communications*, **11**(1), 1–15.

Jiménez, G., Guichet, A., Ephrussi, A., and Casanova, J. (2000). Relief of gene repression by torso rtk signaling: role of capicua in drosophila terminal and dorsoventral patterning. *Genes & development*, **14**(2), 224–231.

Jiménez, G., Shvartsman, S. Y., and Paroush, Z. (2012). The capicua repressor–a general sensor of rtk signaling in development and disease. *Journal of cell science*, **125**(6), 1383–1391.

Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., Mirzaei, H., Goldsmith, E. J., Longgood, J., Pei, J., *et al.* (2012). Cell-free formation of rna granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*, **149**(4), 753–767.

Koppel, D. E. (1974). Statistical accuracy in fluorescence correlation spectroscopy. *Physical Review A*, **10**(6), 1938.

Kremers, G.-J., Goedhart, J., van den Heuvel, D. J., Gerritsen, H. C., and Gadella, T. W. (2007). Improved green and blue fluorescent proteins for expression in bacteria and mammalian cells. *Biochemistry*, **46**(12), 3775–3783.

Lam, Y. C., Bowman, A. B., Jafar-Nejad, P., Lim, J., Richman, R., Fryer, J. D., Hyun, E. D., Duvick, L. A., Orr, H. T., Botas, J., *et al.* (2006). Ataxin-1 interacts with the repressor capicua in its native complex to cause sca1 neuropathology. *Cell*, **127**(7), 1335–1347.

Latchman, D. S. (1996). Transcription-factor mutations and disease. *New England Journal of Medicine*, **334**(1), 28–33.

Lebrecht, D., Foehr, M., Smith, E., Lopes, F. J., Vanario-Alonso, C. E., Reinitz, J., Burz, D. S., and Hanes, S. D. (2005). Bicoid cooperative dna binding is critical for embryonic patterning in drosophila. *Proceedings of the National Academy of Sciences*, **102**(37), 13176–13181.

Lee, S.-H. and Min, K.-J. (2019). Drosophila melanogaster as a model system in the study of pharmacological interventions in aging. *Translational Medicine of Aging*, **3**, 98–103.

Lim, B. and Levine, M. S. (2021). Enhancer-promoter communication: hubs or loops? *Current Opinion in Genetics & Development*, **67**, 5–9.

Lorén, N., Hagman, J., Jonasson, J. K., Deschout, H., Bernin, D., Cella-Zanacchi, F., Diaspro, A., McNally, J. G., Ameloot, M., Smisdom, N., *et al.* (2015). Fluorescence recovery after photobleaching in material and life sciences: putting theory into practice. *Quarterly reviews of biophysics*, **48**(3), 323–387.

Lu, F. and Lionnet, T. (2021). Transcription factor dynamics. *Cold Spring Harbor Perspectives in Biology*, page a040949.

Lucas, T., Tran, H., Perez Romero, C. A., Guillou, A., Fradin, C., Coppey, M., Walczak, A. M., and Dostatni, N. (2018). 3 minutes to precisely measure morphogen concentration. *PLoS genetics*, **14**(10), e1007676.

Ma, X., Yuan, D., Diepold, K., Scarborough, T., and Ma, J. (1996). The drosophila morphogenetic protein bicoid binds dna cooperatively. *Development*, **122**(4), 1195–1206.

Magde, D., Elson, E. L., and Webb, W. W. (1974). Fluorescence correlation spectroscopy. ii. an experimental realization. *Biopolymers: Original Research on Biomolecules*, **13**(1), 29–61.

Mir, M., Reimer, A., Haines, J. E., Li, X.-Y., Stadler, M., Garcia, H., Eisen, M. B., and Darzacq, X. (2017). Dense bicoid hubs accentuate binding along the morphogen gradient. *Genes & development*, **31**(17), 1784–1794.

Mir, M., Stadler, M. R., Ortiz, S. A., Hannon, C. E., Harrison, M. M., Darzacq, X., and Eisen, M. B. (2018). Dynamic multifactor hubs interact transiently with sites of active transcription in drosophila embryos. *Elife*, **7**, e40497.

Mirny, L., Slutsky, M., Wunderlich, Z., Tafvizi, A., Leith, J., and Kosmrlj, A. (2009). How a protein searches for its site on dna: the mechanism of facilitated diffusion. *Journal of Physics A: Mathematical and Theoretical*, **42**(43), 434013.

Mirzoyan, Z., Sollazzo, M., Allocca, M., Valenza, A. M., Grifoni, D., and Bellosta, P.

(2019). Drosophila melanogaster: a model organism to study cancer. *Frontiers in genetics*, **10**, 51.

Mitsis, T., Efthimiadou, A., Bacopoulou, F., Vlachakis, D., Chrousos, G. P., and Eliopoulos, E. (2020). Transcription factors and evolution: An integral part of gene expression. *World Academy of Sciences Journal*, **2**(1), 3–8.

Mueller, F., Wach, P., and McNally, J. G. (2008). Evidence for a common mode of transcription factor interaction with chromatin as revealed by improved quantitative fluorescence recovery after photobleaching. *Biophysical journal*, **94**(8), 3323–3339.

Nüsslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in drosophila. *Nature*, **287**(5785), 795–801.

Okimoto, R. A., Breitenbuecher, F., Olivas, V. R., Wu, W., Gini, B., Hofree, M., Asthana, S., Hrustanovic, G., Flanagan, J., Tulpule, A., *et al.* (2017). Inactivation of capicua drives cancer metastasis. *Nature genetics*, **49**(1), 87–96.

Ong, C., Yung, L.-Y. L., Cai, Y., Bay, B.-H., and Baeg, G.-H. (2015). Drosophila melanogaster as a model organism to study nanotoxicity. *Nanotoxicology*, **9**(3), 396–403.

Patel, A. L., Zhang, L., Keenan, S. E., Rushlow, C. A., Fradin, C., and Shvartsman, S. Y. (2021). A fast-acting transcriptional brake in the drosophila embryo. *submitted to Current Biology*.

Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006).

Engineering and characterization of a superfolder green fluorescent protein. *Nature biotechnology*, **24**(1), 79–88.

Perez-Romero, C. A., Tran, H., Coppey, M., Walczak, A. M., Fradin, C., and Dostatni, N. (2018). Live imaging of mrna transcription in drosophila embryos. In *Morphogen Gradients*, pages 165–182. Springer.

Petkova, M. D., Tkačik, G., Bialek, W., Wieschaus, E. F., and Gregor, T. (2019). Optimal decoding of cellular identities in a genetic network. *Cell*, **176**(4), 844–855.

Petrášek, Z. and Schwille, P. (2008). Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. *Biophysical journal*, **94**(4), 1437–1448.

Porcher, A. and Dostatni, N. (2010). The bicoid morphogen system. *Current biology*, **20**(5), R249–R254.

Prüßing, K., Voigt, A., and Schulz, J. B. (2013). Drosophila melanogaster as a model organism for alzheimer's disease. *Molecular neurodegeneration*, **8**(1), 1–12.

Reiter, L. T., Potocki, L., Chien, S., Gribskov, M., and Bier, E. (2001). A systematic analysis of human disease-associated gene sequences in drosophila melanogaster. *Genome research*, **11**(6), 1114–1125.

Reynolds, N., O'Shaughnessy, A., and Hendrich, B. (2013). Transcriptional repressors: multifaceted regulators of gene expression. *Development*, **140**(3), 505–512.

Rhee, D. Y., Cho, D.-Y., Zhai, B., Slattery, M., Ma, L., Mintseris, J., Wong, C. Y., White, K. P., Celniker, S. E., Przytycka, T. M., *et al.* (2014). Transcription factor networks in drosophila melanogaster. *Cell reports*, **8**(6), 2031–2043.

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

Rose, M., Kurylowicz, M., Mahmood, M., Winkel, S., Moran-Mirabal, J. M., and Fradin, C. (2021). Direct measurement of the affinity between tbid and bax in a mitochondria-like membrane. *International journal of molecular sciences*, **22**(15), 8240.

Schier, A. F. (2009). Nodal morphogens. *Cold Spring Harbor perspectives in biology*, **1**(5), a003459.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., *et al.* (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, **9**(7), 676–682.

Schwille, P. and Ries, J. (2011). Principles and applications of fluorescence correlation spectroscopy (fcs). In *Biophotonics: Spectroscopy, Imaging, Sensing, and Manipulation*, pages 63–85. Springer.

Scott, D. J., Gunn, N. J., Yong, K. J., Wimmer, V. C., Veldhuis, N. A., Challis, L. M., Haidar, M., Petrou, S., Bathgate, R. A., and Griffin, M. D. (2018). A novel ultra-stable, monomeric green fluorescent protein for direct volumetric imaging of whole organs using clarity. *Scientific reports*, **8**(1), 1–15.

Shahbazi, M. N. (2020). Mechanisms of human embryo development: from cell fate to tissue shape and back. *Development*, **147**(14), dev190629.

Shvartsman, S. Y., Coppey, M., and Berezhkovskii, A. M. (2008). Dynamics of maternal morphogen gradients in drosophila. *Current opinion in genetics & development*, **18**(4), 342–347.

Spirov, A., Fahmy, K., Schneider, M., Frei, E., Noll, M., and Baumgartner, S. (2009). Formation of the bicoid morphogen gradient: an mrna gradient dictates the protein gradient.

Sprague, B. L., Pego, R. L., Stavreva, D. A., and McNally, J. G. (2004). Analysis of binding reactions by fluorescence recovery after photobleaching. *Biophysical journal*, **86**(6), 3473–3495.

Stapornwongkul, K. S. and Vincent, J.-P. (2021). Generation of extracellular morphogen gradients: the case for diffusion. *Nature Reviews Genetics*, **22**(6), 393–411.

Tanaka, M., Yoshimoto, T., and Nakamura, T. (2017). A double-edged sword: The world according to capicua in cancer. *Cancer science*, **108**(12), 2319–2325.

Tram, U., Riggs, B., and Sullivan, W. (2001). Cleavage and gastrulation in drosophila embryos. *e LS*.

Tran, H., Desponds, J., Perez Romero, C. A., Coppey, M., Fradin, C., Dostatni, N., and Walczak, A. M. (2018). Precision in a rush: Trade-offs between reproducibility and steepness of the hunchback expression pattern. *PLoS computational biology*, **14**(10), e1006513.

Tran, H., Walczak, A. M., and Dostatni, N. (2020). Constraints and limitations on the transcriptional response downstream of the bicoid morphogen gradient. *Current topics in developmental biology*, **137**, 119–142.

Turing, A. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London*, **237**, 37–72.

Valbuena, F. M., Fitzgerald, I., Strack, R. L., Andruska, N., Smith, L., and Glick, B. S. (2020). A photostable monomeric superfolder green fluorescent protein. *Traffic*, **21**(8), 534–544.

von Hippel, P. H. (1998). An integrated model of the transcription complex in elongation, termination, and editing. *Science*, **281**(5377), 660–665.

Wachsmuth, M. (2014). Molecular diffusion and binding analyzed with frap. *Protoplasma*, **251**(2), 373–382.

Wachsmuth, M., Waldeck, W., and Langowski, J. (2000). Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *Journal of molecular biology*, **298**(4), 677–689.

Wartlick, O., Mumcu, P., Kicheva, A., Bittig, T., Seum, C., Jülicher, F., and Gonzalez-Gaitan, M. (2011). Dynamics of dpp signaling and proliferation control. *science*, **331**(6021), 1154–1159.

Widengren, J., Mets, U., and Rigler, R. (1995). Fluorescence correlation spectroscopy of triplet states in solution: a theoretical and experimental study. *J. Phys. Chem.*, **99**(36), 13368–13379.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. *Journal of theoretical biology*, **25**(1), 1–47.

Yeung, C., Shtrahman, M., and Wu, X.-l. (2007). Stick-and-diffuse and caged diffusion: a comparison of two models of synaptic vesicle dynamics. *Biophysical journal*, **92**(7), 2271–2280.

Zhang, L., Perez-Romero, C., Dostatni, N., and Fradin, C. (2021). Using fcs to accurately measure protein concentration in the presence of noise and photobleaching. *Biophysical Journal*, **120**(19), 4230–4241.