



OPEN

## Fusion and fission events regulate endosome maturation and viral escape

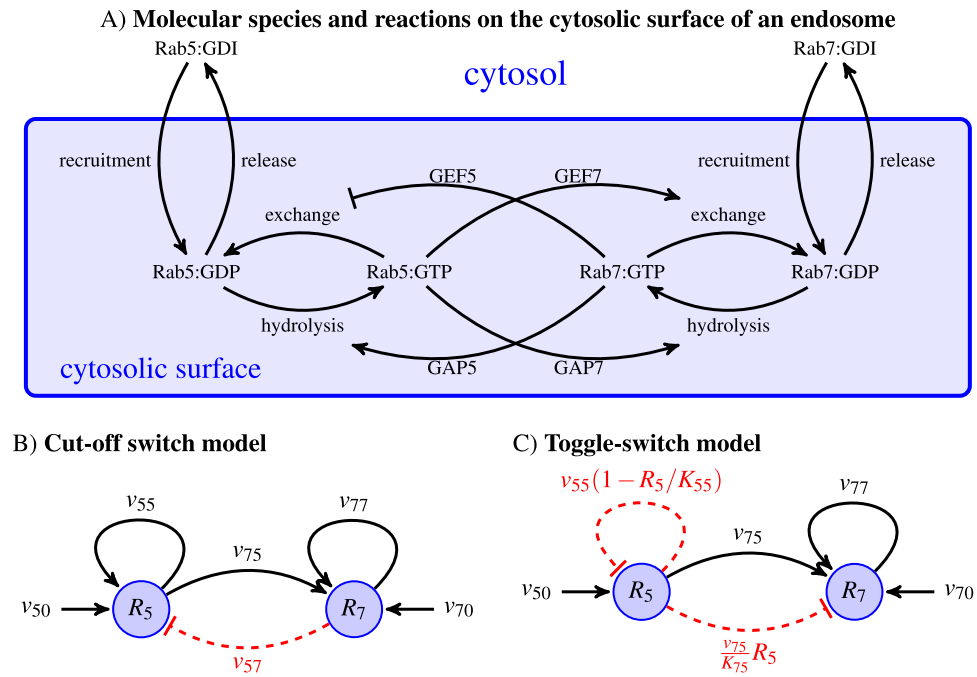
Mario Castro<sup>1,5</sup>✉, Grant Lythe<sup>2</sup>, Jolanda Smit<sup>3</sup> & Carmen Molina-Paris<sup>2,4,5</sup>✉

Endosomes are intracellular vesicles that mediate the communication of the cell with its extracellular environment. They are an essential part of the cell's machinery regulating intracellular trafficking via the endocytic pathway. Many viruses, which in order to replicate require a host cell, attach themselves to the cellular membrane; an event which usually initiates uptake of a viral particle through the endocytic pathway. In this way viruses hijack endosomes for their journey towards intracellular sites of replication and avoid degradation without host detection by escaping the endosomal compartment. Recent experimental techniques have defined the role of endosomal maturation in the ability of enveloped viruses to release their genetic material into the cytoplasm. Endosome maturation depends on a family of small hydrolase enzymes (or GTPases) called Rab proteins, arranged on the cytoplasmic surface of its membrane. Here, we model endosomes as intracellular compartments described by two variables (its levels of active Rab5 and Rab7 proteins) and which can undergo coagulation (or fusion) and fragmentation (or fission). The key element in our approach is the "per-cell endosomal distribution" and its dynamical (Boltzmann) equation. The Boltzmann equation allows us to derive the dynamics of the total number of endosomes in a cell, as well as the mean and the standard deviation of its active Rab5 and Rab7 levels. We compare our mathematical results with experiments of Dengue viral escape from endosomes. The relationship between endosomal active Rab levels and pH suggests a mechanism that can account for the observed variability in viral escape times, which in turn regulate the viability of a viral intracellular infection.

Endosomes are enigmatic organelles that regulate intracellular cargo trafficking<sup>1,2</sup>. These (literally) *inner bodies* are dynamic in movement and have decorated cytoplasmic membranes<sup>3</sup>. They can merge; that is, they can undergo fusion. To avoid confusion, in this manuscript we shall refer to viral RNA release from the endosome to the cytosol as *escape* and to the merging of two endosomes as *fusion*. We note that other authors refer to fusion as the merging of the viral envelope and the endosomal membrane<sup>4</sup>. Endosomes can also break up (fission)<sup>5</sup> or kiss and run<sup>6</sup>, in a random choreography that allows the cell to sort, recycle or degrade its internalised cargo. This endosomal maturation programme requires a dramatic transformation of these organelles: from early endosomes, to late ones, recycling ones, and degrading lysosomes<sup>1,7</sup>.

In recent years, a lot of effort has been devoted to decipher how viral entry and subsequent intracellular viral genome release are related to endosomal maturation and intracellular trafficking<sup>8</sup>. Current consensus suggests that many viruses, independently of their entry pathway, are delivered to the so-called early endosomes, and thus, enter an intricate network of endosomes: from early to late endosomes, until the endosomal pH is low enough to trigger viral membrane fusion and escape to the cytosol<sup>9,10</sup>. In this way, endosomes become one of the main vehicles for viral intracellular trafficking in the infected cell, transporting viral genomic cargo from the plasma membrane to the cytosol<sup>4</sup>. For many viruses, escape is preceded by a sudden drop in the endosomal pH<sup>8,9,11</sup>. This often occurs in late endosomes, where in fact, high levels of the GTPase Rab7 have been shown to be responsible for viral escape<sup>12,13</sup>. This rather interesting feature seems to indicate that by tracking the active Rab decoration of the endosomal *cytoplasmic surface*, one is in fact, also following the maturation history of individual endosomes. For example, Rab5 is a marker of early endosomes and Rab7 of late ones<sup>14</sup>. Thus, as well

<sup>1</sup>Grupo Interdisciplinar de Sistemas Complejos (GISC) and Instituto de Investigación Tecnológica (IIT), Universidad Pontificia Comillas, Madrid, Spain. <sup>2</sup>Department of Applied Mathematics, School of Mathematics, University of Leeds, Leeds, UK. <sup>3</sup>Department of Medical Microbiology and Infection Prevention, University Medical Center Groningen, Groningen, The Netherlands. <sup>4</sup>Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. <sup>5</sup>These authors contributed equally: Mario Castro and Carmen Molina-Paris. ✉email: marioc@comillas.edu; molina-paris@lanl.gov



**Figure 1.** Schematic summary of the biochemical reactions which take place on the cytosolic surface of an endosome and competing hypotheses for Rab5 and Rab7 activation. **(A)** Biochemical reactions that regulate the levels of Rab5 and Rab7 in the cytosol and on the surface of endosomes in inactive (Rab-GDP) and active (Rab-GTP) forms. GDI means GDP-dissociation inhibitor (see Ref.<sup>21,23</sup> for details). Eqs. (12) and (13) (or Eqs. (14) and (15)) encode the key molecular interactions considered in each mathematical model. **(B)** Cut-off switch model of Rab5 to Rab7 switching. For instance, the red dashed line is modelled by the term proportional to  $v_{57} x_7$ . **(C)** Toggle-switch model of Rab5 to Rab7 switching. The increase of Rab7 with time is caused by a reduction of the activation signal due to an inhibitory one. The level of active Rab5 ( $R_5$ ) decreases with time due to a negative feedback of the active Rab7 ( $R_7$ ). Both models have been adapted from Ref.<sup>21</sup>. The rates  $v_{ij}$  will be encoded in the molecular current,  $\bar{J}$ , defined in Eqs. (12) and (13) for the cut-off switch model, and Eqs. (14) and (15) for the toggle-switch model.

as serving as markers of endosome maturation, different Rab proteins have been identified as central regulators of the endosomal transport machinery<sup>15</sup>. Of particular relevance to viral escape, is the fact that the pH of a given endosome has been experimentally linked to its levels of active Rab5 and Rab7 proteins<sup>7,15–17</sup>.

Given the large variability observed in viral escape times<sup>11,13,18</sup>, it is timely to ask ourselves the following questions: (i) is pH alone the main regulator of viral escape events, (ii) is the pH dependence gradual (analogical) or abrupt (digital), (iii) can endosome maturation and its dynamics explain the observed heterogeneity in the distribution of viral escape times, and (iv) what mechanism of endosome maturation is compatible with current empirical observations. Question (i) has already been quantitatively addressed, both experimentally and theoretically<sup>12,19,20</sup>. These studies concluded that a low endosomal pH triggers viral escape. At the biochemical level, a pH drop induces a protein conformational change in the viral envelope protein, which, consequently, drives the fusion between the viral envelope protein and the internal surface of the endosomal membrane. This event initiates the release of viral genetic material into the cytosol, and in turn leads to the virus escaping potential degradation in the lysosome.

In this paper, we aim to provide answers to questions (ii)–(iv) above. We do so by making use of a novel mathematical approach to describe the population dynamics of endosomes containing endocytosed viral particles. The mathematical framework developed and presented here includes the contribution to endosome maturation and dynamics from fusion and fission events, as well as those molecular (biochemical) processes related to active Rab recruitment and endosome acidification. In this way, the method allows one to characterise and quantify the mechanistic details of endosomal maturation with reduced mathematical complexity to enable comparison with experimental data of viral escape.

**The biology of endosomal Rab5 and Rab7 dynamics.** Early endosomes (Rab5-positive) undergo a progressive replacement of Rab5 with Rab7<sup>5</sup>. Rabs are not endocytic cargo but endosome surface proteins which can be found in two conformations: Rab5:GDP and Rab5:GTP, inactive and active, respectively. This replacement process involves several molecules and chemical reactions (see a schematic summary in Fig. 1A, adapted from Ref.<sup>21</sup>). Specifically (see Ref.<sup>21</sup> for details), Rab5-positive endosomes (a signature of early endosomes) gradually become Rab7-positive endosomes (namely, late endosomes)<sup>5</sup>. The complete *endosomal maturation* process then replaces the Rab5 decoration with a Rab7 one. Rab5 and Rab7, guanine nucleotide exchange factors (GEFs),

GTPase-activating proteins (GAPs), and effector molecules interact in a series of biochemical reactions (see Fig. 1A) which take place on the endosomal cytosolic surface<sup>22</sup>. The conversion from Rab5 to Rab7 is, in fact, a consequence of programmed and simultaneous changes in the nucleotide cycle of both Rab5 and Rab7, which shuttle between inactive, GDP-bound, and active, GTP-bound, conformations. GEFs catalyse the exchange of GDP into GTP and GAPs catalyse the hydrolysis of GTP into GDP<sup>23,24</sup>. This series of biochemical reactions can be explained by two different and competing hypotheses (see Fig. 1B,C), as comprehensively described in Ref.<sup>21</sup>. The first one is the so-called *toggle-switch model*. It considers a weak repression of Rab7 (by Rab5), which leads to reduced Rab7 activation (see Fig. 1B). The second one, so-called *cut-off switch model*, assumes that Rab7 activation strongly suppresses that of Rab5 (see Fig. 1C). The authors of Ref.<sup>21</sup> aimed to identify which hypothesis could be best supported by experimental data, and thus, introduced a modular mathematical model (see Appendix C in Ref.<sup>21</sup>). The drawback of the exhaustive model comparison described in Ref.<sup>21</sup> is the need to make use of specific fitting algorithms to parameterise the models with data from Ref.<sup>25</sup>, as well as its inherent model complexity. We avoid this by clearly encoding the biochemical reactions described in Fig. 1 in the specific choices for the *molecular currents* which will describe active Rab5 and Rab7 endosomal decoration. Once the Rab molecular reactions that take place on the endosomal cytosolic surface are at hand, our framework also considers the interactions of endosomes with other nearby endosomes. In particular, endosomes can undergo fusion and fission<sup>26</sup>, or kiss-and-run<sup>6</sup> events. The latter process occurs when Rab5 molecules restrict the complete fusion of an endosome, allowing the exchange of solutes between them, but without complete intermixing of their membranes<sup>6</sup>. Here, we shall only consider complete endosomal fusion and fission events.

**Mathematical framework of endosome dynamics and maturation.** Mathematical models have helped identify some of the mechanisms involved in the endosomal maturation process. For instance, some regulatory proteins, e.g., SNAREs, Rabs, and other coating proteins, have been linked to the formation, degradation and renewal of endosomes. There is consensus on what mechanisms are involved in these processes, although their individual quantitative contribution has remained elusive<sup>17,27</sup>. In this section, we summarise the biology of endosomal Rab dynamics, as well as previous mathematical modelling efforts.

Most mathematical models of endosome dynamics (with the notable exception of the model in Ref.<sup>28</sup> which pioneered the approach developed here) make use of a set of ordinary differential equations (ODEs) to describe the dynamics of endosomal Rab protein decoration. This poses a practical problem since, in all cases, the data used to parameterise the mathematical models is averaged over all the endosomes in a cell. However, as we mentioned above, fusion and fission events couple two nearby endosomes, and hence, endosomal averages might not be appropriate to describe the pH of an individual endosome. To illustrate why traditional ODE models do not capture this variability properly, let us consider an extreme case of two endosomes with pH 4 and 7, respectively, and let us assume that viral escape events occur when the endosomal pH is below 5. The average pH will be 5.5 in this case, so viral escape will not occur, which contradicts the fact that it will in one of the endosomes, but not the other (if no average is considered). Thus, endosomal averaging can smooth out fast changes or significant heterogeneities, and in turn lead to the conclusion that viral escape is a continuous process rather than mediated by a pH threshold<sup>27</sup>. At the other end of the modelling spectrum, the biochemistry of proton exchange has been used to describe the thermodynamics of endosome pH acidification<sup>29</sup>. Also, the authors of Ref.<sup>16</sup> model the acidification of individual endosomes, without considering interactions between them. Finally, it is worth mentioning that, in Ref.<sup>30</sup> the authors made use of a hybrid method, which includes differential equations for active Rab5, inactive Rab5 and the total number of endosomes. Specifically, the model considers the mean number of Rab5 molecules per endosome and assumes that Rab5 exchange and activation between cytosol and endosomes is proportional to the number of endosomes (see Eqs. (S50)–(S53) in Section 4 of the Supplementary Information, which were proposed by the authors of Ref.<sup>30</sup>). Note that Eq. (S50) incorporates the role of fusion and fission events in terms of the mean number of endosomes. Thus, this *mean-field* approach is unable to describe individual endosome dynamics. While this can be a valuable approximation to characterise endosome population dynamics, it fails to capture endosome heterogeneity, which is essential if we want to understand and describe the observed experimental variability. A suitable method to address this challenge would be the use of agent-based models (ABM)<sup>31</sup>, in which each endosome can be simulated, as an agent, and is described in terms of its molecular content (or cargo). However, ABMs are often computationally expensive or not amenable to parameter inference.

Our approach here takes a different route beyond deterministic descriptions of endosomal dynamics at the cellular level<sup>21,25,30,32</sup>, or at the single endosome level<sup>16,17,27</sup>. As we have discussed earlier, the former presumes that Rab dynamics at the cellular scale can be studied from the observation of endosomal averages, and, the latter, focuses on the biochemical reactions that occur on the cytosolic surface of endosomes, and fails to include the interaction between endosomes. To overcome these limitations, we introduce a model inspired by the dynamics of droplet coalescence, also known in the physics literature as the Smoluchowski equation<sup>28,33,34</sup>. Our approach allows one to recover previous phenomenological models<sup>21,30</sup> under certain mathematical assumptions (see Section 3 of the Supplementary Information). These models are based on an integro-differential equation which can describe cells<sup>35–37</sup> or endosomes<sup>28,38</sup>. A difficulty with phenomenological models is their lack of mathematical tractability. To avoid this significant limitation, here, we make use of a formalism that enables one to describe the dynamics of the number of endosomes and their cargo of (active) Rab5 and Rab7 molecules. Furthermore, using the two-dimensional Laplace transform, the formalism naturally leads to analytical expressions for the rate equations (ODEs) of interest. Once the rate equations have been derived, and together with experimental data<sup>10</sup>, we test the validity of these equations and discriminate between two different mechanistic hypotheses<sup>39</sup>. Equipped with this mathematical framework, we can simplify some of the underlying assumptions used in other models (see Section 3 of the Supplementary Information) and perform a model selection analysis (making use

of experimental data). Our interest is to directly emphasise the underlying molecular mechanisms rather than certain mathematical choices; this can be done since our approach avoids *ad hoc* reaction terms based on Hill or logistic functions (see the “Results” section).

Finally, and going back to our initial discussion on endosome acidification, we emphasise that endosome acidification is a complex problem and as yet, not fully understood. Traditionally, theories of organelle acidification assumed thermodynamic equilibrium, but cellular compartments are not in equilibrium during endosome maturation. Modern alternatives propose sophisticated dynamical models<sup>29</sup> to characterise the dynamics of acidification, but they require solving a system of differential equations modelling proton exchange and its interaction with endosomal cations, and involving 22 unknown parameters, which cannot be identified with our experimental data sets. Alternative ways to estimate intracellular pH can be performed with the use of DNA nanomachines<sup>40</sup> or by luminescence markers of acidity acting as cargo transported by the endosome<sup>41</sup>. Since the experimental error can be as large as  $\Delta\text{pH} \sim 1$ , these experiments might not be accurate enough in the case of abrupt pH changes. To overcome these experimental limitations, several approaches have been proposed to understand the interplay between Rab decorations and endosomal pH<sup>16,27</sup>. These approaches, in a nutshell, provide a mechanistic model of proton dynamics and a heuristic interpolation model between early and late endosomes, through an equation of the following kind:

$$\text{pH}(t) = \text{pH}^{\text{early}} + (\text{pH}^{\text{late}} - \text{pH}^{\text{early}}) \frac{\text{Rab7}(t)}{\text{Rab5}(t) + \text{Rab7}(t)}, \quad (1)$$

where  $\text{Rab5}(t)$  and  $\text{Rab7}(t)$  have to be understood at the single-endosome level, not at the single-cell level. We shall make use of this equation to derive the time course of endosomal pH,  $\text{pH}(t)$ , by providing the dynamics of  $\text{Rab5}(t)$  and  $\text{Rab7}(t)$  as derived from our general mathematical framework.

**Boltzmann equation for the endosomal distribution.** In this section, we provide a mathematical description of the biological mechanisms discussed in the previous one. We consider a single cell containing a collection of endosomes. Each endosome is characterised by its active Rab cargo, technically, [Rab5:GTP] and [Rab7:GTP], respectively,  $(x_5, x_7)$  at time  $t$ . The intensity of the fluorescence markers for Rab5 and Rab7 is normalized to the maximum intensity provided by the experimental device. Our variables,  $x_5$  and  $x_7$  are the normalized intensities and, hence, they are proportional to the number of Rab molecules on the membrane. We introduce

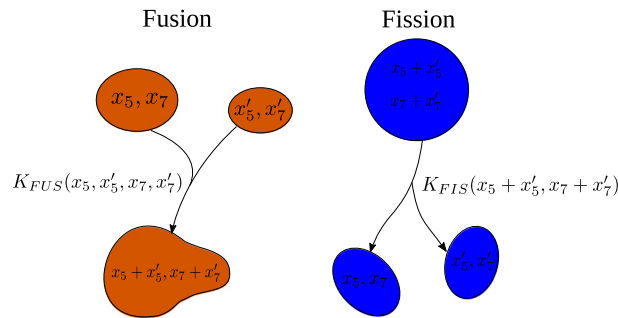
$$n(x_5, x_7; t) \equiv \text{“per-cell endosomal distribution”}. \quad (2)$$

Using  $x_5, x_7$  as defined above instead of the Rab concentration simplifies the mathematical modelling, since we avoid the need to describe endosomal volumes to compute the concentration. The price paid by these assumptions is that we do not include or consider any such details such as the volume, size or shape of endosomes. Equation (2) allows us to define the observable moments of the distribution defined in the Supplementary Information. For example, the first moments of the distribution have been defined in Eqs. (S1)–(S3).

We can now write the evolution equation for  $n(x_5, x_7; t)$ , that we refer to as the Boltzmann equation for the endosomal distribution. In the following section, we explain the meaning of each term in the equation and its biological rationale.

$$\begin{aligned} \frac{\partial n(x_5, x_7; t)}{\partial t} = & \underbrace{\frac{1}{2} \int_0^{x_5} dx'_5 \int_0^{x_7} dx'_7 K_{FUS}(x'_5, x_5 - x'_5, x'_7, x_7 - x'_7) n(x'_5, x'_7; t) n(x_5 - x'_5, x_7 - x'_7; t)}_{\text{fusion (gain term)}} \\ & - \underbrace{n(x_5, x_7; t) \int_0^{+\infty} dx'_5 \int_0^{+\infty} dx'_7 K_{FUS}(x_5, x'_5, x_7, x'_7) n(x'_5, x'_7; t)}_{\text{fusion (loss term)}} \\ & + \underbrace{\int_0^{+\infty} dx'_5 \int_0^{+\infty} dx'_7 K_{FIS}(x_5, x'_5, x_7, x'_7) n(x_5 + x'_5, x_7 + x'_7; t)}_{\text{fission (gain term)}} \\ & - \underbrace{\frac{1}{2} n(x_5, x_7; t) \int_0^{x_5} dx'_5 \int_0^{x_7} dx'_7 K_{FIS}(x'_5, x_5 - x'_5, x'_7, x_7 - x'_7)}_{\text{fission (loss term)}} \\ & - \underbrace{\mu_0 n(x_5, x_7; t)}_{\text{degradation}} - \underbrace{\nabla \cdot \vec{J}(x_5, x_7)}_{\text{Rab5 and Rab7 interactions}} + \underbrace{S_0 \delta(x_5) \delta(x_7)}_{\text{endocytosis of viral particle}}. \end{aligned} \quad (3)$$

Although this precise equation has not been proposed before, simplified versions with fewer biological mechanisms<sup>28</sup> or in fewer dimensions<sup>37,42</sup> have been studied in different contexts.



**Figure 2.** Fusion and fission events depend on the number of Rab5 and Rab7 molecules, not their concentration. Left: two endosomes with different levels of active Rab5 and Rab7,  $(x_5, x_7)$  and  $(x'_5, x'_7)$ , respectively, merge into a new endosome with levels of active Rab5 and Rab7,  $(x_5 + x'_5, x_7 + x'_7)$ , with rate  $K_{FUS}(x_5, x'_5, x_7, x'_7)$ . Right: an endosome with levels of active Rab5 and Rab7  $(x_5 + x'_5, x_7 + x'_7)$  splits into two new endosomes with levels of active Rab5 and Rab7  $(x_5, x_7)$  and  $(x'_5, x'_7)$ , respectively, with rate  $K_{FIS}(x_5, x'_5, x_7, x'_7)$ .

### Mathematical modelling hypotheses.

1. *Endocytosis (endosome generation)*: new endosomes are formed by the invagination of the cell membrane. Newly created endosomes (containing endocytosed viral particles) do not have any Rab5 or Rab7 molecules. Progressively, the endosomes become *decorated* with Rab5 and, subsequently, with Rab7 proteins. Mathematically, endosome generation is described by a source term of the form

$$S(x_5, x_7) = S_0 \delta(x_5) \delta(x_7), \quad (4)$$

where  $S_0$  is a real positive constant with dimensions of number of endosomes per unit time and the symbol  $\delta(\cdot)$  stands for the Dirac delta function.

2. *Endosome degradation/removal*: endosomes can fuse with a lysosome and thus, be removed from the cytosol. In our case, there exists a second form of endosome removal due to viral escape. We assume that the endosomes can be removed at any stage of the maturation process. Hence, we assume a constant removal (or *death*) rate proportional to the number of endosomes, that is independent of its Rab cargo. We write

$$\mu(x_5, x_7) = \mu_0 n(x_5, x_7; t). \quad (5)$$

3. *Fusion (coalescence)*: fusion of two endosomes involves the merging of their membranes and, as a consequence, the Rab5 and Rab7 molecules are shared<sup>5</sup> (see Fig. 2 for a schematic description of this process). In particular, when two endosomes with Rab cargo given by  $(x_5, x_7)$  and  $(x'_5, x'_7)$ , respectively, fuse, they share their cargo, so the total number of Rab cargo of the newly formed endosome will be  $(x_5 + x'_5, x_7 + x'_7)$ . The contribution to the time derivative of  $n(x_5, x_7; t)$  is

$$\begin{aligned} & \frac{1}{2} \int_0^{x_5} dx'_5 \int_0^{x_7} dx'_7 K_{FUS}(x'_5, x_5 - x'_5, x'_7, x_7 - x'_7) n(x'_5, x'_7; t) n(x_5 - x'_5, x_7 - x'_7; t) \\ & - n(x_5, x_7; t) \int_0^{+\infty} dx'_5 \int_0^{+\infty} dx'_7 K_{FUS}(x_5, x'_5, x_7, x'_7) n(x'_5, x'_7; t), \end{aligned} \quad (6)$$

where the first term represents the net *gain* of endosomes with Rab levels  $(x_5, x_7)$  and the second one represents the *loss* of endosomes with total levels  $(x_5, x_7)$  after fusion with other endosomes. The function  $K_{FUS}(x_5, x'_5, x_7, x'_7)$  is referred to as the fusion kernel. It dictates the rate of endosomal fusion, and it clearly depends on the endosomal levels of active Rab molecules.

A central feature of our model is the consideration of both fusion and fission events. Fusion is enhanced in early endosomes so the rate of fusion correlates positively with the levels of Rab5 and negatively with those of Rab7<sup>5</sup>. For simplicity, we assume that the fusion rate is a linearly increasing function of Rab5 and a linearly decreasing function of Rab7. Hence, we propose

$$K_{FUS}(x_5, x'_5, x_7, x'_7) = K_{FUS}^{(0)} + K_{FUS}^{(5)}(x_5 + x'_5) - K_{FUS}^{(7)}(x_7 + x'_7), \quad (7)$$

where  $K_{FUS}^{(k)}$  are constants for  $k = 0, 5, 7$ . As shown in Ref.<sup>43</sup>, endosomes need to be spatially close to merge. As we are not modelling the intracellular endosome spatial location explicitly, the latter equation favours fusion of early endosomes (higher  $x_5$  increases the overall fusion rate) and reduces fusion in late endosomes (it decreases for higher  $x_7$ ). Yet, our model does not preclude fusion events between early and late endosomes<sup>10</sup>; otherwise, our fusion kernel should be multiplied by a Dirac delta function to force fusion only between endosomes with the same levels of  $x_5$  and  $x_7$ .

4. *Fission*: endosomes can divide, thus splitting the amount of Rab5 and Rab7 between the two newly created endosomes. When fission occurs, the total amount of Rab5 and Rab7 is shared randomly between the two



newly created endosomes (see Fig. 2 for a schematic description of this process). Similarly to fusion, fission can be described introducing a kernel function, as follows

$$\int_0^{+\infty} dx_{5'} \int_0^{+\infty} dx_{7'} K_{FIS}(x_5, x'_5, x_7, x'_7) n(x_5 + x'_5, x_7 + x'_7; t) - \frac{1}{2} n(x_5, x_7; t) \int_0^{x_5} dx_{5'} \int_0^{x_7} dx_{7'} K_{FIS}(x'_5, x_5 - x'_5, x'_7, x_7 - x'_7). \quad (8)$$

The first term is the *gain* due to the fission of a larger endosome leading to two endosomes, one of them with Rab levels  $(x_5, x_7)$ . The second one is a *loss* term due to the fission of an endosome with Rab levels  $(x_5, x_7)$ .

Endosomal fission is less well understood than fusion. In Ref.<sup>5</sup> it is suggested that fission occurs randomly at any stage of maturation. Thus, we consider that fission is independent of the number of Rab5 or Rab7 molecules, but that it is not necessarily symmetric (namely, when an endosome splits, the amount of Rab going to each daughter endosome can be different). Mathematically, we propose

$$K_{FIS}(x_5, x'_5, x_7, x'_7) = K_{FIS}^{(0)} f(|x_5 - x'_5|, |x_7 - x'_7|), \quad (9)$$

where, by the symmetric properties of the fission *kernel*<sup>33</sup>, the function  $f$  satisfies the normalisation condition,  $f(0, 0) = 1$  and is symmetric in its arguments. The simplest case one can consider is symmetric fission; namely, we write

$$f(x_5 - x'_5, x_7 - x'_7) = \delta(x_5 - x'_5) \delta(x_7 - x'_7). \quad (10)$$

This choice for  $f$  means that 50% of each cargo is equally shared between daughter endosomes. Relaxing this assumption only affects the higher moments of the distributions (variance, skewness, etc.) but does not change the mean values of Rab5, Rab7 and the number of endosomes. We discuss this issue in greater detail in the Supplementary Information.

The contribution to the time derivative of  $n(x_5, x_7; t)$  is then

$$4K_{FIS}^{(0)} n(2x_5, 2x_7; t) - K_{FIS}^{(0)} n(x_5, x_7; t).$$

5. **Rab5/Rab7 activation/deactivation:** Rab molecules on the endosomal membrane can be activated after prior incorporation of inactive Rab:GDI from the cytosol. A schematic summary of the reactions of Rab activation/deactivation can be found in Fig. 1.

In Ref.<sup>21</sup>, the authors considered two competing hypotheses for Rab5/Rab7 activation/deactivation (see Fig. 1B,C). The first one is the *toggle-switch* model and consists in weak repression of Rab7 by Rab5, described by a logistic term. In the second one, the *cut-off switch* model, Rab7 activation strongly suppresses Rab5. To identify which hypothesis was more compatible with the experimental data, the authors introduced a modular model where certain mechanisms could be explained by making use of different mathematical functions (see Supplementary information 1 of Ref.<sup>21</sup>). The drawback of this type of exhaustive model comparison is that specific fitting algorithms had to be adapted to infer model parameters from the data<sup>25</sup>.

We note that the total number of endosomes does not change with the activation/deactivation of Rab molecules. This fact can be naturally expressed in terms of a conserved quantity. We also note that we model the number of Rab molecules in a given endosome rather than its concentration since the latter one might be affected by fusion and fission events, where the volume or the area of the endosome can significantly change. In the absence of other mechanisms, this conservation law can be expressed in terms of a *molecular current*. The contribution of the dynamics of  $x_5$  and  $x_7$  for a given endosome to the time derivative of  $n(x_5, x_7; t)$  is equal to minus the divergence of a current  $\vec{J}$ :

$$\partial_t n(x_5, x_7; t) + \nabla \cdot \vec{J}(x_5, x_7) = 0,$$

where

$$\nabla \cdot \vec{J}(x_5, x_7) \equiv \partial_5 J_5(x_5, x_7) + \partial_7 J_7(x_5, x_7), \quad (11)$$

with  $\partial_k$  denoting  $\partial_{x_k}$ , for  $k = 5, 7$ . We have also made use of the notation  $\vec{J}(x_5, x_7) = (J_5(x_5, x_7), J_7(x_5, x_7))$ . One must supplement Eq. (11) with *constitutive equations* for the currents,  $J_5(x_5, x_7)$  and  $J_7(x_5, x_7)$ . In this paper we are going to assume that the currents are proportional to the total number of endosomes,  $n(x_5, x_7; t)$ . This implies that  $J_5(x_5, x_7) = v_5(x_5, x_7) n(x_5, x_7; t)$  and  $J_7(x_5, x_7) = v_7(x_5, x_7) n(x_5, x_7; t)$ , where we have introduced the velocities  $v_{5,7}(x_5, x_7)$ . The *velocities*  $v_{5,7}$  are generic functions of  $x_5$  and  $x_7$  that need to be prescribed according to the underlying biology of Rab5/Rab7 activation/deactivation discussed earlier. Following Refs.<sup>21,30</sup>, we assume that both molecules evolve and are coupled to each other (see Fig. 1). In other words, the concrete form of the functions  $v_{5,7}(x_5, x_7)$  will be determined by performing model selection and thus, identifying the mechanisms that are best supported by the experimental data. Rab5 and Rab7 interact via positive/negative feedback loops. We include these molecular interactions in the current,  $\vec{J}$ , and assume a linear dependence on the number of Rab molecules. For the *cut-off switch* model, one has<sup>21</sup>

$$v_5(x_5, x_7) \equiv v_{50} - v_{55} x_5 - v_{57} x_7, \quad (12)$$

$$v_7(x_5, x_7) \equiv v_{70} + v_{75} x_5 - v_{77} x_7, \quad (13)$$

where the choice of the signs in the coefficients  $v_{ij}$  is determined by the network of interactions in Fig. 1. For instance, the inhibition described by the red dashed arrow is captured by the term  $-v_{57} x_7$ . For the *toggle-switch* model one has  $v_{57} = 0^{21}$ . Namely, levels of Rab7 do not affect those of Rab5, and thus, the velocity  $v_5(x_5, x_7)$  does not depend on  $x_7$ . Yet, the model is non-linear, since it includes two logistic terms. These non-linear terms encode inhibitory mechanisms, as for example, the terms proportional to  $v_{55}$  and  $v_{75}$ . For the *toggle-switch* model one has<sup>21</sup>

$$v_5(x_5, x_7) \equiv v_{50} - v_{55} x_5 \left(1 - \frac{x_5}{K_{55}}\right), \quad (14)$$

$$v_7(x_5, x_7) \equiv v_{70} + v_{75} x_5 \left(1 - \frac{x_5}{K_{75}}\right) - v_{77} x_7. \quad (15)$$

The parameters  $K_{55}$  and  $K_{75}$  are *carrying capacities* that encapsulate the inhibitory behaviour of Rab5 in the toggle-switch model. As a consequence, we shall show (see Eq. (21)), that the ODE for the endosomal average of Rab5 does not contain the inhibitory feedback proportional to  $v_{57}$  present in Eq. (18).

Note that, as newly created endosomes have zero levels of Rab5 and Rab7 (see Eq. (4)), the following boundary conditions for the current  $\bar{J}$  must be fulfilled:

$$J_5(0, x_7) = 0 = J_7(x_5, 0). \quad (16)$$

Intuitively, these equations mean that the Rab5-associated current of endosomes with non-zero levels of Rab7 (first equation) and the Rab7-associated current of endosomes with non-zero levels of Rab5 (second equation) have to be zero.

**Equations for the moments of the Boltzmann distribution.** Equation (3) is a non-linear integro-differential equation that is, in principle, analytically intractable. However, it can be simplified under some assumptions that should be carefully scrutinised together with experimental data. In many experimental conditions, only the time evolution of the mean number of molecules of different species (including the total number of endosomes) can be measured. To this end, it is convenient to compute the first moments of the distribution, defined in Eqs. (S1)–(S3).

If we make use of the two-dimensional Laplace transform (for details, see Supplementary Information), we obtain an ODE for the first order moments defined above. For the *cut-off switch* model one can show

$$\frac{dN(t)}{dt} = S_0 + \left( \frac{K_{FIS}^{(0)}}{8} - \mu_0 - K_{FUS}^{(5)} R_5(t) + K_{FUS}^{(7)} R_7(t) \right) N(t) - \frac{1}{2} K_{FUS}^{(0)} N^2(t), \quad (17)$$

$$\frac{dR_5(t)}{dt} = v_{50} N(t) - (v_{55} + \mu_0) R_5(t) - v_{57} R_7(t), \quad \text{and} \quad (18)$$

$$\frac{dR_7(t)}{dt} = v_{70} N(t) + v_{75} R_5(t) - (v_{77} + \mu_0) R_7(t). \quad (19)$$

For the *toggle-switch* model the ODEs for the first moments can be shown to be

$$\frac{dN(t)}{dt} = S_0 + \left( \frac{K_{FIS}^{(0)}}{8} - \mu_0 - K_{FUS}^{(5)} R_5(t) + K_{FUS}^{(7)} R_7(t) \right) N(t) - \frac{1}{2} K_{FUS}^{(0)} N^2(t), \quad (20)$$

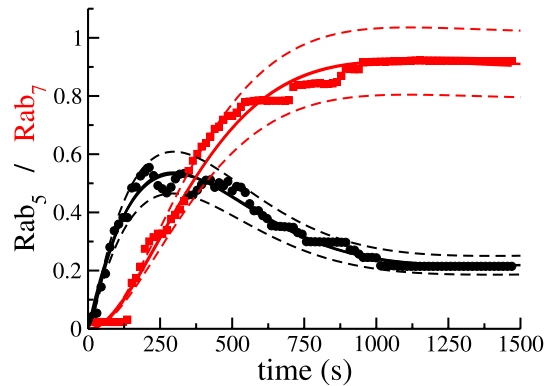
$$\frac{dR_5(t)}{dt} = v_{50} N(t) - \mu_0 R_5(t) - v_{55} R_5(t) \left(1 - \frac{R_5(t)}{K_{55}}\right), \quad \text{and} \quad (21)$$

$$\frac{dR_7(t)}{dt} = v_{70} N(t) + v_{75} R_5(t) \left(1 - \frac{R_5(t)}{K_{75}}\right) - (\mu_0 + v_{77}) R_7(t). \quad (22)$$

We note that as expected, Eqs. (17) and (20) are the same in both models. We also note that Eq. (3) allows one to derive the dynamical non-linear equation of Refs.<sup>21,30</sup> (for further details, see Supplementary Information). In all cases, the initial conditions are  $R_5(0) = R_7(0) = N(0) = 0$ .

## Experimental data

We make use of experimental data obtained by labelling Dengue viral particles (DENV) with the lipophilic fluorescent probe DiD, as previously reported in Ref.<sup>10</sup> and reproduced in Fig. 3. We use the normalised number of *probes* (DENV viral particles) colocalised with endosomal Rab5 and Rab7 (measured by fluorescence) as an estimate of the number of Rab molecules in endosomes with endocytosed DENV. Here, we are not interested in the DENV lifecycle, but the viral particles will serve as markers to track endosomal dynamics since in the experiments viral particles were colocalised with fluorescent markers for endosomal Rab5 and Rab7 molecules. Overall, 51 escape events were analysed to quantify the levels of Rab5 and Rab7 (see Fig. 3). Analysis of those 51



**Figure 3.** Cut-off switch model captures the mean and fluctuations of the experimental data. Comparison of model predictions (solid lines) as described by Eqs. (17)–(19) and the experimental data (black circles and red squares) from Ref.<sup>10</sup>. Solid lines correspond to the mean values of the normalised intensity of Rab5 and Rab7 fluorescence markers. Dashed lines correspond to the mean  $\pm$  one standard deviation, as derived from Eqs. (S26)–(S28) in the Supplementary Information making use of the experimentally calibrated model parameters. The fluctuations predicted by the model capture consistently those of the data.

cases revealed that, although most escape events took place in early endosomes (86%), a non-negligible number of events, 14%, took place in Rab5/Rab7-positive intermediate endosomes<sup>10</sup>. In addition, tracking fluorescently labelled endosomes allowed the authors to show that almost half of the endosomes skipped several steps of the maturation process by merging with existing Rab7-positive endosomes (precisely, 45%)<sup>10</sup>. Finally, 30% of the tracked endosomes underwent fission events at different stages of their trajectory in the cytoplasm. This supports our choice for a constant fission rate, termed *splitting* in Ref.<sup>10</sup>, which does not depend on the stage of maturation of the endosome, as defined by its Rab cargo ( $x_5, x_7$ ). Each data point corresponds to a different experiment (until escape occurs), and thus, our model describes the time distribution of the number of endosomes sampled at random times.

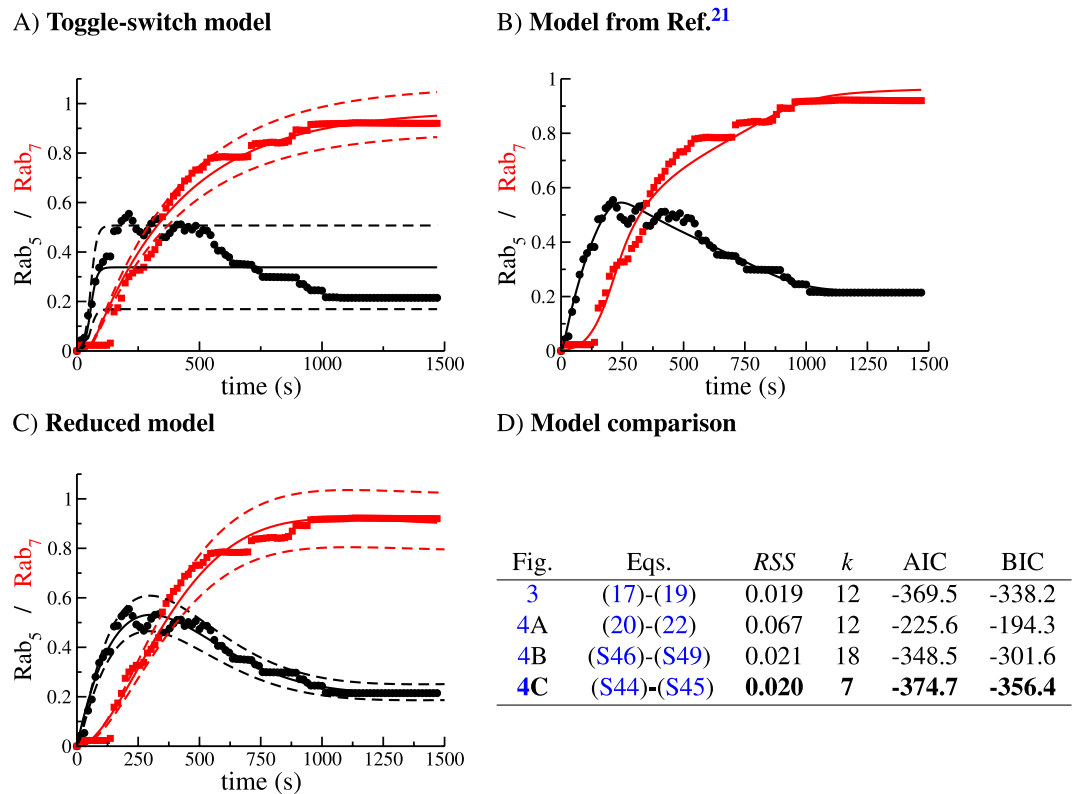
## Results

**Data and mathematical modelling best support the cut-off switch hypothesis.** Figure 3 shows the experimental data (black circles and red squares) from Ref.<sup>10</sup> and the result of fitting the data to the mathematical model (cut-off switch) described by Eqs. (17)–(19) (solid lines). A sensitivity analysis of the cut-off switch model (see Table S2 in Supplementary Information) reveals that the most sensitive parameters are, in order of relevance,  $K_{FUS}^{(0)}$ ,  $K_{FIS}^{(0)}$ ,  $v_{50}$  and  $v_{57}$ , and the least relevant,  $K_{FUS}^{(7)}$ ,  $K_{FUS}^{(5)}$ ,  $\mu_0$  and  $S_0$ . The results from our sensitivity analysis are rather interesting since it shows that fusion and fission are essential to understand the experimental data, but that the corrections to the constant term of the fusion kernel,  $K_{FUS}^{(0)}$ , which are proportional to  $K_{FUS}^{(7)}$  and  $K_{FUS}^{(5)}$  and depend on the Rab cargo, are negligible. Thus, it seems that a simpler model than the cut-off switch can be used to explain the data, as we show in Section 2 of the Supplementary Information (see also A quasi-linear approximation to describe experimental data section below and Fig. 4C). Furthermore, the source term of new endosomes with zero levels of Rab5 and Rab7 does not significantly affect the dynamics of endosomal Rab5 or Rab7. Numerical integration of Eq. (17) shows that, independently of the initial condition,  $N(0)$ , the number of endosomes quickly approaches a steady state (see Fig. S1). Taking into account the order of magnitude of the parameters in Table S1 of the Supplementary Information and the results of the sensitivity analysis, we find that

$$\begin{aligned} \frac{dN(t)}{dt} &= S_0 + \left( \frac{K_{FIS}^{(0)}}{8} - \mu_0 - K_{FUS}^{(5)} R_5(t) + K_{FUS}^{(7)} R_7(t) \right) N(t) - \frac{1}{2} K_{FUS}^{(0)} N^2(t) \simeq 0 \\ \Rightarrow N(t) &\simeq N_{ss} \equiv \frac{K_{FIS}^{(0)}}{4K_{FUS}^{(0)}}, \end{aligned} \quad (23)$$

where we have neglected  $K_{FUS}^{(7)}$ ,  $K_{FUS}^{(5)}$ ,  $\mu_0$  and  $S_0$  and assumed  $N_{ss} \neq 0$ . As a consequence of  $N(t)$  being almost stationary, the terms  $v_{50}N$  and  $v_{70}N$  in Eqs. (18) and (19), respectively, are also almost constant. Our theoretical analysis is consistent with our numerical results: Fig. S1 shows that the number of endosomes,  $N(t)$ , reaches a value close to its steady state after 40 s. Given the best-fit parameters from Table S1 in the Supplementary Information, we conclude that fusion and fission events are more relevant for the dynamics of Rab5 and Rab7 molecules than the rate of endosome generation since we have  $S_0 \ll K_{FUS}^{(0)}N^2$  and  $S_0 \ll K_{FIS}^{(0)}N$ . Finally, we now explore the role of the parameter  $v_{57}$  in the decrease of  $R_5(t)$  at the time of increase of  $R_7(t)$ . This coefficient encodes the inhibitory effect of Rab7 on the dynamics of Rab5. On the one hand, the sensitivity analysis of the cut-off switch model (see Table S2 in the Supplementary Information) gives this parameter a normalised value of 0.7543 and on the other hand, the best-fit parameters from Table S1 in the Supplementary Information, suggest a value of  $4 \times 10^{-3} \text{ s}^{-1}$  for  $v_{57}$ . These two results together show the qualitative and quantitative importance





**Figure 4.** Cut-off switch model (and its reduced version) explains the data more parsimoniously. **(A)** Toggle-switch model, Eqs. (20)–(22). **(B)** Model from Ref.<sup>21</sup> (see also Eqs. (S46)–(S49) in the Supplementary Information). **(C)** Comparison of the sensitivity-based reduced model, Eqs. (S44)–(S45) (see Supplementary Information for details). **(D)** Model comparison based on the Akaike Information Criterion, Eq. (24) and the Bayesian Information Criterion, Eq. (25). The selected model according to the minimum AIC and BIC is the one shown in boldface. Dashed lines represent the mean  $\pm$  the standard deviation, as described by Eqs. (S26)–(S28). We note that the formalism introduced here allows one to predict the variance of the estimated solution (A,C), unlike traditional ODE modelling approaches, where only the mean can be explained (B).

of this parameter in the cut-off switch model, which in turn, and in light of the experimental data, supports the cut-off switch hypothesis, in agreement with Ref.<sup>21</sup>. Note that  $v_{57} = 0$  in the toggle-switch hypothesis, since Rab7 does not affect the dynamics of Rab5. To further test this conclusion, we have also fitted the toggle-switch model, Eqs. (20), (21), (22), to the experimental data. The results are shown in Fig. 4A. It can be concluded that the toggle-switch model cannot explain the decrease of Rab5 observed in the data. We first note that in Eq. (21)  $R_5(t)$  does not depend on  $R_7(t)$ . We then argue that for this model, only fine-tuned mathematical functions of  $R_5$  might explain the decrease of  $R_5$  at late times. Yet, there is biological evidence to support that the Rab5 decrease and the Rab7 decrease are not independent events. Possibly, more sophisticated mathematical models, as those explored in Ref.<sup>21</sup>, might provide better agreement with the experimental data. Still, simply adding more mathematical terms (and so more parameters) to the dynamical equations would obscure our ability to systematically select between plausible biological mechanisms regulating the dynamics of Rab5 and Rab7.

**A quasi-linear approximation to describe experimental data.** Ziegler et al. in Ref.<sup>30</sup> considered Rab5-dependent fusion and fission of endosomes, making use of a differential equation similar to Eq. (17). Del Conte-Zerial et al. in Ref.<sup>21</sup> modelled the conversion of early endosomes into late endosomes, which assumes the dynamical replacement of Rab5 by Rab7 during endosome maturation. Their equations, summarised in the Supplementary Information, require a larger number of parameters than those in the cut-off switch model, Eqs. (17)–(19), but do not include a dynamical equation for the number of endosomes. In Fig. 4B we show the fit of their model to the DENV data. Comparison of Figs. 3 and 4 shows that a model that includes the dynamics of the number of endosomes can explain the molecular mechanisms parsimoniously. Our model does not explicitly consider inactive Rabs ([Rab5-GDP] and [Rab7-GDP]) because their levels remain almost constant (see Supplementary information of Ref.<sup>21</sup>). On the other hand, our framework provides dynamical equations for the variance of the number of molecules, as shown in Fig. 4A,C.

In order to perform quantitative model selection, we show in Fig. 4D the value of the Akaike Information Criterion (AIC) for the three models, defined as:

$$\text{AIC} \equiv n \log(\text{RSS}) + 2k, \quad (24)$$

where RSS is the residual sum of squares,  $n$  the number of points in the data series, and  $k$  the number of parameters in the mathematical model. Similarly, we compute the Bayesian Information Criterion (BIC), defined as

$$\text{BIC} \equiv n \log(\text{RSS}) + k \log n. \quad (25)$$

Both methods quantify the goodness of fit but introduce a *penalty* on the number of parameters (the lower the better). As the RSS is similar in all the models (and  $n = 100$  is the same for all of them), the most decisive factor is the number of parameters,  $k$ , thus, pointing at the reduced model, given by Eqs. (S44) and (S45) (see Fig. 4C). Thus, we conclude that a mathematical model that considers the mean number of endosomes in a quasi-steady state,  $N_{\text{ss}} \equiv \frac{K_{\text{FIS}}^{(0)}}{4K_{\text{FUS}}^{(0)}}$ , and assumes linear dynamics for the evolution of Rab5 and Rab7 (see Eqs. (S44) and (S45)) is the best candidate for future model extensions. Clearly, the toggle-switch model cannot capture the behaviour of the experimental data, as argued in the previous section. This emphasises an important conclusion that can be derived from our analysis: fusion, fission and Rab7 inhibition of Rab5 are the main mechanisms regulating endosome maturation and, in our context, the endosome acidification which drives viral escape.

**Fluctuations explain the variability of pH-driven viral escape.** There exists a simple connection between the levels of endosomal Rab5, Rab7 and endosomal pH given by Eq. (1) (for details, please, see Ref.<sup>16</sup>). In our case, the estimated time evolution of pH is given in Fig. 5A, where we have used the best-fit parameters from Fig. 3 (see Table S1 in the Supplementary Information). The shape of the curve is consistent with previous results<sup>16,29,44</sup>. Since we can compute the second-order moments of the distribution, we can evaluate the standard deviation of the pH (dashed lines in Fig. 5A).

As we mentioned in the Introduction, many intracellular processes are initiated by low (below threshold) values of the pH<sup>10–12</sup>. Yet, as shown in Fig. 5A, the pH fluctuates due to the rich dynamics of endosome maturation and Rab decoration. This variability can only be accounted for if one considers the collective dynamics of the population of endosomes, as described by the distribution  $n(x_5, x_7; t)$ .

The exact distribution for the number of endosomes with a given pH cannot be computed analytically, as it would involve: (i) computing the inverse Laplace transform of  $\hat{n}(z_5, z_7; t)$ , and (ii) computing the marginal distribution

$$n(\text{pH}) = \iint dx_5 dx_7 n(x_5, x_7; t) \delta \left[ \text{pH} - \left( \text{pH}^{\text{early}} + \left( \text{pH}^{\text{late}} - \text{pH}^{\text{early}} \right) \frac{x_7}{x_5 + x_7} \right) \right].$$

Instead, we can exploit this variability to approximate the fluctuations by a normal distribution (Central Limit Theorem). In particular, we can write

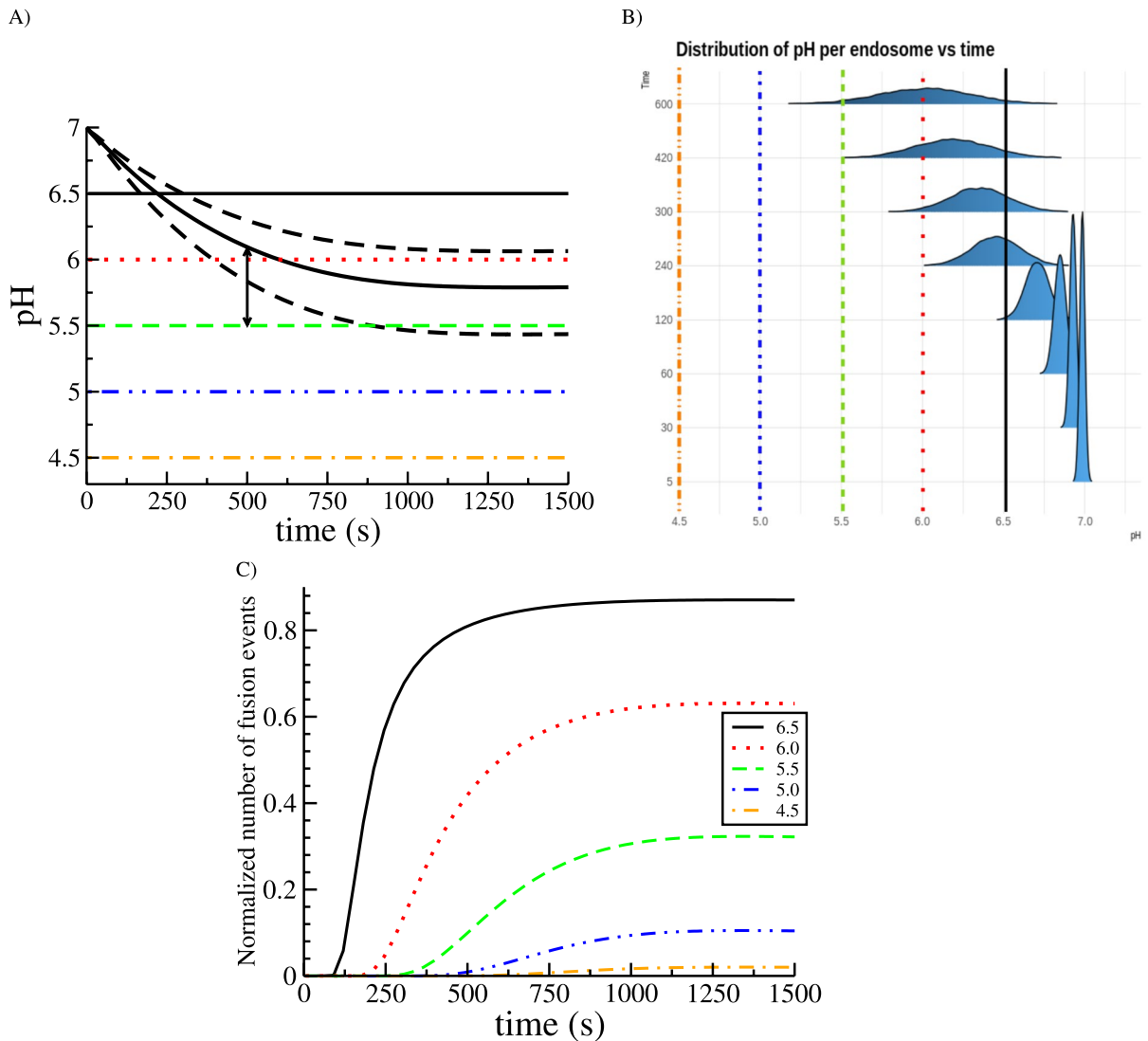
$$\text{pH}(t) \sim \mathcal{N}[\overline{\text{pH}}(t), \sigma_{\text{pH}}(t)], \quad (26)$$

where  $\sim \mathcal{N}$  denotes normally distributed, and  $\overline{\text{pH}}$  and  $\sigma_{\text{pH}}$  are the mean and standard deviation of the pH (solid and dashed lines in Fig. 5A, respectively). In Fig. 5B we show synthetically generated histograms according to Eq. (26) to emphasise the role of the width of the distribution and how it affects the probability of having a pH below a certain threshold. The resulting probabilities are shown in Fig. 5C for different values of the pH threshold. For instance, the green dashed line in Fig. 5C shows that at time 500 s, the probability of finding endosomes with pH below 5.5 is already 10%, although the mean pH is above 6.0 at that time (see the arrows in Fig. 5). This probability represents also the normalised number of viral escape events (or fraction of *successful* events); that is, for a virus and a fiducial (fixed but arbitrary) pH escape threshold, it gives the probability of viral escape. For instance, under ideal conditions, a virus that requires a pH=4.5 to escape would only have a maximal probability of success of 0.022 ( $\sim 2\%$ ). This, in addition to the large mean time to achieve that probability (that would allow the endosome to fuse with the lysosome and, thus, destroy the intracellular virus), would make the infection non-viable.

This result is of great relevance in the understanding of endosomal viral escape events. For instance, in Ref.<sup>10</sup> (whose data we are using in the present work), analysis of the experiments revealed that escape events occurred from 300 s post-entry (viral endocytosis). Moreover, colocalisation of escape events with levels of Rab5 and Rab7 also showed that around 5% of the viral particles escaped from within Rab5-positive early endosomes (with no active Rab7 molecules). This implies that the quick pH drop after those early endosomes merged with more acidic endosomes is ultimately regulating the initiation of viral escape events.

## Discussion and conclusions

Traditional mathematical methods, based on ordinary differential equations, can only tell a part of the story when describing systems with a small number of “particles”, such as endosomes in the present case, and where heterogeneity can be large. Deterministic approaches can be a good approximation if one is interested in average numbers or trends, but exclusively at the cell population level. However, in those cases where the response of a system to small variations of a parameter is abrupt (as in the case of viral escape to a pH threshold), averaging can provide confounding answers or lead to mathematical models requiring a large number of parameters. One of the main results of the present work is that, while detailed models of Rab5 and Rab7 dynamics can be found to fit accurately to experiments, such as in Refs.<sup>21–30</sup>, a description based on individual endosomes, decorated with Rab molecules (defined by the distribution  $n(x_5, x_7; t)$ ), and their interactions (characterised by fusion and fission events), provides a natural link to the underlying biological mechanisms. The mathematical framework proposed here also allows us to characterise the fluctuations beyond the mean number of endosomes or Rab molecules, since higher-order moments can be computed from the endosomal distribution  $n(x_5, x_7; t)$ . As such, we can determine the time course of the variance in the number of Rab molecules and the covariance.



**Figure 5.** Cut-off switch model predicts large endosomal pH fluctuations. **(A)** Time evolution of the mean endosomal pH (thick curved line) computed from Eqs. (17)–(19) and Eq. (1), and mean  $\pm$  standard deviation (computed from Eqs. (S26)–(S28) in Supplementary Information, curved dashed lines). The horizontal colour lines correspond to different pH thresholds that could regulate and trigger virus endosomal escape. **(B)** Similar to **(A)** but showing the sampled distribution of pH using Eq. (26). Note how the distribution broadens with time, thus increasing the probability of crossing a given pH threshold. **(C)** Normalised number of viral escape events quantified as the probability of pH being below a threshold (see colour-coded legend), making use of Eq. (26), namely,  $\mathbb{P}[\mathcal{N}(\bar{pH}, \sigma_{pH}) < \text{threshold}]$ .

We have shown that fusion and fission events regulate the maturation and dynamics of endosomes. Even when we have only considered linear constitutive equations, such as  $K_{FUS}(x_5, x'_5, x_7, x'_7)$  or  $\bar{J}(x_5, x_7)$ , they have allowed us to capture the complexity of the problem at hand. For instance, more detailed models where kiss-and-run<sup>6</sup> or vesicle budding events<sup>45</sup>, decoupled from the rates of fusion/fission, might help to quantify the relative role of those mechanisms. Our approach also sheds some light on the variability of other processes relying on this maturation. As an illustration, as it has been shown for many different viruses<sup>10–12</sup>, viral escape times (after pH drop) are rather heterogeneous, even though endosomal acidification takes place in a gradual manner. For instance, in the case of DENV<sup>10</sup>, it was already reported that some viral particles escaped as early as a few seconds after entry via endocytosis and that, in those cases, fusion with a more acidic endosome preceded that escape event. So, understanding endosome dynamics can be relevant to ascertain the role of different entry pathways in the subsequent fate of viral particles, since different receptors deliver the virus into distinct populations of early endosomes<sup>46</sup>.

Finally, and from a pragmatic perspective, while Eq. (3) has proven useful to study the time evolution of the number of endosomes, and the total number of active Rab5 and Rab7 molecules on the endosomal membrane, it is still a complex system of ODEs, hard to solve analytically. Thus, computational methods aimed to find solutions to these equations can generate rather valuable information. For example, knowledge of the exact distribution of

endosomes and their Rab5 and Rab7 cargo,  $n(x_5, x_7; t)$ , would provide the cell *endosomal pH spectrum*, namely, the number of endosomes with a certain pH,  $n(\text{pH}, t)$ , that could be compared with recent experimental methods aimed to quantify intracellular pH<sup>40</sup>. Also, this formalism can be translated to other contexts or scales:  $n(x, y; t)$  might be seen as the number of cells with a certain expression level of receptors  $x$  and  $y$ . Thus, solving the corresponding Smoluchowski equation for  $n(x, y; t)$  would be a theoretical metaphor of flow cytometry experiments. Another interesting and timely application of our mathematical framework is that of mitochondrial dynamics and interactions. In this case, and in analogy with the endosomes, mitochondria are subject to fusion and fission events modulated by different cargo species (e.g.,  $\text{Ca}^{2+}$ , ATP, reactive oxygen species, mtDNA, etc.)<sup>47</sup>. These applications will be considered in future work and are out of the scope of this manuscript.

Received: 22 October 2020; Accepted: 15 March 2021

Published online: 12 April 2021

## References

- Naslavsky, N. & Caplan, S. The enigmatic endosome—sorting the ins and outs of endocytic trafficking. *J. Cell Sci.* **131**, jcs216499 (2018)
- Tokarev, A. A., Alfonso, A. & Segev, N. Overview of intracellular compartments and trafficking pathways. In *Trafficking Inside Cells*, 3–14 (Springer, 2009).
- Pfeffer, S. R. Rab GTPase regulation of membrane identity. *Curr. Opin. Cell Biol.* **25**, 414–419 (2013)
- Staring, J., Raaben, M. & Brummelkamp, T. R. Viral escape from endosomes and host detection at a glance. *J. Cell Sci.* **131**, jcs216259. <https://doi.org/10.1242/jcs.216259> (2018).
- Rink, J., Ghigo, E., Kalaidzidis, Y. & Zerial, M. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* **122**, 735–749 (2005)
- Duclos, S. *et al.* Rab5 regulates the kiss and run fusion between phagosomes and endosomes and the acquisition of phagosome leishmanicidal properties in RAW 264.7 macrophages. *J. Cell Sci.* **113**, 3531–3541 (2000)
- Huotari, J. & Helenius, A. Endosome maturation. *EMBO J.* **30**, 3481–3500 (2011)
- Flipse, J., Wilschut, J. & Smit, J. M. Molecular mechanisms involved in antibody-dependent enhancement of dengue virus infection in humans. *Traffic* **14**, 25–35 (2013)
- Krishnan, M. N. *et al.* Rab 5 is required for the cellular entry of dengue and West Nile viruses. *J. Virol.* **81**, 4881–4885 (2007)
- van der Schaar, H. M. *et al.* Dissecting the cell entry pathway of dengue virus by single-particle tracking in living cells. *PLoS Pathog.* **4**, e1000244 (2008)
- Chao, L. H., Klein, D. E., Schmidt, A. G., Peña, J. M. & Harrison, S. C. Sequential conformational rearrangements in flavivirus membrane fusion. *Elife* **3**, e04389 (2014)
- van Duijl-Richter, M. K., Blijleven, J. S., van Oijen, A. M. & Smit, J. M. Chikungunya virus fusion properties elucidated by single-particle and bulk approaches. *J. Gen. Virol.* **96**, 2122–2132 (2015)
- Hoorweg, T. E. *et al.* Dynamics of chikungunya virus cell entry unraveled by single-virus tracking in living cells. *J. Virol.* **90**, 4745–4756 (2016)
- Feng, Y., Press, B. & Wandinger-Ness, A. Rab 7: An important regulator of late endocytic membrane traffic. *J. Cell Biol.* **131**, 1435–1452 (1995)
- Mellman, I., Fuchs, R. & Helenius, A. Acidification of the endocytic and exocytic pathways. *Ann. Rev. Biochem.* **55**, 663–700 (1986)
- Lagache, T., Sieben, C., Meyer, T., Herrmann, A. & Holcman, D. Stochastic acidification, activation of hemagglutinin and escape of influenza viruses from an endosome. *Front. Phys.* **5**, 25 (2017)
- Binder, B., Goede, A., Berndt, N. & Holzhütter, H. -G. A conceptual mathematical model of the dynamic self-organisation of distinct cellular organelles. *PLoS ONE* **4**, e8295 (2009)
- Rodenhuis-Zybert, I. A., Wilschut, J. & Smit, J. M. Dengue virus life cycle: Viral and host factors modulating infectivity. *Cell. Mol. Life Sci.* **67**, 2773–2786 (2010)
- Floyd, D. L., Ragains, J. R., Skehel, J. J., Harrison, S. C. & van Oijen, A. M. Single-particle kinetics of influenza virus membrane fusion. *Proc. Natl. Acad. Sci.* **105**, 15382–15387 (2008)
- Ivanovic, T., Choi, J. L., Whelan, S. P., van Oijen, A. M. & Harrison, S. C. Influenza-virus membrane fusion by cooperative fold-back of stochastically induced hemagglutinin intermediates. *Elife* **2**, e00333 (2013)
- Del Conte-Zerial, P. *et al.* Membrane identity and GTPase cascades regulated by toggle and cut-out switches. *Mol. Syst. Biol.* **4**, 206 (2008)
- Gautreau, A., Oguievetskaia, K. & Ungermann, C. Function and regulation of the endosomal fusion and fission machineries. *Cold Spring Harb. Perspect. Biol.* **6**, a016832 (2014)
- Wandinger-Ness, A. & Zerial, M. Rab proteins and the compartmentalization of the endosomal system. *Cold Spring Harb. Perspect. Biol.* **6**, a022616 (2014)
- Vetter, I. R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299–1304 (2001)
- Tashkova, K., Korošec, P., Šilc, J., Todorovski, L. & Džeroski, S. Parameter estimation with bio-inspired meta-heuristic optimization: Modeling the dynamics of endocytosis. *BMC Syst. Biol.* **5**, 159 (2011)
- Vonderheit, A. & Helenius, A. Rab7 associates with early endosomes to mediate sorting and transport of Semliki forest virus to late endosomes. *PLoS Biol.* **3**, e233 (2005)
- Binder, B. & Holzhütter, H. -G. A hypothetical model of cargo-selective rab recruitment during organelle maturation. *Cell Biochem. Biophys.* **63**, 59–71 (2012)
- Foret, L. *et al.* A general theoretical framework to infer endosomal network dynamics from quantitative image analysis. *Curr. Biol.* **22**, 1381–1390 (2012)
- Grabe, M. & Oster, G. Regulation of organelle acidity. *J. Gen. Physiol.* **117**, 329–344 (2001)
- Zeigerer, A. *et al.* Rab5 is necessary for the biogenesis of the endolysosomal system in vivo. *Nature* **485**, 465–470 (2012)
- Castro, M., Lythe, G. & Molina-París, C. The T cells in an ageing virtual mouse. In *Stochastic Processes, Multiscale Modeling, and Numerical Methods for Computational Cellular Biology*, 127–140 (Springer, 2017).
- Zerial, M. & McBride, H. Rab proteins as membrane organizers. *Nat. Rev. Mol. Cell Biol.* **2**, 107–117 (2001)
- Smoluchowski, M. V. Brownsche Molekularbewegung und Koagulation von Kolloidteilchen. *Phys. Zeits.* **17**, 585–599 (1916)
- Meakin, P. Aggregation kinetics. *Physica Scripta* **46**, 295 (1992)
- Lade, S. J., Coelho, M., Tolić, I. M. & Gross, T. Fusion leads to effective segregation of damage during cell division: An analytical treatment. *J. Theor. Biol.* **378**, 47–55 (2015)
- Donovan, G. M. & Lythe, G. T cell and reticular network co-dependence in HIV infection. *J. Theor. Biol.* **395**, 211–220 (2016)
- Cuesta, J. A., Delius, G. W. & Law, R. Sheldon spectrum and the plankton paradox: Two sides of the same coin—a trait-based plankton size-spectrum model. *J. Math. Biol.* **76**, 67–96 (2018)

38. Kalaidzidis, Y. *et al.* Learning microscopic kinetic characteristic of endosomal network by quantitative analysis of snap-shot microscopy images. In *Proceedings of the International Conference on Bioinformatics & Computational Biology (BIOCOMP)*, 1 (The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2013).
39. Ganusov, V. V. Strong inference in mathematical modeling: A method for robust science in the twenty-first century. *Front. Microbiol.* **7**, 1131 (2016)
40. Modi, S. *et al.* A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nat. Nanotechnol.* **4**, 325 (2009)
41. Padilla-Parra, S. *et al.* Quantitative imaging of endosome acidification and single retrovirus fusion with distinct pools of early endosomes. *Proc. Natl. Acad. Sci.* **109**, 17627–17632 (2012)
42. Camacho, J. Scaling in steady-state aggregation with injection. *Phys. Rev. E* **63**, 046112 (2001)
43. Murray, D. H. *et al.* An endosomal tether undergoes an entropic collapse to bring vesicles together. *Nature* **537**, 107 (2016)
44. Brett, C. L., Tukaye, D. N., Mukherjee, S. & Rao, R. The yeast endosomal Na<sup>+</sup> (K<sup>+</sup>)/H<sup>+</sup> exchanger Nhx1 regulates cellular pH to control vesicle trafficking. *Mol. Biol. Cell* **16**, 1396–1405 (2005)
45. Foret, L. & Sens, P. Kinetic regulation of coated vesicle secretion. *Proc. Natl. Acad. Sci.* **105**, 14763–14768 (2008)
46. Lakadamyali, M., Rust, M. J. & Zhuang, X. Ligands for clathrin-mediated endocytosis are differentially sorted into distinct populations of early endosomes. *Cell* **124**, 997–1009 (2006)
47. Farmer, T., Naslavsky, N. & Caplan, S. Tying trafficking to fusion and fission at the mighty mitochondria. *Traffic* **19**, 569–577 (2018)

## Acknowledgements

This work is partially supported by Ministerio de Economía y Competitividad, Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación through Grant PID2019-106339GB-I00 (MC). The research leading to these results has also received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement number PIRSES-GA-2012-317893 (MC, GL and CMP) and from EPSRC under a University of Leeds Impact Acceleration Account number IAA3025 (GL and CMP).

## Author contributions

M.C. and C.M.P. conceived and developed the mathematical methods. M.C. analysed the data and carried out parameter calibration. M.C. and C.M.P. wrote the manuscript. M.C., C.M.P., G.L. and J.S. contributed to the discussions and reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-86877-w>.

**Correspondence** and requests for materials should be addressed to M.C. or C.M.-P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021