A Minimal Self-organisation model of the Golgi Apparatus

Quentin Vagne[a,1], **Jean-Patrick Vrel**[b,1], **Pierre Sens**[b]

***For correspondence:**
pierre.sens@curie.fr (PS)

[1]These authors contributed equally to this work

[a]Center for Systems Biology Dresden, Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany; [b]Laboratoire Physico Chimie Curie, Institut Curie, PSL Research University, CNRS UMR168, 75005, Paris, France

Abstract The design principles dictating the spatio-temporal organisation of eukaryotic cells, and in particular the mechanisms controlling the self-organisation and dynamics of membrane-bound organelles such as the Golgi apparatus, remain elusive. Although this organelle was discovered 120 years ago, such basic questions as whether vesicular transport through the Golgi occurs in an anterograde (from entry to exit) or retrograde fashion are still strongly debated. Here, we address these issues by studying a quantitative model of organelle dynamics that includes: *de-novo* compartment generation, inter-compartment vesicular exchange, and biochemical conversion of membrane components. We show that anterograde or retrograde vesicular transports are asymptotic behaviors of a much richer dynamical system. Indeed, the structure and composition of cellular compartments and the directionality of vesicular exchange are intimately linked. They are emergent properties that can be tuned by varying the relative rates of vesicle budding, fusion and biochemical conversion.

Introduction

The Golgi apparatus is an intracellular organelle at the crossroad of the secretory, lysosomal and endocytic pathways (***Heald and Cohen-Fix, 2014***). One of its most documented functions is the sorting and processing of many proteins synthesized by eukaryotic cells (***Lippincott-Schwartz et al., 2000***). Proteins translated in the endoplasmic reticulum (ER) are addressed to the Golgi, where they undergo post-translational maturation and sorting, before being exported to their final destination. The Golgi itself is composed of distinct sub-compartments, called cisternae, of different biochemical identities (***Stanley, 2011***). From the ER, cargo-proteins successively reside in cisternae of the *cis*, *medial* and *trans*-identities, after which they exit the Golgi via the *trans*-Golgi network (TGN). In some organisms such as the yeast Saccharomyces cerevisiae, cisternae are dispersed throughout the cell (***Suda Yasuyuki and Nakano Akihiko, 2011***) and each cisterna undergoes maturation from a *cis* to a *trans*-identity (***Losev et al., 2006***; ***Matsuura-Tokita et al., 2006***). In most other eukaryotes, cisternae are stacked in a polarized fashion, with cargo-proteins entering via the *cis*-face and exiting via the *trans*-face (***Boncompain and Perez, 2013***). This polarity is robustly conserved over time, despite cisternae constantly exchanging vesicles with each other, the ER and the TGN (***Lippincott-Schwartz et al., 2000***). Pharmacological treatments that alter the structure and stacking of Golgi compartments in mammalian cells affect the processing of certain proteins, and in particular glycosylation (***Hu et al., 2005***; ***Shen et al., 2007***; ***Xiang et al., 2013***; ***Bekier et al., 2017***).

The highly dynamical nature and compact structure of a stacked Golgi makes it difficult to determine how cargo-proteins are transported in an anterograde fashion from the *cis* to *trans*-side of the stack and how this transport is coupled to processing. Several models have been proposed to explain the transport dynamics of Golgi cargo (See sketch in Fig.1 A). They mostly belong to

two classes: one involving stable compartments exchanging components through fusion-based mechanisms and one involving transient compartments undergoing *en-block* maturation and in which fusion mechanisms are dispensable for cargo transport. Historically (***Malhotra and Mayor, 2006***), the "Vesicular transport" model postulated that cisternae have fixed identities and cargo progresses from one cisterna to the next via anterograde vesicular transport, while the "Cisternal Maturation" model postulated that cisternae themselves undergo maturation from the *cis* to the *trans*-identity and physically move through the stack, while Golgi resident proteins remain in the stack by moving toward the *cis*-face by retrograde vesicular transport. Several extensions of these two models have been proposed, including cisternal maturation with tubular connections (***Trucco et al., 2004***; ***Rizzo et al., 2013***), rapid partitioning within Golgi cisternae between processing and exporting sub-compartments (***Patterson et al., 2008***), or the cisternal progenitor model (***Pfeffer, 2010***), in which stable cisternae exchange material by the fusion and fission of sub-compartments defined by their composition of Rab GTPases, which evolve over time through exchange with the cytosol. The strengths and weaknesses of these different models are nicely reviewed in ***Glick and Luini*** (***2011***). Variations of these models, such as the “rim progression” model also exist (***Lavieu et al., 2013***). It is noteworthy that these models do not provide a quantitative analysis of the generation and maintenance of Golgi compartments, nor do they attempt to relate the Golgi structure (number and size of compartments) and transport dynamics. Therefore, much remains unknown regarding the mechanisms that dictate the organisation and dynamics of the Golgi.

Although there is a large diversity in Golgi structures and dynamics among different species, the physiological function of the organelle as a sorting and processing hub is common to all species, suggesting that important mechanisms controlling Golgi dynamics are conserved. Works in evolutionary biology and biophysics have attempted to describe these mechanisms (***Klute et al., 2011***; ***Sens and Rao, 2013***). Different classes of mathematical models have been proposed, from models of vesicle budding and fusion based on rate equations (***Binder et al., 2009***; ***Ispolatov and Müsch, 2013***; ***Mukherji and O’Shea, 2014***) - some specifically including space (***Sachdeva et al., 2016***), to continuous, discrete and stochastic models of protein sorting in the Golgi (***Gong et al., 2010***; ***Vagne and Sens, 2018b***) and computer simulations based on membrane mechanics (***Tachikawa and Mochizuki, 2017***). To account for the ability of the Golgi to reassemble after mitosis (***Wei and Seemann, 2017***), many of these studies have sought to describe the Golgi as a self-organized system (***Binder et al., 2009***; ***Kühnle et al., 2010***). However, the manner in which the kinetics of the Golgi two main functions (namely the transport and biochemical conversion of its components) interplay with one another to yield a robust steady-state has so far received much less attention (***Sachdeva et al., 2016***; ***Vagne and Sens, 2018b***).

Most existing models discuss transport through a pre-existing compartmentalised Golgi structure. The main goal of our model is to explain how this compartmentalised structure and the flux within it can spontaneously emerge through self-organisation. We propose here a model in which both Golgi self-organisation and vesicular transport are solely directed by (local) molecular interactions resulting in composition-dependent budding and fusion. In particular, we neglect the spatial structure of the Golgi, and assume that any two components of the system can interact with one another (see the Discussion section for the relevance of this approximation). Thus, the directionality of vesicular transport is an emergent property of the self-organisation process that co-evolves with the size and number of Golgi sub-compartments, rather than being fixed by arbitrary rules. This approach brings new conceptual insights that highlight the link between the large-scale steady-state organisation of the Golgi, the directionality of vesicular transport, and the kinetics of individual processes at the molecular scale. We show that a wide variety of organelle phenotypes can be observed upon varying the rates of the local processes. In particular, our model uncovers a correlation between the size and composition (purity) of Golgi sub-compartments. Besides, it emphasizes that composition-dependent exchange between compartments yields complex vesicular flux patterns that are neither purely anterograde (as in the “vesicular transport" model) nor purely retrograde (as in the “cisternal maturation” model).

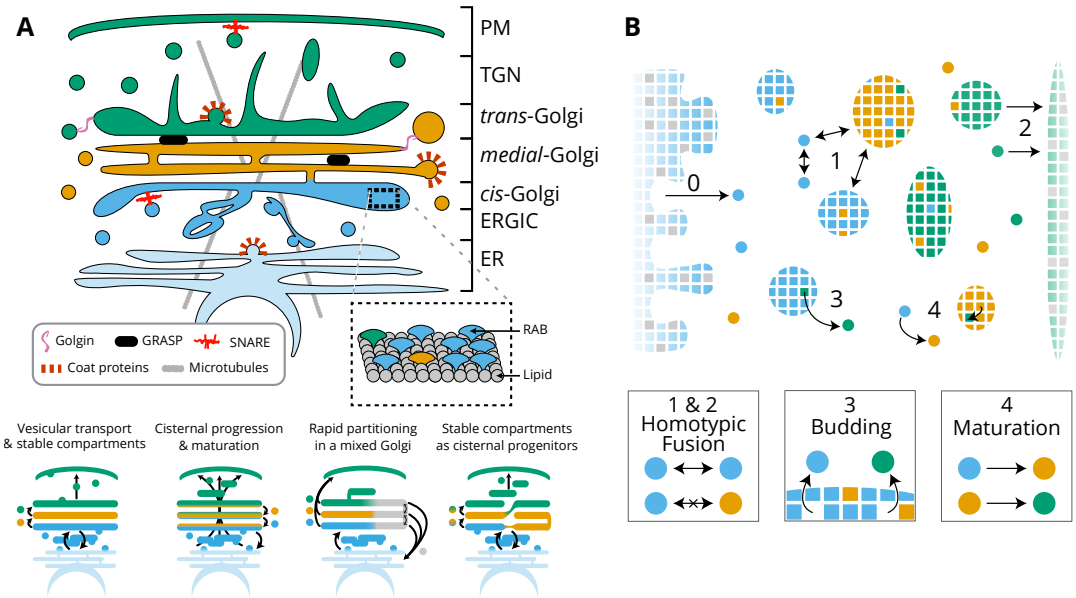


Figure 1. Model of golgi structure and transport (A) Top: classical representation of the structure of the Golgi, showing some of the important protein actors. The three main membrane identities are shown in different colors (*cis*: blue, *medial*: orange, *trans*: green): ER=Endoplasmic Reticulum, ERGIC= ER-Golgi intermediate Compartment, TGN=Trans-Golgi-Network, PM=Plasma Membrane. Bottom: Sketches of the four main models of Golgi transport (see text). (B) Our quantitative model of Golgi self-organisation. The left boundary is the ER, composed of a *cis*-membrane identity, and the right boundary is the TGN, composed of a *trans*-membrane identity. Golgi compartments self-organise *via* three stochastic mechanisms: **Fusion**: (1) All compartments can aggregate using homotypic fusion mechanisms: the fusion rate is higher between compartment of similar identities. (2) Each compartment can exit the system by fusing homotypically with the boundaries. **Budding:** (3) Each compartment larger than a vesicle can create a vesicle by losing a patch of membrane. **Biochemical conversion:** (4) Each patch of membrane undergoes a conversion from a *cis* to a *trans*-identity. New *cis*-vesicles (0) bud from the ER at a constant rate. In the sketch, the boundaries also contain neutral (gray) membrane species that dilute their identity (impact of this dilution in Appendix App.7, page p.39).

Model

The model, shown in Fig.1 and fully described in the Appendix App.1, is a coarse-grained representation of the Golgi Apparatus. The system is discretised at the scale of small vesicles, which define the unit size of the system. Vesicles can fuse together or with existing compartments to form larger compartments and can bud from compartments. Compartments consist of a number of membrane patches (of size unity) with distinct identities. The biochemical identity of a membrane patch is defined in a very broad sense by its composition in lipids and proteins that influence its interaction with other patches, such as Rab GTPases (***Grosshans et al., 2006***) or fusion proteins such as SNAREs (***Cai et al., 2007***). Consistently with the three main biochemical identities reported for Golgi membranes (***Day et al., 2013***), vesicles and membrane patches are of either *cis*, *medial* or *trans*-identity, and can undergo irreversible biochemical conversion from the *cis* to the *trans*-identity. The system is fed by a constant influx J of *cis* vesicles coming from the ER, and vesicles and compartments can fuse homotypically with the ER (itself of *cis* identity) and the TGN (of *trans* identity).

Compartments are defined by their size n (a number of patches) and their composition, the fractions ϕ_i (with $i = cis, medial, trans$) of patches of the three different identities composing it. These quantities are dynamically controlled by three, composition dependent, microscopic mechanisms – budding, fusion and biochemical conversion – as described below.

- **Fusion**: Homotypic fusion – the higher probability of fusion between compartments of similar identities – is a feature of cellular organelles in general (***Antonin et al., 2000***) and the Golgi apparatus in particular (***Pfeffer, 2010***; ***Bhave et al., 2014***). This is controlled here by a fusion

Codes used for the simulation can be found on the UMR168 Curie website, section `team-sens/software`. **Click to access**

rate between compartments proportional to the probability that they present the same identity at the contact site. The total fusion rate for two compartments (a) and (b) with composition $\phi_i^{(a)}$ and $\phi_i^{(b)}$ for each identity i (with i equals *cis*, *medial* and *trans*) is then:

$$K_\mathrm{f} \times \sum_i \phi_i^{(a)} \phi_i^{(b)} \tag{1}$$

where K_f is the fusion rate between two identical compartments. Fusion with the boundaries follow the same rules, with the ER containing a fraction α_ER of *cis* patches and the TGN a fraction α_TGN of *trans* patches. Numerical results described in the main paper are for $\alpha_\mathrm{ER} = \alpha_\mathrm{TGN} = 1$, and the role of the composition of the boundaries is discussed in the App.7.

- **Budding**: Budding is the emission of a vesicle from a larger compartment. Composition-dependent vesicular budding is an important sorting mechanism in cellular traffic (***Bonifacino and Glick, 2004***). Owing to the high specificity of the budding machinery, each vesicle is assumed to be composed of a single membrane identity. The budding flux for any components i (equals *cis*, *medial* or *trans*) present in a compartment is assumed not to depend on the number $n_i = n\phi_i$ of patches of identity i, but on the total size of the compartment:

$$J_{\mathrm{b},i} = K_\mathrm{b} \times n. \tag{2}$$

This budding kinetics corresponds to the saturated regime of a Michaelis-Menten reaction kinetics, and is appropriate to the case where budding actors (*e.g.* coat proteins) bind non-specifically to the membrane of the compartment and find their budding partners (a patch of a particular identity) by diffusion (***Vagne and Sens, 2018b***). An alternative budding scheme where the budding flux is linear with the number of patches of a given identity ($J_{\mathrm{b},i} = K_\mathrm{b} n_i$) is discussed in App.8 on p.41

- **Biochemical conversion**: Each membrane patch is assumed to undergo stochastic biochemical conversion from a *cis* to a *medial* to a *trans*-identity. This local mechanism of identity conversion is motivated by the Rab cascade, during which the membrane identity evolves over time through molecular exchange with the cytosol, as was observed in endosomes (***Rink et al., 2005***) and in the Golgi of Yeast (***Losev et al., 2006***; ***Matsuura-Tokita et al., 2006***). The Rab conversion mechanism has not been demonstrated in mammalian Golgi, but is thought to be universally important for the specificity of membrane trafficking (***Grosshans et al., 2006***), and many of the proteins involved in the Yeast Rab cascade are conserved in mammalian cells (***Klute et al., 2011***). Rab conversion is at the basis of the cisternal progenitor model (***Pfeffer, 2010***), and it is thus of fundamental importance to understand its potential implication on Golgi structure and dynamics with the use of a quantitative model. We stress that our model of biochemical conversion is not limited to Rab GTPases, but can also correspond to the modification of lipid components by enzymes in the Golgi (***McDermott and Mousley, 2016***). This is a local mechanism of identity conversion distinct from the maturation of entire compartments, which is also affected by the dynamics of budding and fusion. To limit the number of parameters, the biochemical conversion rate K_m is the same for the two reactions (cis→medial and medial→trans) and is independent of the concentration. Introducing cooperativity in the conversion process is expected to increase the robustness of the self-organisation process (***Vagne and Sens, 2018a***).

In its simplest form, our model contains only four parameters: the rates of injection, fusion, budding, and biochemical conversion (J, K_f, K_b, K_m). By normalizing time with the fusion rate, we are left with three parameters: $j = J/K_\mathrm{f}$, $k_\mathrm{b} = K_\mathrm{b}/K_\mathrm{f}$ and $k_\mathrm{m} = K_\mathrm{m}/K_\mathrm{f}$. The dynamics of the

system is entirely governed by these stochastic transition rates, and can be simulated exactly using a Gillespie algorithm (***Gillespie, 1977***), described in App.1.

Results

Mean-field description of the system.

The complexity of the system prohibits rigorous analytical calculation. Nevertheless, analytical results can be obtained for a number of interesting quantities in certain asymptotic regimes where simplifying assumptions can be made. We present below some of these derivations.

De-novo formation and steady-state composition in the well-sorted limit

The composition of the system at steady-state is difficult to compute due to the fact that the exit of components from the system, through fusion with the boundaries, depends on the composition of the exiting compartments, which cannot be derived analytically in the general case. This calculation becomes straightforward if the system is well sorted and all compartments are pure. As explained below, this corresponds to the limit of high budding rate: $k_\mathrm{b} \gg k_\mathrm{m}$. In the well-sorted limit, only *cis* components exit through the *cis* boundary and only *trans* components exit through the *trans* boundary, and mean-field equations can be derived for the total amount N_{cis}, N_{medial} and N_{trans} of each species:

$$\frac{\partial N_{cis}}{\partial t} = j - N_{cis}(\alpha_\mathrm{ER} + k_\mathrm{m})\,,\ \frac{\partial N_{medial}}{\partial t} = k_\mathrm{m}(N_{cis} - N_{medial})\,,\ \frac{\partial N_{trans}}{\partial t} = k_\mathrm{m} N_{medial} - \alpha_\mathrm{TGN} N_{trans} \tag{3}$$

With j the injection rate of new *cis*-vesicles in the system and k_m the biochemical conversion rates, both normalized by the fusion rate, and α_ER (α_TGN) is the fraction of *cis* (*trans*) species promoting homotypic fusions with Golgi components in the ER (TGN).

At steady-state, the total amounts of *cis*, *medial* and *trans*-species are fixed by the balance between influx, exchange, biochemical conversion and exit:

$$N_{cis} = N_{medial} = \frac{j}{\alpha_\mathrm{ER} + k_\mathrm{m}}\,,\ N_{trans} = \frac{k_\mathrm{m}}{\alpha_\mathrm{TGN}} \frac{j}{\alpha_\mathrm{ER} + k_\mathrm{m}} \tag{4}$$

To estimate the typical size of compartments, we assume that for each species, a single large compartment of size n interacts through budding and fusion with a number n_v of vesicles of the same identity (so that $N = n + n_\mathrm{v}$). The compartments size for each species then satisfies:

$$\frac{\partial n}{\partial t} = n_\mathrm{v} - k_\mathrm{b} n \quad \xrightarrow[\text{steady state}]{} \quad n = \frac{N}{1 + k_\mathrm{b}} \tag{5}$$

To limit the number of parameters, we fix the average system's size $N = N_{cis} + N_{medial} + N_{trans}$ to a set value $N = 300$ in the main text, which is suitable for Golgi ministacks whose total area is of the order of $1\mu\mathrm{m}^2$ (the area of a mammalian Golgi ribbon is much larger) (***Yelinek et al., 2009***), corresponding to about a few hundreds of vesicles of diameter $\sim 10 - 50\mathrm{nm}$. This is obtained by adjusting the influx to the variation of the other parameters according to:

$$j = \frac{N\ (\alpha_\mathrm{ER} + k_\mathrm{m})\ \alpha_\mathrm{TGN}}{2\alpha_\mathrm{TGN} + k_\mathrm{m}} \tag{6}$$

We show in App.1 Fig.1 that the overall structure of the phase diagrams for the compartments size and composition is robust upon variation of the average system's size.

The de-novo formation of the system can be approached with Eq.3. Starting with an empty system with an influx $J(= K_f j)$ of *cis* vesicles, the system grows stochastically to reach the steady-state size after a time of order $1/K_\mathrm{f}$. The evolution of the size of the system with time obtained from numerical simulations, shown in App.4, Fig.1 for different sets of parameters, agrees with Eq.3. At steady-state, the influx of vesicles is balanced by an outflux of material. The exit kinetics can be quantified by looking at the fate of a pulse of cargo released from the ER at a given time. App.4 Fig.2 shows that the exit kinetics is exponential, in agreement with results of FRAP (***Patterson et al., 2008***) or pulse-chase (***Boncompain et al., 2012***) experiments.

Although Eq.6 is only valid in the well-sorted limit, it permits a satisfactory control of the average system size over the entire range of parameters, with only a 10% variation for small budding rates (App.5, Fig.1A). At steady-state, the system size exhibits stochastic fluctuations around the mean value that depend on the parameters (App.5, Fig.1A). The large fluctuation (about 30%) for low budding rate k_b stem from the fact that compartments are large and transient, as explained below.

Compartment size and composition in the maturation-dominated regime.

For low values of the budding rate k_b, compartments do not shed vesicles and evolve in time independently of one another in a way dictated by a balance between fusion-mediated growth and biochemical maturation. Their size and composition can be approximatively calculated by assuming that compartments grow in size by fusion from a constant pool of vesicles containing the same number n_v of vesicle for all three identities. Calling $n_{cis}(t)$, $n_{medial}(t)$ and $n_{trans}(t)$ the amount of components of the three identities in a given compartment of total size $n = n_{cis} + n_{medial} + n_{trans}$, and neglecting vesicle budding, the mean-field equations satisfied by these quantities are:

$$\partial_t n_{cis} = n_\mathrm{v}\frac{n_{cis}}{n} - k_\mathrm{m} n_{cis} \ , \ \ \partial_t n_{medial} = n_\mathrm{v}\frac{n_{medial}}{n} + k_\mathrm{m}(n_{cis} - n_{medial}) \ , \ \ \partial_t n_{trans} = n_\mathrm{v}\frac{n_{trans}}{n} + k_\mathrm{m} n_{medial}. \tag{7}$$

Starting with a vesicle of *cis* identity for $t = 0$: $n_{cis}(0) = 1$, $n_{medial}(0) = n_{trans}(0) = 0$, the compartment size evolves linearly with time: $n(t) = 1 + n_\mathrm{v} t$, and the composition of each species satisfies:

$$\begin{aligned} n_{cis}(t) =& (1 + n_\mathrm{v} t)e^{-k_\mathrm{m} t} \\ n_{medial}(t) =& (1 + n_\mathrm{v} t)k_\mathrm{m} t\ e^{-k_\mathrm{m} t} \\ n_{trans}(t) =& (1 + n_\mathrm{v} t)\left(1 - (1 + k_\mathrm{m} t)e^{-k_\mathrm{m} t}\right) \end{aligned} \tag{8}$$

The fraction of each species in the compartment is thus independent on n_v, and reads:

$$\phi_{cis}(t) = e^{-k_\mathrm{m} t} \ , \ \phi_{medial}(t) = k_\mathrm{m} t\ e^{-k_\mathrm{m} t} \ , \ \phi_{trans}(t) = \left(1 - (1 + k_\mathrm{m} t)e^{-k_\mathrm{m} t}\right) . \tag{9}$$

These results show that in this regime, a given compartment smoothly evolves from a mostly *cis* to a mostly *trans* identity over a typical time $1/K_\mathrm{m}$. The maximum amount of medial identity is obtain for $k_\mathrm{m} t = 1$ and is $\phi_\mathrm{medial} = 1/e \simeq 0.37$. Therefore, the mechanism does not lead to pure *medial* compartments. These analytical results are confirmed by the full numerical solution of the system in the low budding rate regime shown in App.5, p.34 and discussed in detail below.

Steady-state organisation

The steady-state organisation and dynamics of the system is described in terms of the average size and purity of compartments (introduced here and detailed in the App.2). The stationary size distribution of compartments is a decreasing function of the compartment size, which typically shows a power law decay at small size, with an exponential cut-off at large size (Fig.2A). This is expected for a non-equilibrium steady-state controlled by scission and aggregation (***Turner et al., 2005***) and means that there exist many small compartments and very few large ones. The typical compartment's size is defined as the ratio of the second over the first moments of the size distribution. This corresponds to a size close to half the size of the exponential cut-off (see App.2) beyond which it is unlikely to find a compartment (***Vagne et al., 2015***).

The typical compartment size varies with parameters as shown in Fig.2B. Increasing the ratio of budding to fusion rate k_b decreases the compartment size (***Turner et al., 2005***). The size depends on the biochemical conversion rate k_m in a non-monotonic manner: for a given budding rate, the size is smallest for intermediate values of the biochemical conversion rate ($k_\mathrm{m} \simeq 1$). The conversion rate controls the composition of the compartment and hence their fusion probability. This dependency is well explained by the simple analytical model discussed above (Eqs. 3,4) that approximates the system by three pure compartments interacting with a pool of vesicles (solid lines in Fig.2B)>

The purity of a given compartment is defined as the (normalized) Euclidean distance of the composition of the compartment (the fractions ϕ_i in the different i-identities, i = *cis*, *medial* and

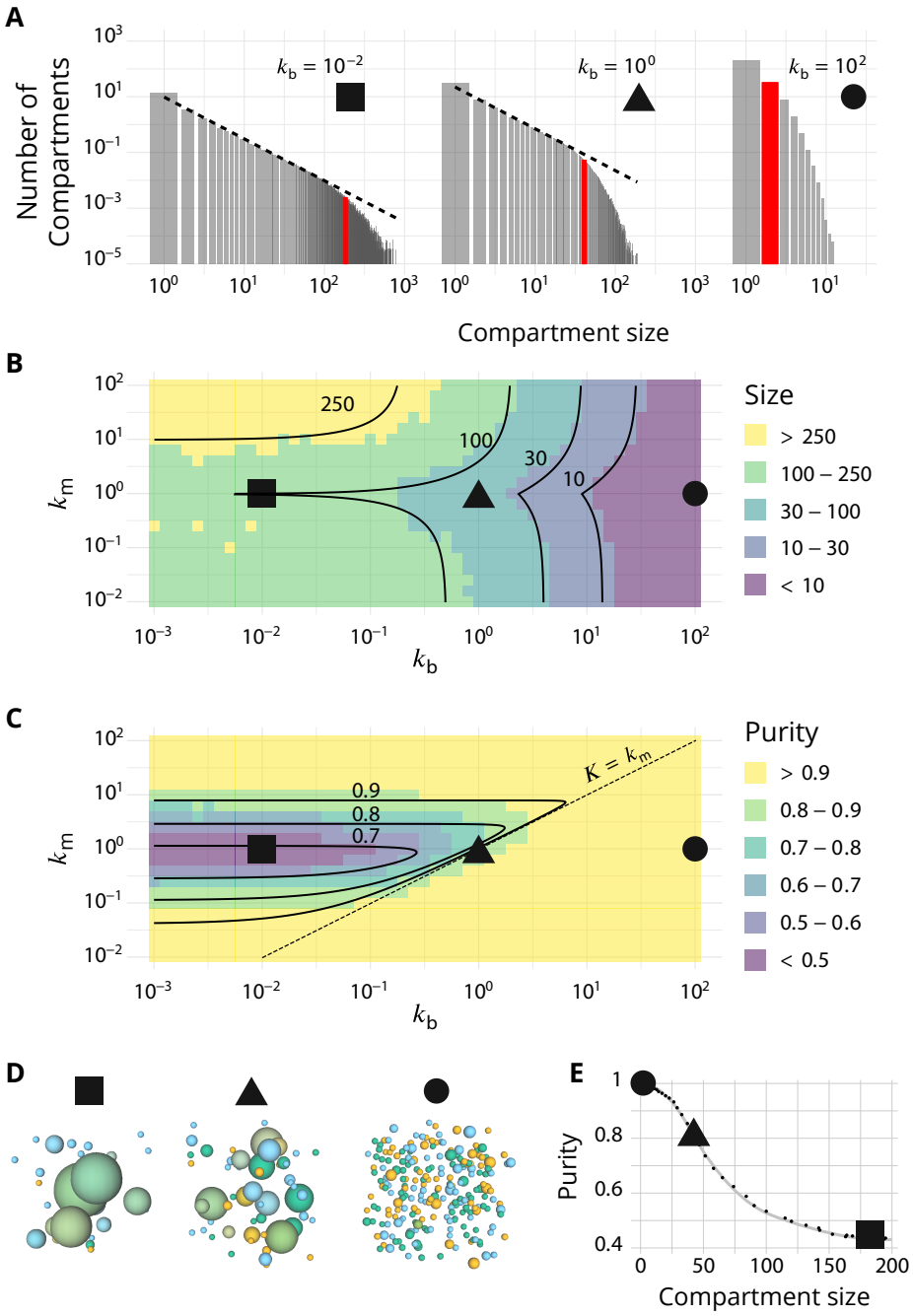


Figure 2. Steady-state of the self-organized model of Golgi apparatus. (A) Size distribution of compartments for a biochemical conversion rate $k_m = 1$ and different values of the budding rate k_b. The red bar shows the characteristic compartment size. (B) Phase diagram of the size of the compartments as the function of k_b and k_m. Black lines: theoretical prediction (Eqs.4,6) (C) Phase diagram of the purity of the system as a function of k_b and k_m. Dashed line is $k_m = k_b$, black lines are a theoretical prediction (see App.3, p.28). (D) System snapshots for $k_m = 1$ showing the *mixed regime* (square - $k_b = 10^{-2}$), the *sorted regime* (triangle - $k_b = 1$), and the *vesicular regime* (circle - $k_b = 10^2$) - see text. (E) Average compartment purity as a function of their average size, obtained by varying the budding rate k_b for $k_m = 1$ (black dots: simulation results – gray line: guide for the eyes). For all panels, $\alpha_{ER} = \alpha_{TGN} = 1$. See also App.5, p.34 for further characterizations of the steady-state organisation, App.7, p.39 for the role of composition of the exit compartment, and App.8, p.41 for results with alternative budding and fusion kinetics.

trans) from a perfectly mixed composition. $P = 0,\ 1/2,\ 1$ correspond respectively to compartments that contain the same amount of the three identities, the same amount of two identities, and a single identity (App.2, page p.25). The purity of the system is the average purity of each compartment weighted by its size, ignoring vesicles. The dependency of the average purity with the parameters is shown in Fig.2C. As for the size of the compartments, it is non-monotonic in the biochemical conversion rate, and regions of lowest purity are found for intermediate biochemical conversion rates $k_{\mathrm{m}} \simeq 1$, when all three identities are in equal amount. The purity increases with the budding rate k_{b} in a sigmoidal fashion (see also Fig.3 A). This can be understood by analyzing the processes involved in mixing and sorting of different identities. In a first approximation, mixing occurs by biochemical conversion and sorting occurs by budding of contaminating species, suggesting a transitions from low to high purity when the budding rate reaches the biochemical conversion rate ($k_{\mathrm{b}} \simeq k_{\mathrm{m}}$). A "purity transition" is indeed observed in an intermediate range of biochemical conversion rates ($0.1 < k_{\mathrm{m}} < 10$, see Fig.2C). It is most pronounced for $k_{\mathrm{m}} \simeq 1$. Beyond this range, the system is dominated by one identity and compartments are always pure. The purity variation can be qualitatively reproduced by an analytical model (App.3) that accounts for the competition between biochemical conversion and budding, but also includes fusion between compartments and with the exits (solid lines in Fig.2C.).

In summary, three types of organisation can be found, mostly controlled by the ratio of budding over fusion rate k_{b}: a *mixed regime* at low k_{b}, where compartments are large and contain a mixture of identities, a *vesicular regime* at high k_{b}, where compartments are made of only a few vesicles and are very pure, and a *sorted regime* for intermediate values of k_{b}, where compartments are both rather large, and rather pure. Snapshots of these three steady-states are shown in Fig.2D. The relationship between average size and average purity of compartments is shown in Fig.2E. The existence of a well-defined intermediate regime with large sorted compartments is promoted by our assumption of "saturated budding" where the budding flux of any identity only depends on the compartment size (Eq.2). If the budding flux is linear with the number of patches of a given identity, the purity transition occurs for larger values of the budding rate k_{b}, where compartments are smaller (see App.8, p.41)).

Vesicular Transport

Vesicular exchange between compartments is quantified by following the dynamics over time of passive cargo injected from the ER. Each cargo in a compartment of size n has a probability $1/n$ to join a vesicle budding from this compartment. When this vesicle fuses with another compartment or a boundary, the composition difference $\Delta\phi_i$ between the receiving and emitting compartment (a number between -1 and 1) is recorded for the three i-identities (i =*cis*, *medial*, *trans*). Averaged over all budding/fusion events, this defines the *enrichment vector* $\vec{E}$, whose three components E_i are normalized for readability so that $\sum_i |E_i| = 1$ (see Fig.1A - App.6, p.37 for non-normalized enrichment). Vesicular flux is predominantly anterograde if $E_{cis} < 0$ and $E_{trans} > 0$, while it is predominantly retrograde if $E_{cis} > 0$ and $E_{trans} < 0$.

We focus on systems containing comparable amounts of each species at steady-state ($k_{\mathrm{m}} = 1$ - see Eq.6). The enrichment in *cis*, *medial* and *trans*-identities are shown in Fig.3B as a function of the budding rate k_{b}. The most striking result is the high correlation between purity of the compartment and the directionality of vesicular transport. For low values of the budding rate ($k_{\mathrm{b}} \ll 1$), compartments are well mixed (purity $\lesssim 1/2$) and the vesicular flux is retrograde, with a gain in *cis*-identities and a loss in *trans*-identities. For high values of the budding rate ($k_{\mathrm{b}} \gg 1$), compartments are pure (purity $\simeq 1$) and the vesicular flux is anterograde, with a gain in *trans*-identities and a loss in *cis*-identities. The archetypal behaviors most often discussed in the literature are thus, within the limits of our model, asymptotic regimes for extreme values of the ratio of budding to fusion rates. Remarkably, the cross over between these two asymptotic behaviors is rather broad ($k_{\mathrm{b}} \sim 0.1 - 1$) and displays a more complex vesicular transport dynamics, mostly oriented toward medial compartments.

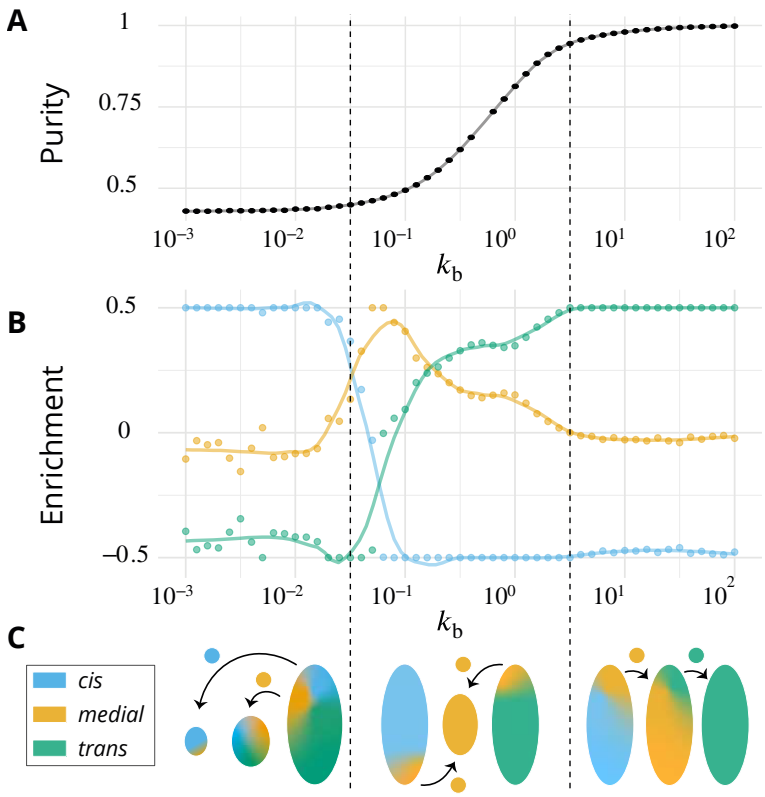


Figure 3. Directionality of vesicular transport, displayed varying the budding rate k_b for a biochemical conversion rate $k_m = 1$ (equal amount of all species in the system). (A) System purity as a function of the budding rate – a cut through the purity phase diagram shown Fig. 2 C. (B) Normalized enrichment in *cis* (blue), *medial* (orange) and *trans* (green) identities between the acceptor and donor compartments during vesicular exchange (solid lines are guides for the eyes). See App.6, p.37 for non-normalized fluxes. (C) Sketches showing the direction of the dominant vesicular fluxes – consistent with data of A and B and data from Fig.1 App.6 – omitting less contributive transports. Before the purity transition (low values of k_b - purity ~ 0.5) the vesicular flux is retrograde. After the transition (high values of k_b - purity ~ 1) the vesicular flux is anterograde. Around the crossover ($k_b \sim 0.1$), the vesicular flux is centripetal and oriented toward *medial*-compartments. The centripetal flux disappears if inter-compartments fusion is prohibited, see App.8, p.41.

In our model organelle, vesicle budding and fusion are biased solely by local compositions, which is subjected to irreversible biochemical conversions. The interplay between these microscopic processes gives rise to an irreversible flux of vesicles across the system. To better understand the directionality of vesicular exchanges, the net vesicular flux leaving compartments with a given composition is represented as a vector field in the compartments composition space in Fig.4. This is shown as a triangle plots in which each point corresponds to a fraction of *cis*, *medial* and *trans* components in a compartment, and the three vertices of the triangle correspond to pure compartments (see App.2). The variation of the vesicular flux with the budding rate k_b (Fig.4) is consistent with that of the composition enrichment in the different identities (Fig.3B). As k_b increases, the flux evolves from being mostly retrograde at low k_b (from *trans*-rich toward less mature compartments) to being anterograde at high k_b (from *cis* to *medial*, and *medial* to *trans*-compartments). In between, the vesicular flux is centripetal toward mixed compartments, leading to a net enrichment in *medial*-identity. The relationship between structure and transport is explained below, and a dissection of this fluxes with respect to the composition of the vesicles is shown in App.6. The net flux leaving a particular region of the triangular phase space is proportional to the total system's size (number of compartments times their size) with this composition, while the flux arriving at a particular region depends on the number of compartments with that composition. Both quantities are shown on Fig.4.

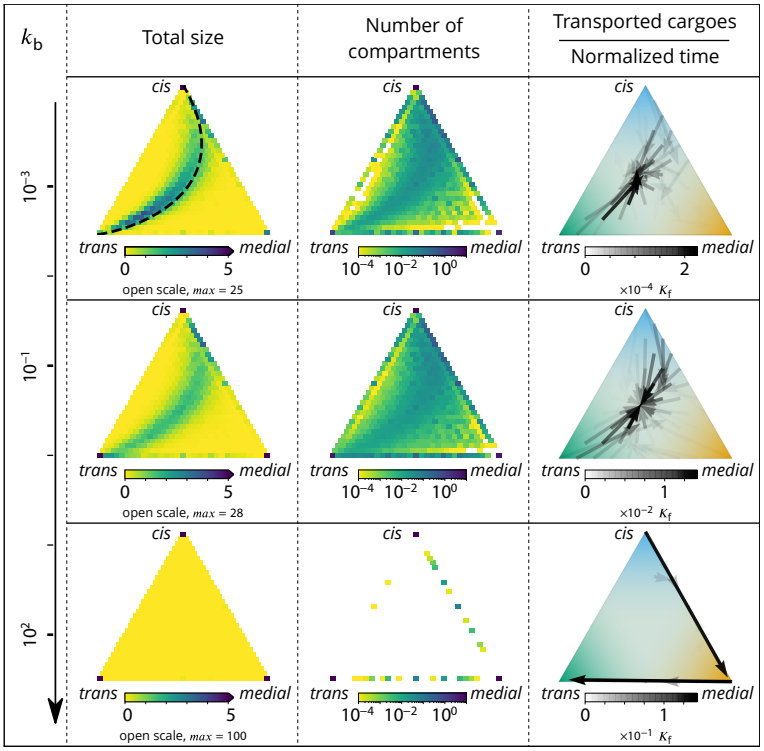


Figure 4. Relationship between the structure of the system and the vesicular fluxes. Distribution of the total size of the system (truncated scale, maximum value in brackets) and number of compartments, and the vesicular flux between them, shown in a triangular composition space. Each point represents a given composition of compartments and the triangle vertices are compartments of pure *cis*, *medial* and *trans* identity. Arrows show the vesicular flux. The base of each arrow is the composition of the donor compartment and the tip is the average composition of receiving compartments (ignoring back fusion). The opacity of the arrows is proportional to the flux of transported cargo going through this path per unit time ($1/K_f$), normalized by the total number of cargo-proteins. The dashed line on the top-left triangle is the theoretical prediction given by Eq.9. See also App.6 p.36 for further characterizations of the vesicular transport, and App.8 p.41 to discuss the impact of other budding and fusion implementations on the vesicular fluxes.

At low budding rate ($k_b \ll 1$ - top row in Fig.4), the system is in the mixed regime, where compartments are large (hundreds of vesicles in size) but mixed (purity $\lesssim 1/2$). The system is dominated by compartment maturation, and the compartment distribution is spread around a path going from purely *cis* to purely *trans*-compartments, consistent with a dynamics where each compartment maturates independently from the other and given by Eq.9. - see top left panel of Fig.4. The majority of the transport vesicles are emitted by *trans*-rich compartments, which concentrate most of the components of the system, leading to a retrograde vesicular flux. Note that although vesicular flux is clearly retrograde in the maturation-dominated regime, as seen from the increase in *cis* identity and the decrease in *trans* identity between emitting and receiving compartment (Fig.3B), there is no net flux of vesicles from *medial*-rich to *cis* compartments in this limit (Fig.4). This differs from the classical "cisternal progression & maturation" mechanism sketched in Fig.1A, and is due to the fact that the compartments that contain the highest fraction of *medial* identity in this regime are well mixed ($\phi_{cis} \simeq \phi_{medial} \simeq \phi_{trans} \simeq 1/3$). The vesicular flux leaving such compartments is equally split into *cis, medial* and *trans* vesicles, which fuse homotypically with compartments enriched in their identity, yielding a vanishing net flux averaged over the three identities. There is nevertheless a strong retrograde flux of *cis* vesicles toward *cis* compartments, as shown in App.6 Fig.1. In the classical "cisternal maturation" mechanism, retrograde transport is needed to recycle *cis*-Golgi components from more mature compartments. Within the present model, this can be achieved by targeting such components to cis vesicles budding from more

mature compartments.

At high budding rate ($k_b \gg 1$ - bottom row in Fig.4), the system is in the vesicular regime, where compartments are fairly pure, but rather small, with a size equivalent to a few vesicles. Compartments are exclusively distributed along the *cis/medial* and *medial/trans*-axes of composition, and accumulate at the vertices of the triangle. Membrane patches that just underwent conversion bud quickly from the donor compartment and fuse with pure compartments with the appropriate *medial* or *trans*-identity, explaining the clear anterograde vesicular transport. This feature is further reinforced under very high budding rate, where the vesicular flux is dominated by budding vesicles undergoing biochemical conversion *before* fusion, which prohibits their back fusion with the donor compartment.

In the intermediate regime (middle row in Fig.4), the distribution of mixed compartments is more homogeneous along the line describing the maturation of individual compartments - both in terms of size and number of compartments. *Cis*-rich and *trans*-rich compartments emit comparable amount of transport vesicles, while large *medial*-rich compartments are absent. This leads to a centripetal vesicular flux with an enrichment in *medial* identity (see Fig.3 B). The lack of large *medial*-rich compartment is due to the fact that they can be contaminated by either *cis* or *trans*-species and fuse with *cis*-rich or *trans*-rich compartments. If inter-compartment fusion is prohibited, the enrichment in medial-identity in the intermediate regime is abolished, with a direct transition from anterograde to retrograde vesicular transport upon increasing k_b (see App.8 and the Discussion section).

A graphical sketch consistent with our numerical observations for the typical size and composition of different compartments and for the dominant vesicular fluxes between them in the different regimes is shown in Fig.3C.

Discussion

The genesis and maintenance of complex membrane-bound organelles such as the Golgi apparatus rely on self-organisation principles that are yet to be fully understood. Here we present a minimal model in which steady-state Golgi-like structures spontaneously emerge from the interplay between three basic mechanisms: the biochemical conversion of membrane components, the composition-dependent vesicular budding, and inter-compartment fusion. As our model does not include space, we do not refer to the spatial structure of a stacked Golgi, but rather to the self-organisation of Golgi components into compartments of distinct identities undergoing vesicular exchange. We show that directional vesicular exchanges between compartments of different identities spontaneously emerge from local biochemical interactions, despite the absence of spatial information. This is in line with recent observations that the Golgi functionality is preserved under major perturbation of its spatial structure resulting from land-locking Golgi cisternae to mitochondria (***Dunlop et al., 2017***). Indeed, cargo transport still proceed, and is merely slowed down by a factor of two (from 20 min. to 40 min.) in land-locked Golgi despite the absence of spatial proximity between cisternae. This suggests that spatial information is less crucial to the Golgi function than biochemical information. Our main result is that the steady-state organisation of the organelle, in terms of size and composition of its compartments, is intimately linked to the directionality of vesicular exchange between compartments through the mechanism of homotypic fusion. Both aspects are controlled by a kinetic competition between vesicular exchange and biochemical conversion.

The results presented in Fig.2 can be used to explain a number of experimental observations. The predicted role of the budding and fusion rates is in agreement with the phenotype observed upon deletion of Arf1 (a protein involved in vesicle budding), which decreases the number of compartments and increases their size, particularly that of *trans*-compartments which seem to aggregate into one major cisterna (***Bhave et al., 2014***). A comparable phenotype has been observed upon mutation of NSF (a fusion protein), which produces extremely large, but transient, *trans*-compartments (***Tanabashi et al., 2018***), thereby increasing the stochasticity of the system. This is consistent with an increase of the fusion rate according to our model (see App.5, p.34 for a

quantification of the fluctuations in our model). Modifying the expression of *VPS74* (yeast homolog of *GOLPH3*) leads to an altered Golgi organisation, comparable to the *ARF1* deletion phenotype (***Iyer et al., 2018***). Both Δ*arf1* and Δ*vps74* present an enlargement of the Golgi cisternae and a disruption of molecular gradients in the system, which we interpret, in the limits of our system, as a lower purity due to slower sorting kinetics following an impairment of the budding dynamics. Importantly, the correlation between the size and purity of the compartments predicted by our model is in agreement with the observation that decreasing the budding rate by altering the activity of COPI (a budding protein) leads to larger and less sorted compartments (***Papanikou et al., 2015***). In yeast again, over-expression of Ypt1 (a Rab protein) increases the transition rate from early to transitional Golgi, and increases the co-localization of early and late Golgi markers (***Kim et al., 2016***). We interpret this as a decrease of the purity of Golgi cisternae upon increasing the biochemical conversion rate by unbalancing the ratio of budding to conversion rates. This suggests that the wild-type Yeast Golgi is closed to the line $k_{\mathrm{m}} = k_{\mathrm{b}}$ shown in Fig.2.

One crucial model prediction is that the size and purity of Golgi compartments should be affected in a correlated fashion by physiological perturbations. This could be tested experimentally by further exploring the phase diagrams of Fig.2B-C by simultaneously varying the ratios of biochemical conversion over fusion rate k_{m} and of budding over fusion rate k_{b}. Compartment purity can be altered without changing their size by acting on k_{m}, while it should be correlated with a change of size by acting on k_{b}. We thus predict that the purity decrease concomitant with the size increase observed in ***Papanikou et al.*** (***2015***) upon impairing COP-I activity (decreasing k_{b}) could be reversed *without a change of cisterna size* upon decreasing the biochemical conversion rate k_{m}, *e.g.* by impairing Ypt1 activity (as done in ***Kim et al.*** (***2016***)). Along the same line, we predict that the decrease of purity observed by ***Kim et al.*** (***2016***) upon Ypt1 over-expression (which increases the early to transitional conversion rate) should not be associated to change of cisternae size (at steady-state) if altering Ypt1 does not modify the budding rate k_{b}. We can predict further that the loss of purity phenotype could be reversed upon increasing the budding rate by over-expressing COP-I, but that this would be accompanied by an decrease of cisternae size. This experimental proposal is represented graphically on the size and purity phase diagrams in App.9 Fig.1, with somewhat arbitrary arrows, to give a feel for the way our model may be used to design experimental strategies.

The directionality of vesicular transport is intimately linked to the steady-state organisation of the organelle; vesicular transport is anterograde when compartments are pure, and it is retrograde when they are mixed. This can be intuitively understood with the notion of "contaminating species". If a compartment enriched in a particular molecular identity emits a vesicle of that identity, the vesicle will likely fuse back with the emitting compartment by homotypic fusion, yielding no vesicular exchange. One the other hand, an emitted vesicle that contains a minority (contaminating) identity will homotypically fuse with another compartment. If budding is faster than biochemical conversion, the contaminating species is the one that just underwent conversion, and its budding and fusion with more mature compartment leads to an anterograde transport. In such systems, compartments are rather small and pure. If biochemical conversion is faster than budding, the contaminating species is the less mature one, leading to a retrograde transport. Such systems are rather mixed with large compartments. Thus, the directionality of vesicular exchange is an emergent property intimately linked to the purity of the compartments. The *medial*-rich compartments are special in that regard: for intermediate values of the purity ($k_{\mathrm{b}} \sim k_{\mathrm{m}}$), they may be contaminated both by yet to be matured *cis*-species and already matured *trans*-species and can fuse both with *cis*-rich and *trans*-rich compartments. If inter-compartment fusion is allowed, large *medial*-rich compartments are relatively scarce, and emit few vesicles. On the other hand, *medial*-vesicles emitted by *cis/medial* and *medial/trans*-compartments may fuse together to form (small) *medial*-rich compartments. For intermediate values of the budding rates, this vesicular flux dominates vesicular exchange and leads to a centripetal vesicular flux towards *medial*-compartments (details in App.6, p.37). If inter-compartment fusion is prohibited, which for instance corresponds to a situation where Golgi cisternae are immobilized (***Dunlop et al., 2017***), *medial*-rich compartments are present at

steady-state, the centripetal vesicular flux is less intense, and the purity transition is accompanied with a direct transition from retrograde to anterograde vesicular flux (App.8, p.41). Note that vesicular transport of cargo through the Golgi is not bound to follow the net vesicular flux between compartments discussed above. Indeed, the net flux is the average of the flux of vesicles of the three identities. A given cargo follows this flux - on average - if it does not interact preferentially with membrane of particular identity. On the other hand, a cargo that is preferentially packaged into vesicles of *trans* identity (for example) will be transported toward *trans*-rich compartment even if the net vesicular flux is mostly retrograde.

Regardless of the details of the model, we find that systems showing a well defined polarity with well sorted cisternae exhibit anterograde vesicular fluxes, whereas systems with mixed compartments exhibit retrograde fluxes. The former dynamics is expected when biochemical conversion is the slowest kinetic process and compartments are long-lived, while the latter is expected when vesicular exchange is slow and the system is composed of transient compartments undergoing individual maturation. One can relate this prediction to the difference in organisation and dynamics between the Golgi of S. cerevisiae and the more organized Golgi of higher organisms such as vertebrates. Maturation of Golgi cisternae has been directly observed in S. cerevisiae (***Losev et al., 2006***), with colocalization of different identity markers within single cisternae (low purity) during maturation, whereas vesicular transport phenotypes have been indirectly observed (***Dunlop et al., 2017***) or inferred through modeling (***Dmitrieff et al., 2013***) in mammalian cells. Consistent with our predictions, Golgi dynamics is one to two orders of magnitude faster in S. cerevisiae Golgi cisternae (as measured by the typical maturation rate of cisternae $\sim 1/\text{min.}$ (***Losev et al., 2006***; ***Matsuura-Tokita et al., 2006***)) than in mammalian cells (as measured by the typical Golgi exit rate of cargo $\sim 1/20$ min. to $\sim 1/40$ min. (***Bonfanti et al., 1998***; ***Hirschberg et al., 1998***; ***Patterson et al., 2008***; ***Boncompain et al., 2012***)). At this point, one may speculate a link between structure and function through kinetics. A well sorted and polarized Golgi is presumably required to accurately process complex cargo. The glycosylation of secreted cargo is key to the interaction of a vertebrate cell with the immune system of the organism it belongs to (***Ryan and Cobb, 2012***), and glycans appear to be more diverse in higher eukaryotes - which also possess a highly organized Golgi - than in unicellular eukaryotes like yeast (***Wang et al., 2017***). The Golgi organisation in S. cerevisiae could thus be the result of an adaptation that has favored fast transport over robustness of processing, leading to a less organized Golgi characterized by cisternal maturation and retrograde vesicular transport. Remarkably, the slowing down of Golgi transport when S. cerevisiae is starved in a glucose-free environment coincides with the Golgi becoming more polarized (***Levi et al., 2010***), strengthening the proposal derived from our model of a strong connection between transport kinetics and steady-state Golgi structure.

The present model is designed to account for crucial dynamical processes at play in the Golgi with a limited number of parameters. It is sufficient to yield robust self-organised structures which can reproduce a number of experimental observations, but cannot be expected to account for the full richness of Golgi dynamics, especially under severe perturbation. The model includes a single kind of exchange mechanism based on vesicles. Exchange could also proceed via the pinching and fusion of larger cisterna fragments (***Pfeffer, 2010***; ***Dmitrieff et al., 2013***), or via tubular connection between compartments (***Trucco et al., 2004***), although the relevance of the latter to Golgi transport is still debated (***Glick and Luini, 2011***). Exchange mechanisms that do not establish stable connections able to relax concentration gradient between compartments by diffusion could be included as extension of our vesicle exchange mechanism (at the expense of introducing additional parameters to characterise their composition and size-dependent nucleation, scission and fusion rates). Stable Golgi tubulation without tubule detachment is observed under brefeldin A treatment, which prevent the membrane association of COPI coat. This eventually leads to the disappearance of the Golgi by quick resorption into the ER whenever a tubular connection is established between them, in a way suggesting a purely physical (tension-driven) mechanism (***Sciaky et al., 1997***). According to our model, preventing COPI vesicle budding should lead to larger

cisternae, as discussed above. The brefeldin A phenotype can be reconciled with our framework by invoking the fact that larger cisternae should have a higher surface to volume ratio and hence a lower membrane tension, decreasing the force required to nucleate membrane tubules and increasing the probability of direct tubular connection with the ER. Such tension effects are clearly beyond the scope of the present model, but this phenotype could be qualitatively reproduced by invoking an increase of the ER fusion parameter α_{ER} with the size of the compartment due to membrane mechanics considerations.

The model does not account for the possibility of lateral partitioning between processing and exporting domains within compartments which could play an important role in regulating cisterna composition and cargo processing (***Patterson et al., 2008***). A detailed kinetic model of organisation based on this concept and supported by extensive quantitative live cell imaging experiments is proposed in (***Patterson et al., 2008***) to account for the distinct composition of the different Golgi cisterna and the kinetics of cargo transport. This rapid partitioning model (RPM) is very different in scope with the model we propose here, since it focuses on the transport of lipids and cargo molecules in a pre-established Golgi apparatus with a prescribed structure, while the present model describes de-novo Golgi formation and maintenance, through self-organised fusion and scission mechanisms between dynamic compartments. The two models share important similarities, such as the fact that no directionality is assumed a priori for intra-cisternal vesicular transport, and that export is allowed from every cisterna, which is crucial to reproduce the exponential kinetics observed in experiments. The RPM is much more precise in terms of the microscopic description of the biochemical processes, and consequently involves many parameters. Combining our simple model of Golgi self-organisation with the detailed description given by the RPM for the kinetics of lipid, enzymes and cargo within the Golgi is the logical next step to push forward our understanding of Golgi self-organisation.

In summary, we have analyzed a model of self-organisation and transport in cellular organelles based on a limited number of kinetic steps allowing for the generation and biochemical conversion of compartments, and vesicular exchange between them. Although we kept the complexity of individual steps to a minimum, our model gives rise to a rich diversity of phenotypes depending on the parameter values, and reproduces the effect of a number of specific protein mutations. We identify the concomitance of a structural transition (from large and mixed to small and pure sub-compartments) and a dynamical transition (from retrograde to anterograde vesicular exchange). Within our model, anterograde vesicular transport is accompanied by many spurious events of vesicle back-fusion. For very high budding rates, compartments are very pure and anterograde transport is dominated by vesicles undergoing biochemical conversion *after* budding from a compartment. In our model, such vesicle biochemical conversion events, which have been described in the secretory pathway as a way to direct vesicular traffic and to prevent vesicles back-fusion (***Lord et al., 2011***), only occur when compartments are very small (high budding rate). Adding composition-dependent feedback involving specific molecular actors to tune budding and fusion rates - for instance to reduce or prevent the budding of the majority species - will displace the purity transition to lower budding rates and thus extend the regime of anterograde transport to larger sub-compartment steady-state sizes. Cisternal stabilizers like GRASP or vesicles careers like Golgins, known to already be present in the ancestor of eukaryotes (***Barlow et al., 2018***), are obvious candidates. Although more complex models, and in particular the inclusion of spatial dependencies, are surely relevant to dynamics of the organelle, the fundamental relationship between kinetics, structure and transport highlighted by our model is a universal feature of the interplay between biochemical conversion and vesicular exchange in cellular organelles.

Acknowledgments

We acknowledge stimulating discussions with Bruno Goud, Cathy Jackson, Grégory Lavieu, Franck Perez and Stéphanie Miserey-Lenkei, and critical reading of the manuscript by Matthew S. Turner.

Competing interests

The authors declare no conflict of interest.

Nomenclature

Here is the nomenclature used in the main text and Appendix sections.

Table 1. System's parameters

K_{b}	Budding rate per patches of membrane in a compartment
K_{m}	biochemical conversion rate for each patch of membrane (*cis* → *medial* → *trans*)
K_{f}	Maximum fusion rate between two compartments
J	Injection rate of *cis*-vesicles from the ER
k_{b}	Budding rate normalized by the fusion rate ($K_{\mathrm{b}}/K_{\mathrm{f}}$)
k_{m}	biochemical conversion rate normalized by the fusion rate ($K_{\mathrm{m}}/K_{\mathrm{f}}$)
j	Injection rate normalized by the fusion rate (J/K_{f})
α	Fraction of active species (driving homotypic fusion) in the ER (*cis*) or the TGN (*trans*)
α_i	When different between boundaries, α in the boundary i (i equals ER or TGN)

Table 2. Compartments' description

ϕ_i	Fraction of a given species i (*cis*, *medial* or *trans*) in a compartment
$\vec{C}$	Composition of a compartment, whose components are the fractions ϕ_i
n	Size of a compartment, (*i.e.* number of membrane patches in the compartment)
n_i	Number of membrane patches of the i identity, in a compartment
n_{id}	Number of different identities (from 1 to 3) in a single a compartment
n_{v}	Number of vesicles surrounding a compartment that share its composition
P	Purity of a compartment
$J_{\mathrm{b},i}$	Budding flux of vesicles decorated by the identity i, from a compartment

Table 3. Steady-state description

c_n	Size distribution of compartments in the system
n_c	Cutoff in the size distribution, namely the typical size of big compartments
N	Total number of membrane patches, in all compartments, in the steady-state system
N_i	Total number of patches of the i species (*cis*, *medial* or *trans*)

References

Antonin W, Holroyd C, Tikkanen R, Höning S, Jahn R. The R-SNARE endobrevin/VAMP-8 mediates homotypic fusion of early endosomes and late endosomes. Molecular Biology of the Cell. 2000 Oct; 11(10):3289–3298.

Barlow LD, Nývltová E, Aguilar M, Tachezy J, Dacks JB. A sophisticated, differentiated Golgi in the ancestor of eukaryotes. BMC Biology. 2018 Mar; 16:27. https://doi.org/10.1186/s12915-018-0492-9, doi: 10.1186/s12915-018-0492-9.

Bekier MEn, Wang L, Li J, Huang H, Tang D, Zhang X, Wang Y. Knockout of the Golgi stacking proteins GRASP55 and GRASP65 impairs Golgi structure and function. Molecular biology of the cell. 2017 10; 28(21):2833–2842. https://www.ncbi.nlm.nih.gov/pubmed/28814501, doi: 10.1091/mbc.E17-02-0112.

Bhave M, Papanikou E, Iyer P, Pandya K, Jain BK, Ganguly A, Sharma C, Pawar K, Austin J, Day KJ, Rossanese OW, Glick BS, Bhattacharyya D. Golgi enlargement in Arf-depleted yeast cells is due to altered dynamics of cisternal maturation. Journal of Cell Science. 2014 Jan; 127(Pt 1):250–257. doi: 10.1242/jcs.140996.

Binder B, Goede A, Berndt N, Holzhütter HG. A conceptual mathematical model of the dynamic self-organisation of distinct cellular organelles. PloS One. 2009 Dec; 4(12):e8295. doi: 10.1371/journal.pone.0008295.

Boncompain G, Divoux S, Gareil N, de Forges H, Lescure A, Latreche L, Mercanti V, Jollivet F, Raposo G, Perez F. Synchronization of secretory protein traffic in populations of cells. Nature Methods. 2012 Mar; 9(5):493–498. doi: 10.1038/nmeth.1928.

Boncompain G, Perez F. The many routes of Golgi-dependent trafficking. Histochemistry and Cell Biology. 2013 Sep; 140(3):251–260. doi: 10.1007/s00418-013-1124-7.

Bonfanti L, Mironov AA, Martínez-Menárguez JA, Martella O, Fusella A, Baldassarre M, Buccione R, Geuze HJ, Mironov AA, Luini A. Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. Cell. 1998 Dec; 95(7):993–1003.

Bonifacino JS, Glick BS. The mechanisms of vesicle budding and fusion. Cell. 2004 Jan; 116(2):153–166.

Cai H, Reinisch K, Ferro-Novick S. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. Developmental Cell. 2007 May; 12(5):671–682. doi: 10.1016/j.devcel.2007.04.005.

Day KJ, Staehelin LA, Glick BS. A Three-Stage Model of Golgi Structure and Function. Histochemistry and cell biology. 2013 Sep; 140(3):239–249. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3779436/, doi: 10.1007/s00418-013-1128-3.

Dmitrieff S, Sens P. Cooperative protein transport in cellular organelles. Physical Review E, Statistical, Nonlinear, and Soft Matter Physics. 2011 Apr; 83(4 Pt 1):041923. doi: 10.1103/PhysRevE.83.041923.

Dmitrieff S, Rao M, Sens P. Quantitative analysis of intra-Golgi transport shows intercisternal exchange for all cargo. Proceedings of the National Academy of Sciences of the United States of America. 2013 Sep; 110(39):15692–15697. doi: 10.1073/pnas.1303358110.

Dunlop MH, Ernst AM, Schroeder LK, Toomre DK, Lavieu G, Rothman JE. Land-locked mammalian Golgi reveals cargo transport between stable cisternae. Nature Communications. 2017; 8(1):432. doi: 10.1038/s41467-017-00570-z.

Gillespie DT. Exact stochastic simulation of coupled chemical reactions. The Journal of Physical Chemistry. 1977 Dec; 81(25):2340–2361. https://doi.org/10.1021/j100540a008, doi: 10.1021/j100540a008.

Glick BS, Luini A. Models for Golgi Traffic: A Critical Assessment. Cold Spring Harbor Perspectives in Biology. 2011 Nov; 3(11). http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3220355/, doi: 10.1101/cshperspect.a005215.

Gong H, Guo Y, Linstedt A, Schwartz R. Discrete, continuous, and stochastic models of protein sorting in the Golgi apparatus. Physical Review E, Statistical, Nonlinear, and Soft Matter Physics. 2010 Jan; 81(1 Pt 1):011914. doi: 10.1103/PhysRevE.81.011914.

Grosshans BL, Ortiz D, Novick P. Rabs and their effectors: Achieving specificity in membrane traffic. Proceedings of the National Academy of Sciences. 2006 Aug; 103(32):11821–11827. http://www.pnas.org/content/103/32/11821, doi: 10.1073/pnas.0601617103.

Heald R, Cohen-Fix O. Morphology and function of membrane-bound organelles. Current Opinion in Cell Biology. 2014 Feb; 26:79–86. doi: 10.1016/j.ceb.2013.10.006.

Hirschberg K, Miller CM, Ellenberg J, Presley JF, Siggia ED, Phair RD, Lippincott-Schwartz J. Kinetic Analysis of Secretory Protein Traffic and Characterization of Golgi to Plasma Membrane Transport Intermediates in Living Cells. The Journal of Cell Biology. 1998 Dec; 143(6):1485–1503. http://jcb.rupress.org/content/143/6/1485, doi: 10.1083/jcb.143.6.1485.

Hu W, Xu R, Zhang G, Jin J, Szulc ZM, Bielawski J, Hannun YA, Obeid LM, Mao C. Golgi Fragmentation Is Associated with Ceramide-induced Cellular Effects. Mol Biol Cell. 2005; 16:1555–1567.

Ispolatov I, Müsch A. A model for the self-organization of vesicular flux and protein distributions in the Golgi apparatus. PLoS computational biology. 2013; 9(7):e1003125. doi: 10.1371/journal.pcbi.1003125.

Iyer P, Bhave M, Jain BK, RoyChowdhury S, Bhattacharyya D. Vps74p controls Golgi size in an Arf1-dependent manner. FEBS letters. 2018 Oct; doi: 10.1002/1873-3468.13266.

Kelly RB. Pathways of protein secretion in eukaryotes. Science (New York, NY). 1985 Oct; 230(4721):25–32.

Kim JJ, Lipatova Z, Majumdar U, Segev N. Regulation of Golgi Cisternal Progression by Ypt/Rab GTPases. Developmental Cell. 2016 Feb; 36(4):440–452. doi: 10.1016/j.devcel.2016.01.016.

Klemm RW, Ejsing CS, Surma MA, Kaiser HJ, Gerl MJ, Sampaio JL, de Robillard Q, Ferguson C, Proszynski TJ, Shevchenko A, Simons K. Segregation of sphingolipids and sterols during formation of secretory vesicles at the trans-Golgi network. The Journal of Cell Biology. 2009 May; 185(4):601–612. doi: 10.1083/jcb.200901145.

Klute MJ, Melançon P, Dacks JB. Evolution and diversity of the Golgi. Cold Spring Harbor Perspectives in Biology. 2011 Aug; 3(8):a007849. doi: 10.1101/cshperspect.a007849.

Kühnle J, Shillcock J, Mouritsen OG, Weiss M. A modeling approach to the self-assembly of the Golgi apparatus. Biophysical Journal. 2010 Jun; 98(12):2839–2847. doi: 10.1016/j.bpj.2010.03.035.

Lavieu G, Zheng H, Rothman JE. Stapled Golgi cisternae remain in place as cargo passes through the stack. eLIFE. 2013; 2:e00558.

Levi SK, Bhattacharyya D, Strack RL, Austin JR, Glick BS. The Yeast GRASP Grh1 Colocalizes with COPII and Is Dispensable for Organizing the Secretory Pathway. Traffic (Copenhagen, Denmark). 2010 Sep; 11(9):1168–1179. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2919637/, doi: 10.1111/j.1600-0854.2010.01089.x.

Lippincott-Schwartz J, Roberts TH, Hirschberg K. Secretory protein trafficking and organelle dynamics in living cells. Annual Review of Cell and Developmental Biology. 2000; 16:557–589. doi: 10.1146/annurev.cellbio.16.1.557.

Lord C, Bhandari D, Menon S, Ghassemian M, Nycz D, Hay J, Ghosh P, Ferro-Novick S. Sequential interactions with Sec23 control the direction of vesicle traffic. Nature. 2011 May; 473(7346):181–186. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3093450/, doi: 10.1038/nature09969.

Losev E, Reinke CA, Jellen J, Strongin DE, Bevis BJ, Glick BS. Golgi maturation visualized in living yeast. Nature. 2006 Jun; 441(7096):1002–1006. https://www.nature.com/nature/journal/v441/n7096/full/nature04717.html, doi: 10.1038/nature04717.

Malhotra V, Mayor S. Cell biology: The Golgi grows up. Nature. 2006 Jun; 441(7096):939–940. https://www.nature.com/articles/441939a, doi: 10.1038/441939a.

Marra P, Salvatore L, Mironov A, Di Campli A, Di Tullio G, Trucco A, Beznoussenko G, Mironov A, De Matteis MA. The biogenesis of the Golgi ribbon: the roles of membrane input from the ER and of GM130. Molecular Biology of the Cell. 2007 May; 18(5):1595–1608. doi: 10.1091/mbc.E06-10-0886.

Marsh BJ, Volkmann N, McIntosh JR, Howell KE. Direct continuities between cisternae at different levels of the Golgi complex in glucose-stimulated mouse islet beta cells. Proceedings of the National Academy of Sciences of the United States of America. 2004 Apr; 101(15):5565–5570. doi: 10.1073/pnas.0401242101.

Matsuura-Tokita K, Takeuchi M, Ichihara A, Mikuriya K, Nakano A. Live imaging of yeast Golgi cisternal maturation. Nature. 2006 Jun; 441(7096):1007–1010. doi: 10.1038/nature04737.

McDermott MI, Mousley CJ. Lipid transfer proteins and the tuning of compartmental identity in the Golgi apparatus. Chemistry and Physics of Lipids. 2016; 200:42–61.

Mukherji S, O'Shea EK. Mechanisms of organelle biogenesis govern stochastic fluctuations in organelle abundance. eLife. 2014 Jun; 3:e02678.

Papanikou E, Day KJ, Austin J, Glick BS. COPI selectively drives maturation of the early Golgi. eLife. 2015 Dec; 4. doi: 10.7554/eLife.13232.

Patterson GH, Hirschberg K, Polishchuk RS, Gerlich D, Phair RD, Lippincott-Schwartz J. Transport through the Golgi Apparatus by Rapid Partitioning within a Two-Phase Membrane System. Cell. 2008 Jun; 133(6):1055–1067. http://www.sciencedirect.com/science/article/pii/S0092867408006181, doi: 10.1016/j.cell.2008.04.044.

Pfeffer SR. How the Golgi works: a cisternal progenitor model. Proceedings of the National Academy of Sciences of the United States of America. 2010 Nov; 107(46):19614–19618. doi: 10.1073/pnas.1011016107.

Presley JF, Cole NB, Schroer TA, Hirschberg K, Zaal KJ, Lippincott-Schwartz J. ER-to-Golgi transport visualized in living cells. Nature. 1997 Sep; 389(6646):81–85. doi: 10.1038/38001.

Rink J, Ghigo E, Kalaidzidis Y, Zerial M. Rab conversion as a mechanism of progression from early to late endosomes. Cell. 2005 Sep; 122(5):735–749. doi: 10.1016/j.cell.2005.06.043.

Rizzo R, Parashuraman S, Mirabelli P, Puri C, Lucocq J, Luini A. The dynamics of engineered resident proteins in the mammalian Golgi complex relies on cisternal maturation. J Cell Biol. 2013; 201:1027–1036.

Ryan SO, Cobb BA. Roles for major histocompatibility complex glycosylation in immune function. Seminars in Immunopathology. 2012 May; 34(3):425–441. doi: 10.1007/s00281-012-0309-9.

Sachdeva H, Barma M, Rao M. Nonequilibrium description of de novo biogenesis and transport through Golgi-like cisternae. Scientific Reports. 2016 Dec; 6:38840. doi: 10.1038/srep38840.

Sciaky N, Presley J, Smith C, Zaal KJ, Cole N, Moreira JE, Terasaki M, Siggia E, Lippincott-Schwartz J. Golgi tubule traffic and the effects of brefeldin A visualized in living cells. The Journal of Cell Biology. 1997 Dec; 139(5):1137–1155.

Sens P, Rao M. (Re)modeling the Golgi. Methods in Cell Biology. 2013; 118:299–310. doi: 10.1016/B978-0-12-417164-0.00018-5.

Shen X, Hong MS, Moss J, Vaughan M. BIG1, a brefeldin A-inhibited guanine nucleotide-exchange protein, is required for correct glycosylation and function of integrin β1A-inhibited guanine nucleotide-exchange protein, is required for correct glycosylation and function of integrin β1. P Natl Acad Sci Usa. 2007; 104:1230–1235.

Stanley P. Golgi glycosylation. Cold Spring Harbor Perspectives in Biology. 2011 Apr; 3(4). doi: 10.1101/cshperspect.a005199.

Suda Y, Kurokawa K, Hirata R, Nakano A. Rab GAP cascade regulates dynamics of Ypt6 in the Golgi traffic. Proceedings of the National Academy of Sciences of the United States of America. 2013 Nov; 110(47):18976–18981. doi: 10.1073/pnas.1308627110.

Suda Yasuyuki, Nakano Akihiko. The Yeast Golgi Apparatus. Traffic. 2011 Dec; 13(4):505–510. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0854.2011.01316.x, doi: 10.1111/j.1600-0854.2011.01316.x.

Tachikawa M, Mochizuki A. Golgi apparatus self-organizes into the characteristic shape via postmitotic reassembly dynamics. Proceedings of the National Academy of Sciences of the United States of America. 2017 May; 114(20):5177–5182. doi: 10.1073/pnas.1619264114.

Takamori S, Holt M, Stenius K, Lemke EA, Grønborg M, Riedel D, Urlaub H, Schenck S, Brügger B, Ringler P, Müller SA, Rammner B, Gräter F, Hub JS, De Groot BL, Mieskes G, Moriyama Y, Klingauf J, Grubmüller H, Heuser J, et al. Molecular anatomy of a trafficking organelle. Cell. 2006 Nov; 127(4):831–846. doi: 10.1016/j.cell.2006.10.030.

Tanabashi S, Shoda K, Saito C, Sakamoto T, Kurata T, Uemura T, Nakano A. A missense mutation in the NSF gene causes abnormal Golgi morphology in Arabidopsis thaliana. Cell Structure and Function. 2018 Feb; doi: 10.1247/csf.17026.

Trucco A, Polishchuk RS, Martella O, Pentima AD, Fusella A, Giandomenico DD, Pietro ES, Beznoussenko GV, Polishchuk EV, Baldassarre M, Buccione R, Geerts WJC, Koster AJ, Burger KNJ, Mironov AA, Luini A. Secretory traffic triggers the formation of tubular continuities across Golgi sub-compartments. Nat Cell Biol. 2004; 6(11):1071–1081.

Turner MS, Sens P, Socci ND. Nonequilibrium Raftlike Membrane Domains under Continuous Recycling. Physical Review Letters. 2005 Oct; 95(16):168301. https://link.aps.org/doi/10.1103/PhysRevLett.95.168301, doi: 10.1103/PhysRevLett.95.168301.

Vagne Q, Sens P. Stochastic Model of Maturation and Vesicular Exchange in Cellular Organelles. Biophys J. 2018; 114:947–957.

Vagne Q, Sens P. Stochastic Model of Vesicular Sorting in Cellular Organelles. Physical Review Letters. 2018 Feb; 120(5):058102. https://link.aps.org/doi/10.1103/PhysRevLett.120.058102, doi: 10.1103/PhysRevLett.120.058102.

Vagne Q, Turner MS, Sens P. Sensing Size through Clustering in Non-Equilibrium Membranes and the Control of Membrane-Bound Enzymatic Reactions. PloS One. 2015; 10(12):e0143470. doi: 10.1371/journal.pone.0143470.

Wang P, Wang H, Gai J, Tian X, Zhang X, Lv Y, Jian Y. Evolution of protein *N*-glycosylation process in Golgi apparatus which shapes diversity of protein *N*-glycan structures in plants, animals and fungi. Scientific Reports. 2017 Jan; 7:40301. https://www.nature.com/articles/srep40301, doi: 10.1038/srep40301.

Wei JH, Seemann J. Golgi ribbon disassembly during mitosis, differentiation and disease progression. Current Opinion in Cell Biology. 2017 Aug; 47:43–51. doi: 10.1016/j.ceb.2017.03.008.

Xiang Y, Zhang X, Nix DB, Katoh T, Aoki K, Tiemeyer M, Wang Y. Regulation of protein glycosylation and sorting by the Golgi matrix proteins GRASP55/65. Nature communications. 2013; 4:1659–1659. https://www.ncbi.nlm.nih.gov/pubmed/23552074, doi: 10.1038/ncomms2669.

Yelinek JT, He CY, Warren G. Ultrastructural study of Golgi duplication in Trypanosoma brucei. Traffic (Copenhagen, Denmark). 2009 Mar; 10(3):300–306. doi: 10.1111/j.1600-0854.2008.00873.x.

Appendix 1

DETAILED MODEL

General Overview

The nomenclature we use in this section is presented in the Tables 1, 2 and 3 in p.15. We propose a coarse-grained, discrete model of the Golgi Apparatus. The smallest element is a vesicle, which defines the unit size of the system. Other compartments are assemblies of fused vesicles; their size is an integer corresponding to an equivalent number of vesicles that have fused to build this compartment. Each unit of membrane is of either *cis*, *medial* or *trans*-identity. A vesicle has a unique identity, and a compartment of size n is a assembly of n discrete membrane patches of different identities.

We let the system self-organizes between two boundaries, the ER and the TGN, that are two biological boundaries of the Golgi (***Presley et al., 1997***; ***Klemm et al., 2009***). This theoretical Golgi is self-organized as we do not impose any topological or functional constraints, such as the number of compartments or the directionality of vesicular trafficking. The evolution of the system is only dictated by three mechanisms that are budding, fusion and biochemical conversion (short description in the main text, complete description bellow). In order to keep the model as simple as possible, we neglect all mechanical and spatial dependencies. This means that we do not take into account motion of compartments in space, or mechanical properties of the membranes. Any two compartments have the possibility to fuse together, with a rate that depends on their relative composition (see below), but not on their size. An alternative model where compartments cannot fuse with one another is presented in App.8 on p.42.

Biochemical conversion

In the Golgi, it is known that cargo (***Kelly, 1985***) and cisternae (***Day et al., 2013***) undergo biochemical modifications over time. Changes in membrane identity are driven by maturation cascades of proteins such Rab GTPases (***Suda et al., 2013***). These biochemical conversions can be directly observed using live microscopy, and have been best characterized in the yeast Golgi (***Matsuura-Tokita et al., 2006***). Biochemical conversion of membrane identity is a complex and possibly multi-step process involving different kinds of enzymes. As we are mostly interested in the interplay between biochemical conversion (which mixes different identities within a single compartment) and sorting mechanisms, we coarse-grain the biochemical conversion events into simple, one step, stochastic processes: *cis→medial* and *medial→trans*. To simplify the model, we choose the same rate K_{m} for both conversions. Each patch of membrane (in a vesicle or a compartment, and whatever the surrounding patches are) has the same transition rate.

Fusion

Intracellular transport strongly relies on vesicular trafficking (***Takamori et al., 2006***). This involves at least two steps: the budding of a vesicle from a donor compartment (described in the next section), and its fusion with a receiving compartment. The fusion process itself requires close proximity between the receiving compartment and the vesicle, followed by the pairing of fusion actors resulting in membrane fusion. In the current model, the rate of encounter of two compartments is constant, and equal to K_{f}, while the actual fusion event depends on the biochemical composition of the fusing compartments. *In vivo*, some of the key proteins involved in the fusion process, such as the SNAREs or tethering proteins are known to closely interact with membrane markers like Rab proteins (***Cai et al., 2007***). This is thought to accelerate fusion events between compartments of similar biochemical

identities and decrease it between compartments of different identities, a process often called homotypic fusion (***Marra et al., 2007***). To take this into account, fusion is modulated by the probability that both compartments exhibit the same identity at the contact site, corresponding to Eq.1 in the main text.

The fusion rate is maximum (K_f) for two identical compartments, and vanishes between two compartments with no common identity. The exchange of vesicles between compartments of different identities (***Glick and Luini, 2011***), or the fusion of compartments of different identities (***Marsh et al., 2004***), are sometimes regarded as heterotypic fusion. Such processes are allowed within our homotypic fusion model, provided the different compartments share at least some patches of similar identities.

Budding

Budding is the formation of a new vesicles from a large compartment. We assume that each budding vesicle is composed of a single membrane identity, which is consistent with the high specificity of the *in vivo* budding machinery (***Bonifacino and Glick, 2004***). In our model, the flux of vesicles budding from a compartment depends on the size and composition of the compartment. Compartments smaller than 2 cannot bud a vesicle. Compartments of size 2 can split into two vesicles. For larger compartments, we consider a general budding flux for vesicles of identity i (for i = *cis*, *medial* and *trans*) of the form:

$$J_{\mathrm{b},i} = K_\mathrm{b} \times n \times f(\phi_i) \tag{10}$$

with n the size of the compartment', K_b a constant and $f(\phi_i)$ a function of the composition of the compartment.

One expects the budding flux to follow a Michaelis-Menten kinetics; linear with the number of patches of a given identity ($f(\phi) = \phi$) at small concentration, and saturating to a constant at high concentration (***Vagne and Sens, 2018b***). This scenario is consistent with the fact that budding proteins interact very dynamically with the membrane, attaching and detaching multiple times before budding a vesicle (***Hirschberg et al., 1998***). Previous theoretical works suggest that vesicular sorting of different membrane species is more efficient in the saturated regime (***Dmitrieff and Sens, 2011***). Since one of the main features of the Golgi is to segregate different biochemical species into different compartments, we assume a very high affinity between membrane patches and their budding partners, setting $f(\phi_\mathrm{i}) = 1$ if $\phi_\mathrm{i} > 0$ and $f(\phi_\mathrm{i} = 0) = 0$ (corresponding to Eq.2 of the main text). The budding flux for a given identity thus depends on the total size of the compartment rather than on the amount of that particular identity. This leads to a total budding flux $K_\mathrm{b} \times n \times n_\mathrm{id}$ for a compartment carrying n_id different identities. The results of simulations with this linear budding scheme are discussed in App.8 on p.41.

Exits

In cells, the Golgi is placed between two intra-cellular structures that are the ER and the TGN (***Presley et al., 1997***; ***Klemm et al., 2009***). Fluxes of material leaving the Golgi thus include retrograde fluxes toward the ER, carrying immature components such as *cis* and *medial*-Golgi enzymes or recycling ER enzymes, and anterograde fluxes toward the TGN of mature components, such as processed cargo that properly underwent all their post-translational maturation steps (***Boncompain and Perez, 2013***). The exiting fluxes are accounted for by allowing the different compartments in the system to fuse with the boundaries. All compartments, from vesicles to the largest ones, can fuse homotypically with the ER or the TGN to exit the system. Thus, these boundaries are modelled as stable compartments, containing a fraction α_ER of *cis* components for the ER and α_TGN of *trans* components for the TGN. These

fraction are simply written α when they are assumed to be the same. This allows immature (*cis*) compartments to undergo retrograde exit, and mature (*trans*) components to undergo anterograde exit. In the main text we focus on the case $\alpha = 1$. Lower values of α reduce fusion with the boundaries and increase the residence time of components in the system. This increases the average size of compartments, and increases fluctuations in the system, as large compartments exiting the system induce large fluctuations in the instantaneous size and composition of the system. The impact of α is studied in some details in App.7 p.39.

Injection of components

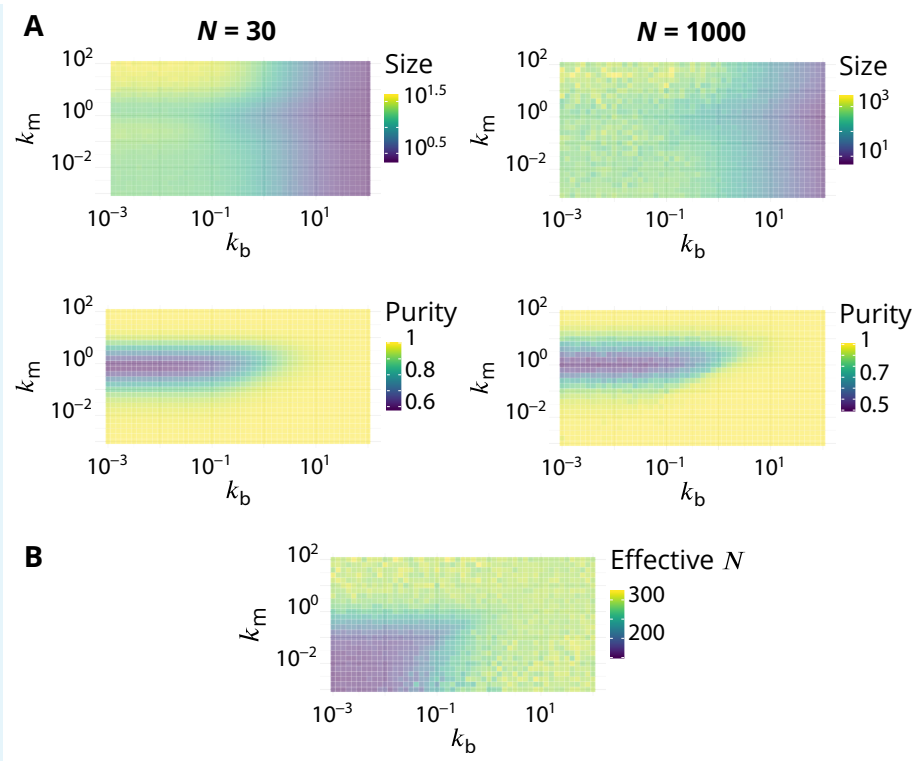
Influx to the Golgi

The Golgi receives material from different compartments like the endosomal network, lyzosomes, etc... (***Boncompain and Perez, 2013***). As we are primarily interested in characterizing the relationship between the structure and dynamics of the Golgi, and in relating these to the rates of individual biochemical conversion and transport processes, we focus here on the secretory role of the Golgi, and only account for the incoming flux of immature components coming from the ER. We define a rate J of injection of *cis*-vesicles in the system. As the parameters of the system are varied, the injection rate is varied as well in order to keep the total system's size to an almost constant value. This constraint is only approximately enforced using Eq.6, which is only strictly valid when all compartments are pure.

The impact of the system's size on the structure of the Golgi is shown on Fig.1A of the current appendix. Increasing the total size increases the number of compartments and hence the total fusion flux between compartments. Consequently, compartments are larger in larger systems, and tend to be (slightly) less pure, as fusion increases mixing. In the main text, we restrict ourselves to a system size of $N = 300$. One should remember that Eq. 6, used to predict the system's size N, assumes that compartments are pure (perfectly sorted). Fig.1B - current appendix, shows that this assumption fails to predict N for systems where the budding rate is low compared to the fusion rate (highly interacting compartments). In this regime, compartments have a great probability to fuse together, creating hybrid *cis/medial/trans*-structures. This makes *medial*-patches sensitive to the interaction with the ER and the TGN, and creates an exit flux for these patches that is not observed in a pure regime. Such systems exhibit a larger exit flux and thus a lower N. Note this is only true for biochemical conversion rates lower than the fusion rate, as the system is saturated with *trans*-patches for high k_{m} (see Eq.4).

Transport of passive cargo-proteins

The direction of vesicular transport is assessed by following the transport of passive cargo injected from the ER as part of incoming *cis*-vesicles. Cargo molecules are passive in the sense that their probability of joining a vesicle is insensitive to the vesicle membrane identity: when a compartment of size n buds a vesicle, each cargo in this compartment has a probability $1/n$ to join the budding vesicle. In the simulation, the number of cargo molecules in the system is kept to a fixed number (typically 20) by injecting a new one each time one leaves the system trough fusion with the boundaries.



Appendix 1 Figure 1. Link between the total size of the Golgi and its steady-state organisation. A. Size dependence of the steady-state organisation of the Golgi (to be compared with Fig.2 in the main text). We look at two different systems that only differ in term of the average total size N, that equals 30 or 1000. Typical size of compartments and average system purity are shown as a function of k_m and k_b. In both cases $\alpha = 1$ (see App.1 p.21). B. Average steady-state total size N, as a function of k_m and k_b, for simulation in which the influx is set to have $N \simeq 300$ (using Eq.6).

Gillespie algorithm and Software implementation

All rates in the system are normalized by the fusion rate: $k_b = K_b/K_f$, $k_m = K_m/K_f$ and $j = J/K_f$. The dynamics of the system can be exactly simulated (to yield the correct trajectories, stochastic noise, etc...) using a Gillespie algorithm (***Gillespie, 1977***), also known as stochastic simulation algorithm. After choosing the initial state of the system (see below), the algorithm is implemented as follows:

1. The rates of all possible events (be it the input of a new vesicle in the system, the fusion between two compartments or between a compartment and one boundary, the budding of a vesicle, or the biochemical conversion of a membrane patch) are calculated. The time interval δt for the next reaction to occur is randomly picked from an exponential distribution with mean $1/R_{TOT}$, where R_{TOT} is the sum of the reaction rates of all possible reactions. The reaction that actually occurs is randomly picked with a probability equal to the rate of this reaction, normalized by R_{TOT}.
2. The state of the system is modified according to the picked reaction. The rates of all events in the modified system are calculated, and the current time of the simulation is incremented by δt.
3. The program comes back to step 1 IF the predefined maximum number of iterations has not been reached, AND there is at least one reaction that has a non-zero rate ($R_{TOT} \neq 0$).

With the Gillespie algorithm, there is no direct dependency between the number of simulation steps and the actual time that is simulated. The physical time between two consecutive

steps decreases with increasing reaction rates. Since we are interested by steady-state quantities, we average over a fixed number of steps rather than a given physical time.

As shown in Appendix 4, the steady-state of the system is reached after a physical time of order $1/K_{\mathrm{f}}$. Depending on the parameters, the steady-state is reached after between 10^3 and 10^6 simulation steps. To characterise the steady-state, we disregards the transient regime and data are recorded after 10^6 steps. The full simulation typically last at least 10^7 steps. This arbitrary amount is a good compromise between computation time and the need to accumulate sufficient amount of data to obtain enough statistics on all the measured quantities. In practice, the time needed to reach steady-state can be shortened by starting the simulation from a vesicular system with the predicted amount of *cis*, *medial* and *trans*-species (with Eq.4).

Because simulation steps have different durations, one should be careful when computing time averages. Two different approaches can be used. Either we weight each configuration by the duration between two consecutive time-steps, or we re-sample data to get a fixed time-step between observations, Δt. We checked that both procedures give the same results, and adopted the second approach for the analysis. As the steady-state of the system is reached after a physical time of order $1/K_{\mathrm{f}}$, we take $\Delta t = 1/K_{\mathrm{f}}$.

Appendix 2

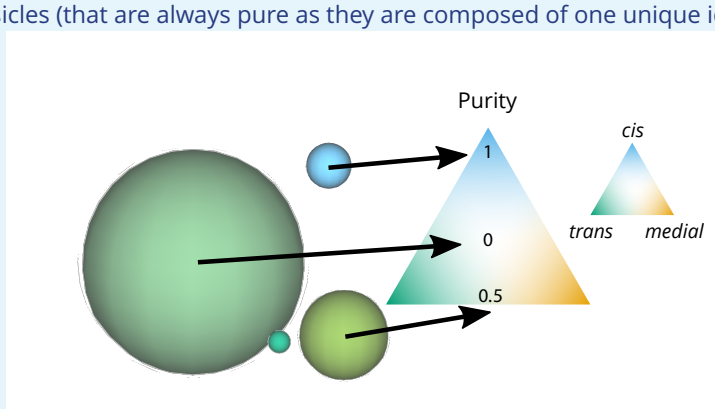
QUANTIFICATION AND STATISTICAL ANALYSIS

Computation of the purity

The purity of a compartment P is defined such that its value is 0 for a perfectly mixed compartment containing the same amount of *cis*, *medial*, and *trans*-species, and it is equal to 1 for a pure compartment containing a single species. With ϕ_i the fraction of the species i in the compartment, the purity P is defined as:

$$P = \sqrt{\frac{3}{2}\sum_i (\phi_i - 1/3)^2} \tag{11}$$

On a triangular composition space where each corner corresponds to a pure compartment, P is the distance from the center of the triangle, see Fig.1 - current appendix. When we show snapshots of the system, each compartment is represented as a sphere whose area is proportional to the compartment size (defined as an equivalent number of vesicles). The composition of a compartment is represented as a color following the color code shown in Fig.1 - current appendix. The purity of the system is a global average (over time and over all compartments) of the purity of individual compartments, weighted by their sizes and ignoring vesicles (that are always pure as they are composed of one unique identity).



Appendix 2 Figure 1. Graphical representation of the size and purity of compartments. Each compartment is represented as a sphere with an area equal to the area of a vesicle (the smallest elements in this snapshot) times the number of vesicles that fused together to create the compartment. The color of the compartment reflects its composition, according to the color code on the triangular composition phase space: 100% *cis* at the top, *medial* at the bottom right, *trans* at the bottom left. The purity of each of these compartments is its distance from a perfectly mixed compartment at the center of the triangle (middle arrow). This distance is normalized so that it equals 1 for a perfectly pure compartment (bottom arrow).

Computation of the typical size

Among other mechanisms, this theoretical organelle self-organizes by budding and fusion of components. In that sense, it is a scission-aggregation structure which should follow the laws dictating the behavior of this class of systems. One of these laws is the fact that the size distribution for small compartments should follow a power-law. Because of the scission, compartments cannot grow indefinitely and the power-law ends by an exponential cutoff (***Turner et al., 2005***; ***Vagne et al., 2015***). Thus, the size distribution c_n of compartments of size n should, on a first approximation, follow this general formulation:

$$c_n \sim \text{Const}\, n^{-\beta} \exp\left(-\frac{n}{n_c}\right) \tag{12}$$

where n_c is the cut-off size and β is an exponent that has been calculated to be $\beta = 3/2$ in a similar system (***Turner et al., 2005***). There are multiple ways to characterize the average size of the distribution, using ratios of moments $k+1$ over k of the distribution:

$$\langle n \rangle_k \equiv \frac{\sum_{n=1}^{\infty} n^{k+1} c_n}{\sum_{n=1}^{\infty} n^k c_n} \tag{13}$$

It turns out that for $n_c \gg 1$, $\sum_{n=1}^{\infty} n^{k-\beta} e^{-n/n_c} \sim 1$ if $1+k-\beta < 0$ and $\sum_{n=1}^{\infty} n^{k-\beta} e^{-n/n_c} \sim n_c^{1+k-\beta}$ if $1+k-\beta > 0$. In order to have $\langle n \rangle_k \sim n_c$, we need to choose the exponent k such that $k > \beta - 1 = 1/2$ in the present case. In the main text, we adopt $k = 1$ and define the characteristic size of the distribution as:

$$\langle n \rangle \equiv \frac{\sum_{n=1}^{\infty} n^2 c_n}{\sum_{n=1}^{\infty} n c_n} \xrightarrow[n_c \gg 1]{} \frac{n_c}{2} \tag{14}$$

As shown in Fig.2A of the main text, the calculated average is in good agreement with simulations' data.

Computation of the mean enrichment

To discriminate between the two models of Golgi's dynamics that we can find in the literature, we need to quantify whether the vesicular transport is anterograde or retrograde. Indeed, the "Vesicular transport" model assumes that cargo is transported sequentially from *cis* to *medial* to *trans*-compartments while resident enzymes remain in place, meaning that the vesicular flux is anterograde. On the other hand, "Cisternal maturation" models assume that cargo remain inside cisternae, and resident enzymes are recycled, thus requiring a retrograde vesicular flux. One way to measure this flux in our simulations is to follow passive cargo molecules and quantify whether they move to more or less mature compartments, as they get carried by vesicles. To do so, we record all events that affect cargo molecules. Every time a compartment buds a cargo, we store the composition of this compartment and compare it to the one in which the vesicle later fuses. Both compositions are a vector $\vec{C}$, with three components that are the fraction ϕ_i (i equals *cis*, *medial* and *trans*) of the donor and acceptor compartments. Defining $\vec{C}_d$ the composition of the donor compartment, and $\vec{C}_a$ the composition of the acceptor, we can compute the enrichment $\vec{E}$ as:

$$\vec{E} = \vec{C}_a - \vec{C}_d \tag{15}$$

The sum of both $\vec{C}$ components equals 1 (as they are fractions of each identities), and the sum of $\vec{E}$ components equals $\vec{0}$, which simply means one cannot gain in fraction of any identity without loosing the same amount of the others. We can now compute $\langle \vec{E} \rangle$ to calculate the mean enrichment in *cis*, *medial* and *trans*-species, between a budding event and the next fusion event.

However, and because of back fusion events (fusion into the same compartment that previously budded the vesicle), $\langle \vec{E} \rangle$ components can be close to 0. This is particularly true for pure, sorted systems. In that case, a budded vesicle has the same identity as its donor compartment, and thus has a great probability to fuse back with the same compartment or one of very similar composition. That is why, non-treated data do not allow to discriminate well between an anterograde and a retrograde regime when the purity (and thus k_b) is high (Fig.1A - App.6, p.37). As we are primarily interested in the sign of these vectors' components, we normalize $\vec{E}$ using the L_1 norm ($\sum_i |E_i| = 1$).

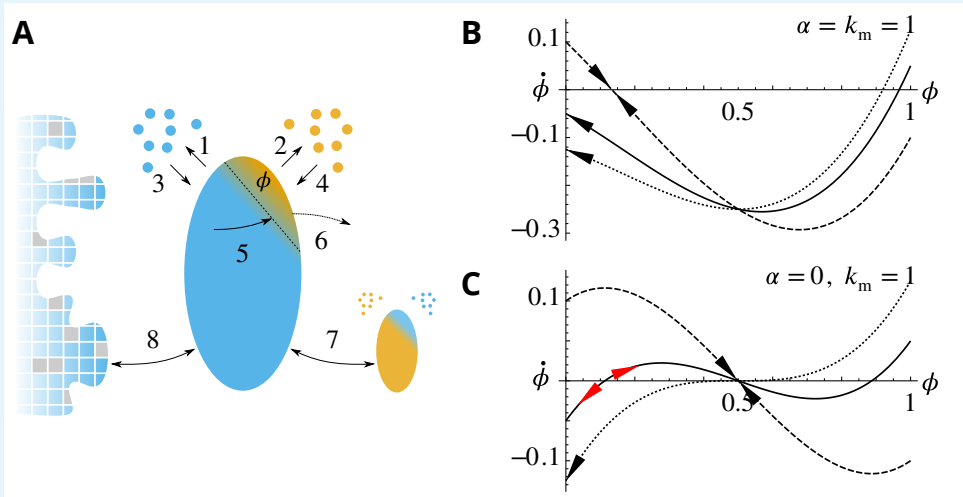
Computation of the vesicular flux vector field

The vesicular flux discussed in the previous section can be directly displayed on the triangle of composition we introduced in Fig.1 - current appendix. Instead of computing $\langle \vec{E} \rangle$ for all $\vec{E}$,

we can bin data with respect to the donor compartment composition (typically triangular bins of size $\Delta\phi = 0.1$) and calculate the average enrichment for each bin of donor compartments. To get rid of back fusion effects in this quantification, we remove the vectors $\vec{E}$ for which all components are smaller (in absolute values) than the binning mesh-grid. For each binned composition we can now compute the mean enrichment vector, and plot this vector on the 2D triangular composition space. To emphasize the dominant fluxes in this vector field, the opacity of vectors is set proportionally to the flux of transported cargo per unit time (normalized by the total amount of cargo-proteins in the system). This can be seen on Fig.4 - main text.

Appendix 3

ANALYTICAL MODEL FOR THE PURITY OF COMPARTMENTS



Appendix 3 Figure 1. Analytical description of the *cis*-compartment purity. Related to Fig.2 - main text. A. Analytical model of the steady-state contamination by a *medial*-species (orange) of a *cis*-compartment (blue). The fraction occupied by the *medial*-species is named ϕ. The eight events ((1)-(8)) impacting ϕ are listed in the text. B-C. Time variation $\dot{\phi}$ as a function of ϕ, with $k_m = 1$ and for different k_b (0.9 dashed line, 1.05 solid line, and $1 + 1/8$ dotted line). ϕ has a stable value when $\partial_\phi \dot{\phi} < 0$ and $\dot{\phi} = 0$ or $\dot{\phi}|_{\phi=0} < 0$ (black arrows). It is potentially unstable when $\partial_\phi \dot{\phi} > 0$ and $\dot{\phi}_{\phi\sim 0} = 0$ (red arrows). B. **For $\alpha = 1$**: $k_b = 0.9$ (dashed) exhibits a stable fixed-point for $\phi > 0$ but with $\phi < 1/2$ as $\alpha > 0$; $k_b = 1.05$ (solid) and $k_b = 1 + 1/8$ (dotted) exhibit a stable concentration for $\phi = 0$. C. **For $\alpha = 0$**: $k_b = 0.9$ (dashed) exhibits a stable fixed-point for $\phi = 1/2$; $k_b = 1.05$ (solid) exhibits an unstable fixed-point near $\phi = 0$; $k_b = 1 + 1/8$ (dotted) does not display this unstable fixed-point as $k_b \geq k_m + 1/8$ (see text).

We want to build a simple model to explain how the purity of the system is affected by the parameters. This is challenging as all events (budding, fusion and biochemical conversion) influence the purity of a compartment. We seek to determine the purity transition, namely the values of k_b and k_m for which a pure system becomes impure.

We make the following approximations, based on the assumption that the system is almost pure: *(i) trans*-compartments are pure as they cannot be contaminated by biochemical conversion. *(ii) cis* and *medial*-compartments share the same purity. *(iii)* The total amount of the different species N_{cis}, N_{medial} and N_{trans} follows the steady-state repartition of *cis*, *medial* and *trans*-species in a pure limit, given by Eq.4. Most of the components of the system are within compartments.

The purity of a compartment - defined in the main text - is $P = \sqrt{\left(3/2 \sum_i (\phi_i - 1/3)^2\right)}$. We consider that *cis*-compartments (*medial*-compartments) are contaminated by an average fraction $\langle\phi\rangle \ll 1$ of *medial*-patches (*trans*-patches), while *trans*-compartments are pure. The purity of *cis* and *medial*-compartments is thus $P_{cis,\ medial} = \sqrt{1 - 3\langle\phi\rangle(1 - \langle\phi\rangle)}$ and $P_{trans} = 1$. Weighting these with the fraction of the different species in the system (from Eq.4) gives

$$\langle P \rangle = \frac{2\alpha_{\mathrm{TGN}} \sqrt{1 - 3\langle\phi\rangle(1 - \langle\phi\rangle)} + k_m}{2\alpha_{\mathrm{TGN}} + k_m} \tag{16}$$

To compute the average contamination $\langle\phi\rangle$, we consider the different events that affect the purity of a *cis*-compartment of size n contaminated by a fraction ϕ of *medial*-patches, surrounded by a number n_v of *cis*-vesicles, and a number n'_v of *medial*-vesicles. These events are sketched in Fig.1 - current appendix, and listed below:

1. Budding of *cis*-patches (increases ϕ)
2. Budding of *medial*-patches (decreases ϕ)
3. Fusion of *cis*-vesicles (decreases ϕ)
4. Fusion of *medial*-vesicles (increases ϕ)
5. Conversion of *cis*-patches (increases ϕ)
6. Conversion of *medial*-patches (decreases ϕ)
7. Fusion with a *medial*-compartment ($\phi \to \frac{1}{2}$)
8. Fusion with the ER ($\phi \to 0$)

The rate $p(i)$ (normalized by the fusion rate) for each mechanism, and the extent $\delta\phi$ to which it affects ϕ, are:

$$p(1) = nk_{\rm b}\,, \qquad \delta\phi(1) = \frac{n\phi}{n-1} - \phi\,; \qquad p(2) = nk_{\rm b}\,, \qquad \delta\phi(2) = \frac{n\phi-1}{n-1} - \phi$$
$$p(3) = n_{\rm v}(1-\phi)\,, \qquad \delta\phi(3) = \frac{n\phi}{n+1} - \phi\,; \qquad p(4) = n'_{\rm v}\phi\,, \qquad \delta\phi(4) = \frac{n\phi+1}{n+1} - \phi$$
$$p(5) = n(1-\phi)k_{\rm m}\,, \qquad \delta\phi(5) = \frac{1}{n}\,; \qquad p(6) = n\phi k_{\rm m}\,, \qquad \delta\phi(6) = -\frac{1}{n}$$
$$p(7) = \phi(1-\phi)\,, \qquad \delta\phi(7) = \frac{1}{2} - \phi\,; \qquad p(8) = \alpha(1-\phi)\,, \qquad \delta\phi(8) = -\phi$$

We can now compute the temporal variation of ϕ: $\dot{\phi} = \sum p(i)\delta\phi(i)$. We concentrate on the limit of large compartments: $n \gg 1$. For simplicity, we also assume that $n \gg (n_{\rm v} - n'_{\rm v})$, so that the net contribution of vesicle fusion to the purity variation, which scales like $(n'_{\rm v} - n_{\rm v})/n$, is negligible. Neglecting stochastic fluctuations, the mean-field evolution of the fraction of contaminating species is:

$$\dot{\phi} = (k_{\rm m} - k_{\rm b})(1-2\phi) + \phi(1-\phi)(1/2 - \phi - \alpha) \tag{17}$$

Below, we discuss separately the case where $\alpha = 1$ (fast fusion with the ER boundary - discussed in the main text), and the case where $\alpha = 0$ (no exit through the ER - discussed in App.7, p.39).

• $\alpha = 1$. A pure compartment; $\phi = 0$, is stable ($\dot{\phi}|_{\phi=0} < 0$) when $k_{\rm m} < k_{\rm b}$. If $k_{\rm m} > k_{\rm b}$, the compartment is contaminated by a stable fraction of *medial*-patches given by the single stable root of Eq.17 (see Fig.1B - current appendix). The purity transition lines shown in Fig.2C of the main text are obtained by inserting this root in Eq.16.

• $\alpha = 0$. If $k_{\rm b} < k_{\rm m}$, the single root of Eq.17 is for $\phi = 1/2$ and is stable. The system is always impure with mixed *cis-medial* compartments. If $k_{\rm b} > k_{\rm m}$, pure compartments ($\phi = 0$) are stable at the mean-field level, but there exist an unstable fixed point for small values of ϕ. Stochastic fluctuations may bring the system passed the unstable fixed point toward the stable mixed solution $\phi = 1/2$. For $k_{\rm b} > k_{\rm m} + \frac{1}{8}$, $\phi = 1/2$ is an unstable fixed point and the only stable solution is for $\phi = 0$ (see Fig.1C - current appendix). The purity transition is thus with the range $k_{\rm m} < k_{\rm b} < k_{\rm m} + \frac{1}{8}$.

A more precise characterization of the transition in the case $k_{\rm b} > k_{\rm m}$ requires to include stochastic effects. The different stochastic events are associated to very different fluctuations amplitudes. Budding and fusion of vesicles and biochemical conversion of membrane patches represent small fluctuations for a large compartment, while fusion between compartments and with the boundaries represent a very large alteration of the purity. In what follows, we propose a simplified treatment of this process, focusing on the case $\alpha = 0$ for simplicity, and neglecting the contribution of vesicle fusion as discussed above. For a compartment of size n contaminated by n_{medial} patches, we identify 2 types of mechanisms (depending on their amplitude). First, a smooth drift in the contamination due to biochemical conversion (with a rate $k_{\rm m} * (n - n_{medial})$) and budding (with a rate $k_{\rm b}n$). Second, a jump in the contamination due to fusion with the neighbor compartment (rate $n_{medial}(n - n_{medial})/n^2$). We compute below the mean contamination, assuming that the *cis*-compartment spends a

time τ_{conta} undergoing small fluctuations $\delta\phi \simeq 0$ close to the pure state, before fusing with a *medial*-compartment of composition $1 - \delta\phi$. The resulting compartment with $\phi \simeq 1/2$ then spends a time τ_{deconta} undergoing a decontamination process, which depends on the balance between budding and biochemical conversion.

We first calculate the average value $\delta\phi$ of the contamination around $\phi = 0$ due to biochemical conversion and budding, disregarding inter-compartment fusion. The probability $\mathbb{P}(n_{medial})$ of finding n_{medial} in the cisterna satisfies:

$$\begin{aligned}\dot{\mathbb{P}}(n_{medial}) = &- \mathbb{P}(n_{medial})(k_{\text{m}}(n - n_{medial}) + k_{\text{b}} \times n) \\ &+ \mathbb{P}(n_{medial} - 1)\, k_{\text{m}}(n - n_{medial} + 1) \\ &+ \mathbb{P}(n_{medial} + 1) k_{\text{b}} \times n\end{aligned} \tag{18}$$

In the limit $n \gg n_{medial}$, this becomes:

$$\dot{\mathbb{P}}(n_{medial})/n = -\mathbb{P}(n_{medial})(k_{\text{m}} + k_{\text{b}}) + \mathbb{P}(n_{medial} - 1)k_{\text{m}} + \mathbb{P}(n_{medial} + 1)k_{\text{b}} \tag{19}$$

The steady-state solution is

$$\mathbb{P}(n_{medial}) = (k_{\text{m}}/k_{\text{b}})^{(n_{medial})} / \sum_{i=0}^{n/2} (k_{\text{m}}/k_{\text{b}})^i = \frac{(k_{\text{b}} - k_{\text{m}})(k_{\text{m}}/k_{\text{b}})^{(n_{medial})}}{k_{\text{b}} - k_{\text{m}}(k_{\text{m}}/k_{\text{b}})^{n/2}} \tag{20}$$

and the average contamination $\delta\phi$ due to budding and biochemical conversion is:

$$\delta\phi = \sum_{i=0}^{n/2} \frac{i}{n} \times \mathbb{P}(i) \simeq \frac{k_{\text{m}}}{n(k_{\text{b}} - k_{\text{m}})}, \text{(for } k_{\text{b}} > k_{\text{m}}) \tag{21}$$

These fluctuations in composition allow fusion between compartments, at a rate $\delta\phi(1 - \delta\phi) \simeq \delta\phi$ (for $\delta\phi \ll 1$). The average waiting time for an inter-compartment fusion event is thus $\tau_{\text{conta}} = 1/\delta\phi$.

After inter-compartment fusion, the compartment is at $\phi = 1/2$, which is an unstable fixed point according to Eq.17. Any small fluctuation in composition leads to a smooth decontamination that satisfies $\dot{\phi} = (k_{\text{m}} - k_{\text{b}})(1 - 2\phi)$, disregarding inter-compartment fusion. The decontamination time τ_{deconta} and the average contamination during the process $\langle\phi\rangle_{\text{deconta}}$, can be estimated by integrating $\dot{\phi}$, from $\phi = 1/2 - 1/n$ to $\phi = 0$:

$$\tau_{\text{deconta}} = \frac{\log(n/2)}{2(k_{\text{b}} - k_{\text{m}})}, \quad \langle\phi\rangle_{\text{deconta}} = \frac{2 + n\left[\log(n/2) - 1\right]}{2n\log(n/2)} \tag{22}$$

We can now calculate $\langle\phi\rangle$ as a temporal average of ϕ:

$$\langle\phi\rangle = \frac{\tau_{\text{conta}} \times \delta\phi + \tau_{\text{deconta}} \times \langle\phi\rangle_{\text{deconta}}}{\tau_{\text{conta}} + \tau_{\text{deconta}}} \tag{23}$$

When $\log(n/2)$ is close to 1 (typically true for the range of n that is interesting here, $n \sim 100$), $\langle\phi\rangle$ can be simplified to:

$$\langle\phi\rangle \simeq \frac{k_{\text{m}} + 2k_{\text{m}}(k_{\text{b}} - k_{\text{m}})}{2k_{\text{m}} + 2n(k_{\text{b}} - k_{\text{m}})^2} \tag{24}$$

Injecting this result into Eq.16 gives the approximate purity boundary for $\alpha_{\text{ER}} = 0$:

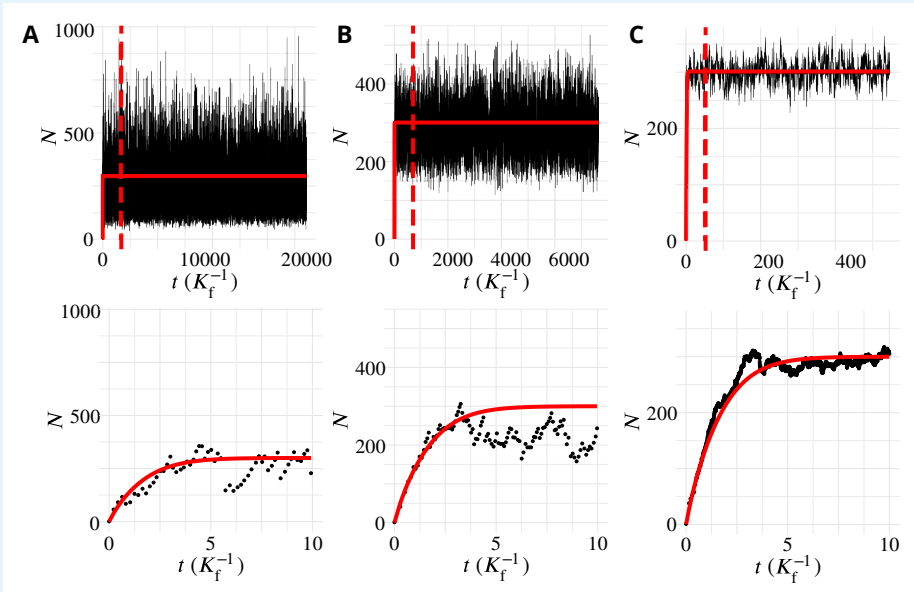
$$\langle P\rangle \simeq 1 - \frac{3\alpha_{\text{TGN}}}{2\alpha_{\text{TGN}} + k_{\text{m}}} \frac{k_{\text{m}} + 2k_{\text{m}}(k_{\text{b}} - k_{\text{m}})}{2k_{\text{m}} + 2n(k_{\text{b}} - k_{\text{m}})^2} \tag{25}$$

Despite numerous simplifications, the model gives good predictions on the purity, both for $\alpha_{\text{ER}} = \alpha_{\text{TGN}} = 1$ (Fig.2 - main text) and for $\alpha_{\text{ER}} = 0$, $\alpha_{\text{TGN}} = 1$ (Fig.1 - App.7, p.39). Depending on the normalized biochemical conversion rate k_{m}, we can discriminate three regions in the purity phase diagram:

- $k_{\mathrm{m}} \gg 1$, the system is saturated in *trans*-species and is thus always pure.
- $k_{\mathrm{m}} \sim 1$, compartment's contamination occurs by biochemical conversion, matured species are sorted by budding. The purity transition occurs for $k_{\mathrm{b}} \simeq k_{\mathrm{m}}$ (compartments are impure system if $k_{\mathrm{b}} < k_{\mathrm{m}}$).
- $k_{\mathrm{m}} \ll 1$, the system is saturated by *cis* and *medial*-compartments, and the transition of purity occurs for $k_{\mathrm{b}} \ll 1$: the fusion rate is dominant in this regime. Thus interaction with the boundaries or other compartments are crucial to understand the transition of purity. For $\alpha_{\mathrm{ER}} = 1$, fusion with the ER decontaminates the system, thus the transition occurs for $k_{\mathrm{b}} < k_{\mathrm{m}}$. The transition becomes independent of k_{b} for very small biochemical conversion rates, where the system is saturated in *cis*-species. For $\alpha_{\mathrm{ER}} = 0$, purifying the system by recycling compartments is not possible, and inter-compartment fusion decreases the purity, thus the transition of purity requires large values of the budding rate $k_{\mathrm{b}} > k_{\mathrm{m}}$, as given by Eq.25.

Appendix 4

DE NOVO FORMATION AND CARGO EXPORT KINETICS



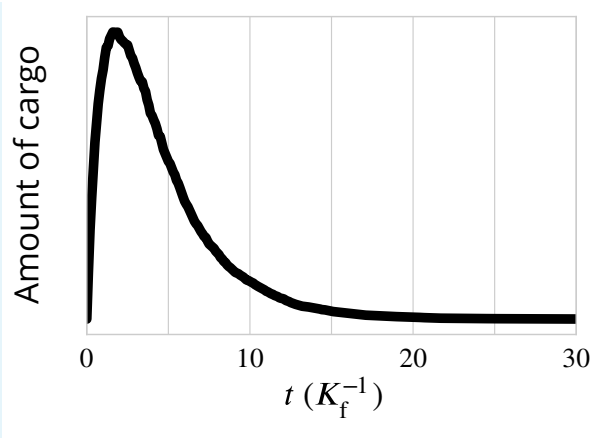
Appendix 4 Figure 1. Transient regime towards the steady-state System's size (total number of membrane patches in the system) as a function of time (in unit $1/K_f$) for three different values of the budding rate (A: $k_b = 10^{-2}$, B: $k_b = 1$, C: $k_b = 10^2$). The top row shows the full simulations, with a vertical dashed bar when the number of simulation steps reaches 10^6. The steady-state data analysed and discussed in this paper are recorded after this time, and the full simulation typically lasted 10^7 time-steps. The bottom row shows the evolution at early time (de-novo formation). The red lines are the analytical solution given by Appendix 4 Eq.3.

Transient evolution toward the steady-state: De novo formation

The evolution of the system's size with time for different values of the dimensionless budding rate k_b in shown in Fig.1 - current appendix. The organelle forms de-novo, and reaches the steady-state after a time of order a few $1/K_f$. The steady-state is reached under any parameter values, but the fluctuations around the steady-state depend on the parameter, as discussed in Appendix 5. The analytic calculation presented in Eqs.3 (red line in Fig.1 above) is very accurate for high budding rate, but gives a 10% error for low budding rate.

Transit kinetics of passive cargo.

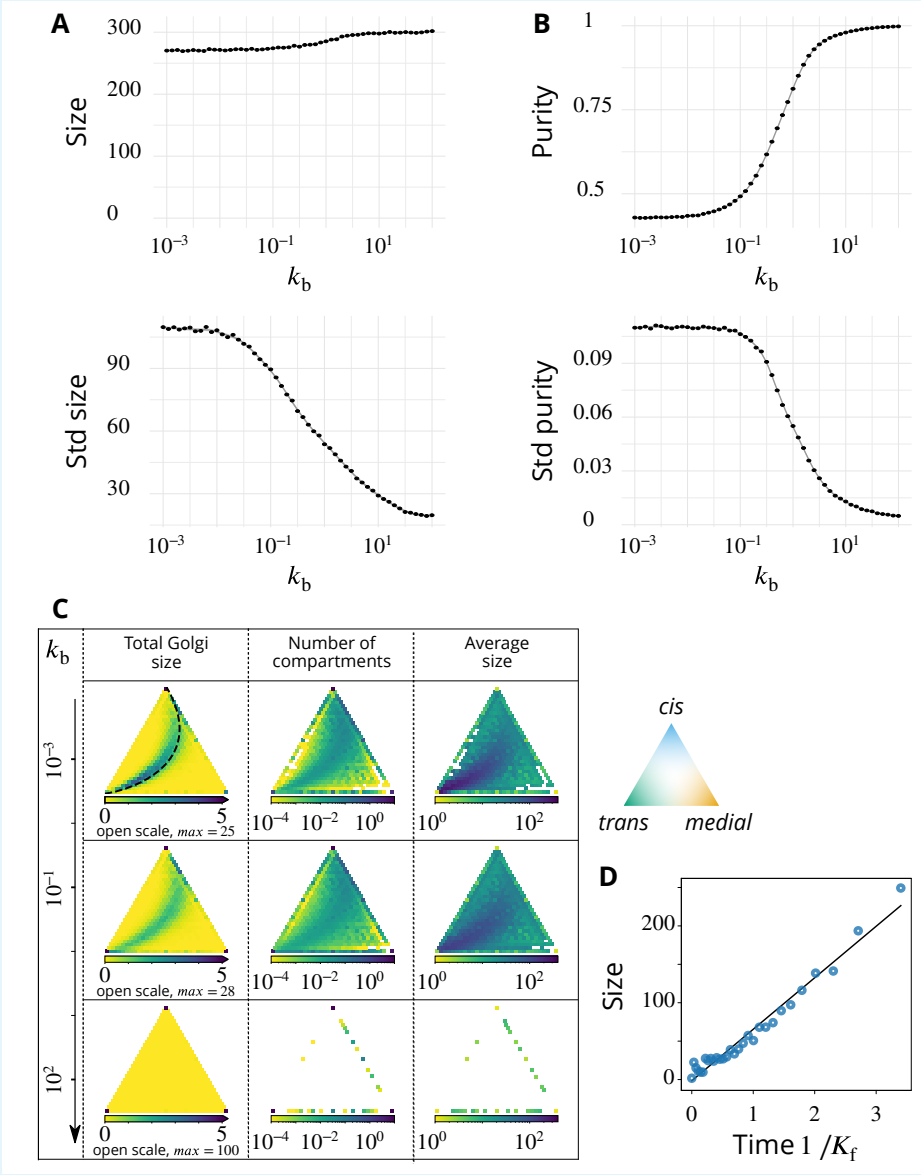
The cargo used as marker of vesicular transport directionality can be used to quantify the transit kinetics of a passive cargo across the self-organised organelle. Fig.2 - current appendix shows the amount of cargo present in the system as a function of time during a typical "pulse-chase" experiment, in which a given number of cargo molecules is allowed to exit the E.R. at $t = 0$, by joining the vesicles constituting the influx j. The amount of cargo in the system initially increases, then decreases in a close-to-exponential fashion. The export kinetics resemble the one observed by iFRAP (Fig.2 i of ***Patterson et al.*** (***2008***)) or using the RUSH technics (Fig.3 e of ***Boncompain et al.*** (***2012***)).



Appendix 4 Figure 2. Exit kinetics of passive cargo. Simulated amount of cargo in the Golgi following a pulse-chase like simulation in which 1000 cargo molecules are trapped in the ER and allowed to be packaged into ER exiting vesicles at $t = 0$. This is expected to mimick the iFRAP experimental setup used in ***Patterson et al.*** (***2008***) and to the RUSH setup used in ***Boncompain et al.*** (***2012***). The decay of the amount of cargo is close to a single-time exponential decay, is agreement with experimental observations for transiting cargo (Fig.2 i of ***Patterson et al.*** (***2008***) and Fig.3 e in ***Boncompain et al.*** (***2012***)). Parameters are $k_b = k_m = \alpha = 1$, corresponding to the "sorted regime" (the triangle of Fig.2) which we expect to be the most reasonable for mammalian Golgi.

Appendix 5

DETAILED organisation



Appendix 5 Figure 1. Steady-state composition and fluctuation of compartments. Related to Fig.2 - main text. (A-B) Characterization of the temporal fluctuations of the system as a function of the budding rate k_b. A. average values of the system total size (top panel) and standard deviation of the size (bottom panel). B. average values of the system purity (top panel) and standard deviation of the purity (bottom panel). C. Distribution of the total size, number of compartments and their average size, as a function of the composition of the compartments, shown in a triangle plot (see App.2) for different values of the (normalized) budding rate k_b. The average size for a given composition is calculated as the average size divided by the average number of compartments. Plots are shown for four different values of k_b, with $k_m = \alpha = 1$. The dashed black line on the top, left panel (low k_b regime) corresponds to the theoretical prediction of Eq.9. D. The evolution of the compartment's size as a function of time, in the low $k_b = 10^{-3}$ limit, agrees with the linear prediction of Eq.7

The fluctuations of the system around its steady-state are characterized by computing the temporal standard deviation of the total system's size and purity, shown in Fig.1A-B - current appendix. Fluctuations decrease as the budding rate k_b increases, owing to the

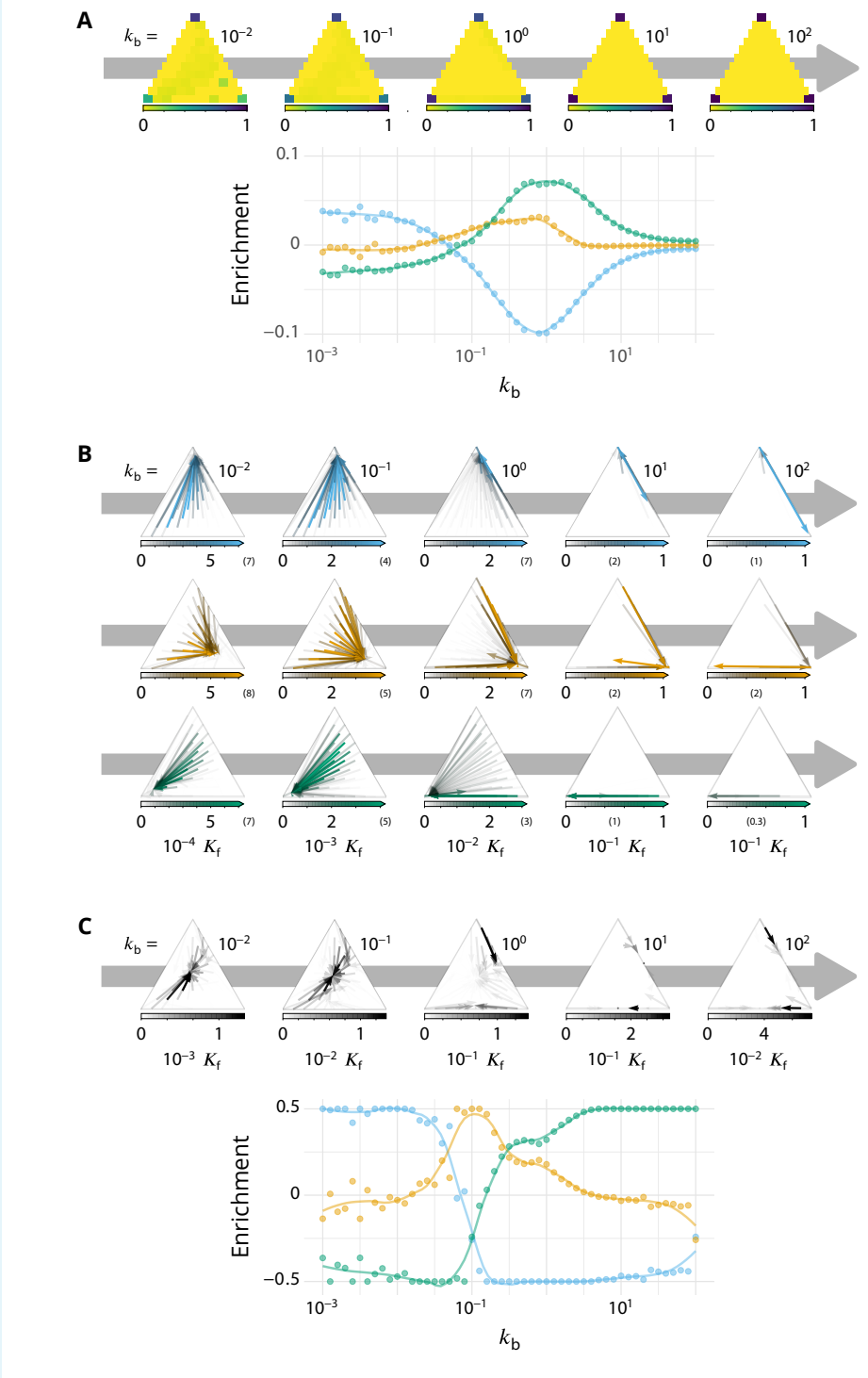
fact that compartments are on average smaller for larger budding rate, so that the removal of a compartment by fusion with the boundaries has a smaller impact on the state of the system. We note in Fig.1A - current appendix, that the total size of the system depends on the budding rate k_b, despite the fact that the influx j is varied according to Eq.6 to limit variations of the system's size. This is due to the fact that Eq.6 is valid only in the limit of pure compartments (large k_b). For smaller values of the budding rate, *medial*-compartments contaminated by *cis* and *trans*-species and may thus exit the system by fusing with the boundaries, which decreases the total size for low k_b. For the regime we are interested in, namely $k_m \sim \alpha \sim 1$, this phenomenon is practically negligible, but its impact on the average size is more severe for lower k_m (App.1, p.23).

We also characterize the variation of the size of compartments and their abundance as a function of their composition. In the main text (Fig.4), variations of the total size (number of membrane patches) and the number of compartments is shown as a function of their composition, in the composition triangular space defined in App.2, p.25. This is obtained by binning the composition space in such a way that the characterization is both relatively precise and statistically significant (binning ϕ with a mesh of $1/30$). The size and number of compartments are averaged over the simulation time (once the steady-state is reached) for each composition bin. These results are reproduced in Fig.1C - current appendix, where the average size of compartment, computed as the ratio of average size over average number of compartments, is also shown.

For low values of the budding rate k_b, both the size and the number of compartments follow the theoretical line obtained from the "cisternal maturation" limit (Eq.9). Compartments also grow linearly in time, in agreement with the linear prediction of Eq.7. Their sizes are shown in Fig.1D - current appendix, as a function of their theoretical lifetime. The sizes are computed by measuring the average compartment size for each composition bin along the theoretical line. The compositions are then used to estimate the lifetime of compartments using Eq.9.

Appendix 6

DETAILED CHARACTERIZATION OF VESICULAR TRANSPORT



Appendix 6 Figure 1. Detailed characterizations of the vesicular transport. Related to Fig.3-4 - main text. $\alpha = k_{\mathrm{m}} = 1$ for all panels. A. **Characterization of vesicle back-fusion** (fusion of vesicles with the same compartment or with a compartment of comparable composition). Top panel: fraction of vesicular flux undergoing back-fusion, as a function of the composition of the budding compartments, for different values of the budding rate k_{b}. Bottom panel: non-normalized enrichment in *cis* (blue), *medial* (orange) and *trans*-species (green) seen by a passive cargo, between a budding and the next fusion event, as a function of k_{b}. B. **Vesicular fluxes of *cis* (blue), *medial* (orange) and *trans* (green) vesicles** for different values of the budding rate k_{b}. The base of each arrow gives the composition of the donor compartment and the tip gives the average composition of receiving compartments - ignoring back-fusion - as in Fig.4, main text. The opacity of the arrows represents the normalized number of cargo (divided by the total number of cargo in the system) transported per unit time (normalized by the fusion rate). C. **Net vesicular flux, excluding vesicle biochemical conversion,** for different values of the budding rate k_{b}. Top panel: vesicular flux is represented as a vector field as in Fig.4, main text. Bottom panel: vesicular flux is represented as the enrichment between the donor and the acceptor compartment, excluding back-fusion, as in Fig.3, main text.

Quantification of the back-fusion

In the main text, we discussed the importance of back-fusion when the budding rate k_{b} increases. To determine the likelihood of a vesicle to fuse back with the compartment it budded from, or a compartment of similar composition, we analyzed the trajectories of passive cargo-proteins.

Back fusion is defined as a vesicular transport event where the donor and acceptor compartments fall within the same composition bin. The fraction of the total vesicular flux leaving a compartment that undergoes back fusion is shown as a function of the composition of compartments in a triangular composition space in Fig.1A (top panel) - current appendix. As expected, back-fusion is dominant near the vertices of the composition triangle, where compartments are pure and mostly bud vesicles of identity similar to their own identity. These vesicles are very likely to fuse back with the donor compartment by homotypic fusion. The back fusion flux increases with the budding rate k_{b}, going from $\sim [90\%,\ 10\%,\ 20\%]$ for respectively [*cis*, *medial* and *trans*]-compartments when $k_{\mathrm{b}} = 10^{-2}$, to $\sim 100\%$ when $k_{\mathrm{b}} = 10^{2}$ for all compartments.

The mean enrichment (composition difference between donor and acceptor compartments) seen by a transport vesicle is plotted in the main text in a normalized fashion. The non-normalized enrichment is plotted in Fig.1A (bottom panel) - current appendix. The mean enrichment for all three identities vanishes when k_{b} is large. This is explained by the contribution of back fusion. When k_{b} is large, the compartments are very pure, and vesicles have a high probability to undergo back fusion, leading to a small value of the net enrichment.

Vesicular transport by composition

Fig.4 of the main text shows the average flux of vesicular transport as a vector field on the triangular composition phase space. The main conclusion is that the vesicular flux is retrograde (toward *cis*-compartments), when the compartments are mixed (low budding rate k_{b}), anterograde when the compartments are pure (high k_{b}), and centripetal (toward compartments of *medial*-identity) for intermediate budding rates. In the main text, only the total flux is described, which is the sum of the fluxes of vesicles of the three different identities. To better understand this net enrichment, the total flux can be dissected into the flux of of *cis*, *medial* and *trans*-vesicles. These results are shown in Fig.1B - current appendix.

For low budding rate ($k_{\mathrm{b}} \ll 1$), compartments are mixed with all three identities, with a composition distributed around a trajectory given by Eq.9 and shown in Fig.1C - current appendix. These mixed compartments emit vesicles of all identities. The *cis*-vesicles predominantly fuse with pure *cis*-compartments, which always exist at steady-state, since pure

cis-vesicles are continuously being injected in the system. On the other hand, *medial* and *trans*-vesicles fuse with *cis/medial* and *cis/trans*-compartments, as pure *medial*-compartments are absent, and pure *trans*-compartments are few and transitory. In this limit, the net enrichment is completely dominated by the retrograde transport of *cis*-vesicles, and is positive in *cis*-identity and negative in *trans*-identity.

For intermediate budding rate ($k_b \sim 0.1 \to 1$), and as k_b increases, components can be efficiently sorted as they mature, since the budding and biochemical conversion rates are of the same order. Medial compartments are rather unstable, as they can fuse with both *cis*-rich and *trans*-rich compartments. Their reformation involves *medial* vesicles leaving *cis*-rich and *trans*-rich compartments to fuse with each other or small *medial*-rich compartments. Fig.1C - current appendix, shows that this is the main contribution to the overall vesicular flux, thereby defining the centripetal net vesicular flux shown in Fig.4 - main text. Note that they primarily exit *trans*-compartments for $k_b \sim 0.1$, and *cis*-compartments for $k_b \sim 1$, denoting a transition from retrograde to anterograde vesicular fluxes.

For high budding rate ($k_b \gg 1$), compartments are very pure, and most budding events lead to vesicle back fusion from the donor compartment (Fig.1A - current appendix). A new phenomenon can be observed, which is the anterograde transport of *cis* and *medial*-vesicles fusing with *medial* and *trans*-compartments, respectively. This is due to vesicle biochemical conversion *after* budding, as discussed below.

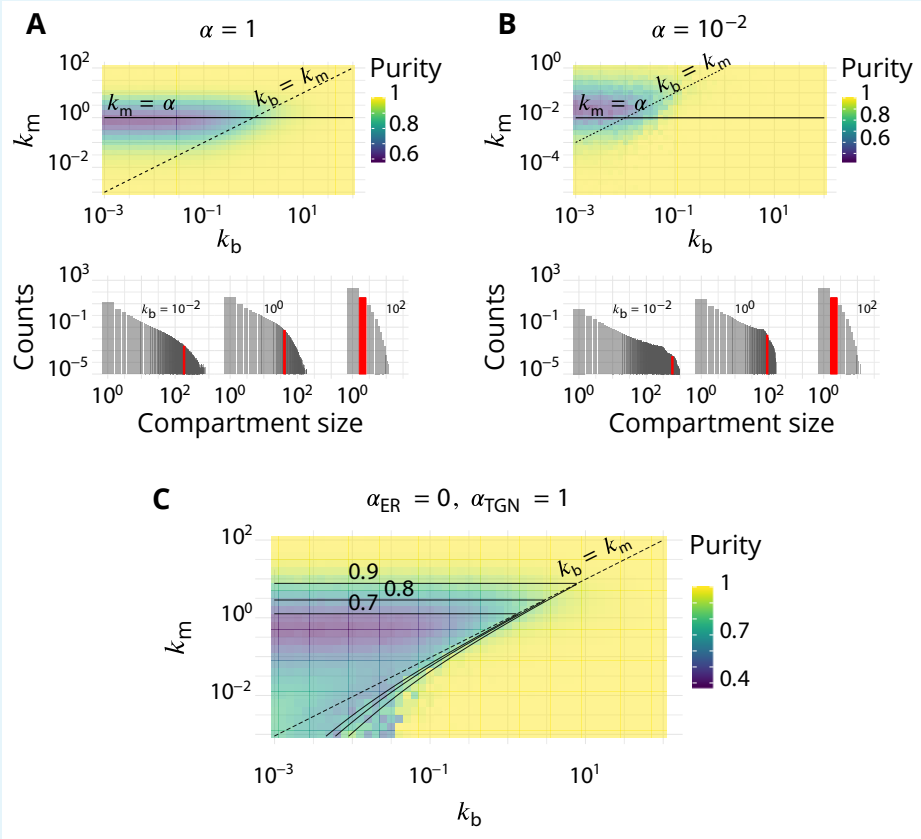
Role of vesicles biochemical conversion

The biochemical conversion of vesicles after their budding from an immature compartment has been presented as a mechanism to prevent back fusion with the donor compartment and to promote anterograde vesicular transport (see Discussion - main text). This mechanism is naturally included in our model, as vesicles have the same biochemical conversion rate than any membrane patch that belong to bigger compartments. One can notice on Fig.1B - current appendix, that there is a major increase of the anterograde *cis*-to-*medial* and *medial*-to-*trans* vesicular fluxes between $k_b = 10$ and $k_b = 10^2$. This is due to the biochemical conversion of the vesicles *after* their budding. The budding vesicle has the same identity than the donor compartment, but undergoes biochemical conversion before undergoing back fusion and fuses with a more mature compartment.

Fig.1C - current appendix, shows the net vesicular flux (for the same simulation than in the main text) ignoring the events where vesicles undergo biochemical conversion after their budding. The net vesicular flux vanishes in the vector field for high budding rates when these events are removed. We interpret this result considering dimers and trimers are dominant when k_b is large. Disregarding vesicle biochemical conversion, such compartments emit as many immature vesicles fusing with immature compartments as mature vesicles fusing with mature compartments, yielding a vanishing net vesicular flux. However, this is not true for the net enrichment which still displays an anterograde signature of the vesicular transport. Indeed, the explanation of the vesicular transport considering the minority identity (see Discussion section in the main text) is still relevant in this regime; compartments larger than 3 vesicles can generate an anterograde transport by budding patches of membrane that just underwent biochemical conversion. Consequently, the net enrichment displays the characteristics of an anterograde vesicular flux even in the absence of vesicle biochemical conversion.

Appendix 7

BOUNDARIES COMPOSITION



Appendix 7 Figure 1. Impact of the composition of the boundaries of the system (ER and TGN) on the steady-state organisation. Related to Fig.2 - main text. (A-B). The fraction α of active species driving homotypic fusion with the boundaries are identical for the ER and the TGN. A. $\alpha = 1$ (same value as in the main text) B. $\alpha = 10^{-2}$. For both cases, the top panel is the purity diagram as a function of the normalized biochemical conversion rate k_m and budding rate k_b. Dashed lines correspond to $k_m = k_b$ and solid lines to $k_m = \alpha$. For both cases, the bottom panel shows some examples of the size distribution for $k_m = \alpha$ and different values of k_b. C. Purity diagram as a function of k_m and k_b when fusion with the ER is abolished: $\alpha_{ER} = 0$ whereas $\alpha_{TGN} = 1$. Solid lines are the predictions of the analytical model developed in App.3.

Our system contains two boundaries: the *cis* face (the ER) and the *trans* face (the TGN). *cis* vesicles are injected via the *cis* face, and compartments may exit the system by fusing with either boundary. The way the different compartments fuse with the boundaries has a strong impact on their lifetime and average size. In the model, this is characterized by the parameter α, which represents the fraction of the boundary composed of species able to elicit homotypic fusion of Golgi's components, *i.e.* the fraction of *cis* species in the ER and of *trans* species in the TGN, and might be different for the two boundaries. In the main text, we focus on the case $\alpha = 1$ (arguably the maximum possible value) for both *cis* and *trans* boundaries. Here, we discuss the impact of this parameter on the structure of the steady-state, and we compare the results of simulations with the analytical model derived in App.3.

Decreasing this parameter has two major impacts on the steady-state organisation: it decreases the biochemical conversion rate for which the steady-state fraction in *cis*, *medial* and *trans*-species are equal and it increases the residence time of compartments in the

system. This increases their size, as they have more time to aggregate, and increases the fluctuations in the system, as larger compartments have a stronger impact each time they exit or fuse together.

The impact of α on the fraction of *cis*, *medial* and *trans*-species in the system is rather straightforward if compartments are pure (Eq.4). In this case, the exit flux can be exactly computed, and Eq.4 shows that the different species are in equal amount at steady-state when $k_{\mathrm{m}} = \alpha$. If $k_{\mathrm{m}} \gg \alpha$, the membrane patches have time to undergo biochemical conversion before exiting the system, which is then dominated by *trans*-species. The purity of the compartment is high, but this is a rather uninteresting limit as the system is dominated by a single species. If $k_{\mathrm{m}} \ll \alpha$, compartments are recycled out of the system before *trans*-membrane patches appear, and the system is dominated by *cis* and *medial*-species, which are of equal amount at steady state if the system is perfectly sorted.

The same conclusion can be reached in the slow budding regime when compartments are not pure. Let's consider a *cis*-rich compartment of size n contaminated by a fraction ϕ_{medial} of *medial*-species, and let's disregard, for simplicity, budding and fusion with other compartments. The number of *medial*-patches increases by biochemical conversion with a flux $k_{\mathrm{m}} n(1-\phi_{medial})$, and decreases by compartment fusion with the boundary, which removes all *medial*-patches with an average flux $\alpha(1-\phi_{medial}) \times n\ \phi_{medial}$. At steady-state, one thus expects $\phi_{medial} \simeq k_{\mathrm{m}}/\alpha$, suggesting that the lowest purity will be observed for $k_{\mathrm{m}} \simeq \alpha$. The results of simulations, shown in Fig.1A-B - current appendix, confirm this prediction.

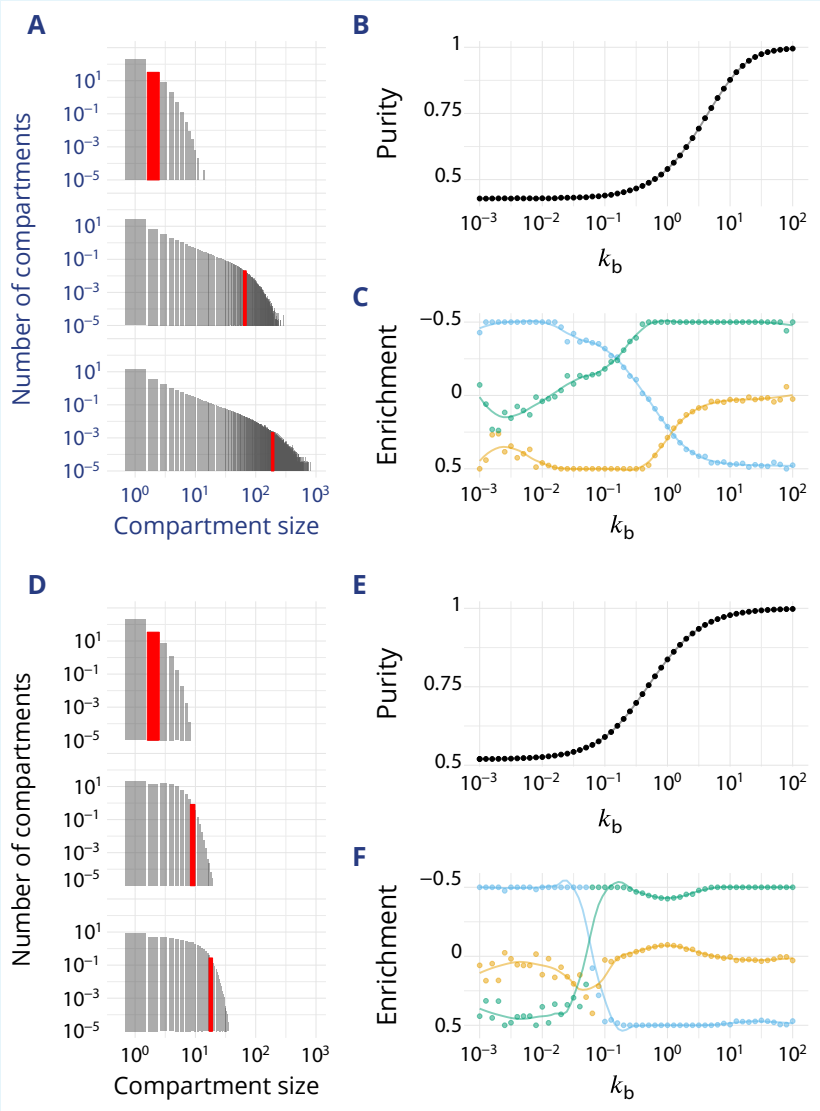
The transition of purity occurs when sorting mechanisms become more efficient than mixing mechanisms. As a first approximation mixing occurs by biochemical conversion and sorting by budding, yielding the prediction that the purity transition occurs when $k_{\mathrm{b}} \simeq k_{\mathrm{m}}$. Additional mixing mechanisms include the fusion of two slightly impure compartments of different compositions. Such slow events become relevant if compartments remain in the system for a long time, *i.e.* if fusion with the boundaries is slow (small values of α). In this case, the transition of purity occurs for $k_{\mathrm{b}} > k_{\mathrm{m}}$. Phase diagram of the purity of the system for a system where $\alpha = 10^{-2}$ can be found in Fig.1B - current appendix. As expected the diagram is centered around $k_{\mathrm{m}} = \alpha$, with a transition of purity between $k_{\mathrm{b}} \sim 10^{-2}$ and $k_{\mathrm{b}} \sim 10^{-1}$ (and thus $k_{\mathrm{b}} > k_{\mathrm{m}}$).

As it tunes the residing time of compartments in the system, α also impact the average size of compartments. Decreasing the value of α increases the residence time and leads to larger compartments, as predicted by Eq.4. Note however that for small budding rates, when compartments are mixed and their exit through the boundary is difficult to estimate, the size distribution deviates from the single-component ideal distribution.

The Golgi being a highly polarized organelle, it could be argued that the parameters controlling the fusion with the *cis* face (ER) and the *trans* face (TGN), α_{ER} and α_{TGN} could be different. An obvious way to reduce the retrograde exit of material from the ER is to reduce α_{ER}. To investigate the role of this asymmetry, we use an extreme regime of $\alpha_{\mathrm{ER}} = 0$ and $\alpha_{\mathrm{TGN}} = 1$. This prevents the recycling of impure compartments by fusion with the ER and broadens the low purity region of the parameter space, as shown in Fig.1C - current appendix, and explained by the analytical model presented in App.3

Appendix 8

OTHER BUDDING/FUSION SCHEMES



Appendix 8 Figure 1. Impact of different budding and fusion kernels on the system's organisation and vesicular transport. Related to Fig.2-3-4 - main text. $k_m = \alpha = 1$. (A-C): Linear budding scheme. (D-F): Inter-compartment fusion is prohibited. A and D: Size distribution of compartment for k_b equals 10^{-2} (top), 1 (middle) and 10^2 (bottom). B and E: Steady-state purity as a function of k_b. C and F: Enrichments in *cis* (blue), *medial* (orange) and *trans*-species (green) during vesicular transport (details of computations of this enrichment presented in App.2, p.26) as a function of k_b.

Linear budding

To verify whether the assumption of saturated budding (Eq.2) has a strong impact on the results, we performed simulations with a linear budding mechanism ($f(\phi) = \phi$ in Eq.10), for which the budding flux for a given identity is linear with the number of patch with this identity ($J_{b,i} = K_b n \phi_i = K_b n_i$). As shown in Fig.1A-B - current appendix, the average compartment size and purity follow the trends discussed in the main text for the saturated budding regime. However, the purity transition occurs for larger k_b, since the total budding flux is merely

proportional to the size of the compartment for linear budding, and is thus smaller than for saturated budding, where it is proportional to the size times the number of species. Saturated budding thus promotes systems that are at the same time well sorted, and with large compartments.

Even though the purity transition is shifted toward higher values of k_b the link between purity and vesicular flux is qualitatively conserved (Fig.1C - current appendix). For low purity, the system is retrograde with an enrichment in *cis*-species and a depletion of more mature species, and for high purity it is anterograde. However the retrograde flux is less marked than with a saturated budding rate (no clear depletion in *trans*-species). Indeed, efficient sorting relies on the capacity to export the minority component out of a compartment. Within the saturated budding scheme, the rate of export is proportional to the size of the compartment and does not depend on the number of minority components to export. Within the linear budding scheme however, it is proportional to the number of minority components. In the low k_b regime, *trans*-rich compartments are large. They emit a large vesicular flux of immature components within the saturated budding scheme, but this flux is much smaller within the linear budding scheme. This explains the qualitative difference between the enrichment curves in the low k_b regime for the saturated budding scheme (large depletion in *trans* identity - Fig.3 - main text) and the linear budding scheme (almost no change in *trans* identity - Fig.1C - current appendix).

Non-fusing compartments

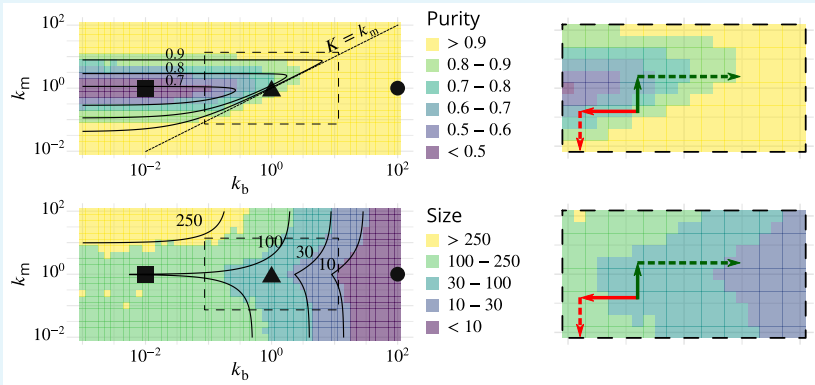
Fusion between compartments is an important ingredient of the model discussed in the main text. It could be argued that the fusion of large cisternae with one another could be a much slower process than fusion involving much smaller transport vesicles. We have performed simulations where fusion between compartments is prohibited if both compartments are larger than a vesicle. Note that in the model, such compartments are still able to fuse with the boundaries, so that the composition of the system can still be predicted by Eq. 4 The results of this model are shown in Fig.1D-F - current appendix.

The size distribution of non-fusing compartment shows a broad peak, instead of the power-law seen in the case where compartments can fuse. As compartments can only grow by vesicle fusion and not by compartments fusion, their size is smaller than in the previous case. The compartments size is results from a balance between the rate at which compartments are recycled (fuse with one boundary) and the rate at which they grow by vesicle fusion. Consequently, the size distribution only weakly depends on k_b, as compartments have very little time to fuse with vesicles before exiting the system (Fig.1-D - current appendix).

As expected, and because we remove the possible interactions between compartments, the centripetal flux we observe between the previously described anterograde and retrograde is abolished. This is particularly visible if we plot the (normalized) mean enrichment in *cis*, *medial* and *trans*-species seen by cargo as they are transported via vesicular fluxes (Fig.1-F - current appendix). As before, the system goes from a purely retrograde flux, characterized by a depletion in *trans*-species and an enrichment in *cis*-species, to an anterograde flux, characterized by an enrichment in *trans*-species and a depletion in *cis*-species. However, the transitional centripetal regime we previously described, characterized by an enrichment in *medial*-species, vanishes. This suggests this particular regime is indeed resulting from the possible fusion between compartments, as it destabilizes *medial*-compartments that have a great probability to fuse with slightly impure *cis* or *trans*-compartments. Thus the transport is anterograde as soon as the purity transition occurs (Fig.1E - current appendix).

Appendix 9

EXPERIMENTAL PROPOSAL



Appendix 9 Figure 1. Possible experimental strategies to test the predicted correlation between size and purity of Golgi cisternae. The left column reproduces Fig.2B-C, showing the phase diagrams for the purity (top) and size (bottom) of Golgi cisternae upon varying the ratio of budding to fusion rates k_b and the ratio of biochemical conversion to fusion rate k_m. The right column shows a zoom of the dashed square area, showing the observed result of physiological perturbations of cisternae purity and size (solid arrows) and prediction for combined perturbations (dashed arrows). Red arrow: observed result of a decrease of the budding rate k_b by impairing COP I (***Papanikou et al., 2015***). Dashed red arrow: prediction for the combined decrease of the biochemical conversion rate k_m by Ypt1 down-regulation. Green arrow: observed result of an increase of the biochemical conversion rate k_m by Ypt1 over-expression (***Kim et al., 2016***). Dashed green arrow: prediction for the combined increase of the budding rate k_b by COP I up-regulation.