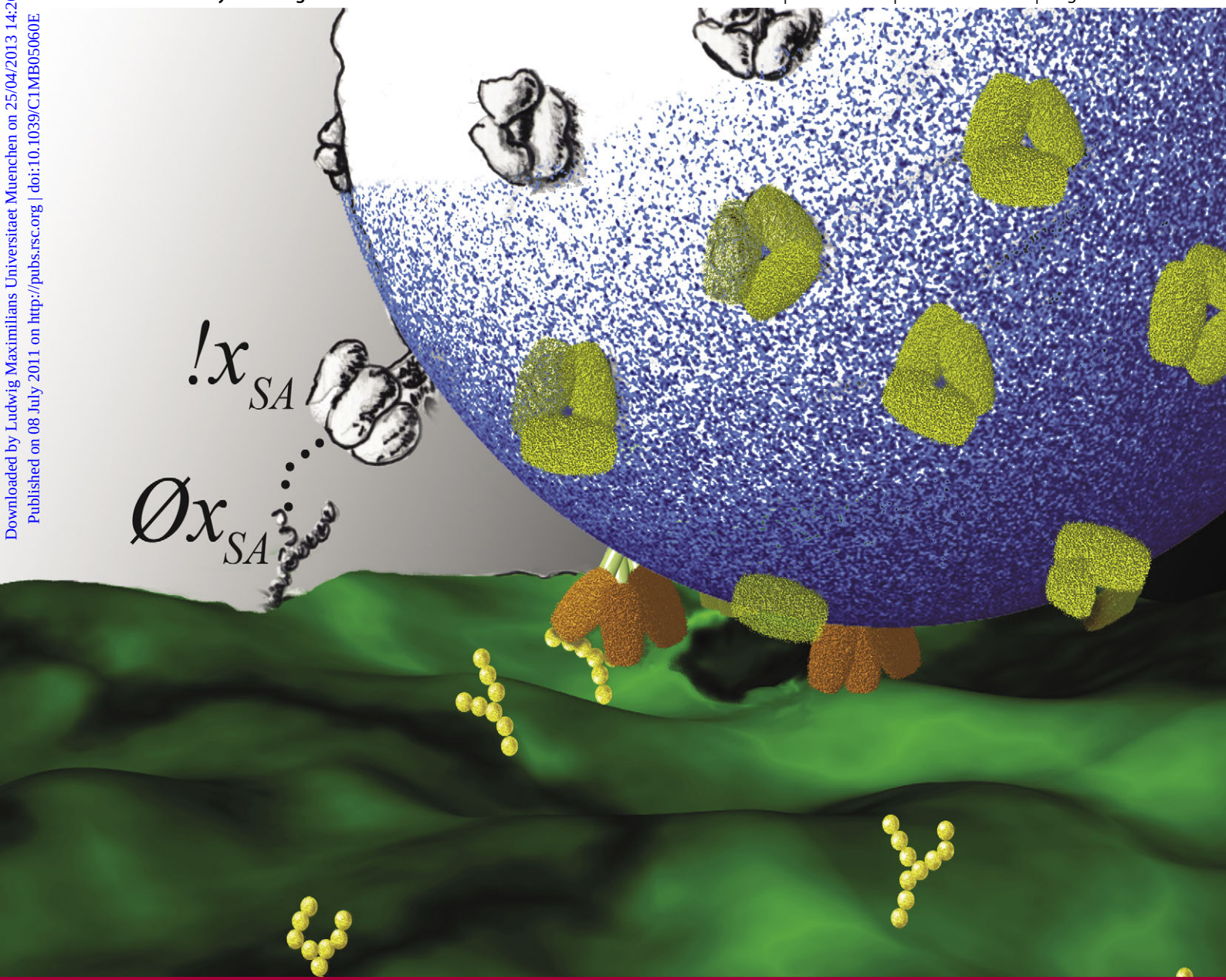
Molecular BioSystems

www.molecularbiosystems.org

Volume 7 | Number 10 | 1 October 2011 | Pages 2723–2910



Themed issue: Computational Biology

ISSN 1742-206X

RSC Publishing

PAPER
Maria Pamela Dobay *et al.*
How many trimers? Modeling influenza virus fusion yields a minimum aggregate size of six trimers, three of which are fusogenic

1742-206X(2011)7:10;1-3

Molecular BioSystems

Dynamic Article Links
View Article Online

Cite this: *Mol. BioSyst.*, 2011, **7**, 2741–2749

www.rsc.org/molecularbiosystems

PAPER

How many trimers? Modeling influenza virus fusion yields a minimum aggregate size of six trimers, three of which are fusogenic†‡

Maria Pamela Dobay,§*[a] Akos Dobay,[bc] Johnrob Bantang[d] and Eduardo Mendoza[ae]

Received 13th February 2011, Accepted 16th June 2011
DOI: 10.1039/c1mb05060e

Conflicting reports in leading journals have indicated the minimum number of influenza hemagglutinin (HA) trimers required for fusion to be between one and eight. Interestingly, the data in these reports are either almost identical, or can be transformed to be directly comparable. Different statistical or phenomenological models, however, were used to analyze these data, resulting in the varied interpretations. In an attempt to resolve this contradiction, we use PABM, a brane calculus we recently introduced, enabling an algorithmic systems biology approach that allows the problem to be modeled in a manner following a biological logic. Since a scalable PABM executor is still under development, we sufficiently simplified the fusion model and analyzed it using the model checker, PRISM. We validated the model against older HA-expressing cell-to-cell fusion data using the same parameters with the exception of three, namely HA and sialic acid (SA) surface densities and the aggregation rate, which were expected to be different as a result of the difference in the experimental setup. Results are consistent with the interpretation that a minimum aggregate size of six HA trimers, of which three undergo a conformational change to become fusogenic, is required for fusion. Of these three, two are free, while one is bound. Finally, we determined the effects of varying the SA surface density and showed that only a limited range of densities permit fusion. Our results demonstrate the potential of modeling in providing more precise interpretations of data.

1 Introduction

Membrane fusion is one of the most fundamental biological processes exhibiting mechanistic similarities across its different forms, from viral and bacterial entry to intracellular fusion.[1,2] Most of what is known regarding viral fusion are derived from structural and mechanistic studies of influenza virus fusion, as mediated by hemagglutinin (HA). Hemagglutinin is a trimeric protein anchored to the viral membrane *via* its C-terminal domain.[3] HA binds host cells through sialic acid (SA)-capped proteins, which are particularly abundant in the respiratory tract, as well as in red blood cells.[4] Binding triggers internalization into an endosome, the acidification of which causes HA trimers to aggregate and to undergo a conformational change that extends its N-terminal fusion peptides and causes it to fuse with the endosome membrane, and release its contents into the host cell cytosol.[3] *In vitro* experiments have expanded the view of HA-mediated fusion through the definition of intermediates between the conformational change step and content mixing. These include the generation of the first fusion pore (FP), through which ions can pass between the virus and target membranes; the lipid channel (LC), which permits the lipids to mix between the two membranes; and the formation of the fusion site (FS), which allows content mixing[5] (Fig. 1).

In an attempt to characterize membrane fusion and its intermediates better, several groups have designed experiments to determine the minimum requirements for the formation of a fusion pore. Viruses or virus-like systems, which have been evolved to efficiently form such pores on the endosome membrane through proteins such as HA, are consequently ideal for such studies. Knowing these minimum fusion requirements is also of interest in artificial gene and drug therapy, where efficient endosomal escape remains one of the main problems. We are particularly interested in designing HA-decorated vectors for gene and drug delivery. These vectors,

[a] *Faculty of Physics and Center for NanoScience, Ludwig-Maximilians-University, Geschwister-Scholl-Platz 1, Munich, Germany. E-mail: maria.pamela.david@physik.uni-muenchen.de; Fax: +49-(0)89-2180-3182; Tel: +49-(0)89-2180-1453*
[b] *Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurstrasse 190, 8057 Zurich, Switzerland*
[c] *Swiss Institute of Bioinformatics, Quartier Sorge Bâtiment Génopode, CH-1015 Lausanne, Switzerland*
[d] *National Institute of Physics, University of the Philippines, Diliman, Quezon City, Philippines*
[e] *Institute of Mathematics, University of the Philippines, Diliman, Quezon City, Philippines*

† Published as part of a Molecular BioSystems themed issue on Computational Biology: Guest Editor Michael Blinov.

‡ Electronic supplementary information (ESI) available. See DOI: 10.1039/c1mb05060e

§ As obtained by Bentz.[5]

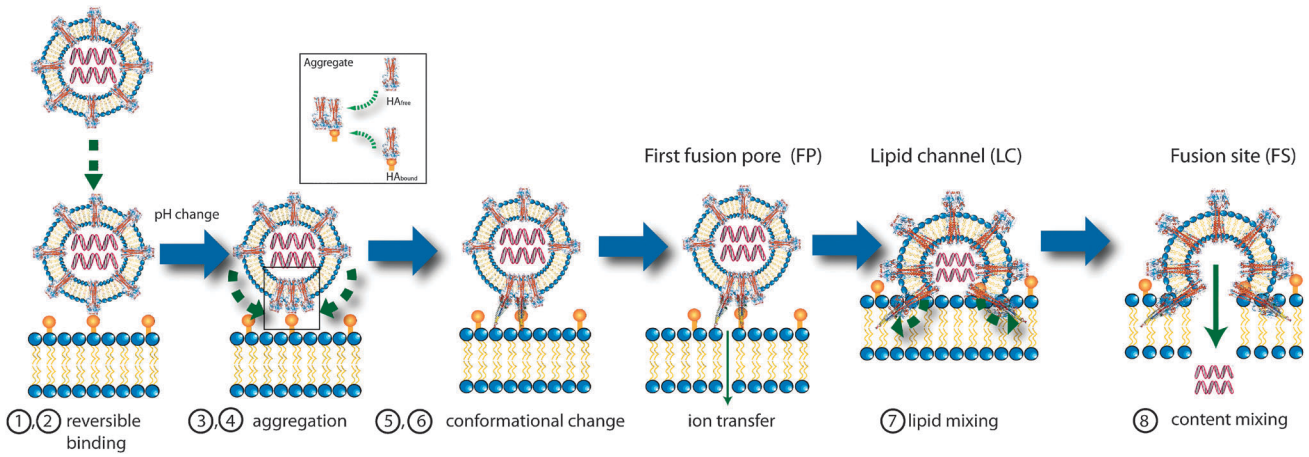


Fig. 1 Steps in HA-mediated viral fusion *in vitro*. An influenza virus binds to sialic acid-capped receptors of the cell through the HA trimers at its surface. Changing the pH triggers the aggregation of other HA trimers at the contact site, as well as a conformational change in a subset of these trimers to form a fusion pore that allows ion exchange between the virus and target. Note that the aggregate can be comprised of both bound and unbound HA trimers (inset). Each of the steps are numbered to correspond to reactions in Section 2.2.

Table 1 Minimum number of hemagglutinin trimers required for fusion as a function of experimental and statistical methods

Paper	ω	q	Virus strain(s)/ cell line(s)	Fusion partner	HA/contact area	SA/contact area	Detection method	Fitting and statistical methods	Observed step
Melikyan *et al.*[12]	8*	n/a	HAb2, GP4f	Planar bilayer with fused RBC	61×10^3–95×10^3	1.4×10^6–7.2×10^6	Time-resolved admittance	Exponential fit	FP
Blumenthal *et al.*[13]	6	n/a	GP4f	RBC	61×10^3	1.4×10^6–7.2×10^6	VFM	Empirical equation based on pore-opening kinetics	LC, FS
Danieli *et al.*[8]	n/a	3	HAb2, GP4f, gp4/6	RBC	37×10^3–479×10^3	1.4×10^6–7.2×10^6	Spectrometry (bulk)	Hill fit	LC
Guenther-Ausborn *et al.*[11]	n/a	1	X-47, A Shangdong	RBC	20–30	74–372	Resonance energy transfer (bulk)	Modified Hill fit	LC, FS
Imai *et al.*[10]	n/a	1	PR/8/1924	RBC	20–30	74–372	VFM	Log–log plots based on HA surface density and fusion rates	LC
Floyd *et al.*[9]	n/a	3	X-31	Planar bilayer with pure GD1a	20–30	74–372	VFM	Γ-fit of frequency *vs.* time distribution of fusion data	FP/LC, FS

which could be even smaller than the viruses itself, should contain at least the minimum number of fusogenic units, but the least number of HAs that would still permit fusion, to reduce their potential immunogenicity.[6] In the case of HA, these minimum requirements refer to the number of aggregated trimers that comprise the fusion pore, ω, which may be comprised of both HA bound to SA (HA_{bound}) and free HA (HA_{free}). A subset of ω, q, undergoes an acidification-mediated conformational change to a final fusogenic form.¶ However, despite the fact that the experiments were directed towards the description of a single phenomenon, the results and interpretations derived from these vary. Table 1 summarizes results from different experimental groups, together with information on the experimental setup used, as well as the statistical methods or phenomenological models used in data analysis, when applicable. Fusion intermediates that are possible to observe with each setup are also indicated. Typically, FP can be observed through conductivity measurements, while LC and FS are typically observed using video fluorescence microscopy (VFM).[7] The results describe the process as either being a cooperative[8,9] or a non-cooperative process.[10,11] Interestingly, a closer analysis of some of the data sets reveals that the results themselves are not so much varied as the analysis. For instance, a superimposition of fusion data from Imai and Floyd, which use comparable experimental setups, reveals that the experimental results are almost identical (Fig. 2). However, the conclusions of the studies are different, with Floyd and his co-authors supporting the idea of $q = 3$ on the basis of a Γ-fit of their data, whereas Imai used an additional set of experiments involving variable fusogenic HA surface densities as the basis for a conclusion of $q = 1$.[9,10]

¶ The convention of using the notations ω and q was taken from Bentz.[5]

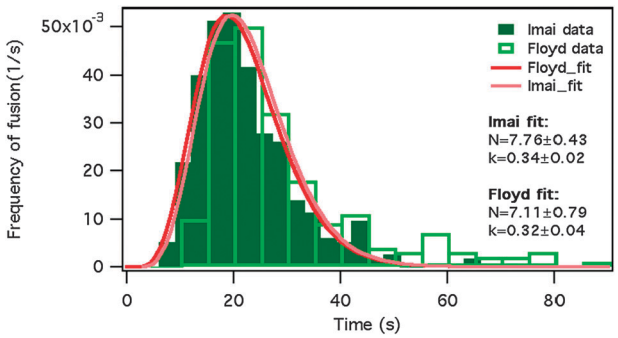


Fig. 2 Superposition of the data reported by Imai *et al.*[10] and Floyd *et al.*[9] indicates that the experimental results are almost identical. Nonetheless, the groups used different statistical methods for analysis. Fitting the Imai data set with an approximation of a Γ function yields a result of the same magnitude as in Floyd *et al.*; if this function is used as a reference for a conclusion, the data of Imai can be interpreted to support a conclusion of $q = 3$.

Consequently, if the Imai data had been interpreted on the basis of a Γ-fit, as the case was in Floyd, then it would have been in support of $q = 3$.

In contrast, there are also some studies where a similar experimental design was used, and that have arrived at the same conclusion that $q = 1$. Nonetheless, the experimental data obtained from these studies were very different, and were likewise analyzed using different techniques (Fig. 3).[10,11] In Fig. 3, the lag time is defined as the time interval between the exposure to low pH and the onset of fusion, and Gunther-Ausborn *et al.* postulated that the reciprocal of the lag time, as well as the initial rate of fusion are directly proportional to the surface density of fusion-competent HA trimers.[11] Obtaining a linear relationship consequently implies that the reaction is first order (*viz.* $q = 1$) with respect to HA, whereas a non-linear relationship would imply $q > 1$. If the Imai data had been interpreted based on this method as shown in Fig. 3, then

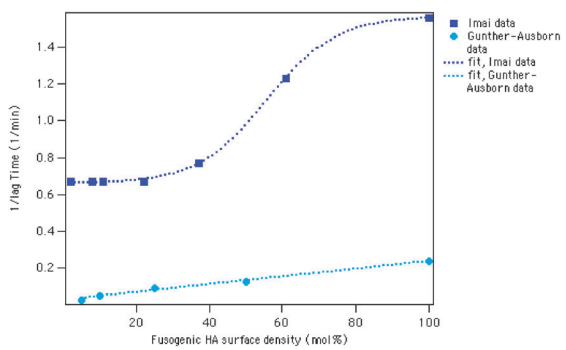


Fig. 3 Superposition of the data reported by Imai *et al.*[10] and Gunther-Ausborn *et al.*[11] using a 1/lag time *vs.* fusogenic HA surface density plot. Gunther-Ausborn *et al.* define the lag time as the interval between sample exposure to low pH and the onset of fusion; they postulated that the relationship between the reciprocal lag time with the fusogenic HA surface density gives the order of the reaction with respect to HA. Their results yield a linear relationship, supporting $q = 1$. If the same analysis had been used on the results of Imai *et al.*, they would have obtained a nonlinear curve that supports $q \neq 1$, instead of $q = 1$.

it would have resulted in a conclusion that $q \neq 1$. From these reports, it is evident that statistical analysis and phenomenological modeling are insufficient to deduce the minimum requirements for HA-mediated fusion.

Here, we attempt to resolve these apparent contradictions through an algorithmic systems biology approach,[14] which mimics the logic of the biological system. In such an approach, biological objects and processes are transformed into objects and instructions in an executable program (Fig. 4). Such an approach is especially suited to the current case, where the knowledge of the steps involved is comparatively extensive. To setup the model, we use PABM, a formal language inspired by membrane processes.[15] PABM permits the representation of biological, membrane-bound objects as dynamic, nested compartments that can merge or split, and from which contents move in and out (Fig. 5A). Changes in compartment topology result from specific interactions of processes on the membranes of compartments; a biological example of such specific interaction is the interaction of a fusion peptide with its target, which precludes fusion (Fig. 5B). Given, however, that a PABM executor is currently under development, we mapped the model to PRISM reactions to check its behavior. Our model yields $\omega = 6$, $q = 3$, where q is comprised of two free and one bound trimers. Apart from providing a possible resolution to the contradictions arising from data analysis, we were able to perform *in silico* experiments of previously untested scenarios, specifically, the effect of varying the surface density of SA. Our model yields a range of SA surface densities at which fusion can still occur. This might be able to explain the pathology of influenza in non-respiratory tract tissue and also be used as a criterion for determining if some individuals have a selective advantage against influenza. These results demonstrate the potential of algorithmic systems biology approaches in data interpretation and predictive modeling.

2 Methods

2.1 Computational modeling in PRISM

PRISM is a probabilistic and symbolic model checker which permits the analysis of all possible behaviors of the system.[16] Apart from the advantages provided by its model checking feature, it also has a simulation engine.[17] Given that the *in vitro* reactions we wish to model include one strictly membrane-related event, it is possible to map all events to biochemical-type reactions. We chose PRISM because it has the combined model checking and simulation features. The simulation feature permits us to quickly perform a sanity check of the system behavior, and to adjust initial parameter estimates. Model checking then allows us to explore all possible states and transitions, and allows us to determine if a certain property holds for a system.[16] Furthermore, it allows us to evaluate the effects of parameter changes on the probability of having fusion events. Finally, there have been numerous precedents for the use of PRISM in the modeling and analysis of biological pathways, including a previous mapping from another model involving compartments.[18–20]

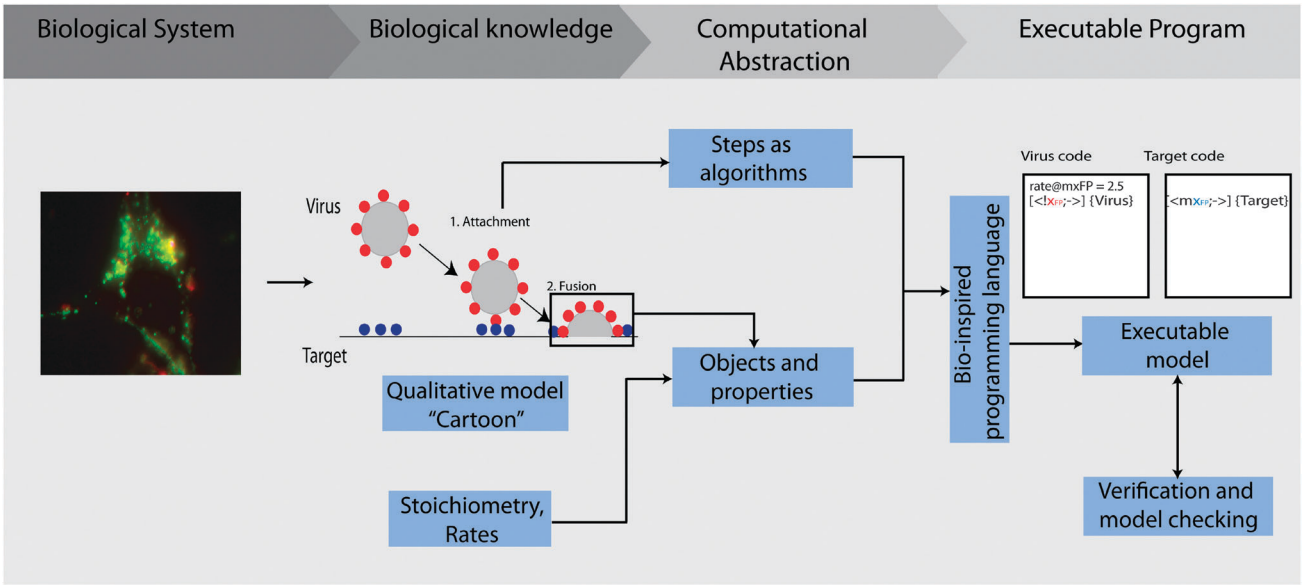


Fig. 4 Modeling workflow based on an algorithmic approach. Biological systems, which are described in terms of qualitative models ("cartoons"), as well as reaction stoichiometries and rates, are abstracted as objects, properties and algorithms that can be coded using a suitable language and executed. The behavior and reliability of the model can be evaluated through model checking and verification, respectively.

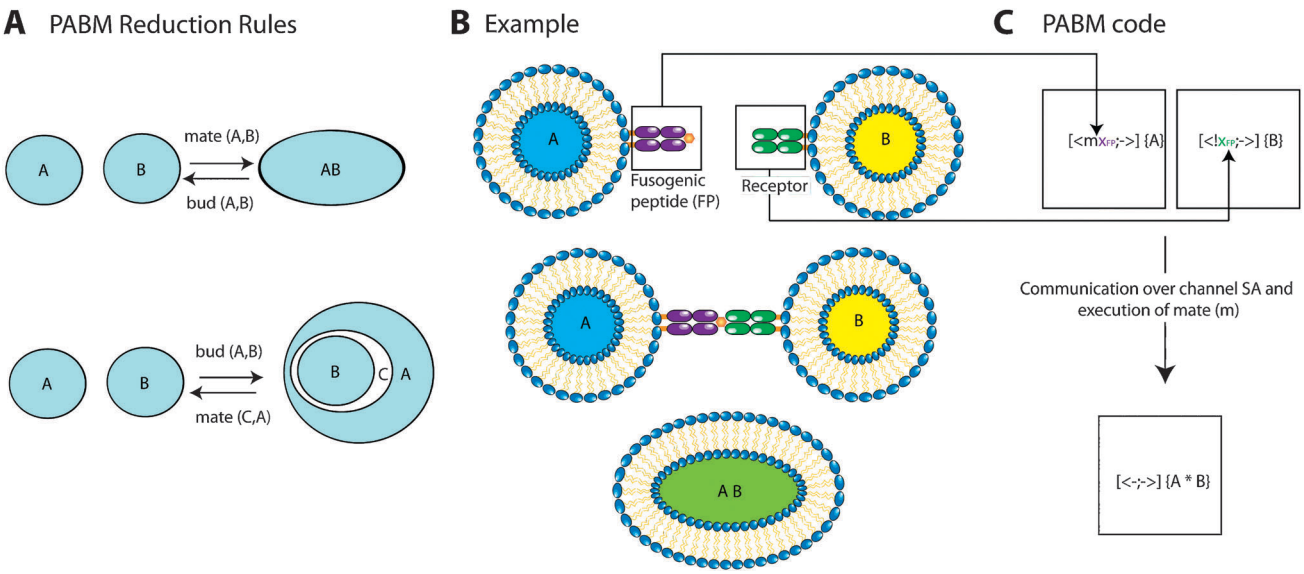


Fig. 5 An overview of PABM. PABM is a formalism that addresses the need to intuitively express biological processes involving membranes. The basic operations of PABM on compartments, known as reduction rules, are fusion (mate) and fission (bud) (A). These rules are implemented in response to specific communications between actions on membranes. Actions define which compartments can interact, as well as the fusion and fission capabilities of the membranes they are associated with. (B) In a simple biological example, a fusogenic peptide FP on the surface of a biological system A may be represented as an action with an instruction for mate (designated mx_{FP} in PABM code). Following its interaction with another system B that has the appropriate receptor for FP (designated $!x_{FP}$ in PABM code), mate is executed, and both the membranes and contents of A and B mix. Note the one-to-one correspondence between the biological system and PABM code, where the objects corresponding to A and B are likewise combined (C).

2.2 Biochemical processes

The biological processes associated with the HA-mediated fusion setups are shown in Fig. 1, where each of the reactions described below are indicated. These reactions were initially formulated in the PABM formalism, then mapped into the guarded commands required in PRISM; the properties analyzed, namely the probability of having Virus_{FP} and Virus_{LC} at time t for virus–cell and cell–cell fusion setups were expressed in continuous stochastic logic (CSL)[16]

(1) Virus binding and unbinding

$$\text{HA}_{\text{Free}} + \text{SA}_{\text{Unbound}} \xrightarrow{k_b} \text{HA}_{\text{Bound}} + \text{SA}_{\text{Bound}} \quad (1)$$

$\text{HA}_{\text{Bound}} + \text{SA}_{\text{Bound}}$ are counted as $\text{HA}_{\text{Bound}} \cdot \text{SA}_{\text{Bound}}$

$$\text{HA}_{\text{Bound}} \cdot \text{SA}_{\text{Bound}} \xrightarrow{k_{ub}} \text{HA}_{\text{Free}} + \text{SA}_{\text{Free}} \quad (2)$$

(2) pH-induced HA aggregation and conformational change

If $i \geq 1$, where i is the number of HA_{Bound} and k is the number of HA_{Free}:

$$i\mathrm{HA}_{\mathrm{Bound}} \xrightarrow{k_a} (j+1)\mathrm{HA}_{\mathrm{Bound,Aggregate}}, (i-1)\mathrm{HA}_{\mathrm{Bound}} \quad (3)$$

$$k\mathrm{HA}_{\mathrm{Free}} \xrightarrow{k_a} (l+1)\mathrm{HA}_{\mathrm{Free,Aggregate}}, (k-1)\mathrm{HA}_{\mathrm{Free}} \quad (4)$$

where j and l are the number of bound and free HA trimers in an aggregate, respectively, and where j and $l = 0$ at the start of the simulation. Both aggregation reactions are preceded by a synchronization guard of rate 1.0 to ensure that all anchor points between the fusing membranes have been established prior to HA clustering. If $HA_{Bound,Aggregate} \geq \min HA_{Bound,Aggregate}$ and $HA_{Free,Aggregate} \geq \min HA_{Free,Aggregate}$ and $HA_{Bound,Aggregate} + HA_{Free,Aggregate} \geq \omega$, where $\min HA_{Bound,Aggregate}$ and $\min HA_{Free,Aggregate}$ are user-defined and ω is equal to the minimum aggregate size:

$$j\mathrm{HA}_{\mathrm{Bound,Aggregate}} \xrightarrow{k_{f,bound}} (m+1)\mathrm{HA}_{\mathrm{Bound,Fusogenic}}, (j-1)\mathrm{HA}_{\mathrm{Bound,Aggregate}} \quad (5)$$

$$l\mathrm{HA}_{\mathrm{Free,Aggregate}} \xrightarrow{k_{f,free}} (n+1)\mathrm{HA}_{\mathrm{Free,Fusogenic}}, (l-1)\mathrm{HA}_{\mathrm{Bound,Aggregate}} \quad (6)$$

where m and $n = 0$ at the start of the simulation. Note that there is a concurrent addition to the number of trimers in an aggregate and a subtraction from the corresponding pool of trimers that were previously not associated with any aggregate. This is denoted by the comma on the right-hand side of the equation.

(3) Fusion pore (FP), lipid channel (LC) and fusion site (FS) formation

If $HA_{Bound,Fusogenic} \geq \min HA_{Bound,Fusogenic}$ and $HA_{Free,Fusogenic} \geq \min HA_{Free,Fusogenic}$ and $HA_{Bound,Fusogenic} + HA_{Free,Fusogenic} \geq q$, where $\min HA_{Bound,Fusogenic}$ and $\min HA_{Free,Fusogenic}$ are user-defined and q is a subset of ω that undergoes a conformational change:

$$\mathrm{Virus}_{\mathrm{FP}} \xrightarrow{k_{lc}} \mathrm{Virus}_{\mathrm{LC}} \quad (7)$$

$$\mathrm{Virus}_{\mathrm{LC}} \xrightarrow{k_{fs}} \mathrm{Virus}_{\mathrm{FS}} \quad (8)$$

where $Virus_{FP}$, $Virus_{LC}$ and $Virus_{FS}$ represent virus particles containing a fusion pore, a lipid channel and a fusion site, respectively.

2.3 Model assumptions and parameter estimates

2.3.1 Inclusion of binding step. Fusion experiments involve pre-binding of viruses or HA-expressing cells to the target membrane, making the virus binding reactions appear unnecessary. However, pre-binding does not prevent additional binding events from taking place in the gap between the pre-binding step and the pH drop.[21] Furthermore, the explicit representation of bound and unbound HA trimers is necessary for determining the subset of bound HA trimers in ω and q.

2.3.2 HA and SA surface densities. The estimate of the number of HA trimers/virion was taken from independent reports by Imai *et al.*, Saitakis and Gizeli and Taylor *et al.*[10,22,23] Other parameters, such as the HA and SA surface density at the contact area (Table 1) were obtained from information in the original papers, as well as estimates in a previous modeling paper.[24] We first performed simulations using these values; we then took the final HA : SA values obtained for successful fusions within the expected time scale and used this ratio in model checking. Given computing constraints in the model checker of PRISM, where values of the order of a hundred molecules for this model result in an out-of-memory error, we scaled down both the HA and SA values to the order of 15 and 5, respectively, reflecting the average 74 HA : 30 SA ratio that results in successful fusion.

2.3.3 Initial parameter estimates. Most rates for each of these transitions, with the exception of k_b and k_a, are either not available in the literature, or could not be estimated from literature values (Table 2). Initial parameter estimates were derived from known rates of diffusion,[8] which presumably affects the aggregation rate, k_a, as well as predicted rates of binding, k_b.[25] Initial values for the acid-induced conformational change, k_f, and fusion pore formation, k_{fp}, were based on parameters obtained from fits reported by Bentz.[5] In the case of k_f, we make a distinction between $k_{f,bound}$ and $k_{f,free}$ to allow us to test the cases $k_{f,bound} \ll k_{f,free}$ and $k_{f,bound} = 0$, given that there is no conclusive experimental evidence regarding the ability or inability of bound HA molecules to undergo a conformational change.[26,27] Nonetheless, if it is able to undergo the conformational change, it could be reasonably expected to be slower.[26] Consequently, we have assumed that it has a rate 1/100 of the original k_f value. For simplicity, only forward reactions were considered, although reactions (1)–(7) are known to be reversible.

Since PRISM does not have a built-in parameter optimization toolkit, derived parameters were obtained using different combinations of parameter ranges; these ranges were chosen based on preliminary runs evaluating the model behavior when a single parameter is varied, while the others are held constant

Table 2 Model parameters and rates

Reaction	Parameters	Initial rate/s^{-1}	Fitted rate/s^{-1}
Binding	74 HA, 223 SA for virus–cell fusion experiments; 15 HA, 5 SA for model-checking 18 HA, 446 SA to 30 HA, 446 SA for cell–cell fusion experiments; 15 HA, 5 SA for model-checking	0.2[25]	34.81
Unbinding	n/a	n/a	0.0001–0.25
HA aggregation	ω from 5 to 9	830[8,25]	100, virus–cell, 0.00225–0.00765, cell–cell
Conformational change	q from 1 to 3	—	6.25
Transition to FP	n/a	—	0.9025
Transition to LC	n/a	n/a	1.1025
Transition to FS	n/a	n/a	n/a

(data not shown). The data reported by Imai *et al.* were initially fitted; the set of parameters associated with the best fits, with the exception of the aggregation (presumed slower), were then used in fitting the data reported by Melikyan *et al.*

A summary of model parameters and model-derived rates is presented in Table 2.

2.4 Reducing the solution space

To determine ω, q, and the individual states (bound or unbound) of each trimer within ω and q, the user-defined parameters min $HA_{Bound,Aggregate}$, min $HA_{Free,Aggregate}$, min $HA_{Bound,Fusogenic}$ and min $HA_{Free,Fusogenic}$ can be varied to reflect all possible cases. For ω, we initially tested the values ranging from 5 to 9 (*viz.* 6 ± 1 and 8 ± 1). It is assumed that an aggregate can be comprised of both bound and free HA particles. We also assumed that q can include both bound and free HA, though $k_{f,free}$ is significantly faster than $k_{f,bound}$. Taken together, and eliminating cases that are not biologically plausible (*viz.* cases where none of the HA molecules in ω are bound) the different combinations result in a total of roughly 235 test cases (39 possible combinations for min $HA_{Bound,Aggregate}$ + min $HA_{Free,Aggregate}$ yielding a value from 5 to 9, each considered in the context of an average of 6 possible cases of min $HA_{Bound,Fusogenic}$ + min $HA_{Free,Fusogenic}$ for the range from 1–3). Finally, to eliminate even more unlikely scenarios, we used the fastest reaction for each test set; for $\omega = 8$, for example, the fastest reaction occurs when min $HA_{Free,Aggregate} = 7$, and $q = 1$, where the only trimer undergoing a conformational change is free. Using this strategy, we initially determined the most probable value of ω, then used these values for determining q.

2.5 Parameter sensitivity analysis

Sensitivity analysis was performed to estimate the reliability of the model predictions. Here, we used local sensitivity analysis adapted for the stochastic case. Briefly, parameter values were changed one at a time, while keeping the rest fixed. Sensitivity indices S_a, which represent the sensitivity of the output to a change in each parameter P_i, were calculated based on the standard:[28]

$$S_a = \frac{\partial Y}{\partial P_i} \quad (9)$$

where ∂Y is calculated as the changes in the output resulting from the substitution of the reference parameter with new parameters in incremental ratios, and ∂P_i is the difference between the reference and the new parameter. We compared the full distributions of outputs obtained from model checking for each of the parameters in order to account for changes in the shape of the distribution.

2.6 Variable SA experiments

For the variable SA experiment, we changed the values of the effective HA : SA surface density at the contact area from an original estimated value of 3.0, to values between 0.75 and 15.0, while holding the values of ω and q, obtained using the procedure in Section 2.4, constant.

3 Results

3.1 A minimum aggregate size of six trimers is required for the fusion pore

To determine ω, we initially took the range of 5–9 trimers as possible minima required for the transitions in eqn (5) and (6) to occur. As shown in Fig. 6 a requirement for $\omega \geq 8$ and above would not be able to account for the observed fusion kinetics, although this does not mean that aggregates of this size would not result in fusion. Rather, it simply indicates that majority of the fusion events would have to involve complexes of a smaller size. Presumably, it would require more time to assemble an aggregate of this size. In contrast, $\omega = 6$ closely fits the data. Furthermore, this size is consistent with electron microscopy-based approximations of the pore size formed by the so-called HA rosettes, which can be generated by solubilizing the viral membrane with detergent then sparsely redistributing them across synthetic liposomes.[29]

3.2 A minimum of three trimers in the fusion pore have to undergo a conformational change to become fusogenic

Using the results described in Section 3.1, we simultaneously varied the values of min $HA_{Bound,Fusogenic}$ and min $HA_{Free,Fusogenic}$ to determine q (Fig. 7). A value of $q = 3$, comprised of one bound and two free HA trimers, fits the data. Interestingly, different values of q do not significantly affect the level of

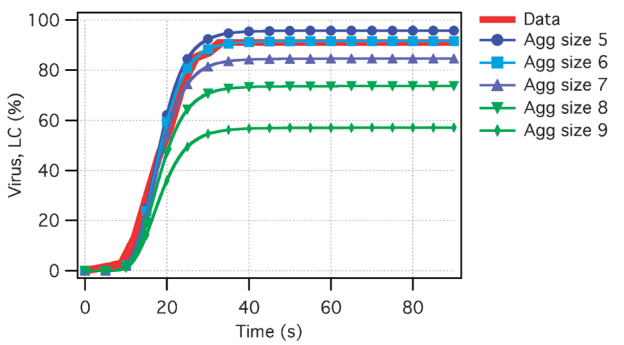


Fig. 6 Simulation of fusion data reported by Imai and Floyd (red) with the assumption that ω is 6 or 8, with at least one bound HA trimer in each case. In the case of $\omega = 8$, fusion is still observed, but does not fit the data.

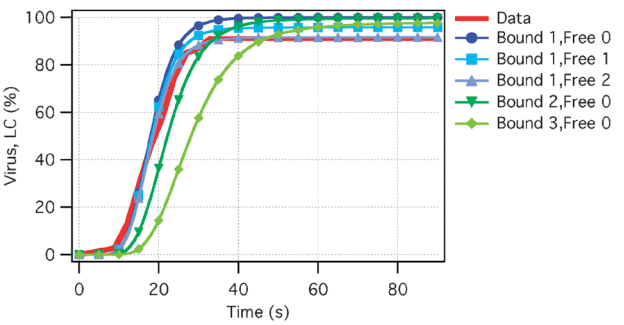


Fig. 7 Dependence of fusion kinetics on q for $\omega = 6$. Of all the possible combinations, an aggregate comprised of at least one bound and five free HA trimers, of which three (one bound, two free) must undergo a conformational change, best describes the data.

fusion as much as ω, but instead causes a shift in the time at which saturation is reached.

3.3 Model validation using cell–cell fusion experiments

We next tried to verify our results by using the predictions in Sections 3.1 and 3.2 to fit the data reported by Melikyan *et al.*,[12] which were obtained using cell–cell fusion measurements, where HA-expressing cells are used instead of viruses. Cell–cell fusion experiments differ from virus–cell fusion experiments in terms of the HA and SA surface densities at the contact area. These are also characterized by slower kinetics because of the lower HA surface density, as well as the presence of other proteins that can influence the fusion kinetics.[8,11] Nonetheless, all mechanisms starting from the point where the aggregate is assembled (3 and 4, Fig. 1) are identical.[5] It should thus be possible to capture the behavior of both HA-mediated virus–cell and cell–cell fusion using a single model. For this, we varied k_a, which is presumably slower. However, due to the memory constraints in PRISM, we had to scale down values of HA and SA, such that the effective HA : SA surface density ratio is maintained (Table 2), instead of using the actual values indicated in Table 1. The best fits for both data sets are still $\omega = 6$, with k_a values ranging from 0.00225–0.00765 s^{-1} (Fig. 8). Other cell–cell fusion measurement data[8,13] were presumed transformable to allow direct comparison with the data reported by Melikyan *et al.*, and were no longer modeled in this paper.[7]

3.4 Parameter sensitivity analysis

Following model validation, we performed sensitivity analysis on estimated parameters. The highest sensitivity index was obtained for k_f, which is consistent with the presumed role of HA conformational change as a rate-limiting step.‖[5] Changing the other rates without changing any of the non-estimated parameters, such as HA and SA surface density, does not materially affect the results (Fig. 9). The relatively low parameter sensitivities are indicative of the robustness of the model.

3.5 Effect of SA surface density on fusion kinetics

Another application that we found for the model is to check the effect of SA surface density on influenza fusion. For this, we used the parameters obtained for the virus–cell fusion setup and assumed $\omega = 6$ and $q = 3$ (1 bound, 2 free), while varying the HA : SA surface density between 0.75 to 15.0.** Of these concentrations, only HA : SA ratios between 5.0 and 2.5 resulted in at least 90% fusion (Fig. 10). At HA : SA concentrations lower than 1.67, fusion decreases dramatically, with almost no fusion occurring at HA : SA = 1.25 and below. This decrease in fusion efficiency is a necessary consequence of the predicted requirement for at least two free trimers in q; with more SA molecules available, the incidence of HA binding would be higher, and it would presumably require more time for the fusogenic complex to be assembled,

‖ This was not assumed *a priori* in our model.
** Experiment involving virus–cell fusion experiments have an average HA : SA surface density ratio of 3.0.

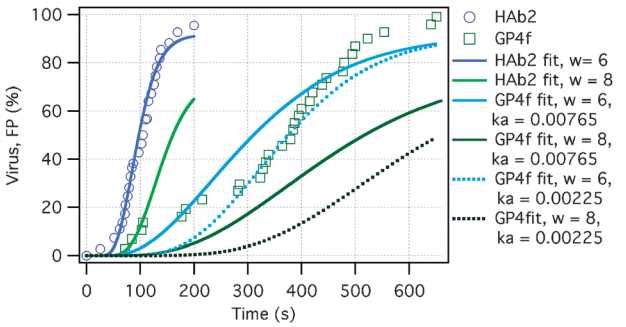


Fig. 8 Simulation of the Melikyan fusion data using parameters obtained from the Imai data, with the exception of HA and SA surface densities, and the k_a value, which was presumed to be slower than in virus–cell fusion experiments. k_a values between 0.00225–0.00765 s^{-1}, which are approximately of order 10^4 slower than k_a values for virus–cell fusion data, were obtained. The results of the model with $\omega = 8$, $q = 2$, as reported by Bentz,[5] are shown for comparison.

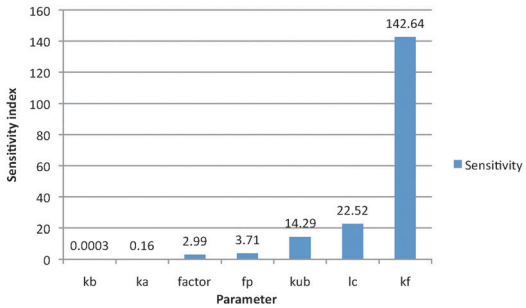


Fig. 9 Sensitivity indices of estimated parameters. The model is sensitive to changes in k_f, the rate of HA conformational change. Parameters tested were rates of binding (k_b), unbinding, (k_{ub}), aggregation (k_a), fusion pore formation (k_{fp}) and lipid channel formation (k_{lc}), as well as the factor by which k_f is decreased when bound HA undergo a conformational change.

if this has not yet been physically prevented by bound molecules at the contact site (Fig. 11).

These results are partly contrary to those obtained by Schreiber *et al.*, who predicted that a higher surface density of SA (*viz.* receptor density) is slightly more efficient than increasing the HA concentration in accelerating the fusion process.[25] It is true that for HA : SA ratios between 15.0 and 5.0, the fusion process is accelerated, and that the extent of fusion increases. However, at HA : SA ratios lower than 2.5, the effect is reversed. A probable reason for this discrepancy is their omission of the HA conformational change requirement for fusion. Furthermore, our results are consistent with earlier studies that have observed lower incidences of fusion when fusion partners with an extremely high SA content were used.[30]

Finally, the predictions may be significant with respect to the pathology of infection of certain types of influenza, which are not limited to tissue in the respiratory tract, but have also been observed in cells in the brain, lymph nodes, liver, kidney, spleen and intestine, which express SA receptors in an appreciable number. However, there was no infection in the esophagus, heart and bone marrow, even if both esophageal and cardiac tissues are in closer proximity to the respiratory tract than the kidney,[4] presumably due to the unavailability or insufficiency of

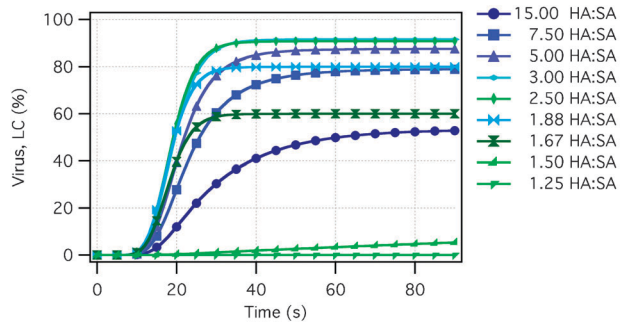


Fig. 10 SA surface density affects viral fusion efficiency. For the variable SA experiments, $\omega = 6$ and $q = 3$ were kept constant. The HA : SA surface density ratios were then varied from 15.0 down to 0.75. Only values between 5.0 and 2.5 HA : SA resulted in at least 90% fusion; between 1.50 and 0.75 HA : SA, no appreciable fusion is expected to occur, given that most of the HA molecules would be bound, and the requirement for two free HA trimers in q is unlikely to be met.

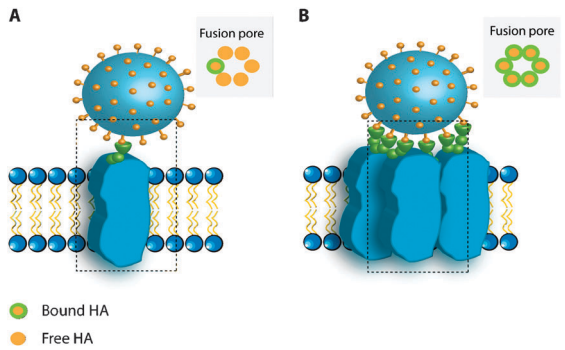


Fig. 11 Influence of SA surface density on fusion-permissive (A) and non-permissive (B) pore formation. Dotted lines represent the contact area in which the fusion pore is formed. The insets show the pores that are created when the SA surface density is lower (A); a higher surface density of SA leads to more bound HA trimers, making it difficult for the requirement of at least two free HA molecules to occur within a complex to be fulfilled.

SA receptors. It is also possible that the variation of SA surface density among individuals confer selective advantages against influenza. For instance, RBCs of thalassemia patients and diabetes patients have been reported to have lower sialic acid content.[31,32] Given that most fusion experiments were performed using red blood cell (RBC) ghosts, the predicted effect may be tested by performing fusion studies using RBC ghosts from patients with these diseases; artificial model membranes containing varying concentrations of purified glycophorin, the main sialylated protein of the RBC, could also be used for verifying our predictions.

4 Discussion

Determining the minimum requirement for a virus to create a fusion pore would provide important insights into the first line of influenza pathogenesis. Knowledge of this minimum requirement would also have interesting applications in drug and artificial gene delivery vector design, where endosomal escape has remained a perennial problem.[33] Several groups have worked on determining values of ω and q for the past 20 years, starting with measurements on HA-trimer expressing cells, and later, on virus-like or virus particles, once an adequate visualization technology was available. However, the experimental setups designed for this purpose were very different, and have all been documented to have an effect on fusion kinetics.[7,34,35] Furthermore, the statistical models and phenomenological methods used in data analysis were also widely varied. It is consequently not surprising that the data obtained appear to be very different at a first glance and that the conclusions derived from them appear contradictory.

A possible solution would be to create a model of the processes, for which consistent parameters could be obtained for at least one virus–cell and one cell–cell fusion experiment. Previous efforts to model the process used mass action kinetics to describe the fusion intermediates starting from the conformational change within the HA aggregate that leads to FP formation.[5,24] The formation of the HA aggregate is not included in the model as a step explicitly, since they have assumed that this is not a rate-limiting step. ω is instead estimated using a nucleation model. The first of the two models[5] yielded a value of $\omega = 8$, and a value of $q = 2$ or 3. A succeeding paper[24] that builds on this model by analyzing additional cell–cell fusion experiments yields $q = 2$, with the assumption that $\omega = 8$.

Here, we try a different approach where we create a stochastic model that includes aggregate formation explicitly. The inclusion of the aggregation step is necessary if we want to derive a parameter set for a model that can fit both virus–cell and cell–cell fusion data. Apart from the fact that the HA surface densities in viruses and cells are different, only the aggregate formation rate, k_a, no matter how fast it is compared to the rate of conformational change, k_f, is the only other thing that can vary between the two setups. All the other steps, from the formation of q to FS, should be the same. In fact, it is noted in an earlier paper of Bentz[36] that the aggregation step in HA-expressing cells appears to be an unfavorable, probably highly reversible event. This is in stark contrast to the step in viruses, where HA trimers might even be almost pre-aggregated.[36] Furthermore, it would be necessary to know the states of HA in the aggregate (*viz.* bound or unbound) if we want to know which of these participate in the formation of q fusogenic units.

Our model is sensitive to both ω and q, with the extent of fusion being dependent on the aggregation step. On one hand, this dependence on ω even for virus–cell fusion setups can appear counter-intuitive, since the density and relative proximity of the trimers on a virus surface could make them practically pre-aggregated. However, if one thinks of it as a reaction at least in 2D, then it could be that constructing a fusogenic aggregate might be slightly more complicated based on how many bound molecules are there at the contact area to begin with. In such a case, the dependence of the kinetics on ω could be explained. The extension of this model to a lattice, as reported by Schreiber *et al.*[25] would be particularly useful in tackling such a question.††

†† This article is not discussed in detail since its purpose is not so much as to determine the smallest fusogenic unit as it is to present a new technique for approaching the problem.

There are discrepancies between the rates obtained using fits from this model and those in that reported by Bentz,[5] which could probably be naturally expected from the fact that the model structure and the assumptions held are different. The fit generated by Bentz for the GP4f data is, of course, superior to the fit that we have obtained, but this might have been partly due either to the overly scaled-down approximation of the number of HA and SA molecules, or to an overfitting of potentially noisy data. Nonetheless, the ability of the model to closely capture both virus–cell and the general behavior of the cell–cell fusion experiments, while keeping the parameters that are expected to be constant, is promising. In the future, a more complete comparison of the two models, towards which modeling the virus–cell fusion experiments using the methods of Bentz would be a first step, would be particularly interesting. It would likewise be interesting to factor in the involvement of HA trimers outside the fusion site in fusion pore expansion.[37] Finally, we are working on creating an experimental setup to verify either of the predictions. In the advent of technologies that permit the manipulation of individual molecules with nanometre precision, it would not be so remote to conceptualize a nanoparticle with a defined number of hemagglutinin trimers at its surface. Coupled with microscopy that allows the tracking of individual HA trimers,[38] such a technique should be able to settle the question of the minimal fusion requirements definitively, while functioning as a litmus test for the significance of the results obtained from modeling processes of this scale.

Acknowledgements

MPD thanks the Deutscher Akademischer Austausch Dienst for her PhD scholarship. AD is supported by the Swiss National Foundation of Science grant number 31003A-125457. The authors wish to thank Masaki Imai and Daniel Floyd for graciously providing their original data, and Thomas Ligon for his valuable input in validating the model and the manuscript.

References

1 T. Söllner, *Curr. Opin. Cell Biol.*, 2004, **16**, 429–435.
2 M. Barocchi, V. Masignani and R. Rappuoli, *Nat. Rev. Microbiol.*, 2005, **3**, 349–358.
3 J. J. Skehel and D. C. Wiley, *Annu. Rev. Biochem.*, 2000, **69**, 531–569.
4 Y. Piwpankaew, Y. Monteerarat, O. Suptawiwat, P. Puthavathana, M. Uipresertkul and P. Auewarakul, *APMIS*, 2010, **118**, 895–902.
5 J. Bentz, *Biophys. J.*, 2000, **78**, 227–245.
6 Y. Krishnamachari, S. M. Geary, C. D. Lemke and A. K. Salem, *Pharm. Res.*, 2011, **28**, 215–236.
7 A. Mittal, E. Leikina, J. Bentz and L. Chernomordik, *Anal. Biochem.*, 2002, **303**, 145–152.
8 T. Danieli, S. Pelletier, Y. Henis and J. White, *J. Cell Biol.*, 1996, **133**, 559.
9 D. Floyd, J. Ragains and J. Skehel, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 15382–15387.
10 M. Imai, T. Mizuno and K. Kawasaki, *J. Biol. Chem.*, 2006, **281**, 12729–12735.
11 S. Gunther-Ausborn, P. Schoen and I. Bartoldus, *J. Virol.*, 2000, **74**, 2714–2720.
12 G. B. Melikyan, W. D. Niles and F. S. Cohen, *J. Gen. Physiol.*, 1995, **106**, 783–802.
13 R. Blumenthal, D. Sarkar and S. Durell, *J. Cell Biol.*, 1996, **135**, 63–71.
14 J. Fisher and T. Henzinger, *Nat. Biotechnol.*, 2007, **25**, 1239–1249.
15 M. David, J. Bantang and E. Mendoza, *Trans. Comput. Syst. Biol.*, 2009, **11**, 164–186.
16 M. Kwiatkowska, G. Norman and D. Parker, *Symbolic Systems Biology*, *Probabilistic Model Checking for Systems Biology*, Jones and Bartlett, 2010, pp. 31–59.
17 T. Pronk, E. de Vink, D. Bosnacki and T. Breit, *Proc. MTCoord 2007*, Paphos, 2007.
18 M. Kwiatkowska, G. Norman and D. Parker, *ACM SIGMETRICS Performance Evaluation Review*, 2008, vol. 35, pp. 14–21.
19 F. Romero-Campero, M. Gheorghe, L. Bianco, D. Pescini, M. Pérez-Jiménez and R. Ceterchi, *Membrane Computing*, Springer, Berlin/Heidelberg, 2006, vol. 4361, pp. 477–495.
20 J. Krivine, R. Milner and A. Troina, *Proc. of MFPS'08, 24th Conference on the Mathematical Foundations of Programming Semantics*, 2008, pp. 73–96.
21 M.-C. Giocondi, F. Ronzon, M. Nicolai, P. Dosset, P.-E. Milhiet, M. Chevalier and C. L. Grimellec, *J. Gen. Virol.*, 2010, **91**, 329.
22 H. Taylor, S. Armstrong and N. Dimmock, *Virology*, 1987, **159**, 288–298.
23 M. Saitakis and E. Gizeli, *Eur. Biophys. J.*, 2010, **40**, 209–215.
24 A. Mittal and J. Bentz, *Biophys. J.*, 2001, **81**, 1521–1535.
25 S. Schreiber, K. Ludwig, A. Herrmann and H. Holzhütter, *Biophys. J.*, 2001, **81**, 1360–1372.
26 H. Ellens, J. Bentz, D. Mason, F. Zhang and J. White, *Biochemistry*, 1990, **29**, 9697–9707.
27 B. Millar, L. Calder, J. Skehel and D. Wiley, *Virology*, 1999, **257**, 415–423.
28 A. B. Massada and Y. Carmel, *Ecol. Modell.*, 2008, **213**, 463–467.
29 R. Ruigrok, N. Wrigley, L. Calder, S. Cusack, S. Wharton, E. Brown and J. Skehel, *EMBO J.*, 1986, **5**, 41.
30 D. Alford, H. Ellens and J. Bentz, *Biochemistry*, 1994, **33**, 1977–1987.
31 I. Kahane, E. Ben-Chetrit, A. Shifter and E. Rachmilewitz, *Biochim. Biophys. Acta, Biomembr.*, 1980, **596**, 10–17.
32 K. Z. Boudjeltia, M. Piagnerelli, P. Piro, D. Bastin, E. Carlier, P. Lejeune and M. Vanhaeverbeek, *Crit. Care*, 2000, **4**, 17.
33 K. Douglas, *Biotechnol. Prog.*, 2008, **24**, 871–883.
34 A. Puri, F. Booy, R. Doms, J. White and R. Blumenthal, *J. Virol.*, 1990, **64**, 3824.
35 V. I. Razinkov, G. B. Melikyan and F. S. Cohen, *Biophys. J.*, 2008, **77**, 3144–3151.
36 J. Bentz, *Biophys. J.*, 1992, **63**, 448–459.
37 E. Leikina, A. Mittal, M. Cho, K. Melikov, M. Kozlov and L. Chernomordik, *J. Biol. Chem.*, 2004, **279**, 26526.
38 S. Hess, T. Gould, M. Gudheti, S. Maas, K. Mills and J. Zimmerberg, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 17370.