Electronic Supplementary Material (ESI) for Molecular BioSystems

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

A PABM model of the *in vitro* virus fusion setups

The following PABM model represents the steps observed in *in vitro* virus fusion setups [1]:

$$HA \equiv \langle !x_{FS}.!x_{LC}.!x_{FP}.(.\varnothing x^{i}_{f,b}.\varnothing x^{j}_{agg,b} + x^{k}_{f}.\varnothing x^{l}_{agg}).!x_{SA}\,;\,-\rangle \tag{1}$$

$$SA \equiv \langle m x_{FS}.\varnothing x_{LC}.\varnothing x_{FP}.!x_{f/f,b}.!x_{agg/agg,b}.\varnothing x_{SA}\,;\,-\rangle \tag{2}$$

where i, j, k and l are the user-defined requirements for $min\,HA_{Bound,Fusogenic}$, $min\,HA_{Bound,Aggregate}$, $min\,HA_{Free,Fusogenic}$ and $min\,HA_{Free,Aggregate}$, respectively. The reacting systems are defined as:

$$[\,HA^{x}\,]\!(\,Virus\ content\,)\!\circ[\,SA^{y}\,]\!(\,Cell\ content\,)\!) \tag{3}$$

where x corresponds to the number of HA molecules and y corresponds to the number of SA molecules at the contact area between the virus and the cell. The only first step possible would be an interaction (binding) between the virus and the target membrane through the channel x_{SA}, resulting in the following configuration:

$$HA_b \equiv \langle !x_{FS}.!x_{LC}.!x_{FP}.\varnothing x^{i}_{f,b}.\varnothing x^{j}_{agg,b}\,;\,-\rangle \tag{4}$$

$$\left[\,HA_b|HA^{x-1}\,\right]\!(\,Virus\ content\,)\!\circ\left[\,m x_{FS}.\varnothing x_{LC}.\varnothing x_{FP}.!x_{f/f,b}.!x_{agg/agg,b}|SA^{y-1}\,\right]\!(\,Cell\ content\,)\!) \tag{5}$$

Note that the parallel composition of HA_b and HA permits a choice between an aggregation reaction involving a b or a f HA. An aggregation reaction involving HA_b would yield an HA'_b in the form of:

$$HA'_b \equiv \langle !x_{FS}.!x_{LC}.!x_{FP}.\varnothing x^{i}_{f,b}.\varnothing x^{j-1}_{agg,b}\,;\,-\rangle \tag{6}$$

whereas an aggregation reaction involving f HA would yield an HA' in the following form:

$$HA' \equiv \langle !x_{FS}.!x_{LC}.!x_{FP}.\varnothing x^{k}_{f}.\varnothing x^{l-1}_{agg}\,;\,-\rangle \tag{7}$$

Aggregate formation strictly precedes any conformational change, given that HA molecules that undergo this conformational change prior to aggregation are inactivated. After $j+l$ aggregation reactions, the reacting systems will now have the following form:

$$HA_{f,b} \equiv \langle !x_{FS}.!x_{LC}.!x_{FP}.\varnothing x^{i}_{f,b}\,;\,-\rangle \tag{8}$$

$$HA_{f,f} \equiv \langle !x_{FS}.!x_{LC}.!x_{FP}.\varnothing x^{k}_{f}\,;\,-\rangle \tag{9}$$

$$\left[\,HA_{f,b}|HA_{f,f}|HA^{x-(j+l)}\,\right]\!(\,Virus\ content\,)\!\circ \tag{10}$$
$$\left[\,(m x_{FS}.\varnothing x_{LC}.\varnothing x_{FP}.!x_{f/f,b})^{j+l}|SA^{y-(j+l)}\,\right]\!(\,Cell\ content\,)\!)$$

[1] The notations are based on the syntax reported in M. David, J. Bantang and E. Mendoza, *Transactions on Computational Systems Biology XI*, 2009, 164 - 186.

Given that $j + l > i + k$, it is certain that the $i \oslash x_{f,b}$ and $k \oslash x_f$ actions in $HA_{f,b}$ and $HA_{f,f}$ will be executed. The first fusion pore (FP) can be formed between two systems in the following state:

$$HA_{f,b'} = HA_{f,f'} \equiv \langle !x_{FS}.!x_{LC}.!x_{FP}\,;\, - \rangle \tag{11}$$

$$\left[\, HA_{f,b'} | HA_{f,f'} | HA^{x-(j+l)} \,\right] (\!| \, Virus\ content \, |\!) \circ \left[\, (mx_{FS}.\oslash x_{LC}.\oslash x_{FP} | SA^{y-(j+l)} \,\right] (\!| \, Cell\ content \, |\!) \tag{12}$$

which can be alternately written as:

$$\left[\, !x_{FS}.!x_{LC}.!x_{FP} | HA^{x-(j+l)} \,\right] (\!| \, Virus\ content \, |\!) \circ \left[\, (mx_{FS}.\oslash x_{LC}.\oslash x_{FP} | SA^{y-(j+l)} \,\right] (\!| \, Cell\ content \, |\!) \tag{13}$$

The lipid channel (LC) can then be formed:

$$\left[\, !x_{FS}.!x_{LC} | HA^{x-(j+l)} \,\right] (\!| \, Virus\ content \, |\!) \circ \left[\, mx_{FS}.\oslash x_{LC} | SA^{y-(j+l)} \,\right] (\!| \, Cell\ content \, |\!) \tag{14}$$

Finally, the formation of the fusion site (FS) leads to membrane and content mixing of the two systems:

$$\left[\, !x_{FS} | HA^{x-(j+l)} \,\right] (\!| \, Virus\ content \, |\!) \circ \left[\, mx_{FS} | SA^{y-(j+l)} \,\right] (\!| \, Cell\ content \, |\!) \tag{15}$$

$$\left[\, SA^{y-(j+l) | HA^{x-(j+l)}} \,\right] (\!| \, Cell\ content \circ Virus\ content \, |\!) \tag{16}$$

Note that although the reactions in an *in vitro* fusion setup do not require the representation of compartments, a PABM model could be easily expanded to the *in vivo* scenario where reactions 7 to 16 occur inside the endosome. The representation of content mixing between the virus and the cell would also be straightforward in the sense that no additional variables representing the location of the contents would be needed.

B PRISM model

```
module fusion
virus_fp : [0..VIR] init 0;
virus_lc : [0..VIR] init 0;
virus_fused : [0..VIR] init 0;

//fusogenic HA
ha_b : [0..VIR*HA_f] init 0;
ha_f : [0..VIR*HA_f] init VIR*HA_f;
complex : [0..VIR*SA] init 0;
ha_ba: [0..VIR*HA_f] init 0;
ha_fa: [0..VIR*HA_f] init 0;
ha_bc : [0..VIR*HA_f];
ha_fc : [0..VIR*HA_f];

//SA
sa_f : [0..SA] init SA;
sa_b : [0..SA] init 0;

x: [0..14] init 0;
[time] x<LONGWAIT -> (x'=min(x+1,14));

//binding reactions
[binding] ha_f>0&ha_b<(VIR*HA_f)&sa_f>0&sa_b<SA&
  (ha_f-1)>=0&(ha_b+1)<=HA_f*VIR&(sa_f-1)>=0&(sa_b+1)<=SA ->
b*ha_f*sa_f : (ha_f'=ha_f-1)&(ha_b'=ha_b+1)&
                        (sa_b'=sa_b+1)&(sa_f'=sa_f-1)&(complex'=complex+1);

//associated virus binding reactions
[unbinding] ha_b>0&(ha_b-1)>=0&sa_b>0&(sa_b-1)>=0&
ha_f<=(VIR*HA_f)&(ha_f+1)<=(VIR*HA_f)&(sa_f+1)<=SA&
complex>0&(complex-1)>=0 ->
ub*complex : (ha_b'=ha_b-1)&(sa_b'=sa_b-1)&(sa_f'=sa_f+1)&
     (ha_f'=ha_f+1)&(complex'=complex-1);

//aggregation can only occur after something is b; reactions must be coupled
[aggregation_f]  x=LONGWAIT&ha_b>0&ha_f>0&(ha_fa+1)<=HA_f*VIR&(ha_f-1)>=0 ->
ka*ha_f : (ha_f'=ha_f-1)&(ha_fa'=ha_fa+1);

[aggregation_b]  x=LONGWAIT&ha_b>0&(ha_ba+1)<=HA_f*VIR&(ha_b-1)>=0 ->
ka*ha_b : (ha_b'=ha_b-1)&(ha_ba'=ha_ba+1);

//conformational change
[conformational_change_b] ha_ba>=min_b_agg&
(ha_ba-1)>=0&(ha_bc+1)<=HA_f*VIR ->
kf*ha_ba/factor : (ha_bc'=ha_bc+1)&(ha_ba'=ha_ba-1);

[conformational_change_f] ha_fa>=min_f_agg&
(ha_fa-1)>=0&(ha_fc+1)<=HA_f*VIR ->
kf*ha_fa : (ha_fc'=ha_fc+1)&(ha_fa'=ha_fa-1);

[ffp] virus_fp=0&virus_lc=0&ha_bc>= min_b_cc & ha_fc>= min_f_cc & (virus_fp+1)<=VIR ->
fp : (virus_fp'=virus_fp+1);

[lc] virus_fp=1 & (virus_lc+1)<=VIR&(virus_fp-1)>=0 ->
lc : (virus_lc'=virus_lc+1)&(virus_fp'=virus_fp-1);
```

```
endmodule

rewards "VIRUS_FP"
true : virus_fp;
endrewards

rewards "VIRUS_LC"
true : virus_lc;
endrewards
```

CSL property checking:

```
const double t;
R{"VIRUS_FP"}=? [ I=t ]
R{"VIRUS_LC"}=? [ I=t ]
```