

Fragmenting the early EGFR pathway (solution)

Jérôme Feret

January 28, 2010

1 The model

We consider the following rule set:

- r01: $EGF(r^1), EGFR(l^1, r) \longrightarrow EGF(r), EGFR(l, r)$
- r02: $EGF(r), EGFR(l, r) \longrightarrow EGF(r^1), EGFR(l^1, r)$
- r03: $EGF(r^2), EGF(r^1), EGFR(l^2, r), EGFR(l^1, r) \longrightarrow EGF(r^3), EGF(r^2), EGFR(l^3, r^1), EGFR(l^2, r^1)$
- r04: $EGF(r^3), EGF(r^2), EGFR(l^3, r^1), EGFR(l^2, r^1) \longrightarrow EGF(r^2), EGF(r^1), EGFR(l^2, r), EGFR(l^1, r)$
- r05: $EGF(r^3), EGF(r^2), EGFR(l^3, r^1), EGFR(Y68_u, l^2, r^1) \longrightarrow EGF(r^3), EGF(r^2), EGFR(l^3, r^1), EGFR(Y68_p, l^2, r^1)$
- r06: $EGFR(Y68_p) \longrightarrow EGFR(Y68_u)$
- r07: $EGF(r^3), EGF(r^2), EGFR(l^3, r^1), EGFR(Y48_u, l^2, r^1) \longrightarrow EGF(r^3), EGF(r^2), EGFR(l^3, r^1), EGFR(Y48_p, l^2, r^1)$
- r08: $EGFR(Y48_p) \longrightarrow EGFR(Y48_u)$
- r09: $EGFR(Y48_p^1, r^-), SHC(Y7_u, pi^1) \longrightarrow EGFR(Y48_p^1, r^-), SHC(Y7_p, pi^1)$
- r10: $SHC(Y7_p, pi^-) \longrightarrow SHC(Y7_u, pi^-)$
- r11: $SHC(Y7_p, pi) \longrightarrow SHC(Y7_u, pi)$
- r12: $EGFR(Y68_p), GRB2(a, b) \longrightarrow EGFR(Y68_p^1), GRB2(a^1, b)$
- r13: $EGFR(Y68_p^1), GRB2(a^1, b) \longrightarrow EGFR(Y68_p), GRB2(a, b)$
- r14: $EGFR(Y68_p), GRB2(a, b^-) \longrightarrow EGFR(Y68_p^1), GRB2(a^1, b^-)$
- r15: $EGFR(Y68_p^1), GRB2(a^1, b^-) \longrightarrow EGFR(Y68_p), GRB2(a, b^-)$
- r16: $EGFR(Y68_p^1), GRB2(a^1, b), SOS(d) \longrightarrow EGFR(Y68_p^2), GRB2(a^2, b^1), SOS(d^1)$
- r17: $EGFR(Y68_p^2), GRB2(a^2, b^1), SOS(d^1) \longrightarrow EGFR(Y68_p^1), GRB2(a^1, b), SOS(d)$
- r18: $GRB2(a, b), SOS(d) \longrightarrow GRB2(a, b^1), SOS(d^1)$
- r19: $GRB2(a, b^1), SOS(d^1) \longrightarrow GRB2(a, b), SOS(d)$
- r20: $GRB2(a^1, b), SHC(Y7_p^1, pi), SOS(d) \longrightarrow GRB2(a^2, b^1), SHC(Y7_p^2, pi), SOS(d^1)$
- r21: $GRB2(a^2, b^1), SHC(Y7_p^2, pi), SOS(d^1) \longrightarrow GRB2(a^1, b), SHC(Y7_p^1, pi), SOS(d)$

- r22: $GRB2(a^1, b), SHC(Y7_p^1, pi^-), SOS(d) \longrightarrow GRB2(a^2, b^1), SHC(Y7_p^2, pi^-), SOS(d^1)$
- r23: $GRB2(a^2, b^1), SHC(Y7_p^2, pi^-), SOS(d^1) \longrightarrow GRB2(a^1, b), SHC(Y7_p^1, pi^-), SOS(d)$
- r24: $EGFR(Y48_p), SHC(Y7_u, pi) \longrightarrow EGFR(Y48_p^1), SHC(Y7_u, pi^1)$
- r25: $EGFR(Y48_p^1), SHC(Y7_u, pi^1) \longrightarrow EGFR(Y48_p), SHC(Y7_u, pi)$
- r26: $EGFR(Y48_p), SHC(Y7_p, pi) \longrightarrow EGFR(Y48_p^1), SHC(Y7_p, pi^1)$
- r27: $EGFR(Y48_p^1), SHC(Y7_p, pi^1) \longrightarrow EGFR(Y48_p), SHC(Y7_p, pi)$
- r28: $EGFR(Y48_p), GRB2(a^1, b), SHC(Y7_p^1, pi) \longrightarrow EGFR(Y48_p^2), GRB2(a^1, b), SHC(Y7_p^1, pi^2)$
- r29: $EGFR(Y48_p^2), GRB2(a^1, b), SHC(Y7_p^1, pi^2) \longrightarrow EGFR(Y48_p), GRB2(a^1, b), SHC(Y7_p^1, pi)$
- r30: $EGFR(Y48_p), GRB2(a^2, b^1), SHC(Y7_p^2, pi), SOS(d^1) \longrightarrow EGFR(Y48_p^3), GRB2(a^2, b^1), SHC(Y7_p^2, pi^3), SOS(d^1)$
- r31: $EGFR(Y48_p^3), GRB2(a^2, b^1), SHC(Y7_p^2, pi^3), SOS(d^1) \longrightarrow EGFR(Y48_p), GRB2(a^2, b^1), SHC(Y7_p^2, pi), SOS(d^1)$
- r32: $EGFR(Y48_p^1), GRB2(a, b), SHC(Y7_p, pi^1) \longrightarrow EGFR(Y48_p^2), GRB2(a^1, b), SHC(Y7_p^1, pi^2)$
- r33: $EGFR(Y48_p^2), GRB2(a^1, b), SHC(Y7_p^1, pi^2) \longrightarrow EGFR(Y48_p^1), GRB2(a, b), SHC(Y7_p, pi^1)$
- r34: $GRB2(a, b), SHC(Y7_p, pi) \longrightarrow GRB2(a^1, b), SHC(Y7_p^1, pi)$
- r35: $GRB2(a^1, b), SHC(Y7_p^1, pi) \longrightarrow GRB2(a, b), SHC(Y7_p, pi)$
- r36: $GRB2(a, b^-), SHC(Y7_p, pi) \longrightarrow GRB2(a^1, b^-), SHC(Y7_p^1, pi)$
- r37: $GRB2(a^1, b^-), SHC(Y7_p^1, pi) \longrightarrow GRB2(a, b^-), SHC(Y7_p, pi)$
- r38: $EGFR(Y48_p^2), GRB2(a, b^1), SHC(Y7_p, pi^2), SOS(d^1) \longrightarrow EGFR(Y48_p^3), GRB2(a^2, b^1), SHC(Y7_p^2, pi^3), SOS(d^1)$
- r39: $EGFR(Y48_p^3), GRB2(a^2, b^1), SHC(Y7_p^2, pi^3), SOS(d^1) \longrightarrow EGFR(Y48_p^2), GRB2(a, b^1), SHC(Y7_p, pi^2), SOS(d^1)$

We start with the following initial state:

$$\begin{aligned}
&2000 * (EGFR(l, r, Y48_u, Y68_u)) \\
&2000 * (EGF(r)) \\
&2000 * (GRB2(a, b)) \\
&1000 * (SOS(d)) \\
&1000 * (SHC(pi, Y7_u))
\end{aligned}$$

2 First intuitions about the model

- (i) Compute the contact map for this model.
 \implies (with the *KappaFactory*)

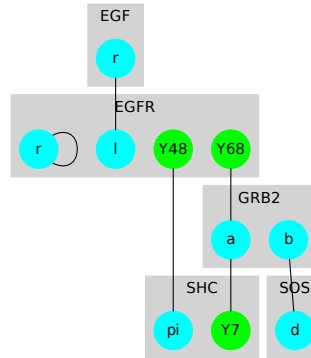


Fig. 1. Contact map

1. Import (Menu 'File' → 'Import' → 'KappaFile') the file *egfr.ka* in the Factory;
2. Select the tab 'Contact Map';
3. Click on the button 'Generate Contact Map'.

⇒ (with *complx* and *graphviz*)

1. Put *egfr.ka* in the same repository than *complx*;
2. Prompt `./complx --reset-all egfr.ka --do-high-res-contact-map --output-high-res-contact-map-dot contact_map.dot`;
3. Prompt `dot -Tpdf contact_map.dot -o contact_map.pdf`;
4. The contact map in Fig. 1 is now in the file *contact_map.pdf*.

□

In this model, some receptors *EGFR* recruit some proteins *SOS*. This can be done in two different ways:

- a receptor *EGFR* can be bound to a protein *GRB2* that is bound to a protein *SOS* (short arm);
- a receptor *EGFR* can be bound to a protein *SHC* that is bound to a protein *GRB2*, that is bound to a protein *SOS* (long arm).

(ii) Compute the *stories* that describe how a receptor *EGFR* can recruit a protein *SOS*.

⇒ (with the *KappaFactory*)

1. Select the tab 'Stories';
2. Check the two rules with the name 'short' and the name 'long';
3. Select the option 'Weak' compression;

4. Increase the number of events to 1000000 and the number of iterations to 50;
5. Click on the button 'Run stories'.

⇒ (with `simplx` and `graphviz`)

1. Put `egfr.ka` in the same repository than `simplx`;
2. Prompt `./simplx --cflow egfr.ka --iteration 50 --event 1000000 --dot-output`;
3. Prompt `ls h*.dot`;
4. For any obtained file, say `h0.dot`, prompt `dot -Tpdf h0.dot -o h0.pdf`, so as to convert the story in pdf.

We give in Fig. 2 the two possible stories (since the simulation is stochastic, it may happen that only one story is found during a set of simulations). □

(iii) Compute the number of reachable species.

There are 356 reachable species.

⇒ (with the `KappaFactory`)

1. Select the tab 'Reachables';
2. Check the box 'Enumerate Complexes';
3. Click on the button 'Generate Reachables'.

⇒ (with `complx`)

1. Put `egfr.ka` in the same repository than `complx`;
2. Prompt `./complx egfr.ka --reset-all --enumerate-complexes --output-reachable-complexes species.txt`;
3. Prompt `tail species.txt`.

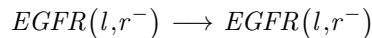
□

(iv) Is it possible for a receptor *EGFR* to be bound to another receptor *EGFR* while not being bound to a ligand *EGF* ?

No, because on the first hand a receptor can bind another receptor only if these two receptors are activated (*r03*) and on the second hand the bond between a ligand and a receptor can be released only if the dimerisation site is free (*r01*).

Let us use the engine to prove this fact.

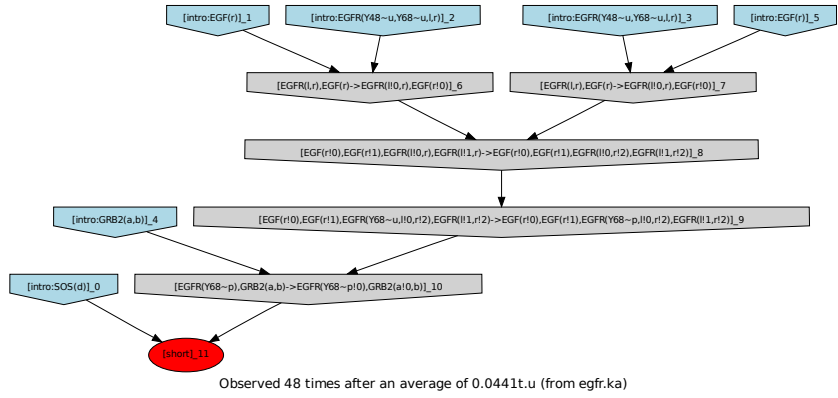
Let us add the following silent rule:



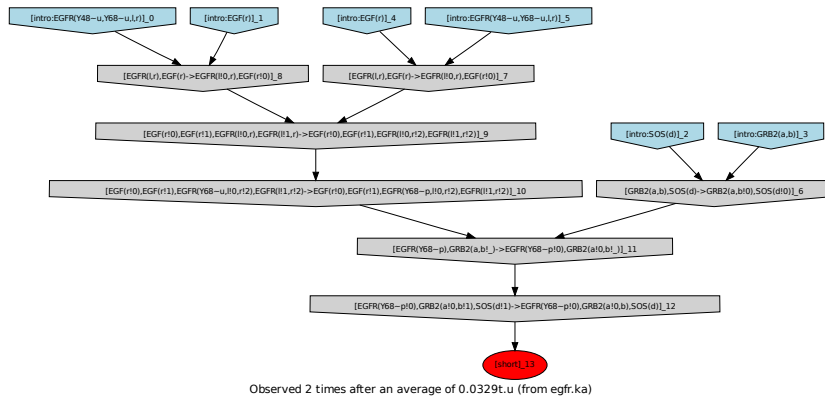
to the kappa file `egfr.ka`.

⇒ (with the `KappaFactory`)

1. Import (Menu 'File' → 'Import' → 'KappaFile') the file `egfr.ka` in the Factory;
2. Select the tab 'Compression';



(a) short arm



(b) long arm

Fig. 2. Stories

3. Select the option 'quantitative' compression;
4. Click on the button 'Perform Compression'.

The added rule is displayed in red which means that its left hand side never occurs at run time (which was what we wanted to prove).

⇒ (with `complx`)

1. Prompt `./complx egfr.ka > Log`;
2. Prompt `grep applied Log`;

The analyzer has proved that the silent rule can never be applied.

3 Internal coarse graining

(i) Use the software to compute the reduced differential systems that is associated with our rule set. How many fragments are there?

⇒ (with `complx`)

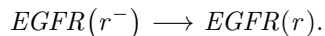
1. Prompt `./complx egfr.ka --reset-all --do-ODE --output-scheme egfr --final-time 6 ;`
2. Prompt `ls *kappa_ODE*`;
3. Prompt `less egfr_kappa_ODE.system_size.m.`

There are 83 fragments.

□

(ii) Argue that the rule `r04` can be replaced with a simpler rule without modifying the behavior of the model.

According to the contact map (Fig. 1), the site r of a receptor can only be bound to the site r of another receptor. Moreover, according to question (2.iv), whenever the site r of a receptor is bound, the site l of the same receptor is also bound (to a ligand according to the contact map (Fig. 1)). As a consequence, whenever the site r of a receptor is bound, we have two receptors bounds to two ligands and a bond between the two receptors. Thus the rule can be replaced with the following one:



□

(iii) Use the software to compress the rule set.

⇒ (with the `KappaFactory`)

1. Select the tab 'Compression';
2. Select the option 'quantitative' compression;
3. Click on the button 'Perform Compression'.

This output cannot be bootstrapped into the factory.

⇒ (with `complx`)

```
1. Prompt ./complx --reset-all --compute-quantitative-compression egfr.ka
--output-quantitative-compression egfr_compressed.ka
```

A new file `egfr_compressed.ka` is created. This file contains the following compressed rules:

```
rc01:  $EGF(r) , EGFR(l) \longrightarrow EGF(r^1) , EGFR(l^1)$ 
rc02:  $EGFR(l^-,r) \longrightarrow EGFR(l,r)$ 
rc03:  $EGFR(l^-,r) , EGFR(l^-,r) \longrightarrow EGFR(l^-,r^1) , EGFR(l^-,r^1)$ 
rc04:  $EGFR(r^-) \longrightarrow EGFR(r) @0.5$ 
rc05:  $EGFR(Y68_u^?,r^-) \longrightarrow EGFR(Y68_p^?,r^-)$ 
rc06:  $EGFR(Y68_p) \longrightarrow EGFR(Y68_u)$ 
rc07:  $EGFR(Y48_u^?,r^-) \longrightarrow EGFR(Y48_p^?,r^-)$ 
rc08:  $EGFR(Y48_p) \longrightarrow EGFR(Y48_u)$ 
rc09:  $EGFR(Y48^1,r^-) , SHC(Y7_u^?,pi^1) \longrightarrow EGFR(Y48^1,r^-) , SHC(Y7_p^?,pi^1)$ 
rc10:  $SHC(Y7_p,pi^-) \longrightarrow SHC(Y7_u,pi^-)$ 
rc11:  $SHC(Y7_p,pi) \longrightarrow SHC(Y7_u,pi)$ 
rc12:  $EGFR(Y68_p) , GRB2(a,b) \longrightarrow EGFR(Y68_p^1) , GRB2(a^1,b)$ 
rc13:  $EGFR(Y68^1) , GRB2(a^1,b) \longrightarrow EGFR(Y68) , GRB2(a,b)$ 
rc14:  $EGFR(Y68_p) , GRB2(a,b^-) \longrightarrow EGFR(Y68_p^1) , GRB2(a^1,b^-)$ 
rc15:  $EGFR(Y68^1) , GRB2(a^1,b^-) \longrightarrow EGFR(Y68) , GRB2(a,b^-)$ 
rc16:  $EGFR(Y68^1) , GRB2(a^1,b) , SOS(d) \longrightarrow EGFR(Y68^2) , GRB2(a^2,b^1) , SOS(d^1)$ 
rc17:  $EGFR(Y68^1) , GRB2(a^1,b^-) \longrightarrow EGFR(Y68^1) , GRB2(a^1,b)$ 
rc18:  $GRB2(a,b) , SOS(d) \longrightarrow GRB2(a,b^1) , SOS(d^1)$ 
rc19:  $GRB2(a,b^-) \longrightarrow GRB2(a,b)$ 
rc20:  $GRB2(a^1,b) , SHC(Y7^1,pi) , SOS(d) \longrightarrow GRB2(a^2,b^1) , SHC(Y7^2,pi) , SOS(d^1)$ 
rc21:  $GRB2(a^1,b^-) , SHC(Y7^1,pi) \longrightarrow GRB2(a^1,b) , SHC(Y7^1,pi)$ 
rc22:  $GRB2(a^1,b) , SHC(Y7^1,pi^-) , SOS(d) \longrightarrow GRB2(a^2,b^1) , SHC(Y7^2,pi^-) , SOS(d^1)$ 
rc23:  $GRB2(a^1,b^-) , SHC(Y7^1,pi^-) \longrightarrow GRB2(a^1,b) , SHC(Y7^1,pi^-)$ 
rc24:  $EGFR(Y48_p) , SHC(Y7_u^?,pi) \longrightarrow EGFR(Y48_p^1) , SHC(Y7_u^?,pi^1)$ 
rc25:  $SHC(Y7_u^?,pi^-) \longrightarrow SHC(Y7_u^?,pi)$ 
rc26:  $EGFR(Y48_p) , SHC(Y7_p,pi) \longrightarrow EGFR(Y48_p^1) , SHC(Y7_p,pi^1)$ 
rc27:  $SHC(Y7_p,pi^-) \longrightarrow SHC(Y7_p,pi)$ 
rc28:  $EGFR(Y48_p) , GRB2(a^1,b) , SHC(Y7^1,pi) \longrightarrow EGFR(Y48_p^2) , GRB2(a^1,b) , SHC(Y7^1,pi^2)$ 
rc29:  $GRB2(a^1,b) , SHC(Y7^1,pi^-) \longrightarrow GRB2(a^1,b) , SHC(Y7^1,pi)$ 
rc30:  $EGFR(Y48_p) , GRB2(a^1,b^-) , SHC(Y7^1,pi) \longrightarrow EGFR(Y48_p^2) , GRB2(a^1,b^-) , SHC(Y7^1,pi^2)$ 
rc31:  $GRB2(a^1,b^-) , SHC(Y7^1,pi^-) \longrightarrow GRB2(a^1,b^-) , SHC(Y7^1,pi)$ 
rc32:  $GRB2(a,b) , SHC(Y7_p,pi^-) \longrightarrow GRB2(a^1,b) , SHC(Y7_p^1,pi^-)$ 
rc33:  $GRB2(a^1,b) , SHC(Y7^1,pi^-) \longrightarrow GRB2(a,b) , SHC(Y7,pi^-)$ 
```

- rc34: $GRB2(a,b), SHC(Y7_p,pi) \longrightarrow GRB2(a^1,b), SHC(Y7_p^1,pi)$
rc35: $GRB2(a^1,b), SHC(Y7^1,pi) \longrightarrow GRB2(a,b), SHC(Y7,pi)$
rc36: $GRB2(a,b^-), SHC(Y7_p,pi) \longrightarrow GRB2(a^1,b^-), SHC(Y7_p^1,pi)$
rc37: $GRB2(a^1,b^-), SHC(Y7^1,pi) \longrightarrow GRB2(a,b^-), SHC(Y7,pi)$
rc38: $GRB2(a,b^-), SHC(Y7_p,pi^-) \longrightarrow GRB2(a^1,b^-), SHC(Y7_p^1,pi^-)$
rc39: $GRB2(a^1,b^-), SHC(Y7^1,pi^-) \longrightarrow GRB2(a,b^-), SHC(Y7,pi^-)$

□.

(iv) Explain the kinetic factor of the rule r04 in the compressed rule set.

Let us denote by k the kinetic rate of the rule r04 and by k' the kinetic rate of the rule rc04. Let x be a mixture. We know that the number of embeddings of the left hand side of r04 into the mixture x and the number of embeddings of the left hand side of rc04 into the mixture x are the same (this is a property of quantitative compression). Let us denote this number by $n(x)$ (we assume that $n(x) > 0$). Then, we want these two rules to have the same activity, which gives the following equations:

$$\frac{k \times n(x)}{|Aut(lhs(r04))|} = activity = \frac{k' \times n(x)}{|Aut(lhs(rc04))|}$$

Then:

$$\frac{k \times n(x)}{2} = \frac{k' \times n(x)}{1}$$

And:

$$k' = \frac{k}{2}.$$

□

(v) Use the software to compute the reduced differential systems that is associated with compressed rule set. How many fragments are there?

⇒ (with complx)

1. Prompt `./complx egfr_compressed.ka --reset-all --do-ODE --output-scheme egfr_compressed --final-time 6;`
2. Prompt `ls *compressed_kappa_ODE*;`
3. Prompt `less egfr_compressed_kappa_ODE.system.size.m.`

There are 38 fragments.

□

The reduction used the annotated contact map that is given in Fig. 3.

(vi) Motivate each solid/strong edge and each covering class.

– *solid edges:*

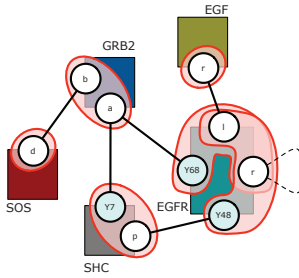


Fig. 3. annotated contact map.

- the edge between *EGF* and *EGFR* is solid because of the rule *rc02* (it can be released by a rule testing something else);
 - the edge between *EGFR* and *GRB2* is solid because of the rule *rc13* (or equivalently *rc15*, *rc16*, or *rc17*);
 - the edge between *EGFR* and *SHC* is solid because of the rule *rc09* (or equivalently *rc25*, *rc29*, or *rc31*);
 - the edge between *GRB2* and *SOS* is solid because of the rule *rc17* (or equivalently *rc19*, *rc21*, or *rc23*);
 - the edge between *SHC* and *GRB2* is solid because of the rule *rc20* (or equivalently *rc21*, *rc22*, *rc23*, *rc28*, *rc29*, *rc30*, *rc31*, *rc33*, *rc35*, *rc37*, *rc39*);
- covering classes:
- *EGFR*
 - * any class containing the site *l* should contain the site *r* because of the rule *rc02*;
 - * any class containing the site *r* should contain the site *l*, because of the rule *rc03*;
 - * any class containing the site *Y48* should contain the site *r*, because of the rule *rc07* (or equivalently *rc09*);
 - * any class containing the site *Y68* should contain the site *r*, because of the rule *rc05*;
 - *GRB2*
 - * any class containing the site *a* should contain the site *b*, because of the rule *rc12* (or equivalently *rc13*, *rc14*, *rc15*, *rc16*, *rc28*, *rc29*, *rc30*, *rc31*, *rc32*, *rc33*, *rc34*, *rc35*, *rc36*, *rc37*, *rc38*, or *rc39*);
 - * any class containing the site *b* should contain the site *a*, because of the rule *rc16* (or equivalently *rc17*, *rc18*, *rc19*, *rc20*, *rc21*, *rc22*, *rc23*);
 - *SHC*
 - * any class containing the site *Y7* should contain the site *pi*, because of the rule *rc09* (or equivalently *rc10*, *rc11*, *rc20*, *rc21*, *rc22*, *rc23*, *rc32*, *rc33*, *rc34*, *rc35*, *rc36*, *rc37*, *rc38*, or *rc39*);

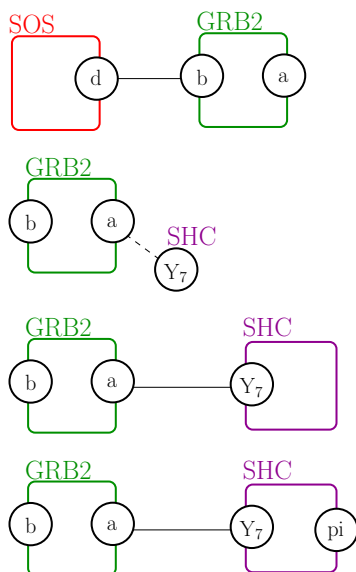


Fig. 4. fragment candidates

* any class containing the site pi should contain the site $Y7$, because of the rule $rc24$ (or equivalently $rc25$, $rc26$, $rc27$, $rc28$, $rc29$, $rc30$, or $rc31$).

That is all the constraints that we have.

□

(vii) Which subspecies in Fig. 4 are fragments with respect to the annotated contact map given Fig. 3 ?

1. The first one because (i) the bond is solid in the annotated contact map as well and (ii) the set $\{a, b\}$ forms a covering class;
2. the second one is not a fragment, because the edge between the site a of $GRB2$ and the site $Y7$ of SHC is solid in the annotated contact map;
3. the third one is not a fragment because the set $\{Y7\}$ is not a covering class;
4. the fourth one is a fragment because (i) the bond is solid in the annotated contact map as well, (ii) the set $\{a, b\}$ forms a covering class, and (iii) the set $\{Y7, pi\}$ forms a covering class.

□

(viii) Explain why, with the initial rule set, the edge between the site r of the agent $EGFR$ and itself, should be solid/strong.

The edge should have been solid due to the rule $r04$ (or equivalently $r05$, or $r07$). All these rules have been simplified by the rule compression.

(ix) Plot the concentration of the proteins *SOS* that are attached to a receptor according to the simulation time. □

⇒ (with octave (with the package odepkg) and gnuplot)

1. Prompt octave `egfr_compressed_kappa_ODE_system.m`;
2. Prompt gnuplot `egfr_compressed_kappa_ODE.gplot`.

The file `egfr_compressed_kappa_ODE.data`, in Fig. 5 is created. □

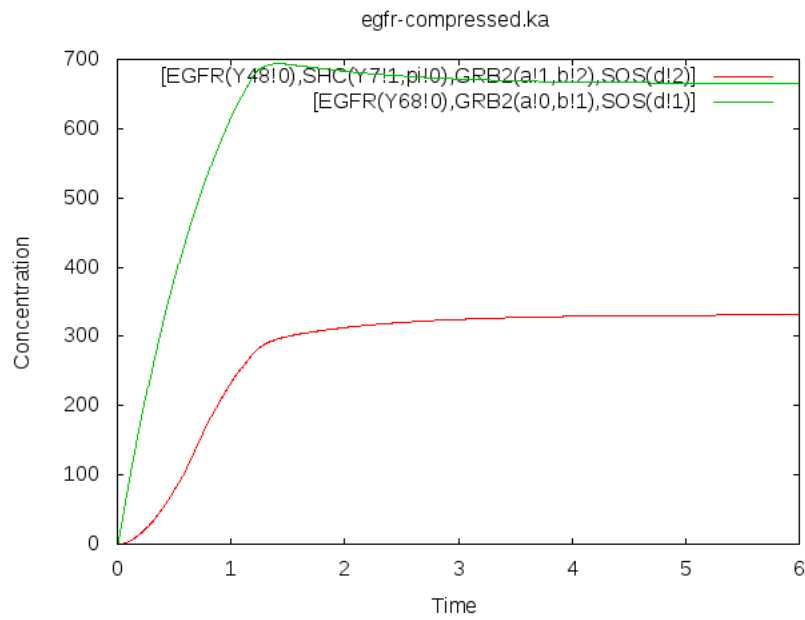


Fig. 5. Concentrations of SOS attached to the membrane with respect to simulation time. □