HYPE

Hybrid method for parameter estimation
In biochemical models

La question …

Les réseaux biologiques sont dynamiques.

Expérimentation :
- déterminer l'ensemble des molécules et des relations
- mesure de certaines molécules en fonction du temps

L'ensemble des molécules et des relations permettent de trouver la topologie du réseau.

Un grand nombre de paramètres restent inaccessibles à l'expérience.

Il faut estimer ces paramètres à partir des mesures dynamiques et de la topologie du réseau : estimation de paramètres.

La question …

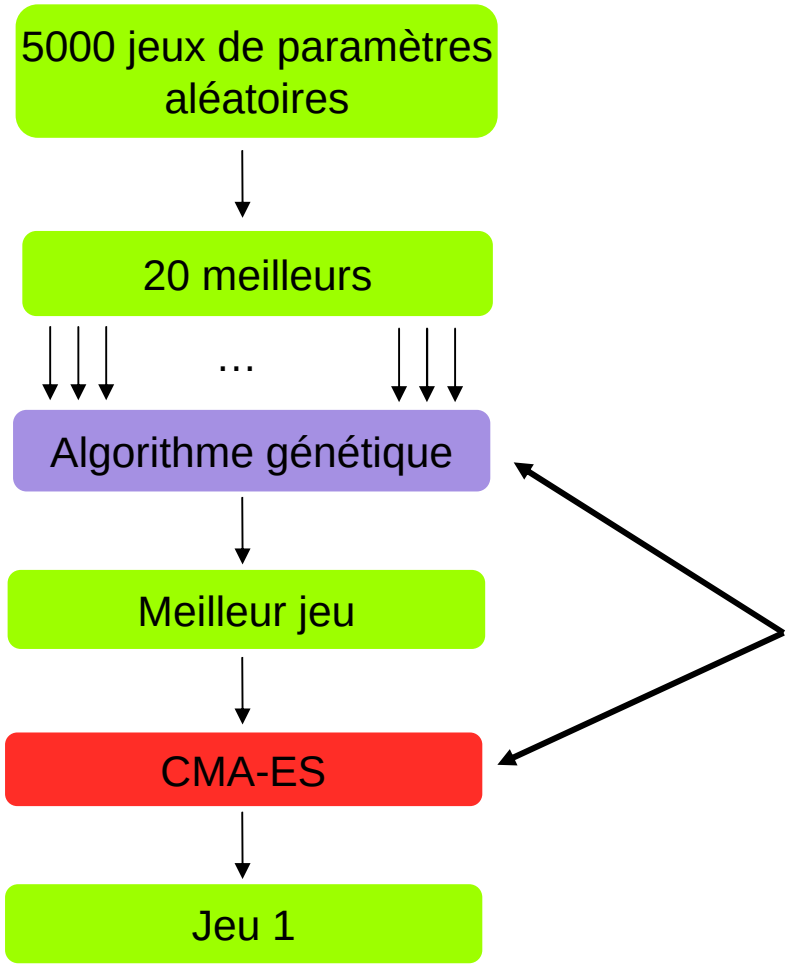
Pour paramétrer les modèles biologiques, nous avons besoin d'une méthode d'estimation des paramètres :

• rapide, pour pouvoir explorer différentes topologies de modèles

• robuste, on veut être surs que le jeu de paramètre trouvé est le bon, et qu'il est unique

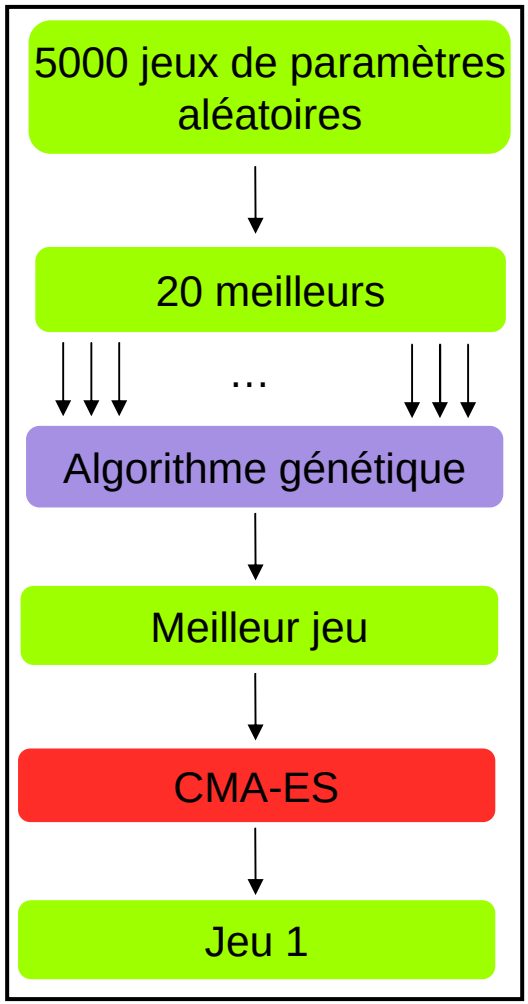
• flexible, pour pouvoir utiliser différents types de données : dose-réponses, séries temporelles, mesures relatives, etc.

Méthode d’optimisation …



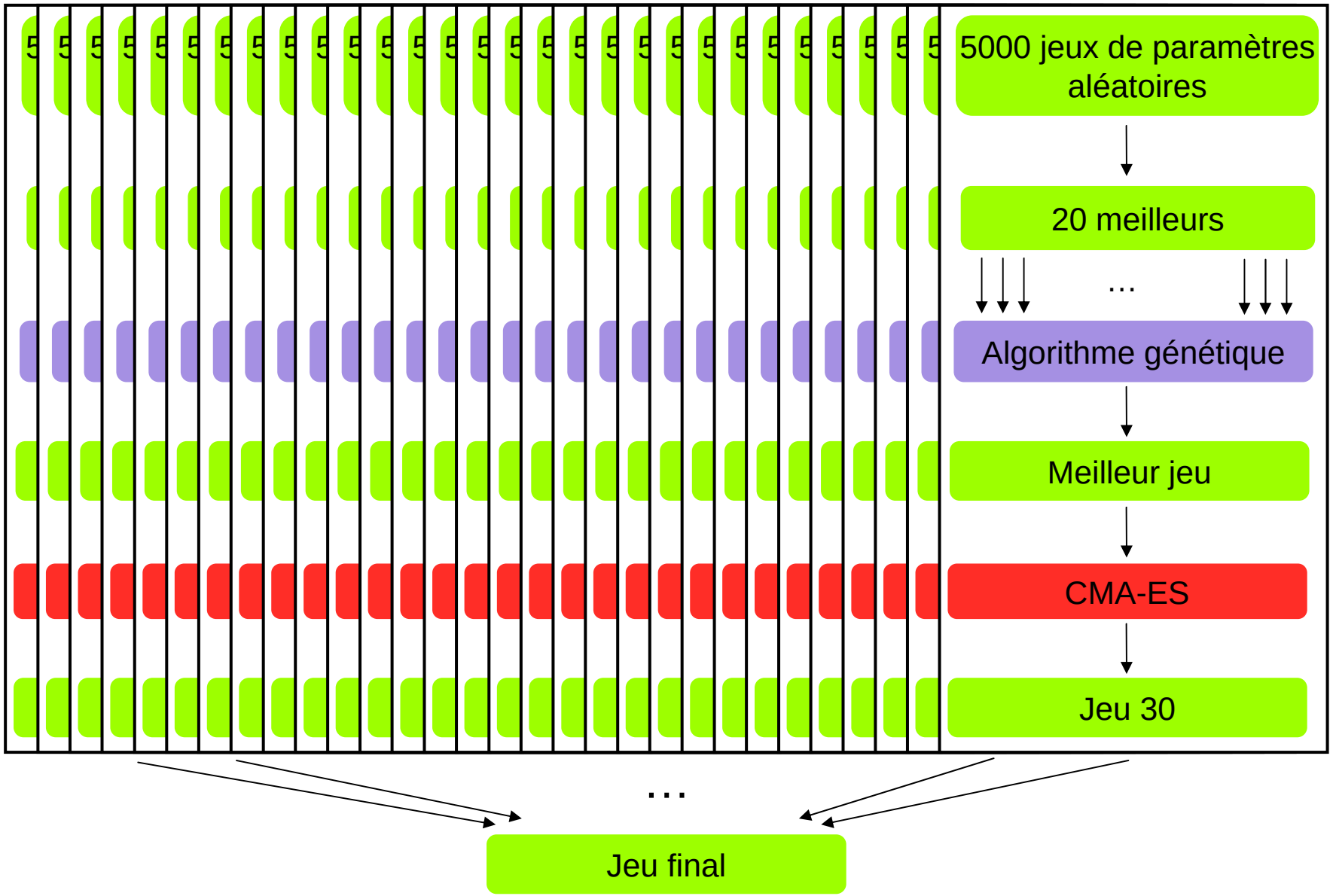
2 méthodes d’optimisation
- algorithme génétique : recherche globale
- CMA-ES : recherche locale

Méthode d’optimisation …



1 run

Méthode d'optimisation …



Fonction objectif …

La fonction objectif est ce que l'on cherche à minimiser.

$$f(P) = \sum_{m \in M} \left[\frac{1}{nt^m} \sum_{t \leq nt^m} \frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)} . \left(1 - exp \left(-\frac{1}{2} \frac{\frac{\left| d_m(t) - s_m^P(t) \right|^2}{d_m(t)}}{\sigma_m(t)^2} \right) \right) \right]$$

Fonction objectif …

La fonction objectif est ce que l'on cherche à minimiser.

$$f(P) = \frac{1}{n_m} \sum_{m \in M} \left[\frac{1}{nt^m} \sum_{t \leq nt^m} \frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)} . \left(1 - exp \left(-\frac{1}{2} \frac{\frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)}^2}{\sigma_m(t)^2} \right) \right) \right]$$

Ecart entre la valeur simulée
et la valeur attendue en %

Fonction objectif …

La fonction objectif est ce que l'on cherche à minimiser.

$$f(P) = \frac{1}{n_m} \sum_{m \in M} \left[\frac{1}{nt^m} \sum_{t \leq nt^m} \frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)} . \left(1 - exp \left(-\frac{1}{2} \frac{\frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)}^2}{\sigma_m(t)^2} \right) \right) \right]$$

Somme sur tous les points
pour un observable

Ecart entre la valeur simulée
et la valeur attendue en %

Fonction objectif …

La fonction objectif est ce que l'on cherche à minimiser.

$$f(P) = \frac{1}{n_m} \sum_{m \in M} \left[\frac{1}{nt^m} \sum_{t \leq nt^m} \frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)} . \left(1 - exp \left(-\frac{1}{2} \frac{\frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)}^2}{\sigma_m(t)^2} \right) \right) \right]$$

Divisé par le nombre de points pour cet observable

Somme sur tous les points pour un observable

Ecart entre la valeur simulée et la valeur attendue en %

Fonction objectif …

La fonction objectif est ce que l'on cherche à minimiser.

$$f(P) = \frac{1}{n_m} \sum_{m \in M} \left[\frac{1}{nt^m} \sum_{t \leq nt^m} \frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)} . \left(1 - exp \left(-\frac{1}{2} \frac{\frac{\left| d_m(t) - s_m^P(t) \right|^2}{d_m(t)}}{\sigma_m(t)^2} \right) \right) \right]$$

Somme sur les observables

Divisé par le nombre de points pour cet observable

Somme sur tous les points pour un observable

Ecart entre la valeur simulée et la valeur attendue en %

Fonction objectif …

La fonction objectif est ce que l'on cherche à minimiser.

$$f(P) = \frac{1}{n_m} \sum_{m \in M} \left[\frac{1}{nt^m} \sum_{t \leq nt^m} \frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)} . \left(1 - exp \left(-\frac{1}{2} \frac{\frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)}^2}{\sigma_m(t)^2} \right) \right) \right]$$

Divisé par le nombre d'observables

Somme sur les observables

Divisé par le nombre de points pour cet observable

Somme sur tous les points pour un observable

Ecart entre la valeur simulée et la valeur attendue en %

Fonction objectif …

La fonction objectif est ce que l'on cherche à minimiser.

$$f(P) = \frac{1}{n_m} \sum_{m \in M} \left[\frac{1}{nt^m} \sum_{t \leq nt^m} \frac{\left| d_m(t) - s_m^P(t) \right|}{d_m(t)} \left(1 - exp\left(-\frac{1}{2} \frac{\frac{\left| d_m(t) - s_m^P(t) \right|^2}{d_m(t)}}{\sigma_m(t)^2} \right) \right) \right]$$

Divisé par le nombre d'observables

Somme sur les observables

Divisé par le nombre de points pour cet observable

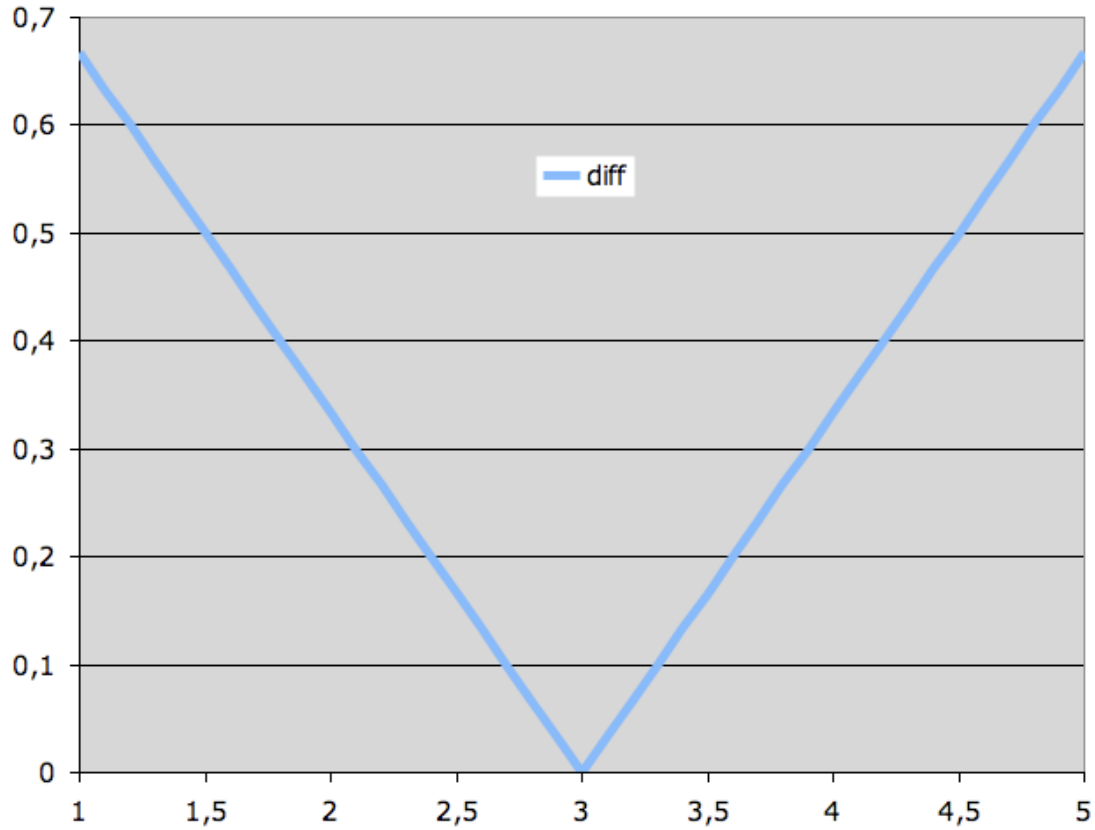
Somme sur tous les points pour un observable

Ecart entre la valeur simulée et la valeur attendue en %

Prise en compte de l'écart-type

Fonction objectif …

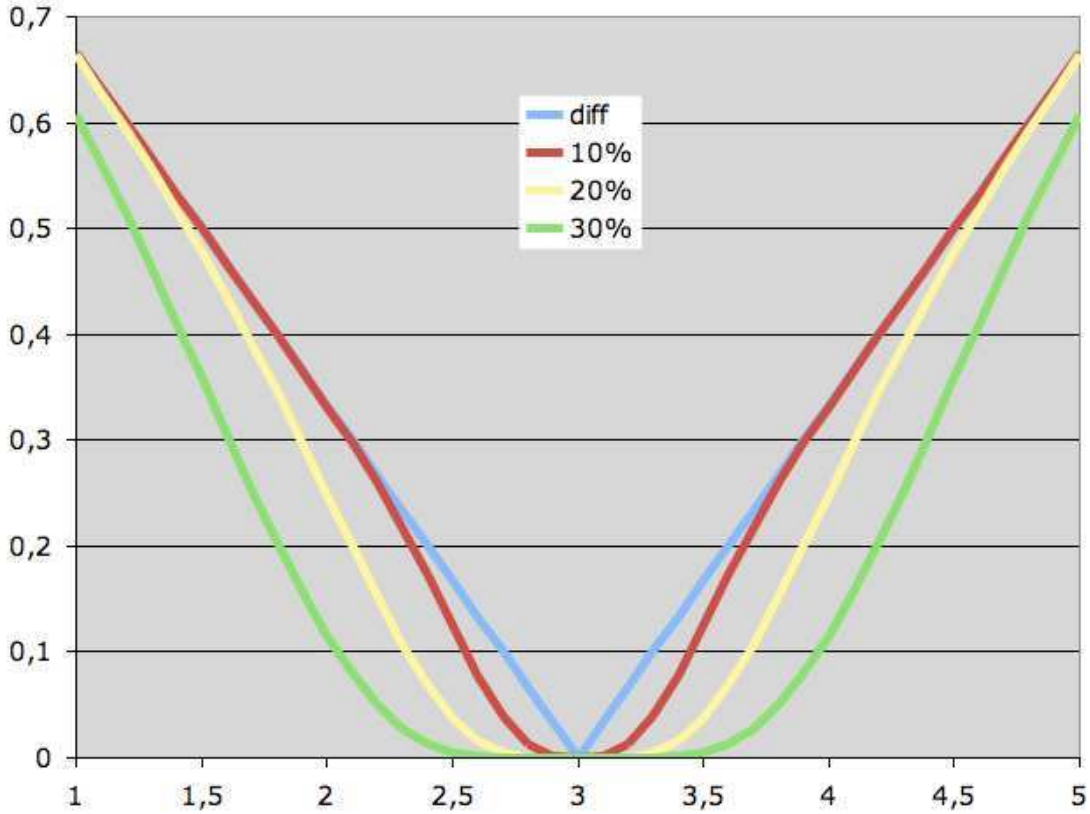
$$\frac{\left|d_m(t) - s_m^P(t)\right|}{d_m(t)}$$



Fonction objectif …

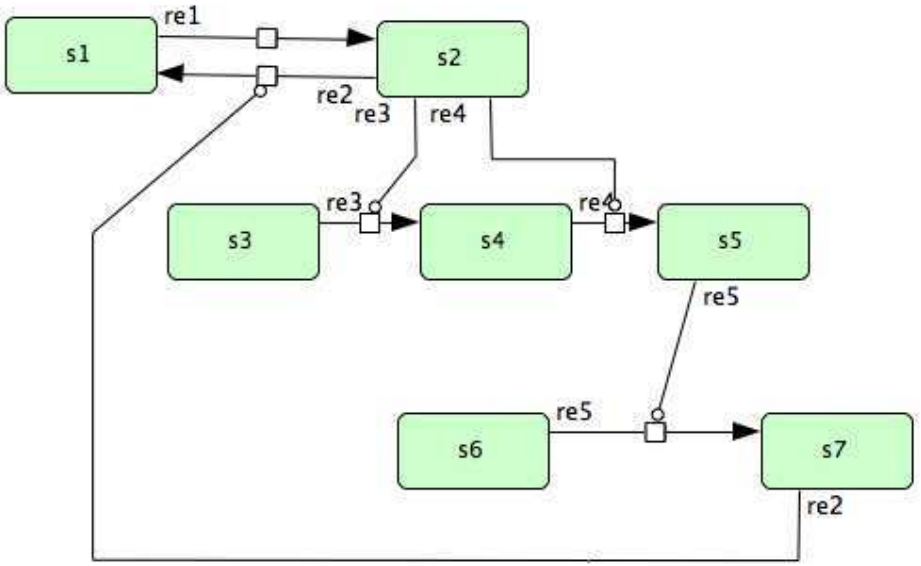
$$\frac{\left|d_m(t) - s_m^P(t)\right|}{d_m(t)}$$

$$\frac{\left|d_m(t) - s_m^P(t)\right|}{d_m(t)} \cdot \left(1 - exp\left(-\frac{1}{2}\frac{\frac{\left|d_m(t) - s_m^P(t)\right|}{d_m(t)}^2}{\sigma_m(t)^2}\right)\right)$$



Modèles de test

4 équations
8 paramètres



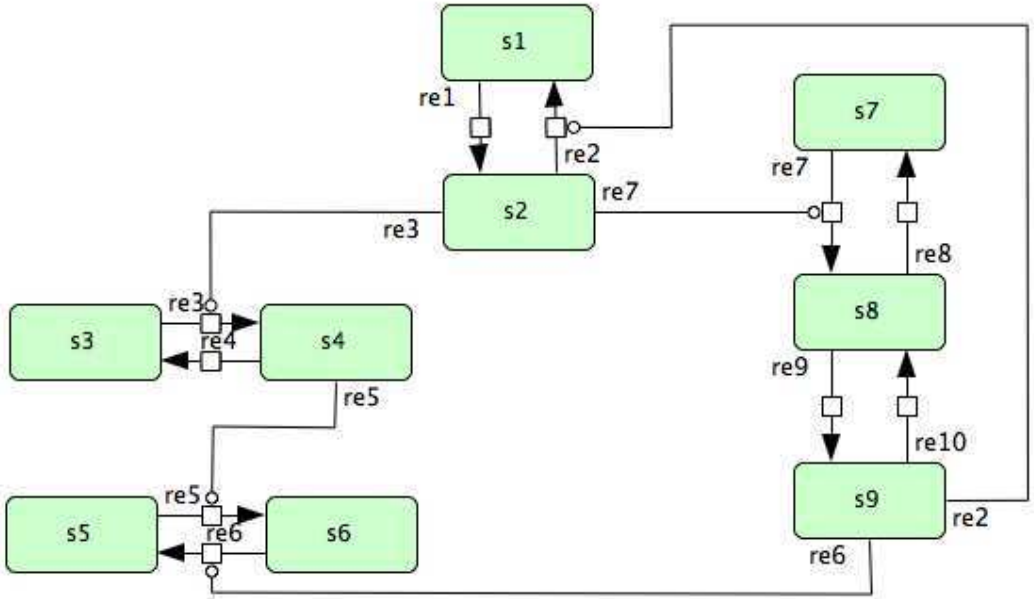
Modèle 1
tous les paramètres à 1

Modèle 2
paramètres de 2.10^{-3} à $1,28.10^{2}$
(5 logs)

Modèle 3
paramètres de 5.10^{-7} à 1.10^{5}
(12 logs)

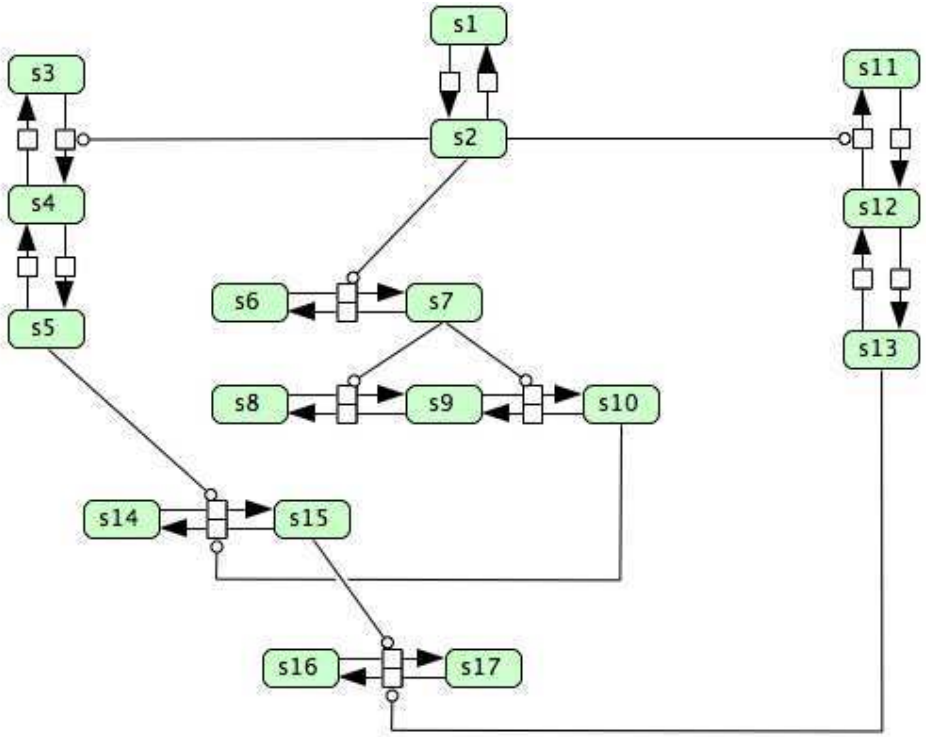
Modèles de test

Modèle 4
5 équations
14 paramètres



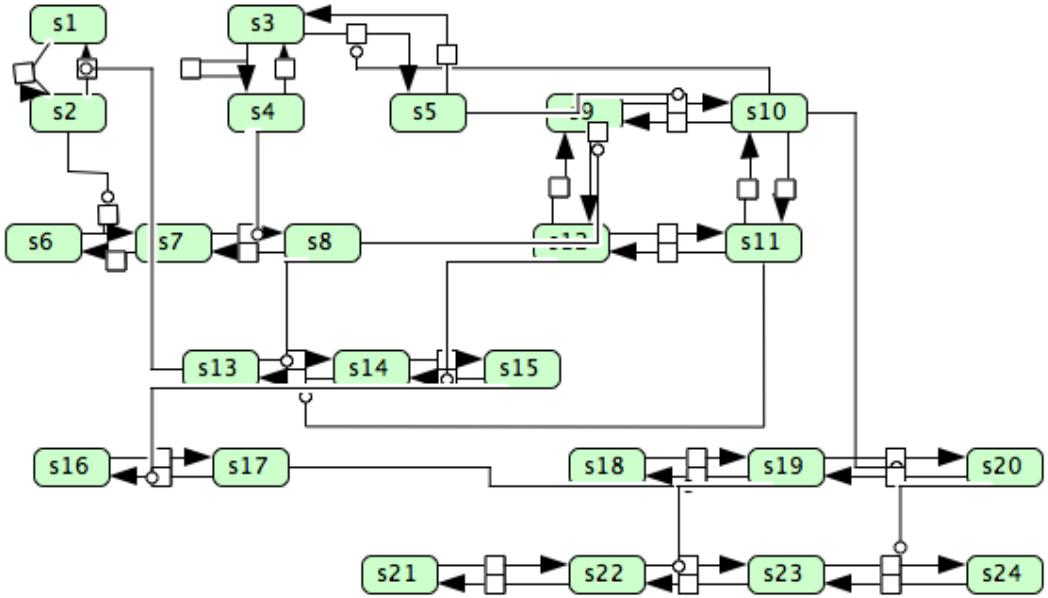
Modèles de test

Modèle 5
10 équations
27 paramètres



Modèles de test

Modèle 6
16 équations
42 paramètres



Modèles de test

Paramètres théoriques

Intégration ODE

Concentrations aux points de temps choisis

Ajout d'une perturbation (*pert*)

Donnée « expérimentales »

HYPE

Estimation des paramètres

Erreur observée

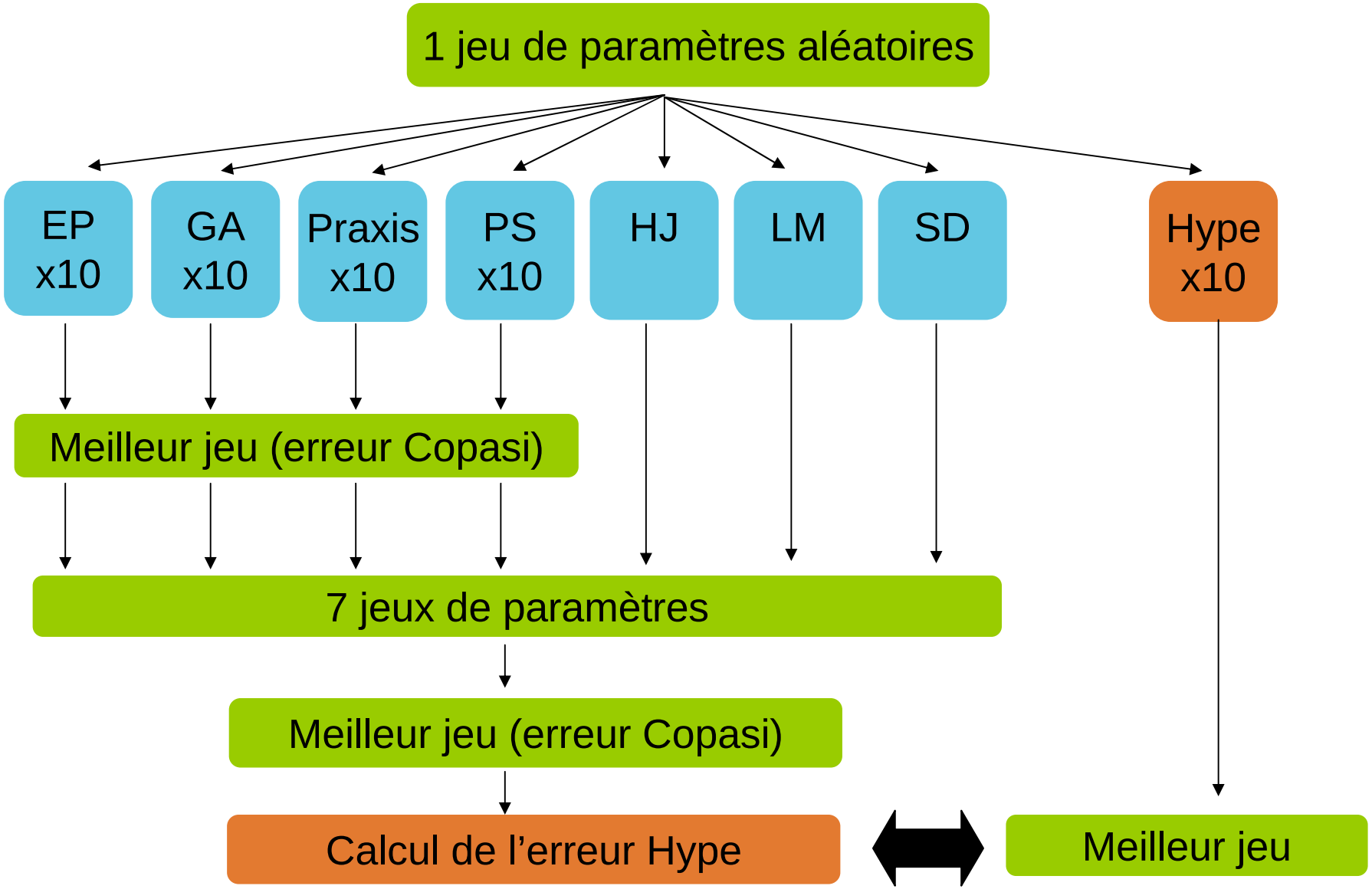
Erreur réelle

Modèles de test

Meilleur jeu sur 30 runs

	Pert = 0	Pert = 10%	Pert = 20%
Modèle 1	0	0.0043	0.0096
Modèle 2	0	0.0053	0.0277
Modèle 3	0	0.0056	0.0441
Modèle 4	0	0.0012	0.0097
Modèle 5	0.0014	0.0078	0.0319
Modèle 6			0.0314

Comparaison avec Copasi

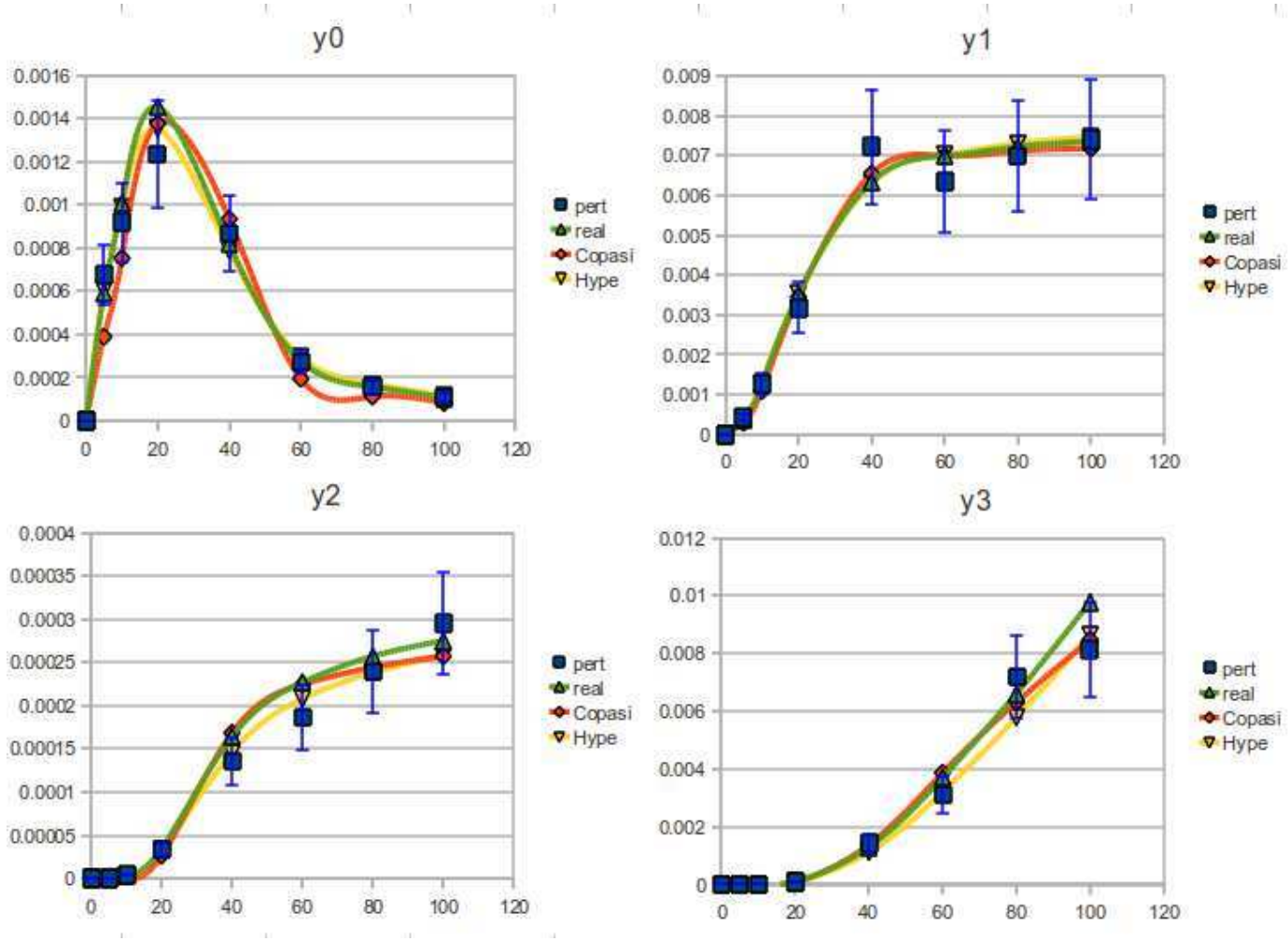


Comparaison avec Copasi

	M1	M2	M3	M4	M5
Evolutionnary programming	0.0323	6.0341	1.7981	1.3813	0.0816
Genetic algorithm - SR	0.0254	0.9625	0.7261	3.2374	0.1762
Hooke-Jeeves	0.0343	1.1516	3.7091	1.8991	0.0827
Levenberg-Marquardt	0.0316	0.4133	0.9943	1.6996	0.0779
Praxis	0.0316	1.4917	19.895	1.2437	0.0829
Particle swarm	0.0316	0.9818	0.2895	1516	0.0696
Steepest descent	0.0276	3.2476	16.223	1.2995	0.9442
Best Copasi	0.0316	0.4133	0.2895	1.6996	0.0696
Best Hype	0.0096	0.0277	0.0447	0.0103	0.0317

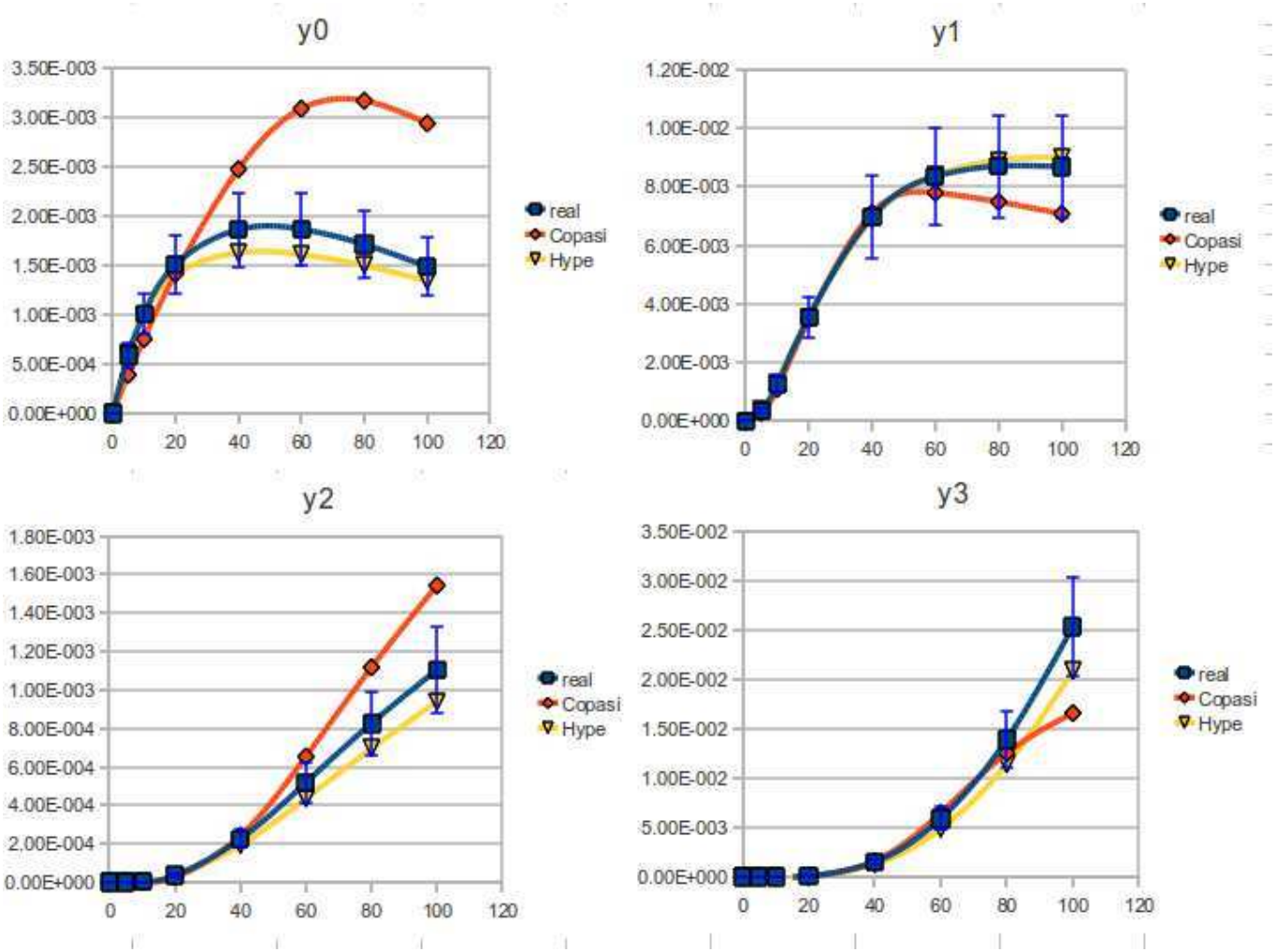
Comparaison avec Copasi

Simulation dans les conditions contrôle



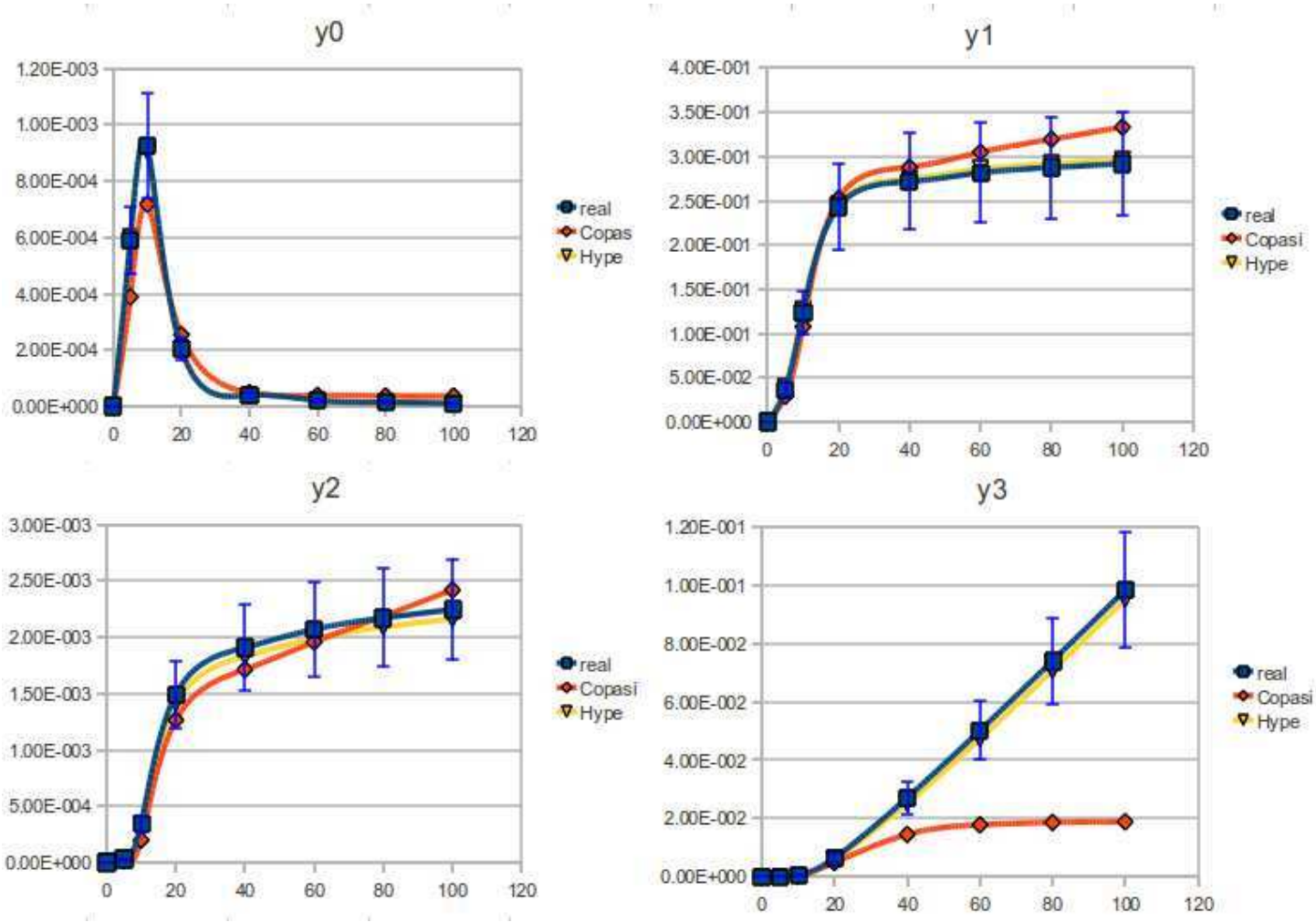
Comparaison avec Copasi

Simulation dans des conditions perturbées : k1 / 100

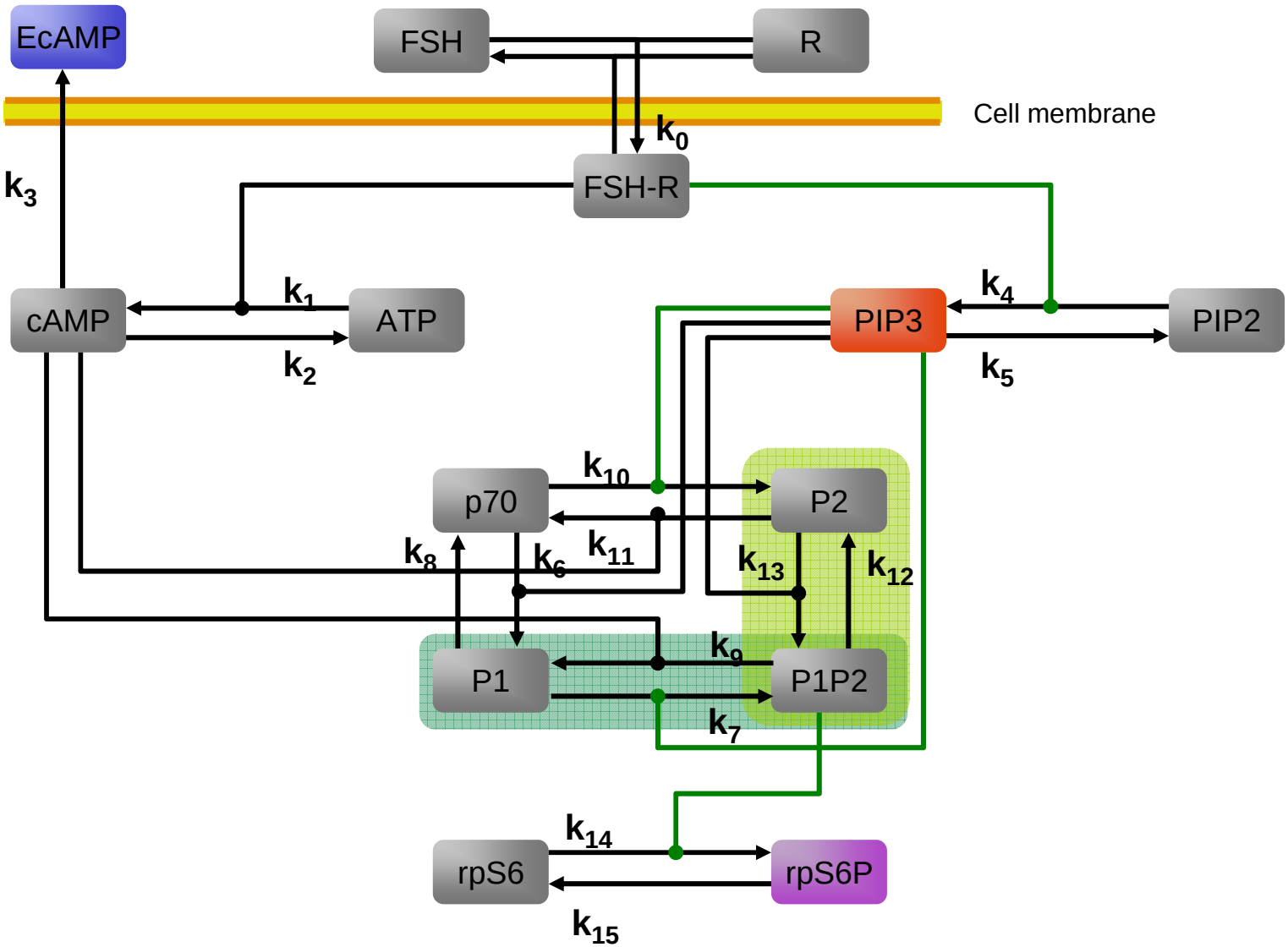


Comparaison avec Copasi

Simulation dans des conditions perturbées : k6 * 100

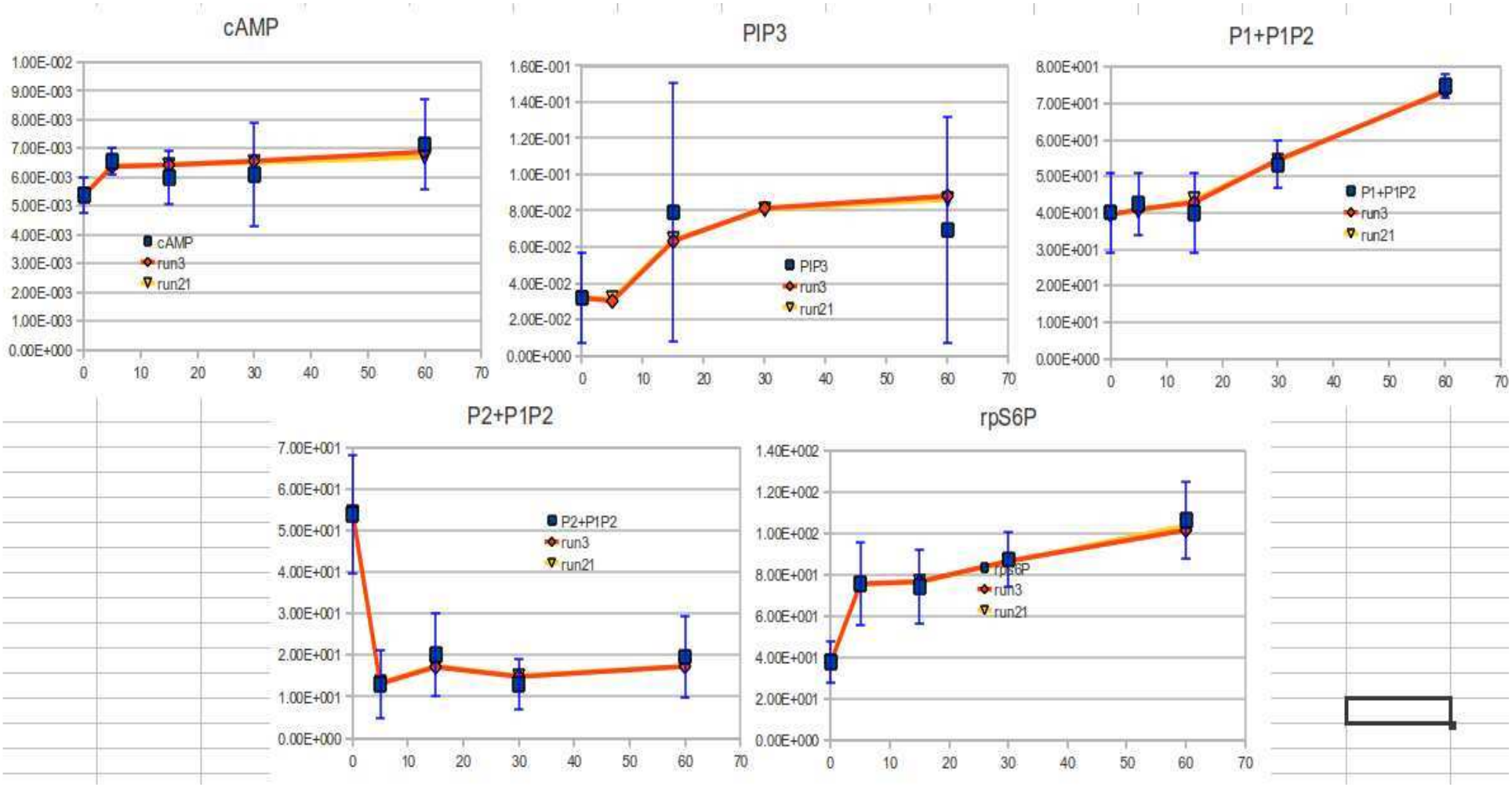


Régulation de la traduction



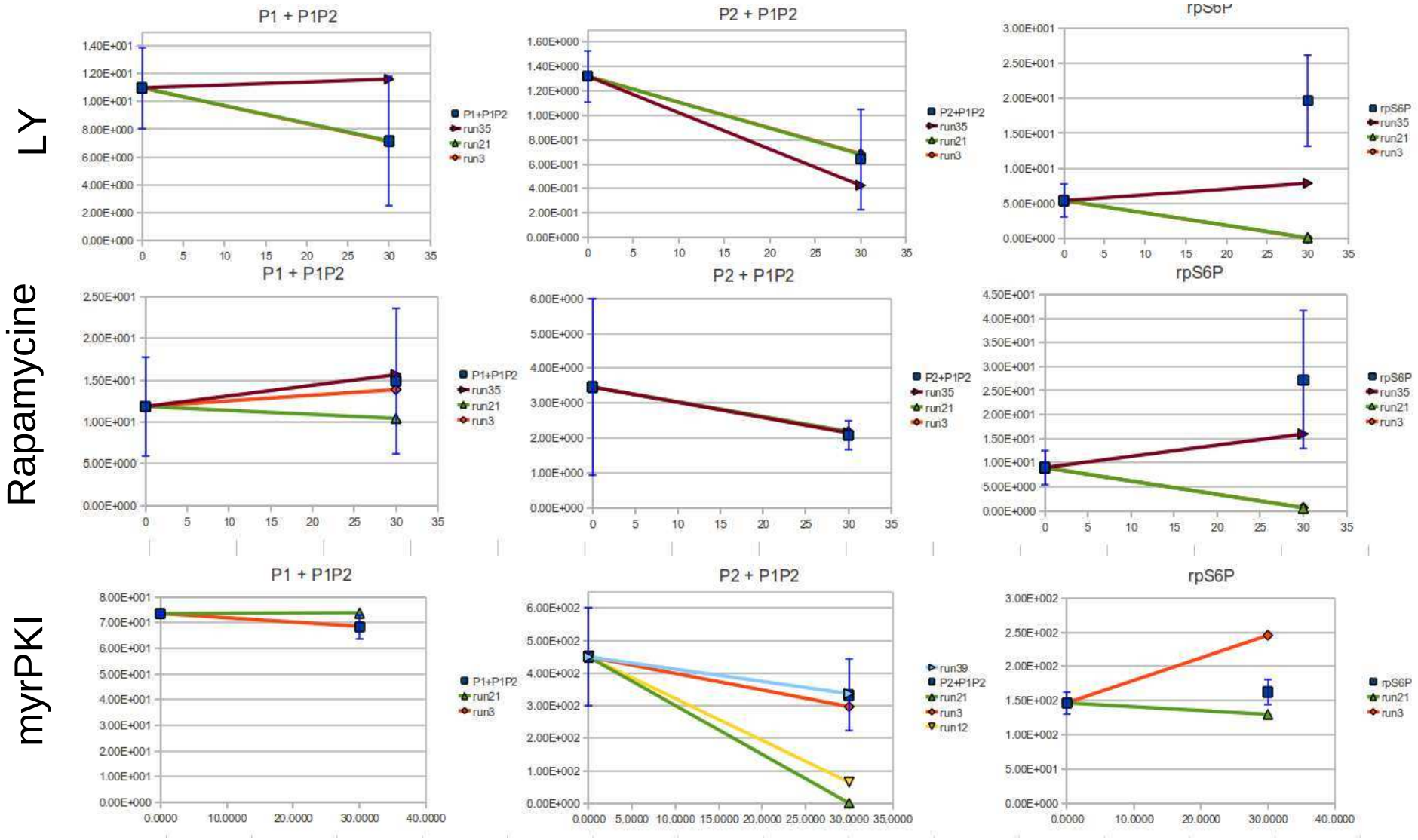
Régulation de la traduction

Erreur 0.0041

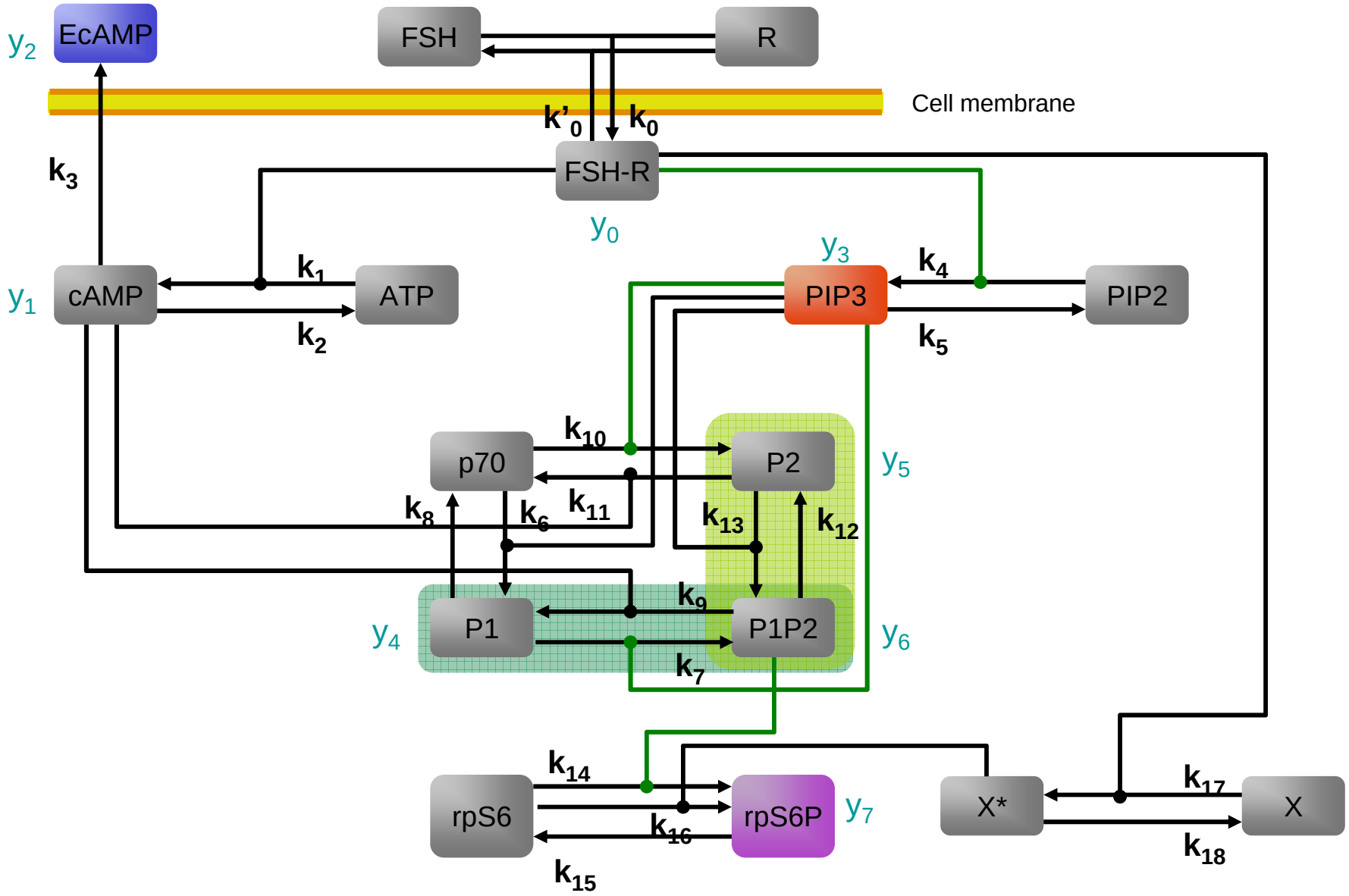


Régulation de la traduction

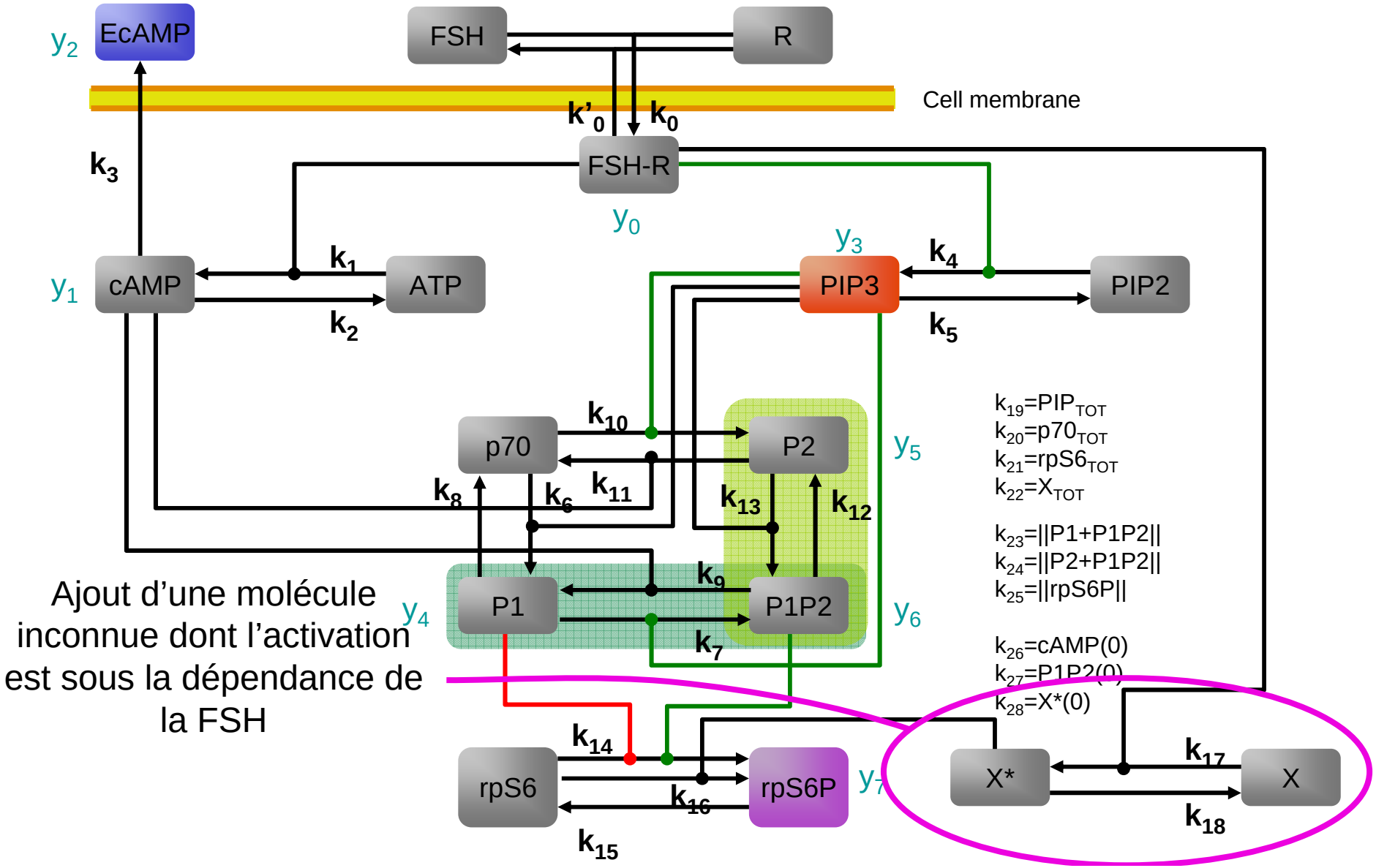
Mais… erreur = 0.3334



Régulation de la traduction

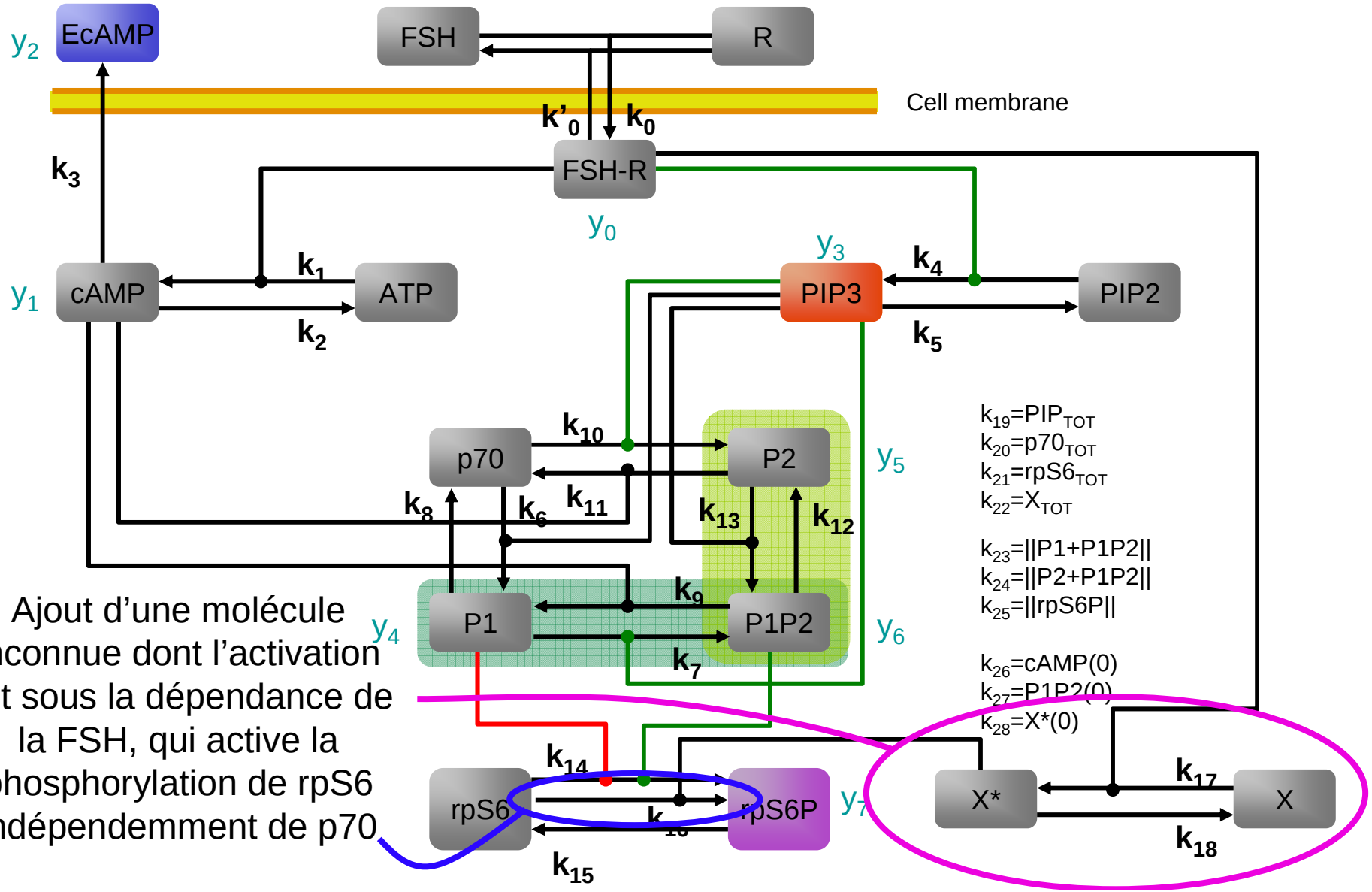


Régulation de la traduction



Ajout d'une molécule inconnue dont l'activation est sous la dépendance de la FSH

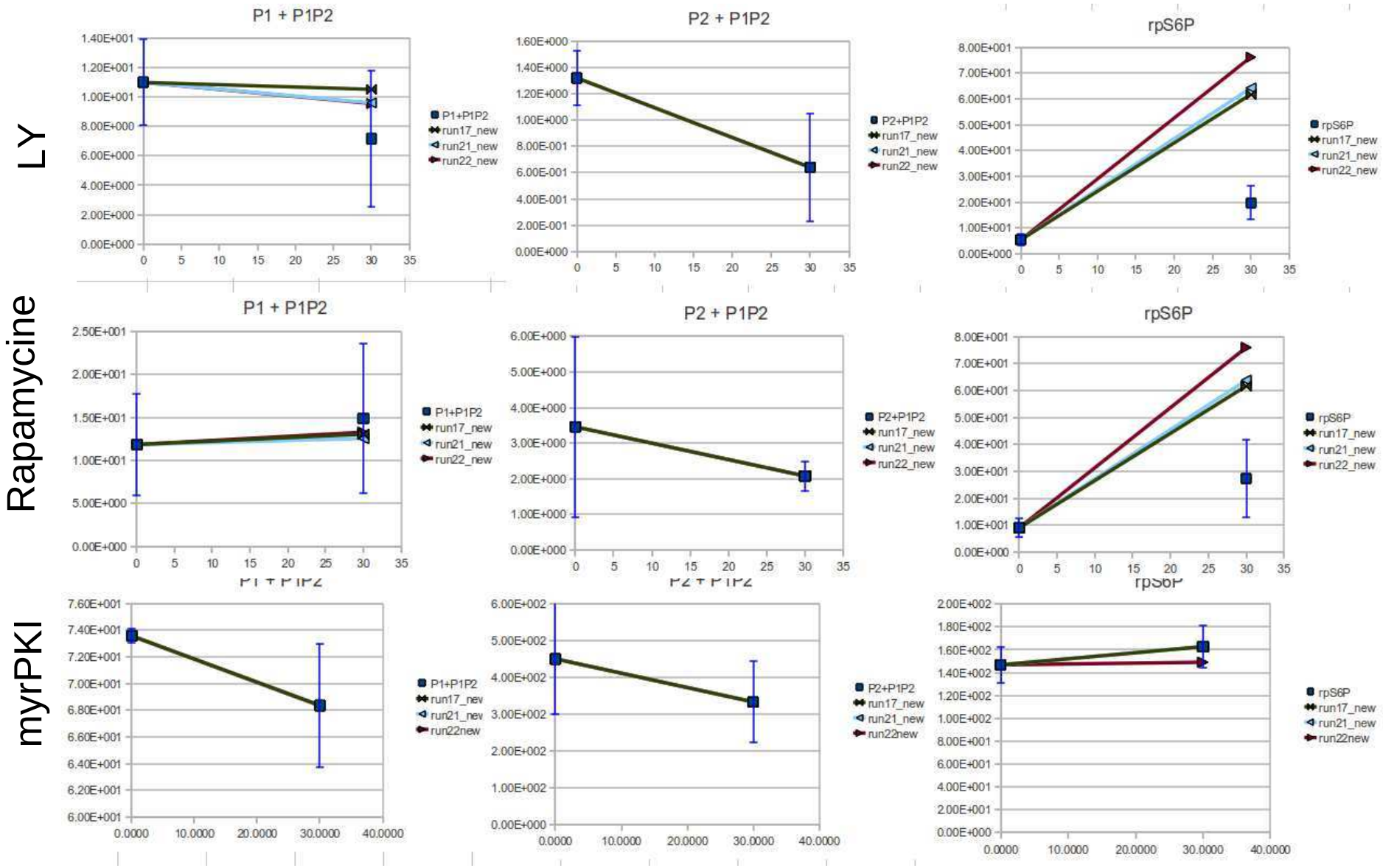
Régulation de la traduction



Ajout d'une molécule inconnue dont l'activation est sous la dépendance de la FSH, qui active la phosphorylation de rpS6 indépendemment de p70

Régulation de la traduction

erreur = 0.2458



Remerciements

Equipe BIOS

Domitille Heitzler
Eric Reiter
Pascale Crépieux
Astrid Musnier
Kelly Leon

LRI, Orsay

Jérôme Azé
Nikolaus Hansen