



**HAL**  
open science

# Modélisation dynamique des mécanismes de signalisation induits par l'hormone folliculo-stimulante et l'angiotensine

Domitille Heitzler

► **To cite this version:**

Domitille Heitzler. Modélisation dynamique des mécanismes de signalisation induits par l'hormone folliculo-stimulante et l'angiotensine. Bio-Informatique, Biologie Systémique [q-bio.QM]. Université François Rabelais - Tours, 2011. Français. NNT: . tel-00847767v2

**HAL Id: tel-00847767**

**<https://theses.hal.science/tel-00847767v2>**

Submitted on 24 Jul 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# UNIVERSITÉ FRANÇOIS RABELAIS DE TOURS



École Doctorale SST

Équipe de recherche : Biologie et Bioinformatique des Systèmes de Signalisation (BIOS)  
UMR 6175 INRA - CNRS - Université de Tours, PRC

**THÈSE** présenté par :

**Domitille HEITZLER**

soutenue le : 7 janvier 2011

pour obtenir le grade de : Docteur de l'université de Tours

Discipline : Sciences de la Vie et Santé

**Modélisation dynamique des mécanismes de signalisation  
cellulaire induits par l'hormone folliculo-stimulante et  
l'angiotensine**

THÈSE DIRIGÉE PAR :

Éric REITER

Frédérique CLÉMENT

DR INRA, Nouzilly

DR INRIA, Rocquencourt

RAPPORTEURS :

Stéphane DALLE

Christine FROIDEVAUX

CR1 INSERM, HDR, Montpellier

Prof. Université Paris Sud 11, Orsay

JURY :

Frédérique CLÉMENT

Stéphane DALLE

Anne DUITTOZ

Geneviève DUPONT

François FAGES

Christine FROIDEVAUX

Éric REITER

DR INRIA, Rocquencourt

CR1 INSERM, HDR, Montpellier

Prof. Université de Tours

Chercheur Qualifié, ULB, Bruxelles, Belgique

DR INRIA, Rocquencourt

Prof. Université Paris Sud 11, Orsay

DR INRA, Nouzilly



*à Vincent, ma fille et mon grand-père*

---

# Remerciements

Tout d'abord, je remercie mon directeur de thèse, Eric Reiter, pour son aide, son ouverture d'esprit et la perspicacité qu'il a montrée envers le projet de modélisation dans lequel ma thèse s'est inscrite. Je le remercie aussi pour son soutien auprès de la direction. François Houllier, Benoît Malpaux, Florian Guillou et Thierry Magallon, directeur général délégué, chef du département Phase, directeur de l'unité PRC et directeur adjoint de l'unité PRC, je vous remercie d'avoir renouvelé mon contrat afin de terminer ma thèse dans de bonnes conditions et de poursuivre mon projet en C qui me tient tellement à coeur.

Je remercie mes co-directeurs, Frédérique Clément et François Fages pour le sujet vraiment tri-disciplinaire dont ils sont, avec Eric, les créateurs, leur accompagnement autant que la liberté qu'ils m'ont laissée. Vincent Piketty qui m'a initiée aux expériences biologiques et plus particulièrement à la technique FRET (Fluorescence Resonance Energy Transfer). Anne Poupon et son arrivée miraculeuse dans l'équipe. Pascale Crépieux et Astrid avec qui j'ai publié mon premier modèle. Florence d'Alché-buc et Nicolas Brunel pour leurs conseils, leurs connaissances en apprentissage de paramètres et leur recul.

Mille mercis à mon époux qui m'a toujours écoutée, a su m'accompagner et me soutenir tout au long de mon apprentissage de la programmation mais aussi trouver les mots doux nécessaires au quotidien. Sans oublier Noémie et ses grands sourires au petit matin, mon petit frère et ses plaisanteries, ma cousine Amandine loin en distance et tellement proche, Albine preuve vivante qu'avec courage et optimisme, on arrive à tout, Maud autant organisatrice de barbecue d'exception que baby-sitter dévouée, Brigitte à l'écoute quelque soit le moment, Pauline à qui il arrive toujours l'anecdote la plus drôle, Marianne pour qui, même sans moyen de locomotion Tours est à deux pas, mon papy inégalable et mes parents géniaux. Je citerai aussi les amis qui m'ont apporté durant ces trois années en compréhension, soutiens et sourires : Lilianne, Sylvie et Clotilde, Marie, Karen, Amina et Samia, Christine et Nathalie, Noémie, Morris et Chris, Anne, Nicolas, Sabine et Joseph, Sandrine, Rémy, mais aussi les personnes extérieures qui ont éclairé mon chemin vers la réussite : Marie-Pierre C., Anne C., Laurence C., Jeanne C., Louis-Marie P. et Françoise M..



# Résumé

La signalisation cellulaire induite par les récepteurs à sept domaines transmembranaires (R7TMs) contrôle les principales fonctions physiologiques des mammifères. Chez l'humain, ces récepteurs représentent la plus grande famille de gènes et sont la cible de plus de 30% des médicaments. Ces R7TMs initient de larges réseaux d'interactions complexes et variées responsables des effets physiologiques. Au cours de cette thèse, nous avons modélisé dynamiquement certains de ces réseaux afin d'en comprendre le fonctionnement et de prédire des données inaccessibles expérimentalement. Notre modélisation a utilisé le formalisme des équations différentielles ordinaires (ODE) en assimilant à chaque espèce une variable et à chaque constante cinétique un paramètre.

L'équipe BIOS s'intéresse plus particulièrement au récepteur de l'hormone folliculo-stimulante (RFSH) dont la signalisation a un rôle déterminant au sein de la fonction de reproduction. En parallèle, le récepteur angiotensine de type 1A ( $RAT_{1A}$ ), régulant la tension, est largement étudié et s'inscrit comme un R7TM modèle. C'est ainsi que les deux modélisations réalisées au cours cette thèse détaillent la signalisation du RFSH au sein des cellules de Sertoli et la signalisation du  $RAT_{1A}$  dans des cellules HEK293. Le modèle FSH est centré sur les interactions de la PKA et de la PI3K sur la phosphorylation de la p70S6K. Il rend compte de la signalisation FSH au sein des cellules de Sertoli en prolifération et différenciées en conditions contrôles, PKA inhibées et PI3K inhibées. Les données expérimentales disponibles ont été traduites en concentrations molaires et ont permis d'ajuster les paramètres de l'ODE. Ce modèle explique comment, par de simples déplacements d'équilibres et de régulations, les mêmes effecteurs peuvent avoir deux comportements très différents et induire des réponses biologiques adaptées.

Le second modèle porte sur l'activation des MAP Kinase ERK par le  $RAT_{1A}$ . Comme pour le précédent, les données ont servi à l'ajustement des paramètres du modèle qui donne l'évolution des différentes espèces en concentrations molaires. Ce modèle comporte deux voies de signalisation induisant les ERK : la voie dépendante des protéines G rapide et transitoire ainsi que la voie dépendante des  $\beta$ -arrestines plus lente et soutenue. Le modèle permet de rendre compte de l'ensemble des données disponibles. Il nous a également permis de mieux comprendre la dynamique et la régulation de ce réseau de signalisation.

Enfin, pour les deux modèles nous avons été contraints de rechercher certains paramètres manquants nécessaires. Ces ajustements paramétriques peuvent être réalisés à l'aide de différents outils. Dans une troisième partie de la thèse, nous avons donc comparé plusieurs algorithmes d'optimisation à une méthode que nous avons nous-mêmes développée. Cette nouvelle méthode hybride s'avère performante et adaptée à la paramétrisation des réseaux



---

de signalisation.

**Mots clés :** Signalisation intracellulaire, R7TM, FSH, angiotensine, modèles mathématiques, ODEs, optimisation continue.

# Abstract

Seven transmembrane receptor (7TMR) signaling controls main physiological functions of mammals. In humans, 7TMR represent the largest family of genes and are targeted by more than 30 % of drugs. 7TMR initiate large and complex networks responsible for physiological effects. In this thesis, we dynamically modelled these networks to understand their mechanisms of functioning and predict experimentally unreachable data. Our modeling, based on ODE formalism, treats each species as a variable and each kinetic rate as a parameter.

Our team is focused on FSHR-induced networks that have critical role in reproduction. In parallel, the  $AT_{1A}R$ , which regulates blood pressure, is widely studied and considered as a 7TMR model. The two models produced during this thesis detail FSHR signaling in Sertoli cells and  $AT_{1A}R$  signaling in HEK293 cells. The FSH model focuses on two pathways that regulate p70S6K phosphorylation : PKA-dependent and PI3K-dependent. This model reports FSH signaling within Sertoli cells at proliferating and differentiating stages and when PKA or PI3K are inhibited. This model explains how with tiny movements of regulations, same effectors can have two different biological responses which lead to appropriate responses.

The second model involves the hormone angiotensin and is focused on the MAP Kinase ERK activation by the receptor. The experimental data as well as « western blot » data provided were translated into molar concentrations and used to adjust the parameters of the ODE. This model has two signaling pathways inducing ERK : G-dependent which is fast and transient as opposed to  $\beta$ -arrestin-dependent which is slow and sustained. The model captures all the available data. Its analysis also led us to identify new regulatory features within the signaling network.

In the third part, we developed a new hybrid method for parameter optimization. We compared this method with a panel of already existing ones using a set of dynamical models. Our novel hybrid method is efficient and well-adapted to parametrization of signaling networks.

**Keywords :** Intracellular signaling, 7TMR, FSH, angiotensin, mathematic modeling, ODEs, continuous optimization.

---

# Table des matières

Glossaire	19
Avant-propos	23
<b>I Introduction</b>	<b>25</b>
<b>1 Introduction</b>	<b>27</b>
1.1 Les R7TMs et leur signalisation . . . . .	27
1.1.1 Structure et classification . . . . .	27
1.1.2 Signalisation . . . . .	30
1.1.2.1 La voie G . . . . .	30
1.1.2.2 Désensibilisation . . . . .	33
1.1.2.3 Mouvements intracellulaires des R7TMs . . . . .	34
1.1.2.4 Signalisation indépendante des protéines G . . . . .	35
1.1.2.5 Importance du ligand . . . . .	41
1.2 Le récepteur de la FSH . . . . .	46
1.2.1 Réseaux de signalisation de la FSH . . . . .	47
1.2.1.1 Voie $G_{\alpha s}$ /AC/AMPc/PKA . . . . .	48
1.2.1.2 Voie $G_{\alpha s}$ /PI3K/PKB . . . . .	49
1.2.1.3 Autres voies dépendantes de $G_{\alpha s}$ et $G_{\alpha i}$ . . . . .	50
1.2.1.4 Signalisation indépendante des protéines G . . . . .	50
1.2.1.5 Désensibilisation du signal FSH . . . . .	51
1.3 Le récepteur de type 1A de l'angiotensine . . . . .	54
1.3.1 Signalisation de l'AT <sub>1A</sub> . . . . .	55
1.3.1.1 Signalisation dépendante des protéines G . . . . .	55
1.3.1.2 Voies $\beta$ -arrestines . . . . .	59
1.3.2 Conclusion . . . . .	61
1.4 Vers une biologie systémique des R7TMs . . . . .	63

## TABLE DES MATIÈRES

---

1.4.1	Vers une biologie systémique de la signalisation des récepteurs couplés aux protéines G : Défis et attentes . . . . .	64
1.4.2	Conclusion . . . . .	76
1.5	Modèles dynamiques . . . . .	76
1.5.1	Modèles ODE - État de l'art . . . . .	78
1.6	Réduction de modèle . . . . .	82
1.7	Ajustement Paramétrique - État de l'art . . . . .	85
1.7.1	Méthodes déterministes . . . . .	87
1.7.2	Méthodes stochastiques globales et locales . . . . .	89
1.7.3	Méthodes évolutionnaires inspirées de la biologie . . . . .	92
1.7.4	Stratégie d'Évolution et Adaptation de la Matrice de Covariance . . . . .	96
	<b>Objectifs scientifiques</b>	<b>99</b>
	<b>II Matériels et Méthodes</b>	<b>101</b>
	<b>2 Bases expérimentales et mathématiques</b>	<b>103</b>
2.1	Expériences biologiques et Interprétations . . . . .	103
2.1.1	Dosage ELISA . . . . .	103
2.1.2	FRET : d'un quotient de fluorescence à une variable . . . . .	103
2.1.3	Des Blots aux Puces : une miniaturisation des procédés . . . . .	106
2.2	Bases mathématiques et algorithmiques . . . . .	106
2.2.1	Résolution de l'équation différentielle . . . . .	106
2.2.1.1	La démarche de base . . . . .	107
2.2.1.2	Runge Kutta . . . . .	109
2.2.1.3	Méthodes explicite, implicite et semi-implicite . . . . .	110
2.2.1.4	Pas d'intégration auto-adaptatif . . . . .	110
2.2.1.5	Rosenbrock . . . . .	111
2.2.2	Méthodes de recherches de paramètres . . . . .	111
2.2.2.1	Gradient Conjugué . . . . .	111
2.2.2.2	Important Sampling et Population Monte Carlo . . . . .	112
2.2.2.3	Algorithme génétique . . . . .	113
2.2.2.4	CMAES : Covariance Matrix Adaptation Evolution Strategy	113
2.2.3	Logiciels et environnements de programmation disponibles . . . . .	116
2.2.3.1	Biocham et la logique temporelle . . . . .	116

<b>III Résultats</b>	<b>119</b>
<b>3 Résultats et articles</b>	<b>121</b>
3.1 Modélisation dynamique du réseau de signalisation du RFSH . . . . .	121
3.1.1 La p70S6K une kinase clé pour l'activation de la traduction des protéines par la FSH . . . . .	121
3.1.2 L'apport d'un modèle dynamique . . . . .	122
3.1.3 Données complémentaires . . . . .	141
3.1.4 Conclusion et perspectives . . . . .	141
3.2 Modélisation dynamique du réseau de signalisation du RAT <sub>1A</sub> . . . . .	143
3.2.1 Propriétés émergentes dans la transduction du signal par un R7TM révélées en combinant modélisation et expériences . . . . .	146
3.2.2 Conclusion et perspectives . . . . .	203
3.3 Mise au point et évaluation d'une méthode hybride d'optimisation paramétrique	204
3.4 Conclusion . . . . .	206
<b>IV Discussion et Perspectives</b>	<b>207</b>
<b>Annexes</b>	<b>217</b>
<b>A Méthode RPPA</b>	<b>217</b>
<b>B Modèles avec Biocham</b>	<b>227</b>
B.1 Petit mémo de la syntaxe BIOCHAM . . . . .	227
B.2 Modèle détaillé de réseau de signalisation Angiotensine . . . . .	228

## TABLE DES MATIÈRES

---

# Liste des tableaux

1.1	Les R7TMs et leurs propriétés vis-à-vis des $\beta$ -arrestines . . . . .	38
1.2	Interactants des $\beta$ -arrestines au sein du réseau angiotensine . . . . .	61
1.3	Correspondance entre les réseaux de signalisations et le formalisme ODE . .	76



LISTE DES TABLEAUX

---

# Liste des figures

1.1	Structure d'un R7TM et diversité des ligands . . . . .	28
1.2	Famille de R7TM . . . . .	29
1.3	Cycle de la protéine G hétérotrimérique . . . . .	31
1.4	R7TM et $\beta$ -arrestines . . . . .	35
1.5	Double rôle des $\beta$ -arrestines . . . . .	36
1.6	Internalisation via les $\beta$ -arrestines . . . . .	37
1.7	Signalisation $\beta$ -arrestines . . . . .	39
1.8	Cinétiques différentielles voie G / voie $\beta$ -arrestine . . . . .	40
1.9	Phosphorylations dépendantes des $\beta$ -arrestines . . . . .	42
1.10	Signalisation qui diffère en fonction du ligand . . . . .	43
1.11	Efficacité des ligands suivant les différentes voies de signalisation . . . . .	45
1.12	Parallèle Ovaire - Testicule . . . . .	47
1.13	Structure du RFSH . . . . .	48
1.14	Cinétiques $GpERK$ et $bpERK$ . . . . .	51
1.15	Réseau de signalisation de la FSH . . . . .	53
1.16	Système Rénine Angiotensine Aldostérone . . . . .	55
1.17	Structure du $RAT_{1A}$ . . . . .	56
1.18	Oxydase NAD(P)H . . . . .	57
1.19	Signalisation de l'oxydase NAD(P)H . . . . .	58
1.20	Cinétiques ERK . . . . .	60
1.21	Réseau de signalisation du $RAT_{1A}$ . . . . .	62
1.22	Evolutions des modèles « Ternary Complex » . . . . .	79
1.23	Réduction de la cascade MAPK . . . . .	83
1.24	Visualisation des Normes . . . . .	85
1.25	Illustration pour la fonction objectif . . . . .	86
1.26	Méthode du Gradient et problème de minima locaux . . . . .	88
1.27	Illustration méthode de « smoothing » . . . . .	89
1.28	Illustration méthode de Stochastique . . . . .	90

## LISTE DES FIGURES

---

1.29	Illustration de la méthode du « Recuit Simulé » . . . . .	93
1.30	Représentation de la matrice de covariance . . . . .	94
1.31	Illustration de l'algorithme génétique . . . . .	95
2.1	Dosage . . . . .	104
2.2	Protocole FRET . . . . .	105
2.3	Expérience FRET . . . . .	105
2.4	Protocole Blot et Puce . . . . .	107
2.5	Méthode d'Euler et Erreur à chaque pas . . . . .	109
3.1	Formes de la p70S6K . . . . .	123
3.2	Propriétés QFLTL issus des données FRET . . . . .	144

# Glossaire

- 5-HTR : 5-Hydroxytryptamine Receptor
- AA : Acide Arachidonique
- AC : Adénylate Cyclase
- ACE : Angiotensin-Aonverting Enzyme
- AMPc : Adénosine Monophosphate cyclique
- AMP : Adénosine Monophosphate
- AP2 : dynamine
- ARA2 : Antagonistes du Récepteur à l'Angiotensine II
- ATP : Adénosine Triphosphate
- BAD : Bcl2 Antagonist of cell Death
- calmodulin : CALcium MODULated proteIN
- CDC2 :Cell Division Cycle 2
- CDK : Cell Division protein Kinase
- CFP : Cyan Fluorescent Protein
- C/EBP $\beta$  : CCAAT-box/enhancer-binding protein  $\beta$
- CMLV : Cellule Musculaire Lisse Vasculaire
- CREB : cAMP Response (CRE)-Binding protein
- CREM : cAMP Response (CRE) Modulator
- DAG : Diacyl Glycerol
- *dpp* : days *post partum*
- DGK : Diacylglycerol Kinase
- eCG : equine Chorionic Gonadotropin
- EGF : Epidermal Growth Factor
- eIF : Eukaryotic translation initiation factor
- Epac : Exchange factor directly activated by cAMP
- ERK : Extracellular signal Regulated Kinase
- FLNA : Filamin-A
- FRET : Fluorescence Resonance Energy Transfert
- FSH : Follicle-Stimulating Hormone
- GDP : Guanosine nucleotide diphosphate
- GIPs : G protein interacting proteins
- GIRK : G protein coupled Inwardly Rectifying K channels
- GnRH : Gonadolibérine ou Gonadotropin Releasing Hormone
- GRK : G protein coupled Receptor Kinase
- GTP : Guanosine nucleotide triphosphate

## LISTE DES FIGURES

---

- HEK293 : Human Embryonic Kidney 293 cells
- HIF-1 : Hypoxia-Inductible Factor 1
- IEC : Inhibiteur de l'Enzyme de Conversion de l'angiotensine
- IP3 : Inositol trisphosphate
- JNK : c-Jun N-terminal
- kDa : kilo Dalton, 1kDa = 1.00794 uma (unité de masse atomique)  
= 1.00794 · 1.66054 · 10<sup>-27</sup> kg  
= 1.67372 · 10<sup>-27</sup> kg
- MAPK : Mitogen-Activated Protein Kinase
- MEK : Dual specificity mitogen-activated protein kinase kinase
- Mnk : MAP kinase-interacting serine/threonine-protein kinase
- MPP3 : MAGUK p55 subfamily member 3
- mTOR : Target of Rapamycin
- NF-κB : nuclear factor-kappa B
- OSE : Ovarian Surface Epithelium
- P11 : S100A10 ou calpactin 1 light chain ou annexin 2 light chain
- p47phox : p47 phagocyte oxidase
- PA : Phosphatidic Acid
- PC : Phosphatidylcholine
- PDE : Phosphodiésterases
- PDZ : PSD-95-disc large-zonula occludens
- PH : Pleckstrin Homology
- PI3K : Phosphoinositide 3-kinase
- PIP2 : Phosphatidylinositol 4,5-bisphosphate
- PIP3 : Phosphatidylinositol (3,4,5)-trisphosphate
- PKA : Protéine Kinase A (resp. B, C)
- PLA<sub>2</sub> : Phospholipase A 2
- PLD : Phospholipase D
- PP2C : Protein phosphatase 2C
- protéine G : Guanine nucleotide binding protein
- PTEN : Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase
- PTP : Protein Tyrosine Phosphatases
- p90RSK : p90 ribosomal S6 kinase (effecteur en aval de ERK1/2)
- R7TMs : Récepteurs à sept domaines Trans-Membranaires
- Raf : Raf proto-oncogene serine/threonine-protein kinase
- Rap : Rap GTPase-activating protein
- Ras : Ras guanine nucleotide exchange factor
- RAT<sub>1A</sub> : récepteur angiotensine de type 1A
- RCPGs : Récepteurs Couplés aux protéines G
- RGS : Regulator of G-protein signalant
- RhoA : Transforming protein RhoA
- ROS : reactive oxygen species
- RV2 : récepteur vasopressine 2
- SII : [sarcosine1,Ile4,Ile8]Ang II
- SF1 : steroidogenic factor 1

- 
- src : Proto-oncogene tyrosine-protein kinase Src
  - StAR : Steroïdogenic Acurate Regulatory protein
  - Stat3 : Signal transducer and activator of transcription 3
  - TAK : TGF-beta activated protein kinase
  - TAO : thousand-and-one amino acids
  - TSH : Thyroid-Stimulating Hormone
  - USF : upstream stimulating factor
  - VDCC : Voltage Dependent Calcium Current
  - YFP : Yellow Fluorescent Protein
- 
- ASA : Adaptative Simulated Annealing
  - CMA Covariance Matrix Adaptation
  - CMAES : Covariance Matrix Adaptation Evolution Strategy
  - CSA Cumulative Step-Size Adaptation
  - CTL : Computation Tree Logic
  - GA : Genetic Algorithm
  - LFGN : Loi Forte des Grands Nombres
  - LTL : Linear Time Logic
  - ODE : Ordinary Differential Equation
  - PMC : Population Monte Carlo
  - PSO : Particle Swarm Optimization
  - QFLTL : quantifier free LTL
- 
- $\frac{dX(t)}{dt} = f(t, X(t), P)$  : l'ODE
  - $b_i$  : coefficients de Runge Kutta accompagnants les  $rk_i$
  - $CI = (c_1, c_2, \dots, c_{NEQ})$  : conditions initiales de l' ODE
  - $f_{obj}$  : fonction objectif
  - $h$  : taille du pas de temps de l'intégration numérique
  - $K = (k_1, k_2, \dots, k_{NCIN})$  : paramètres cinétiques de l'ODE
  - $n \in \mathbb{N}$  : un entier
  - $NCIN$  : nombre de paramètre cinétique
  - $Nd$  : nombre de noyau de transition de l'algorithme population monte Carlo
  - $Nenf$  : nombre de jeux "enfant" au sein de l'algorithme génétique
  - $NEQ$  : nombre d'équation de l'ODE
  - $Niter$  nombre d'itération effectuée au sein des algorithmes population monte Carlo et génétique
  - $NOR$  : nombre paramètres de norme associées aux données
  - $Norm = (norm_1, norm_2, \dots, norm_{NOR})$  : paramètres de norme associées aux données
  - $NPR$  nombre de paramètre de l'ODE
  - $Npar$  : nombre de jeux "parent" au sein de l'algorithme génétique
  - $Npop$  : taille de l'échantillon de jeux de paramètres pour les méthode de Monte

- Carlo.
- $NPR$  : nombre de paramètres recherchés
  - $NQT$  : nombre de quantités totales fixées dans l'ODE
  - $Obs(X, Norm)$  : l'*observable* ou la fonction dépendante des variables du modèle qui est directement comparable aux données
  - $p$  : paramètre d'une ODE à une dimension
  - $P = (K, CI, QT)$  : les paramètres de l'ODE et par abus de langage :
  - $P = (p_1, p_2, \dots, p_{NPR})$  : vecteur des paramètres recherchés
  - $QT = (q_1, q_2, \dots, q_{NQT})$  : quantités totales de l' ODE
  - $rk_i$  : fonctions intermédiaires servant aux méthodes d'intégration numérique de Runge Kutta et Rosenbrock
  - $t_n$  : temps d'intégration numérique après  $n$  itérations
  - $x(t)$  : variable d'une ODE à une dimension
  - $X(t) = (x_1(t), x_2(t), \dots, x_{NEQ})$  : variables de l' ODE
  - $x_n$  : valeur intégrée de la variable au temps  $t_n$
  - $X_n$  : valeur intégrée du vecteur des variables au temps  $t_n$
  - $Y$  : les données expérimentales

# Avant-propos

Aujourd'hui, 30% des médicaments ciblent des récepteurs à sept domaines transmembranaires (R7TM) et les études pharmacologiques ne se penchent que très peu sur cette échelle microscopique. Or, chaque R7TM ou récepteur couplé aux protéines G (RCPG) est à la tête d'un vaste réseau de signalisations intracellulaires. Quels sont les mécanismes précis activés par les médicaments qui engendrent, entre autres, les effets souhaités ? Pourrait-on être plus optimal, minimaliste et naturel en sélectionnant les voies de signalisation intracellulaire, en modulant les hormones du système déjà présentes ? Pourrait-on avec une connaissance plus précise des réseaux de signalisation soigner plus largement et proposer de nouvelles alternatives ?

Mathématicienne de base, j'ai été vivement intéressée par l'application des outils mathématiques qui me passionnent : les systèmes dynamiques en vue de comprendre ces réseaux intracellulaires. Faire un minimum pour une efficacité maximum sans polluer le voisinage fait partie de la philosophie de base des démonstrations mathématiques : les théorèmes se basent sur les hypothèses minimales pour démontrer un résultat maximal.

L'impact au sein de la biologie de la reproduction humaine, dont la FSH (Hormone Folliculo-Stimulante) est une hormone clé m'a fortement intéressée. En effet, 25% des couples rencontrent des problèmes de fertilité et les traitements FSH donnent des résultats variables couplés d'effets secondaires. Parallèlement, les moyens de contraception comme la pilule qui représentent 60% de la couverture contraceptive [Aubin and Menninger, 2009] sont basés sur des stéroïdes gonadiques inhibant la synthèse et la sécrétion de FSH et LH (Hormone Lutéinisante) [Heffner, 2003]. En outre, la FSH a récemment été impliquée de manière ectopique dans la vascularisation de tumeurs cancéreuses [Radu et al., 2010], ce qui fixe l'importance indéniable de cette hormones au sein de la santé humaine.

Dans ce contexte là, j'ai voulu relever le défi de modéliser mathématiquement le réseau de signalisation FSH. Établir un modèle FSH permet, dans un premier temps, de rassembler les connaissances et de pointer les interactions inconnues au sein du réseau afin de proposer différentes structures possibles. Dans un second temps, j'ai tenu à réaliser ce modèle de manière dynamique, afin que le modèle intègre les compétitions, les recyclages et les rétrocontrôles de la transduction du signal. Alors que les expériences biologiques observent chaque quantités séparément et sont incapables de quantifier ces interactions, le modèle dynamique apporte une vision globale de l'ensemble des espèces liées entre elles. Une fois le modèle calibré sur les données, il offre compréhension et prédiction sur l'ensemble de du réseau ainsi que sur les mécanismes de signalisation. L'ensemble des conditions imaginables et possibles est alors simple à réaliser : il suffit de simuler le modèle en modi-



## LISTE DES FIGURES

---

fiant les paramètres associés à la perturbation voulue.

Première partie

**Introduction**



# Chapitre 1

## Introduction

### 1.1 Les R7TMs et leur signalisation

Les récepteurs membranaires sont des senseurs précis du milieu extérieur, ils transmettent les messages extra-cellulaires à l'intérieur de la cellule. Ce mécanisme de communication permet de nombreuses régulations physiologiques. Parmi ces récepteurs, les *R7TMs* (*Récepteurs à sept domaines Trans-Membranaires*) aussi appelés *RCPGs* (*Récepteurs Couplés aux protéines G*) ont une place particulière puisqu'ils représentent la plus grande famille de gènes chez l'humain avec plus de 800 membres [Kroeze et al., 2003; Fredriksson et al., 2003]. Les R7TMs sont la cible de divers ligands comprenant les photons, les ions, les phéromones, les acides aminés, les peptides et les hormones glycoprotéiques (Figure 1.1). Leur importance indéniable est appuyée par leur place privilégiée au sein de la pharmacologie. Parmi les quelques centaines de cibles des médicaments utilisés actuellement, 30 à 50 % sont des R7TMs [Drews, 2000; Overington et al., 2006]. L'étude de la signalisation intracellulaire induite par ces récepteurs représente un enjeu majeur afin de comprendre et d'élargir les traitements actuels.

#### 1.1.1 Structure et classification

Les R7TMs ont tous en commun leur structure composée de trois parties distinctes :

- La partie extra-cellulaire, appelée *N-terminal* ou *ectodomaine*.
- Les sept segments transmembranaires hydrophobes. Ces segments sont reliés entre eux par trois boucles extra-cellulaires et trois boucles intracellulaires.
- La partie *C terminale* qui est au sein du cytoplasme. Elle contient la majorité des sites de phosphorylations.

Avec les boucles intracellulaires, le C terminal permet des interactions et des associations qui initient la signalisation intracellulaire. Plus précisément, pour certains R7TMs la protéine G (Guanine nucleotide binding protein) vient se fixer au sein de la deuxième boucle intracellulaire, au niveau de la suite des trois acides aminés Aspartate, Arginine, Tyrosine (ou DRY en notation alphabétique). Des mutations de ce motif altèrent totalement la voie de signalisation dépendante des protéines G [Wei et al., 2003]. Cette spécificité des R7TMs a initié le classement qui suit.

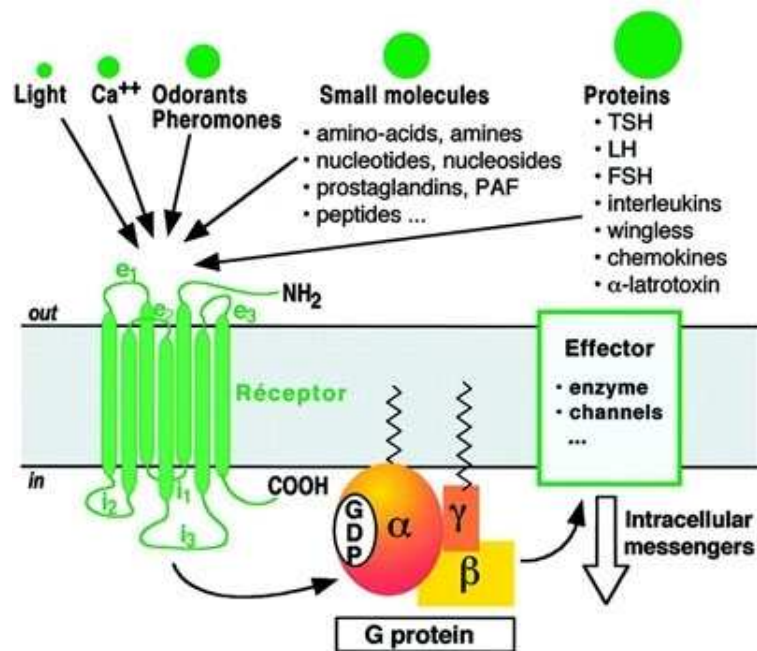


FIGURE 1.1 – Un R7TM fait le lien entre le milieu extra- et intracellulaire grâce à sa structure bien particulière et reçoit des messages de nature très variée [Bockaert and Pin, 1999].

Chez les mammifères, les R7TMs sont répartis en 6 familles selon leur homologie de séquence et de structure [Bockaert and Pin, 1999]. Cette classification est représentée à la Figure 1.2.

**Famille 1 : Récepteurs comportant le motif DRY dit de type rhodopsine** La première de ces familles et la plus large comprend 662 membres, elle est aussi appelée famille des récepteurs de type rhodopsine. Au sein de cet ensemble, on distingue 3 sous-groupes :

- la sous-famille 1a qui réunit les R7TMs activés par des ligands de faible masse moléculaires tels le récepteur  $\beta 2$ -adrénergique qui est la cible de l’adrénaline et permet une réponse au stress, le récepteur angiotensine  $AT_{1A}$  participant à la régulation de la pression cardiaque, le récepteur de la rhodopsine absorbant les photons et l’ensemble des récepteurs odorants. Le site de liaison du ligand se situe au coeur des passages transmembranaires.
- la sous-famille 1b qui contient les R7TMs dont le ligand se fixe entre le N-terminal et les passages transmembranaires tels les récepteurs à chélokinines impliqués dans la régulation des fonctions immunitaires.
- la sous-famille 1c au sein de laquelle on observe des R7TMs avec un très long N-terminal. Les hormones qui ciblent ce type de récepteur sont les volumineuses glycoprotéines LH, FSH impliquées dans la reproduction et l’hormone stimulant la sécrétion des hormones thyroïdiennes (TSH) ayant des effets sur la température du corps ainsi que sur le développement neuronal du fœtus. Ces hormones s’attachent

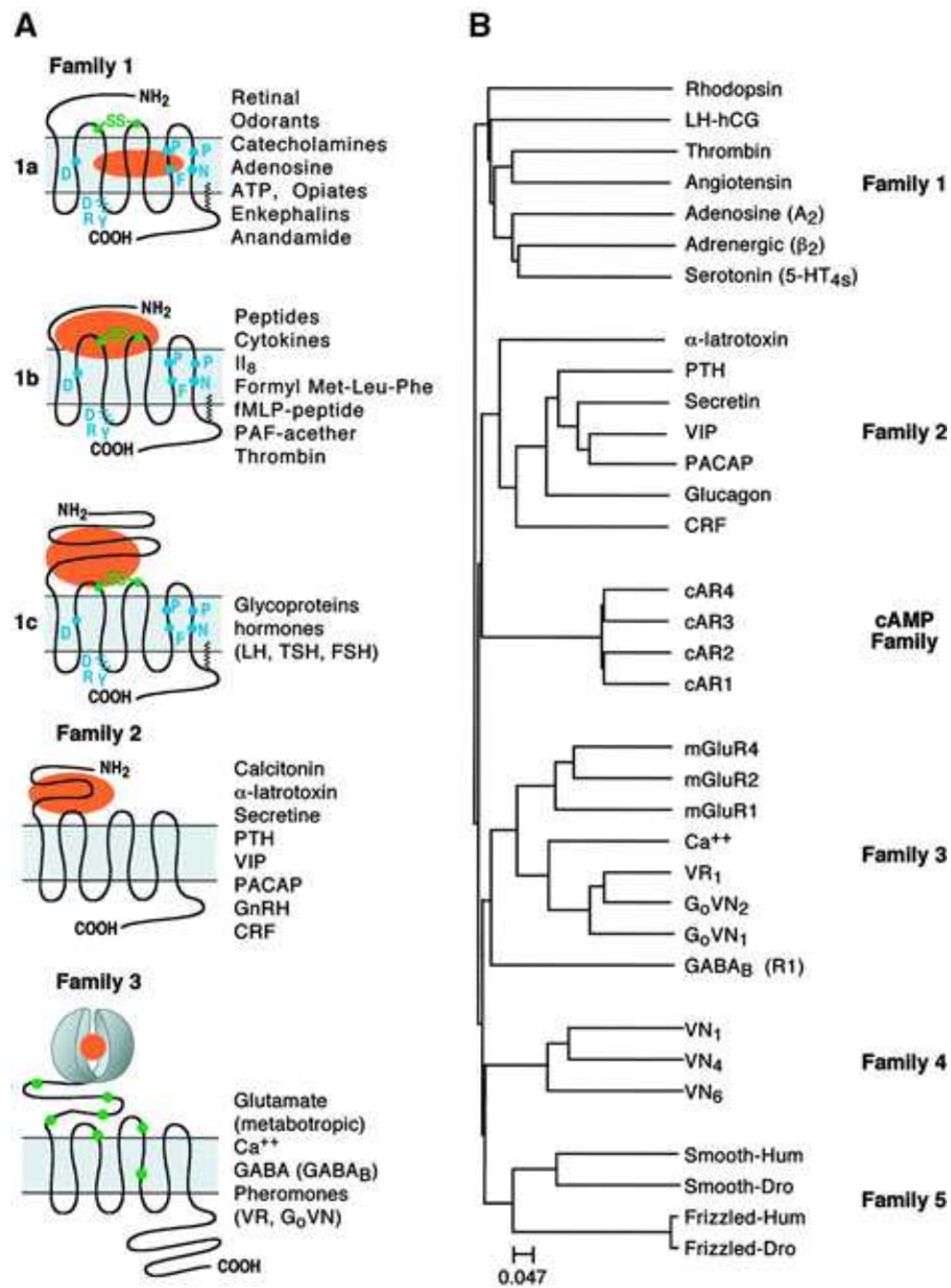


FIGURE 1.2 – Classification R7TM [Bockaert and Pin, 1999]. (A) Les trois principales familles sont les récepteurs de type rhodopsine, les récepteurs à sécrétine et les récepteurs au glutamate et aux phéromones. Pour ces familles la caractérisation du N-terminal ainsi que les sites de liaison au ligand ont été identifiés. (B) Représentation schématique de l'arbre phylogénétique reliant les 6 familles de R7TM.

au sein de l'ectodomaine ainsi que sur la première et la deuxième boucles extra-cellulaires du récepteur.

**Famille 2 : Récepteurs de type sécrétine** La seconde famille aussi appelée famille des récepteurs à sécrétine se caractérise par un large domaine N-terminal extra-cellulaire. Les ligands de ces récepteurs s'arriment comme pour la famille 1c au niveau des domaines extra-cellulaires (domaine N-terminal et boucles extra-cellulaires). Au sein de cette famille on notera le récepteur à sécrétine qui entraîne la sécrétion d'enzyme digestive au sein du pancréas. Cette famille contient aussi le récepteur cible de la Gonadolibérine (GnRH), une neuro-hormone responsable de la synthèse et de la sécrétion de la FSH et de la LH.

**Famille 3 : Récepteurs au glutamate et aux phéromones** La troisième famille se caractérise par son important ectodomaine en forme de pince. Cette famille contient les récepteurs à glutamate, impliqués dans le processus de mémorisation ainsi que des récepteurs à phéromones responsables de sensations de goût.

**Autres récepteurs** Les trois dernières familles totalisent seulement 79 récepteurs. Ce sont la famille des récepteurs aux facteurs d'accouplement chez les champignons, la famille des récepteurs à AMPc caractérisées chez *Dictyostelium discoideum*, une amibe et la famille des récepteurs « frizzled » et « smoothed », impliqués entre autre dans la polarité cellulaire, le développement embryonnaire et la formation des synapses.

Pour tous ces R7TMs, lorsque le ligand vient se fixer sur son récepteur, celui-ci change de conformation. Les passages transmembranaires bougent et permettent le recrutement d'effecteurs.

## 1.1.2 Signalisation

### 1.1.2.1 La voie G

En particulier, le récepteur interagit avec les protéines G. Elles lient une molécule de GDP (Guanosine nucleotide diphosphate) qui peut être échangée pour une molécule de GTP (Guanosine nucleotide triphosphate) pour déclencher ou inhiber des réactions biochimiques dans la cellule.

**Cycle de la protéine G hétérotrimérique** Cette protéine G est hétérotrimérique, ses trois sous-unités sont nommées  $\alpha$ ,  $\beta$  et  $\gamma$ . La molécule de GDP/GTP est liée à la sous-unité  $\alpha$ . Lorsque la protéine G est recrutée au récepteur, le GDP est échangé par le GTP puis le trimère se scinde en deux parties  $G\alpha$  et  $G\beta\gamma$  qui se détachent du récepteur pour activer de multiples voies de signalisation. Ces protéines G peuvent se recycler et interagir de nouveau avec le récepteur (Figure 1.3). Pour cela, le GTP est désactivé en GDP par des protéines de régulation. Une fois l'état inactif de la sous unité  $\alpha$  retrouvé, elle se ré-assemble avec les deux autres sous unités. La protéine G trimérique peut ainsi effectuer des cycles d'activation/désactivation. On notera que l'hypothèse de dissociation des sous-unités de la protéine G est soumise à controverse. [Galés et al., 2006] suppose que le

R7TM est préalablement complexé à la protéine G et que l'activation du R7TM engendre un changement d'organisation des sous-unités permettant la sortie du GDP et découvrant les sites actifs des sous-unités  $G\beta\gamma$ .

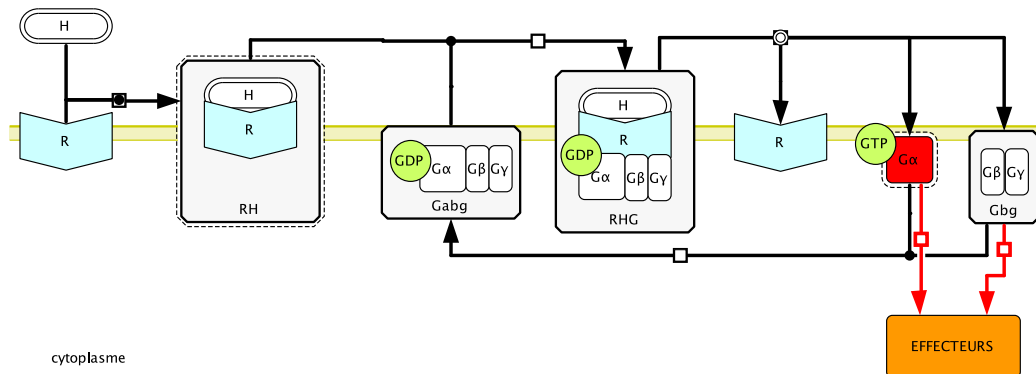


FIGURE 1.3 – Lorsque le récepteur est activé, la protéine G s'y complexe, puis se détache en deux parties : la sous-unité  $G\alpha$  et les sous-unités  $G\beta\gamma$ . Pour être de nouveau recrutée au récepteur, la protéine G hétérotrimérique se reforme.

Outre ces questions d'association/dissociation de la protéine G, il existe plusieurs types de sous-unités (16  $G\alpha$ , 5  $G\beta$  et 14  $G\gamma$  chez l'homme) [Milligan and Kostenis, 2006]. Bien que toutes les combinaisons trimériques ne soient pas possibles, une multitude de possibilités existent et augmentent d'autant les modulations du signal en nature, en cinétique et en quantité. Par la suite nous étudierons les voies de signalisation engendrées par les quatre groupes de protéines  $G\alpha$  :  $\alpha_s$ ,  $\alpha_i$ ,  $\alpha_q$  et  $\alpha_{12/13}$ .

**Voie Gas/AC/AMPc** La protéine Gas active principalement l'adénylate cyclase (AC). L'AC induit ensuite la formation d'adénosine monophosphate cyclique (AMPc) à partir d'adénosine triphosphate (ATP). L'AMPc peut être dégradée par des PDEs (Phosphodi-estérases) en adénosine monophosphate (AMP) ou activer des kinases effectrices comme la PKA (Protéine Kinase A). Pour activer la PKA, l'AMPc se complexe aux deux sous-unités régulatrices de la PKA, ce qui a pour effet de libérer les deux sous-unités catalytiques qui vont alors phosphoryler de nombreux effecteurs tels la p70S6K, Rap et le facteur de transcription CREB (cAMP Response Element).

D'autres voies de signalisation en aval de l'AMPc existent indépendamment de la PKA. En particulier une voie admet Epac (Exchange factor directly activated by cAMP) comme intermédiaire. L'Epac se complexe à l'AMPc et induit l'échange de GDP en GTP sur la petite protéine G Rap. Au sein du réseau de régulation des contractions cardiaques par le récepteur  $\beta_2$ -adrénergique, Rap activé par Epac active la PLC $\epsilon$  (Phospholipase C) [Cheng et al., 2008]. L'Epac peut aussi déclencher la cascade MAPK (Mitogen-Activated Protein Kinase) Raf/MEK/ERK et ces kinases peuvent à leur tour activer le facteur CREB [Neves et al., 2002]. Epac est ainsi responsable de divers réponses cellulaires telles la division, la différenciation ou l'hypertrophie.



**Voie  $G\alpha_i$  et  $G\alpha_o$**  La fonction la plus connue des protéines G contenant une sous-unité  $\alpha_i$  est d'inhiber la voie  $G\alpha_s/AC$ . Cependant, au sein des neurones, le R7TM cannabinoïde CB1 (Famille 1) se couple à  $G\alpha_o/i$  et déclenche l'excroissance de neurites via cette voie. Dans ce cas, une fois activée, la protéine  $G\alpha_o/i$  active rap1 qui induit à son tour l'échange de GDP en GTP des protéines Ra1 et Rac. Ces deux GTPases vont ensuite activer Src et c-Jun N-terminal (JNK), deux kinases en amont du facteur de transcription Stat3 [He et al., 2006].

**Voie  $G\alpha_q/PLC/PIP/PKC$  et  $PIP/Ca$**  La sous-unité  $G\alpha_q$  active la PLC. Cette PLC va catalyser la dissociation du phosphatidylinositol 4,5-bisphosphate (PIP2) en inositol trisphosphate (IP3) et en diacylglycerol (DAG). L'IP3 a pour effet l'ouverture des canaux ioniques et un relargage de calcium du réticulum endoplasmique dans le cytoplasme. L'élévation de la concentration en calcium au sein du cytoplasme engendre l'ouverture de canaux membranaires impliquant une entrée de calcium extra-cellulaire supplémentaire. Le calcium induit, en lien avec le DAG, l'activation de la protéine kinase C (PKC) qui va phosphoryler plusieurs éléments dont Ras. Ras phosphorylé déclenche la cascade MAPK Raf/MEK/ERK [Milligan and Kostenis, 2006]. Enfin, ce calcium entraîne l'activation de la calmodulin (CALcium MODULated proteIN) qui active à son tour la CamKII (Calmoduline Kinase II). Une des cibles de la CamKII est le facteur de transcription CREB.

**Voie  $G\alpha_{12-13}/RhoA/ROCK$**  Cette quatrième famille de protéine G active la petite protéine G RhoA par un échange GDP/GTP. Une fois active, cette protéine RhoA se lie à la Rho Kinase (ROCK) et l'active. ROCK phosphoryle des protéines associées au cytosquelette et active la phosphatase PTEN (Phosphatase and Tensin homolog). PTEN déphosphoryle le phosphatidylinositol (3,4,5)-trisphosphate (PIP3) en PIP2 en opposition à l'action de la phosphoinositide 3-kinase (PI3K). PTEN induit aussi l'activation de la PLC qui scinde de PIP2 en IP3 et DAG. Au sein de cette famille, la  $G\alpha_{12}$  a une activité supplémentaire : elle favorise la désactivation de Ras et donc se comporte comme inhibiteur de la cascade Raf/MEK/ERK [Jiang et al., 1998].

**Voies émanant du complexe  $G\beta\gamma$**  Le rôle habituel du complexe  $\beta\gamma$  est d'inhiber l'activité de la sous-unité  $G\alpha$  en induisant le retour au complexe hétérotrimérique inactif. Cependant, le complexe  $G\beta\gamma$  a une signalisation propre, il est capable d'activer des enzymes telles que la PLC ou la PI3K et de réguler l'ouverture des canaux ioniques GIRK (G protein coupled Inwardly Rectifying K channels) et VDCC (Voltage Dependent Calcium Current) [Wettschureck and Offermanns, 2005]. Le complexe  $G\beta\gamma$  provoque aussi l'activation des ERK1/2 de manière indirecte et spécifique aux R7TMs couplés à  $G\alpha_i$  [Barthet, 2007].

Ces différentes voies déclenchées par les protéines G ont longtemps été considérées comme l'unique signalisation des R7TMs, d'où le nom RCPG. Nous verrons par la suite que même si ces voies G ont été impliquées dans de nombreuses réponses cellulaires, d'autres effecteurs existent et montrent une importance croissante. Mais tout d'abord, nous allons aborder les acteurs de la désensibilisation de la voie G.

### 1.1.2.2 Désensibilisation

Une fois le message hormonal transmis sous forme d'activations intracellulaires, les effets de l'hormone sont stoppés. La désensibilisation est un mécanisme nécessaire au bon fonctionnement des réseaux de communication cellulaire. D'une part, le récepteur doit se dissocier de l'hormone pour être de nouveau prêt à recevoir le messenger suivant. D'autre part, l'activation des différents effecteurs en aval du récepteur doit être régulée sans quoi les quantités d'effecteurs augmenteraient sans limite. Pour cela, la désensibilisation comprend la destruction ou le recyclage du récepteur mais aussi la maîtrise des acteurs intracellulaires résultant souvent de multiples rétrocontrôles au sein des réseaux de signalisation.

**Neutralisation de l'activité GTPase de la protéine G** En aval du récepteur, certains acteurs de la désensibilisation s'opposent à l'activation des protéines G. C'est le cas des protéines RGS (Regulator of G-protein signaling) qui vont accélérer la conversion du GTP en GDP sur la sous-unité  $\alpha$ . En favorisant le retour du GDP sur la sous-unité  $\alpha$  de la protéine G, les RGS désactivent la protéine G et stoppent ainsi toute signalisation émanant de cette sous-unité [Traynor, 2010]. Différemment, au niveau du récepteur des phosphorylations peuvent aussi suspendre les interactions avec les protéines G.

**Phosphorylation du récepteur** En particulier, le récepteur peut induire des kinases dépendantes des seconds messagers comme la PKA et la PKC. Ces kinases peuvent ensuite phosphoryler des récepteurs afin d'en atténuer le signal. Parfois ces rétrocontrôles s'effectuent sur d'autres R7TMs également exprimés dans la cellule. On parlera dans ce cas de désensibilisation *hétérologue*. On y opposera la notion de désensibilisation *homologue* qui ne touche que les récepteurs préalablement activés par l'agoniste. Alors que la PKA et la PKC permettent les deux types de désensibilisation selon le système, d'autres protéines, les GRKs (G protein-coupled receptor kinase), n'agissent que de manière homologue [Lefkowitz, 1998].

Les GRKs sont des protéines ayant une activité sérine/thréonine kinase, il en existe sept types différents dont quatre qui sont impliquées dans la régulation de nombreux R7TMs et présentes dans tous les tissus : GRKs 2, 3, 5 et 6 [Reiter and Lefkowitz, 2006]. Les GRKs interagissent avec les R7TMs activés par leur ligand et les phosphorylent sur des résidus sérine et thréonine localisés sur leurs boucles intracellulaires et leur extrémité carboxy-terminale. Ces phosphorylations modulent la liaison du récepteur avec les différents effecteurs et s'opèrent que lorsque les GRKs sont à la membrane. Les GRKs s'arriment à la membrane par leur queue C-terminale mais selon leur structure et leurs modifications post-traductionnelles, ce n'est pas leur localisation première [Penela et al., 2003]. Ainsi, les GRKs 5 et 6 sont constitutivement localisées à la membrane plasmique. En revanche, les GRKs 2 et 3 sont cytoplasmiques et transitoirement recrutées à la membrane lorsque le récepteur est activé. Leur domaine d'homologie pleckstrine (PH) leur confère une grande affinité pour le dimère  $G\beta\gamma$  libéré après activation du R7TM mais aussi pour différentes molécules de signalisation comme le PIP2, le PIP3, la PKC. Pour le RFSH et le  $RAT_{1A}$ , seules les GRKs 2 et 3 font partie de la voie de désensibilisation. Pour le  $RAT_{1A}$ , en inhibant les GRKs 2 et 3, la réponse ERK est augmentée tandis qu'en inhibant le deuxième couple de GRKs (5 et 6) la réponse ERK est diminuée. Les GRKs 5 et 6 participent à une

autre voie de signalisation indépendante des protéines G que nous détaillerons par la suite [Kim et al., 2005; Ahn et al., 2004a].

**Désensibilisation dépendante des  $\beta$ -arrestines** La phosphorylation du récepteur par les GRKs va induire le recrutement d'arrestines.

Les arrestines sont des protéines d'échafaudage qui contiennent plusieurs domaines d'interaction répartis sur deux lobes [Hirsch et al., 1999]. Selon l'orientation relative des deux lobes de la protéine, l'accessibilité des sites d'interaction est modulée. Au sein des cellules non-stimulées, les arrestines sont localisées dans le cytoplasme en conformation inactives. Leur activation coïncide avec leur recrutement au R7TM activé au niveau de la membrane plasmique. La liaison de l'arrestine au récepteur implique un changement de conformation des deux acteurs permettant le recrutement de clathrines qui engendrent l'internalisation du récepteur [Han et al., 2001].

Il existe 4 types d'arrestines, les arrestines 1 et 4 sont exclusivement impliquées dans la signalisation visuelle, l'arrestine 1 se complexant au récepteur rhodopsine une fois phosphorylé par la réception d'un photon [Krupnick et al., 1994]. Les deux autres arrestines, sont plus communément appelées  $\beta$ -arrestines 1 et 2 car elles ont été identifiées comme intervenant dans la désensibilisation du récepteur  $\beta_2$ -adrénergique [Lohse et al., 1990]. Les  $\beta$ -arrestines sont exprimées dans l'ensemble des cellules. Le *knock out* du gène codant pour la  $\beta$ -arrestine 1 et le *knock out* du gène codant pour la  $\beta$ -arrestine 2 ne perturbent pas le développement chez la souris [Conner et al., 1997; Bohn et al., 1999] tandis que la délétion des deux gènes est létale [Kovacs et al., 2009].

En se complexant aux récepteurs phosphorylés, les  $\beta$ -arrestines bloquent le site dédié aux protéines G (Figure 1.4). Le récepteur ne peut plus engendrer la signalisation G-dépendante ; il est désensibilisé pour la voie G.

D'autre part, les  $\beta$ -arrestines peuvent se coupler aux PDEs ou aux diacylglycerol kinases (DGKs) [Perry et al., 2002; Nelson et al., 2007] pour les récepteurs  $\beta_2$ -adrénergique et M1 muscarinique. Ces études révèlent que les  $\beta$ -arrestines activent les PDEs ou les DGKs en s'y complexant. Puis, ces complexes sont entraînés au récepteur et bloquent les interactions avec la protéine G mais pas seulement. En se complexant aux PDEs et aux DGKs, les  $\beta$ -arrestines désensibilisent doublement la voie G, elles masquent le site d'interaction du récepteur avec la protéine G et active les protéines qui dégradent des seconds messagers (PDEs pour l'AMPC et DGKs pour le DAG) des voies G (Figure 1.5).

### 1.1.2.3 Mouvements intracellulaires des R7TMs

**Internalisation** L'endocytose est déclenchée par l'association des  $\beta$ -arrestines aux adaptines (AP-2). Puis, les clathrines en lien avec les adaptines vont cerner le R7TM lié aux  $\beta$ -arrestines [Laporte et al., 2000]. Ces protéines structurelles, les clathrines, vont constituer l'enveloppe de la vésicule contenant le R7TM. Un autre élément, la dynamine intervient pour faciliter la formation de la vésicule recouverte de clathrines. Les  $\beta$ -arrestines favorisent ce rapprochement des clathrines avec la dynamine en ayant des sites attractifs pour les deux. (Figure 1.6). Plus précisément, la  $\beta$ -arrestine 2 a été reconnue plus efficace dans ce processus que la  $\beta$ -arrestine 1. En effet, la  $\beta$ -arrestine 1 peut être phosphorylée par

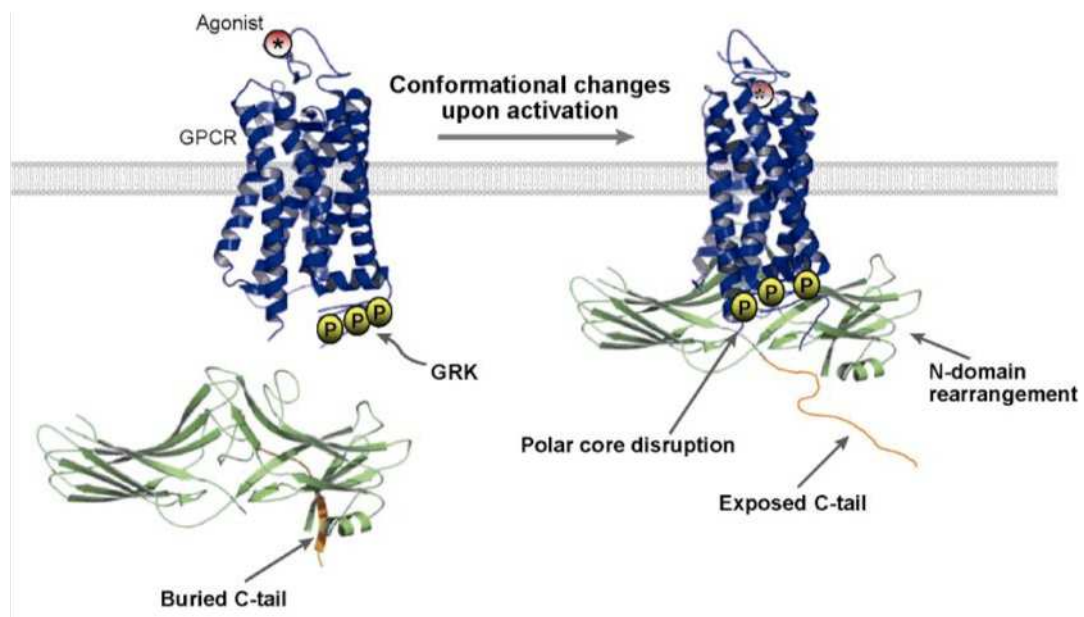


FIGURE 1.4 – Vision en trois dimensions du R7TM et ainsi que de l'arrestine [Moore et al., 2007]. L'arrestine couvre les sites d'interaction du récepteur et prend le relais de la signalisation en changeant de conformation. Ces modifications incluent la rupture du noyau polaire et les réarrangements des domaines terminaux de l'arrestine. Ces changements renforcent la liaison au récepteur et démasquent les sites de liaison des arrestines pour les clathrine et les AP2.

src (Proto-oncogene tyrosine-protein kinase Src) sur un site proche de celui utilisé pour s'associer aux adaptines. Cette phosphorylation gêne alors l'interaction des  $\beta$ -arrestine 1 avec les AP-2 et serait responsable de l'efficacité moindre de cet isoforme [Marion et al., 2007]. D'autres procédés d'internalisation existent tels l'endocytose via les cavéoles. Les cavéoles sont activées lorsque PKC et src les phosphorylent. De même que les clathrines, elles s'associent à la dynamine pour provoquer l'endocytose du R7TM mais ne font intervenir ni les  $\beta$ -arrestines ni les clathrines.

Une fois internalisés les récepteurs peuvent poursuivre leur voie de signalisation avec la  $\beta$ -arrestines (c'est le sujet de la section suivante), être recyclés à la membrane ou adressés au lysosome pour être dégradés.

#### 1.1.2.4 Signalisation indépendante des protéines G

Depuis une dizaine d'années, il est apparu de plus en plus clairement que les R7TMs ont recours à des mécanismes de transduction du signal qui ne font pas intervenir les protéines G hétérotrimériques. Dans la partie qui suit, nous passerons en revue les principaux mécanismes décrits en insistant particulièrement sur le rôle des  $\beta$ -arrestines.

**Recrutement des  $\beta$ -arrestines** Les  $\beta$ -arrestines sont initialement présentes au sein du cytoplasme sous leur forme phosphorylée. Lors de leur migration à la membrane,

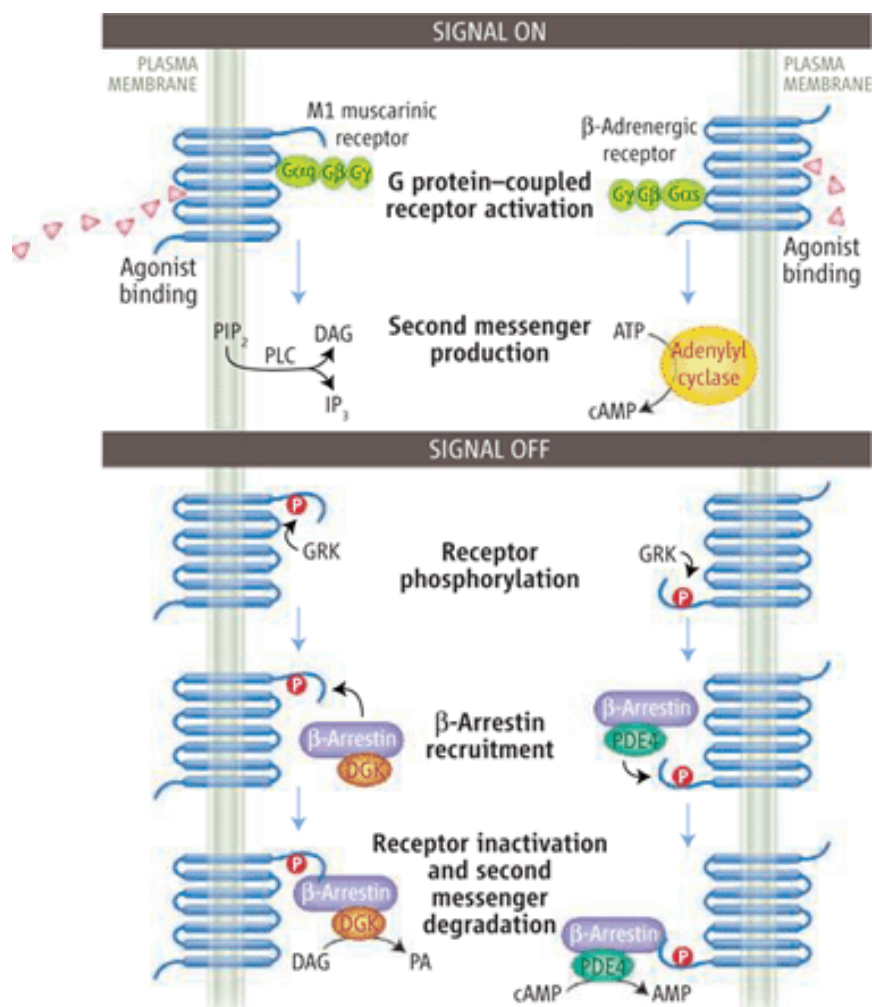


FIGURE 1.5 – Double rôle des  $\beta$ -arrestines [Grady, 2007]. Les  $\beta$ -arrestines désensibilisent le récepteur et activent par recrutement des protéines de régulation des seconds messagers tels les PDEs contre l'AMPc et les DGKs contre le DAG.

elles sont déphosphorylées sur le site  $\{Ser412\}$  pour la  $\beta$ -arrestine 1 et sur les sites  $\{Thr383, Ser361\}$  pour la  $\beta$ -arrestine 2 [DeWire et al., 2007]. Selon le récepteur, elles se lient de manière transitoire ou stable. Cette spécificité a d'ailleurs conduit à la distinction de deux classes de R7TMs. Les récepteurs de classe A interagissent transitoirement avec les  $\beta$ -arrestines et préférentiellement avec l'isoforme 2 [Oakley et al., 2000]. L'interaction récepteur -  $\beta$ -arrestine étant rapide, cela implique que l'internalisation du récepteur est brève et immédiatement suivie de son recyclage à la membrane. Des R7TMs de classe A sont par exemple les récepteurs  $\beta$ -2 et  $\alpha$ 1B -adrénergiques, le récepteur opiacé  $\mu$  (Famille 1a) responsable de l'effet euphorisant des opiacés, le récepteur endothéline A (Famille 1) lié à la vasoconstriction des cellules musculaires lisses et le récepteur à dopamine D1A (Famille 1) qui est impliqué dans le contrôle de la motricité. Les R7TMs de classe B interagissent de manière soutenue avec la  $\beta$ -arrestine et ont la même affinité envers les deux isoformes.

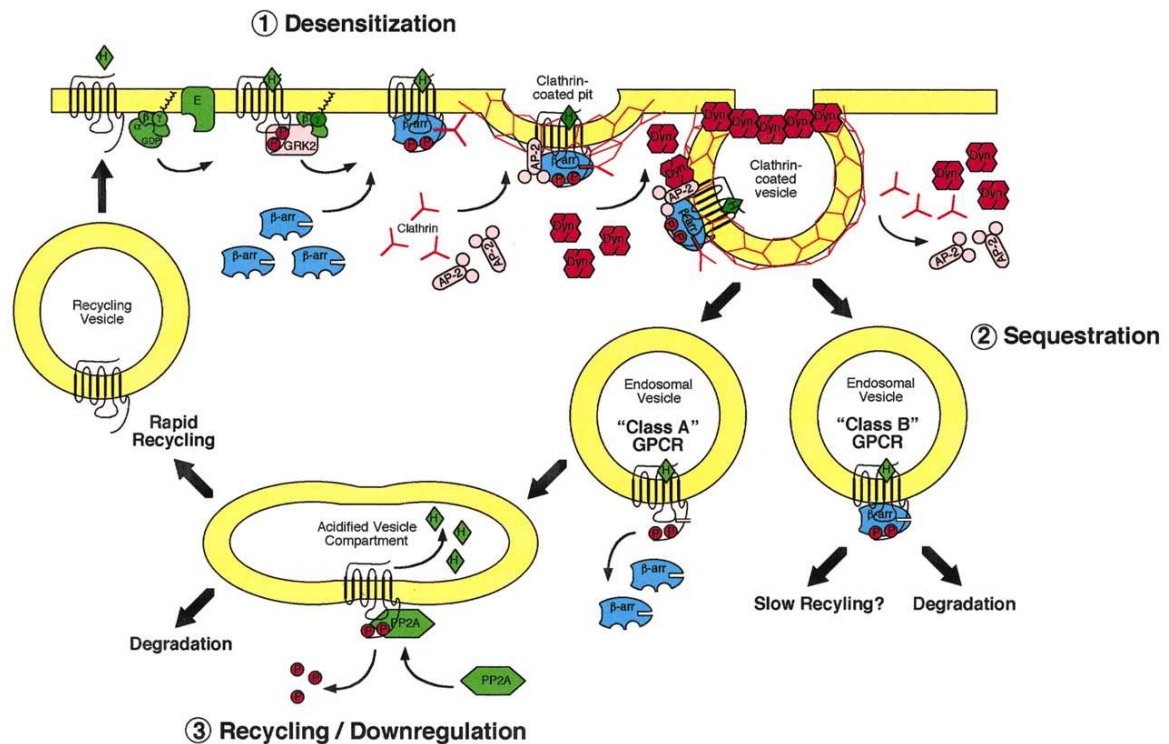


FIGURE 1.6 – Internalisation via les  $\beta$ -arrestines [Luttrell and Lefkowitz, 2002]. 1/ Désensibilisation. Après phosphorylation du R7TM par les GRKs, la  $\beta$ -arrestine se fixe et empêche le couplage entre le récepteur et les protéines G hétérotrimériques, stoppant ainsi la signalisation dépendante des protéines G (Effecteurs E). 2/ Internalisation. La  $\beta$ -arrestine est une protéine d'échafaudage, elle se lie aux clathrines et aux adaptines (AP-2) ce qui engendre l'endocytose du complexe récepteur- $\beta$ -arrestine. 3/ Recyclage/Dégradation. Selon la classe du R7TM, la séquestration au sein du cytoplasme va être brève ou prolongée. Au sein du cytoplasme le R7TM est recyclé ou dégradé.

Ils sont internalisés et maintenus dans le cytoplasme au sein de vésicules endosomales d'où ils peuvent engendrer entre autre un signal ERK persistant [Luttrell et al., 2001; Ahn et al., 2004a]. Au sein de la classe B se trouvent notamment le récepteur angiotensine  $AT_{1A}$  (Famille 1a), le récepteur FSH (Famille 1c), le récepteur à neurotensine (Famille 1) intervenant dans la régulation de la glycémie ainsi que dans la stimulation des sécrétions des hormones hypophysaires LH et prolactine. Le récepteur vasopressine 2 (Famille 1) qui en présence de son hormone a un effet antidiurétique fait aussi partie de la classe B ainsi que le récepteur neurokinine NK-1 (Famille 1), dont les ligands antagonistes pourraient constituer une nouvelle classe de médicaments aux effets antidépresseurs [Varty et al., 2003], anti-anxiolytiques et antiémétiques. Une vision synthétique de quelques R7TMs et leurs interactions avec les  $\beta$ -arrestines est donnée dans le Tableau 1.1.

**Signalisation dépendante des  $\beta$ -arrestines** Classiquement, les R7TMs stimulent les protéines G hétérotrimériques, qui contrôlent la production de seconds messagers

Récepteur	type de recrutement $\beta$ -arrestine	signalisation ERK via les $\beta$ -arrestines	type de signalisation $\beta$ -arrestine régulant les ERKs
AT <sub>1A</sub> wt	classe B	oui	réciproque
AT <sub>1A</sub> $\Delta$ 324	classe A	oui	N.D.
FSH wt	classe B	oui	co-dépendante
$\beta_2$ Adrénérique wt	classe A	oui	co-dépendante
$\beta_2$ Adrénérique GRK <sup>-</sup>	-	non	-
$\beta_2$ Adrénérique PKA <sup>-</sup>	-	-	-
PAR2 wt	classe B	oui	co-dépendante
PAR2 C-tail coupée	classe A	non	N.A.
V2R	classe B	oui	réciproque
PTH1R	classe B	oui	co-dépendante
CCR7	classe B	oui	N.D.
PAR 1	N.D.	oui	réciproque

TABLE 1.1 – Les R7TMs et leurs propriétés vis à vis des  $\beta$ -arrestines [Luttrell and Lefkowitz, 2002; DeWire et al., 2007]. La signalisation  $\beta$ -arrestine est dite réciproque lorsque seule une des isoformes prend part à la signalisation tandis qu'elle est dite co-dépendante lorsque les deux isoformes participent à la signalisation. Notations : N.A. : non applicable. N.D. : non déterminé.

et l'ouverture de canaux ioniques à la membrane plasmique tandis que les  $\beta$ -arrestines désensibilisent les R7TMs. Cependant, des données récentes indiquent le recrutement, l'activation, et l'échafaudage des complexes de signalisation cytoplasmiques via les deux adaptateurs moléculaires  $\beta$ -arrestines 1 et 2.

En particulier, les  $\beta$ -arrestines sont impliquées dans la signalisation ERK (Tableau 1.1). Contrairement à la signalisation ERK dépendante des protéines G qui est rapide et transitoire, le signal ERK via les  $\beta$ -arrestines est un peu plus tardif mais soutenu (Figure 1.8). Comme au sein de la voie de désensibilisation, la phosphorylation du R7TM par les GRKs améliore, voire joue le rôle de prérequis au recrutement des  $\beta$ -arrestines. Ainsi pour plusieurs R7TMs, la phosphorylation du récepteur par les GRKs 2 et 3 (GRK 2 uniquement pour le RFSH) améliore le recrutement des  $\beta$ -arrestines en vue d'une désensibilisation du R7TM. En revanche, les phosphorylations par les GRKs 5 et 6 ont été reconnues comme soutien de la signalisation ERK dépendante des  $\beta$ -arrestines [Ren et al., 2005; Kim et al., 2005; Kara et al., 2006]. Les ERK phosphorylés ne sont qu'un exemple des signaux intracellulaires induits par les  $\beta$ -arrestines [Lefkowitz and Shenoy, 2005].

En 2007, une étude par spectrométrie de masse a été réalisée sur des cellules HEK293 où les deux isoformes de  $\beta$ -arrestines ainsi qu'un R7TM, l'AT<sub>1A</sub>, ont été transfectés. Avec ou sans stimulation par l'angiotensine, [Xiao et al., 2007] ont détecté 337 protéines interagissant avec les  $\beta$ -arrestines. Parmi ces protéines, 71 interagissent seulement avec  $\beta$ -arrestine 1, 164 seulement  $\beta$ -arrestine 2 et 102 avec les deux  $\beta$ -arrestines. Cette étude relate aussi que les cibles des  $\beta$ -arrestines sont présentes partout dans la cellule. Ainsi, 53% des cibles des  $\beta$ -arrestines sont dans le cytoplasme, 31 % sont au sein du noyau, 5% à la membrane plasmique, mais on retrouve aussi des interactants dans les mitochondries, le réticulum

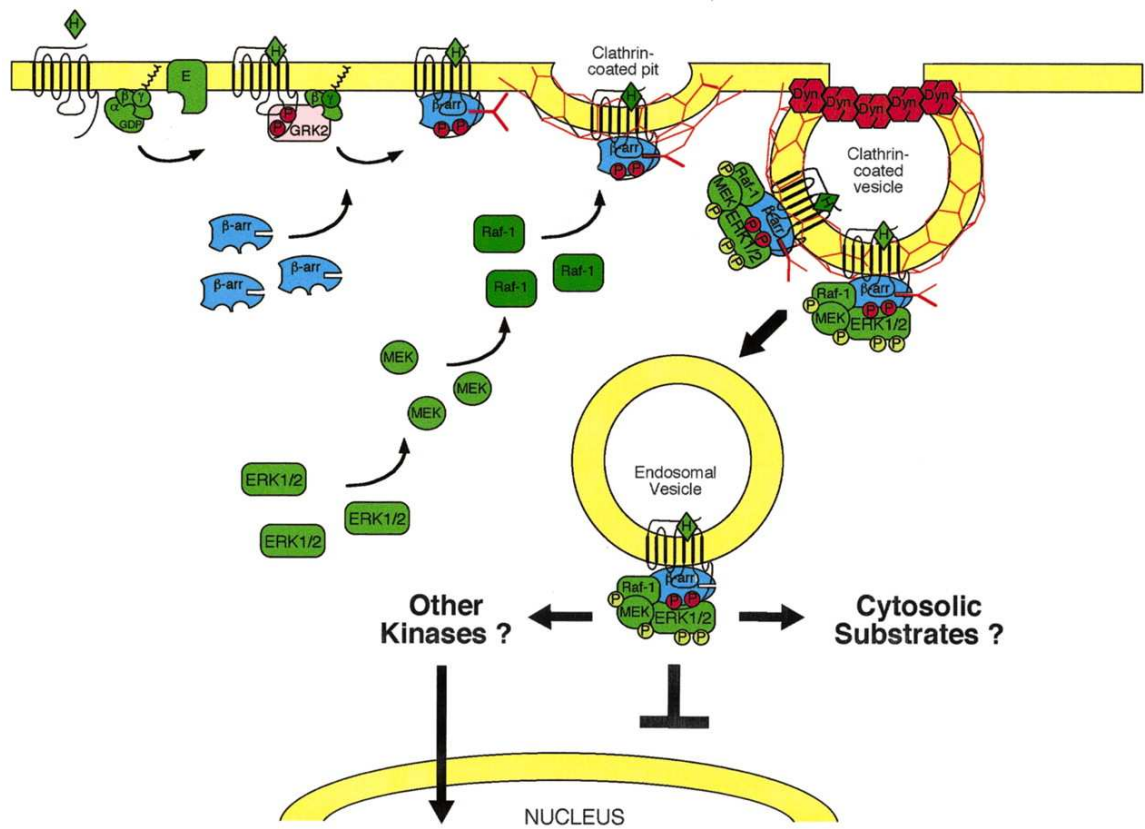


FIGURE 1.7 – Rôle supposé des  $\beta$ -arrestines dans l'activation et le ciblage des MAP kinases [Luttrell and Lefkowitz, 2002]. La liaison de la  $\beta$ -arrestine au R7TMS actif induit l'assemblage d'un complexe MAP kinase. Les complexes récepteur- $\beta$ -arrestine-ERK sont internalisés au sein du cytoplasme. Ces ERKs ne sont pas transportés au sein du noyau et n'induisent pas la prolifération cellulaire. Cependant, l'activation de ERK1/2 au sein de l'échafaudages des  $\beta$ -arrestines peut favoriser la phosphorylation de différents substrats ainsi que l'activation de facteurs transcriptionnels.

endoplasmique et l'appareil de Golgi. Enfin, cette étude confirme que les  $\beta$ -arrestines ont un rôle important en signalisation cellulaire et transduction du signal puisqu'un quart des protéines identifiées appartient à cette catégorie. En particulier, cette étude confirme que les  $\beta$ -arrestines agissent comme des adaptateurs et permettent la formation de complexes fonctionnels comprenant les membres des familles MAPK ERK, p38, JNK3 ainsi que les partenaires des phosphatases PP2A et MKP3 [Luttrell and Lefkowitz, 2002; DeWire et al., 2007] et (Figure 1.7).

Cette étude va au delà des interactions déjà connues et identifie au sein de l'interactome des  $\beta$ -arrestines des protéines impliquées dans un large éventail de processus cellulaires comprenant la transduction du signal (p160ROCK, MEKK1, TAK1 et PI3-kinase), l'organisation cellulaire et le mouvement (p160ROCK, SCY1-like2 et PAK-7), la régulation du cycle cellulaire et prolifération (Wee1A kinase, PAK-7, CDK et PP2C), la synthèse et



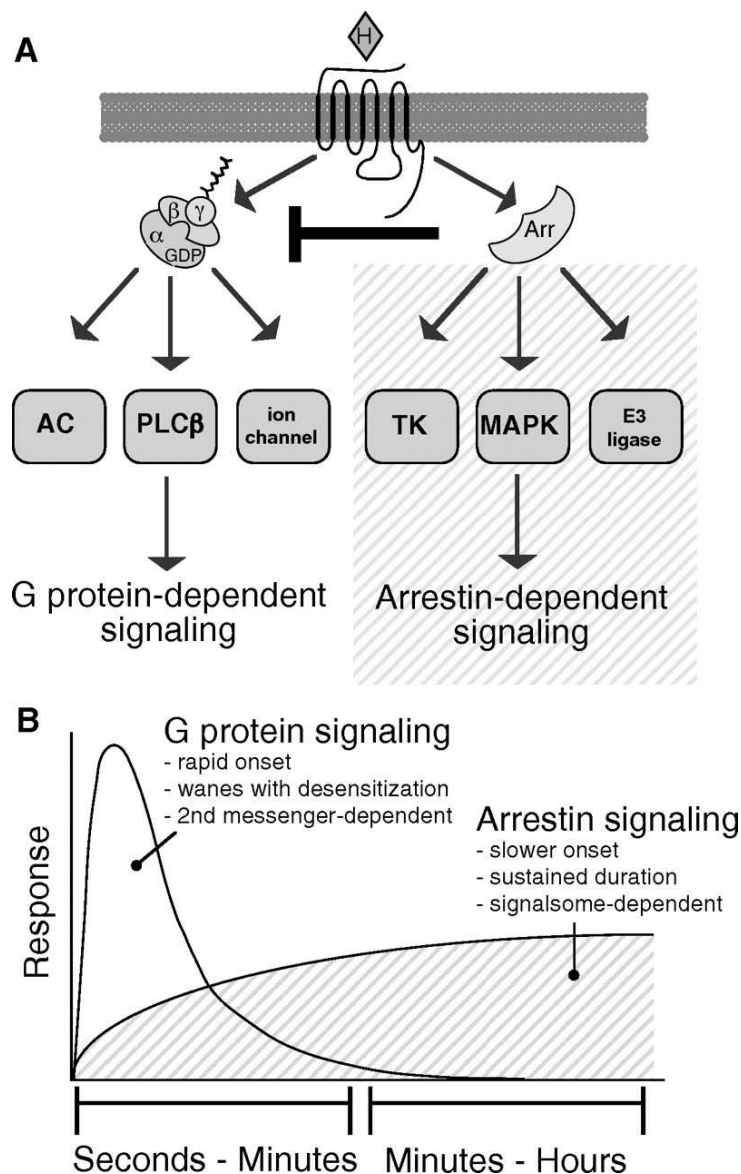


FIGURE 1.8 – Pluridisciplinarité des R7TMs et cinétiques différentielles entre la voie G et la voie  $\beta$ -arrestine [Luttrell and Gesty-Palmer, 2010]. A/ Illustration de deux voies largement répandues parmi les R7TMs. D'un côté, la protéine G se couple au R7TM et induit l'activation de l'AC, de la PLC $\beta$  et des canaux ioniques. De l'autre les  $\beta$ -arrestines jouent un double rôle : elles désensibilisent la voie G et induisent des signaux qui leurs sont propres. B/ Les cinétiques de ses deux voies sont opposés. Alors que l'activation de la voie G est immédiate et transitoire, la signalisation dépendante des  $\beta$ -arrestines est plus tardive mais prolongée. Ceci s'applique en particulier à la phosphorylation des ERKs, on appellera *GpERK* la composante ERK rapide dépendante des protéine G et *bpERK* l'autre dépendante des  $\beta$ -arrestines.

le traitement d'acide nucléique (NIK, MEKK1 et SRPK2), le stress cellulaire (PP2C), et l'apoptose (PAK-7 et la PI3-kinase) [Xiao et al., 2007].

Récemment, la même équipe a réalisé une analyse globale des phosphorylations engendrées par la signalisation dépendante des  $\beta$ -arrestines en aval du récepteur AT<sub>1A</sub> [Xiao et al., 2010]. Pour observer la signalisation dépendante des  $\beta$ -arrestines, ils ont utilisé un ligand sélectif (SII) qui est incapable d'activer la signalisation dépendante de la protéine G [Wei et al., 2003] (voir la section suivante pour la définition des ligands biaisés). La signalisation  $\beta$ -arrestines engendre ainsi l'augmentation significative (plus de 30%) de 171 phosphoprotéines et la diminution de 53 autres après stimulation du récepteur AT<sub>1A</sub> par SII. Avec ses résultats, un réseau a été construit en couplant les interactions déjà caractérisées et un algorithme d'inférence [Xiao et al., 2010]. Ce réseau couvre un large éventail de fonctions cellulaires, signalisation MAPK et PI3K/AKT, la réorganisation du cytosquelette, l'adhérence cellulaire, la communication entre cellules, le cycle cellulaire et le développement, voir (Figure 1.9).

**Autres voies de signalisation G et  $\beta$ -arrestines indépendantes** Les protéines G et  $\beta$ -arrestines ne sont pas les seules espèces intracellulaires avec lesquelles les R7TMs interagissent. Ces protéines sont rassemblées sous le nom de GIPs (G protein interacting proteins) [Bockaert et al., 2004]. Par exemple, les récepteurs à sérotonine (5-HT<sub>1A</sub>, Famille 1), s'associent avec PTEN (Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase) au niveau de la troisième boucle intracellulaire du récepteur 5-HT<sub>2C</sub> [Bockaert et al., 2010]. Cette complexation induit la déphosphorylation du récepteur. Un peptide pénétrant inhibant cette interaction a été testé comme médication potentielle au traitement de l'addiction aux drogues évitant entre autre l'anxiété comme effets secondaires. Les récepteur 5-HT<sub>2C</sub> interagissent aussi avec les domaines PDZ (PSD-95-disc large-zonula occludens) des protéines MPP3 (MAGUK p55 subfamily member 3) ce qui a pour effet de stabiliser le récepteur à la membrane plasmique et d'empêcher sa désensibilisation. Enfin [Bockaert et al., 2010] liste aussi une interaction du R7TM 5-HT<sub>2C</sub> avec P11 (P11 : S100A10 ou calpactin 1 light chain ou annexin 2 light chain) ayant pour action l'adressage du récepteur à la membrane et un effet antidépresseur.

D'autres protéines membranaires ou cytosoliques peuvent lier les R7TMs et ou être activées en aval d'un R7TM, c'est le cas de mTOR et 14-3-3. Ces deux protéines sont particulièrement impliquées dans la signalisation intracellulaire, comme  $\beta$ -arrestines, elles sont au coeur de « carrefours » de signalisation. La 14-3-3 est capable de se lier à une multitude de protéines en ciblant des motifs sérine/thréonine. [Jin et al., 2004] recense des interactions avec 0.6% du génome humain dont 44% sont des protéines de transduction du signal. MTOR, est quant à elle une protéine aux multiples sites de phosphorylation qui, selon les combinaisons activées, engendre différentes voies de signalisation. La signalisation des R7TMs est vaste et variée, beaucoup de données sont disponibles, cependant les principales voies de signalisations connues ont été énoncées ci-dessus.

### 1.1.2.5 Importance du ligand

La signalisation intracellulaire et ses régulations sont des mécanismes complexes. À cela, s'ajoute la multitude de ligands qui peuvent être reconnus par un même R7TM

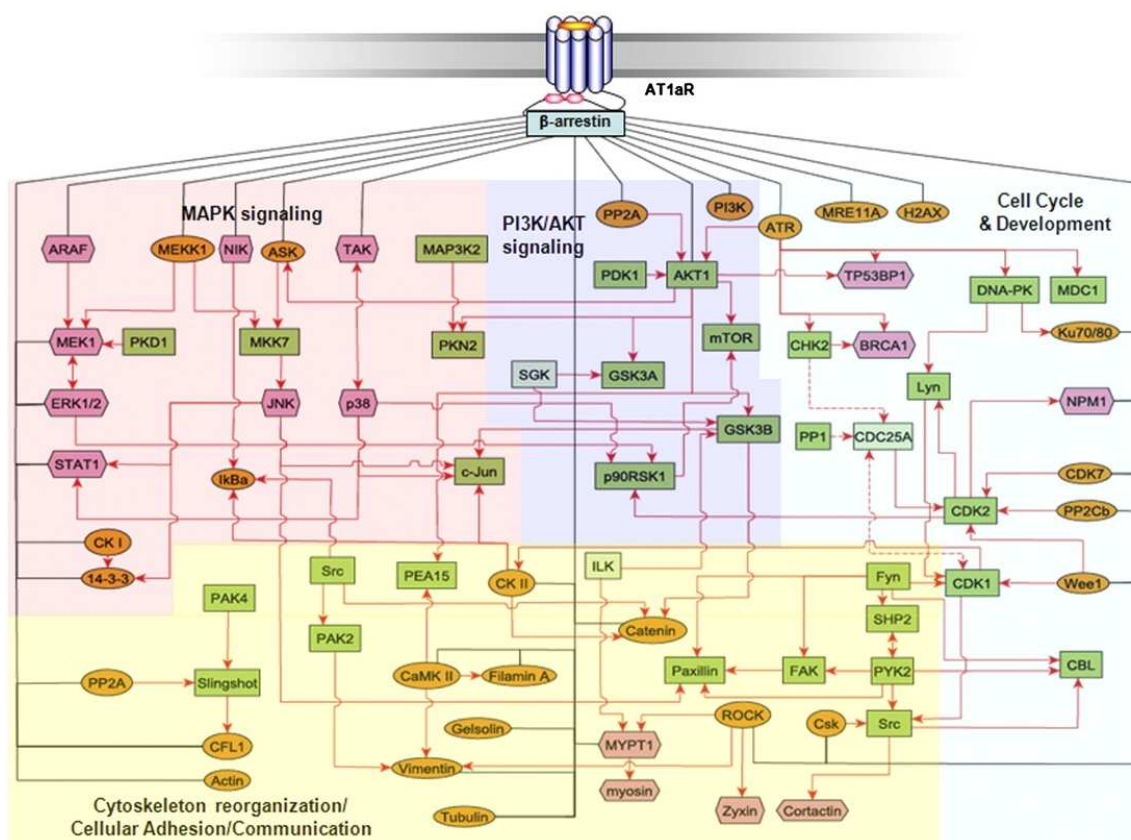


FIGURE 1.9 – Réseau inféré du phosphoprotéome dépendants des  $\beta$ -arrestines ainsi que des réactions déjà connues et détaillées dans la littérature [Xiao et al., 2010]. Dans les rectangles verts, les protéines identifiées dans le phosphoprotéome dépendant des  $\beta$ -arrestines. Dans les ellipses orange, les protéines qui interagissent directement avec les  $\beta$ -arrestines. Dans les hexagones roses, les protéines présentes dans le phosphoprotéome et l'interactome des  $\beta$ -arrestines. Dans les rectangles gris, les nœuds de régulations connus mais absents des deux études. Les flèches rouges représentent les régulations ; les lignes noire, les interactions avec les  $\beta$ -arrestine ; les lignes en pointillés, les régulations des protéines carrefour.

et engendrer différents signaux. C'est un des principes des médicaments qui copient un messenger ou en créent un nouveau en vue de modifier la signalisation du récepteur cible et d'obtenir les effets thérapeutiques souhaités. Une première classification des ligands a été établie en fonction du récepteur cible mais aussi en fonction de l'affinité qu'ils ont pour le récepteur ainsi que de l'intensité de la réponse comparée à celle du ligand endogène. Ce concept conduit à quatre groupes : les agonistes complets qui miment exactement l'activité du ligand endogène, les agonistes partiels qui donnent une réponse plus faible, les antagonistes qui ne produisent aucun signal mais occupent le récepteur et les agonistes inverses diminuant la réponse basale.

**Nouveaux concepts** La classification proposée ci-dessus ne tient compte que d'une réponse intracellulaire alors que, comme nous venons de le voir les R7TMs engendrent plusieurs voies de signalisation. Un ligand va être jugé agoniste car il reproduit l'augmentation d'une espèce activée par le ligand endogène. Qu'en est-il des autres espèces moléculaires présentes au sein des différentes voies de signalisation ? Récemment, plusieurs études ont montré qu'un agoniste n'implique pas forcément une activation équivalente de l'ensemble des voies de signalisations couplées au récepteur (Figure 1.10). Les ligands peuvent activer sélectivement chaque voie avec une efficacité propre : agoniste pour la première, antagoniste pour une seconde et agoniste inverse pour une dernière par exemple [Kenakin, 2002; Galandrin and Bouvier, 2006; Violin and Lefkowitz, 2007; Urban et al., 2007].

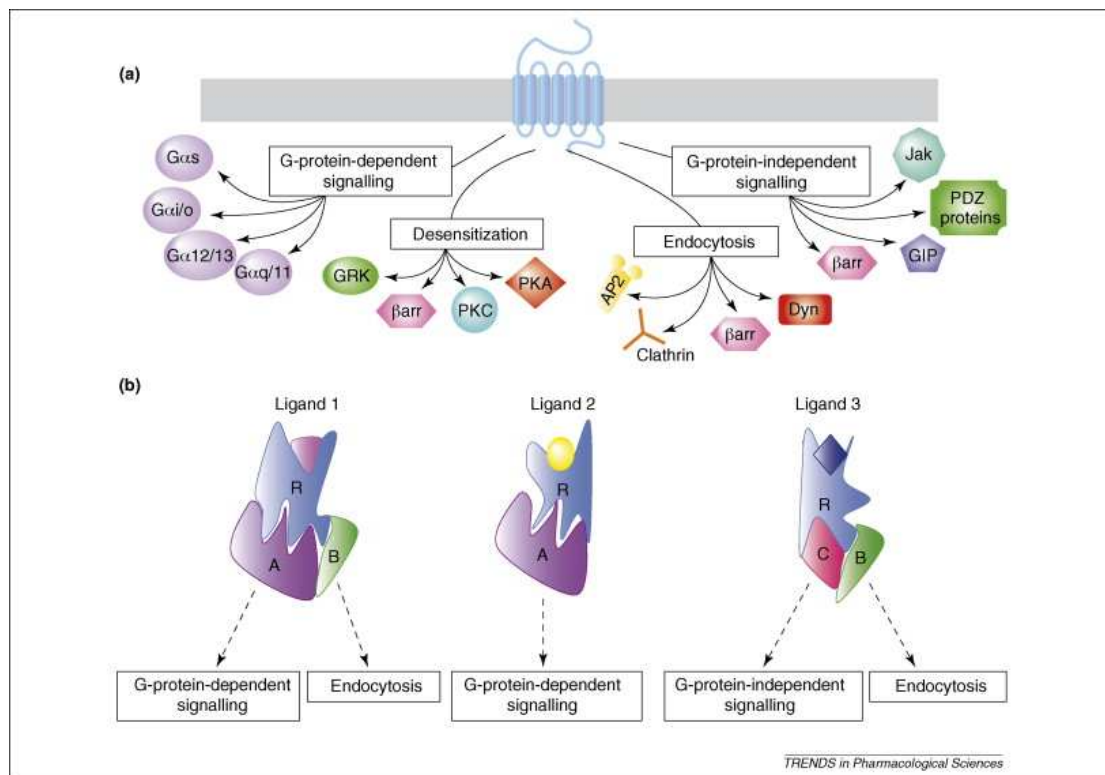


FIGURE 1.10 – Diversité de signalisation R7TM [Galandrin et al., 2007]. a/ Représentation schématique des voies effectrices engendrées par l'activation d'un R7TM ainsi que le type d'activité contrôlée. De telles activités incluent la modulation des voies de signalisation dépendantes ou indépendantes de la protéine G, la désensibilisation et l'endocytose. b/ Représentation schématique de l'efficacité ligand-biaisée, où les trois types de ligands peuvent induire et / ou stabiliser les conformations des récepteurs différentes et promouvoir des réponses distinctes.

**Exemples de ligands biaisés** Les premiers ligands biaisés analysés comme tels ont été dirigés vers le récepteur muscarinique (CHRm1). Ces ligands sont des analogues de

l'acétylcholine, ils ont été testés dans différents modèles animaux pour la maladie d'Alzheimer. Certains ont montré des effets positifs sur les processus mnémotechnique et peuvent être considérés comme traitement potentiel. Alors que le ligand endogène induit le couplage aux protéines  $G_{\alpha s}$  et  $G_{\alpha q}$  pour produire de l'AMPc et de la PLC, les ligands biaisés empêchent le couplage à  $G_{\alpha s}$  et favorisent le couplage à  $G_{\alpha q}$ . Ils activent donc de façon biaisée la voie de signalisation dépendante de la protéine  $G_{\alpha q}$  [Fisher et al., 1993].

Depuis, plusieurs ligands biaisés ont été repérés ciblant différents R7TMs. C'est le cas du  $\beta$  bloquant propranolol, qui agit sur le récepteur  $\beta 2$ -adrénergique comme un agoniste inverse au niveau du couplage à  $G_{\alpha s}$  et donc inhibe la production d'AC et de PKA tandis qu'il agit en tant qu'agoniste partiel au niveau d'une production de ERK G-indépendante qui engendre la transcription du gène CRE [Baker et al., 2003] et (Figure 1.11). [Azzi et al., 2003] ont montré que ces ERKs dépendaient des  $\beta$ -arrestines. Le propranolol est donc plus précisément un agoniste inverse pour la voie G et un agoniste partiel de la voie  $\beta$ -arrestine. Ce type de ligands n'est pas propre au récepteur  $\beta 2$ -adrénergique. SR121463B, un agoniste inverse du récepteur vasopressine 2 pour l'AC permet toujours la production de ERK1/2 et cela aussi a été montré dépendant des  $\beta$ -arrestines.

Des effets similaires ont été observés pour deux ligands ciblant le R7TM H3-histamine. Le premier, le proxyfan, est un agoniste partiel pour l'activité MAPK, le recrutement de la protéine  $G_{\alpha s}$  à la membrane ainsi que la diminution du signal AMPc et un agoniste inverse pour la production d'acide arachidonique. Le deuxième, le ciproxifan, est un antagoniste pour l'activité MAPK, un agoniste inverse pour le recrutement de la protéine  $G_{\alpha s}$  à la membrane, pour le signal AMPc qu'il augmente ainsi que pour la production d'acide arachidonique [Gbahou et al., 2003]. Nous citerons aussi les ligands TICP (Tyr-Ticpsi [CH<sub>2</sub>-NH]Cha-Phe-OH) et ICI174864 qui dirigés vers le R7TM  $\delta$ -opioïd, diminuent la quantité AMPc tout en maintenant la réponse ERK [Audet et al., 2005].

**FSH et Angiotensine II** Enfin, les récepteurs FSH et AT<sub>1A</sub> admettent au moins deux voies de signalisation indépendantes : dépendante des protéines G et dépendantes des  $\beta$ -arrestines (détails dans les section suivantes). Pour ces deux R7TMs, il existe aussi des ligands biaisés. C'est le cas du thiazolidinone pour le RFSH dont un des analogues agit en tant qu'agoniste de la voie  $G_{\alpha s}$  et de la voie  $G_{\alpha i}$  ce qui a des effets mixtes sur la signalisation [Arey et al., 2008]. Très récemment au sein de l'équipe BIOS deux résultats mettent en évidence des activations biaisées du récepteur dues au ligand. Ainsi, [Wehbi et al., 2010b] ont montré qu'une LH équine partiellement déglycosylée adressée au récepteur FSH au sein de cellules HEK activait préférentiellement la voie dépendante des  $\beta$ -arrestines. Et, [Wehbi et al., 2010a] met en lumière les propriétés de signalisation de certains anticorps naturellement produit par des brebis en réponse à une stimulation gonadotrophine chorionique équine (eCG) en vue d'une insémination artificielle. En liaison avec certains anticorps, la FSH endogène a une activité biaisée. In vivo, les anticorps étudiés donnaient de meilleurs taux de réussite à l'insémination.

Pour le R7TM à l'angiotensine II, il existe une multitude de médicaments ciblant le RAT<sub>1A</sub> en vue de soigner l'hypertension. Et, comme pour le RFSH, le RAT<sub>1A</sub> admet au moins deux voies de signalisation indépendantes : dépendante des protéines G et dépendante des  $\beta$ -arrestines. [Wei et al., 2003] propose un ligand sélectif de la voie

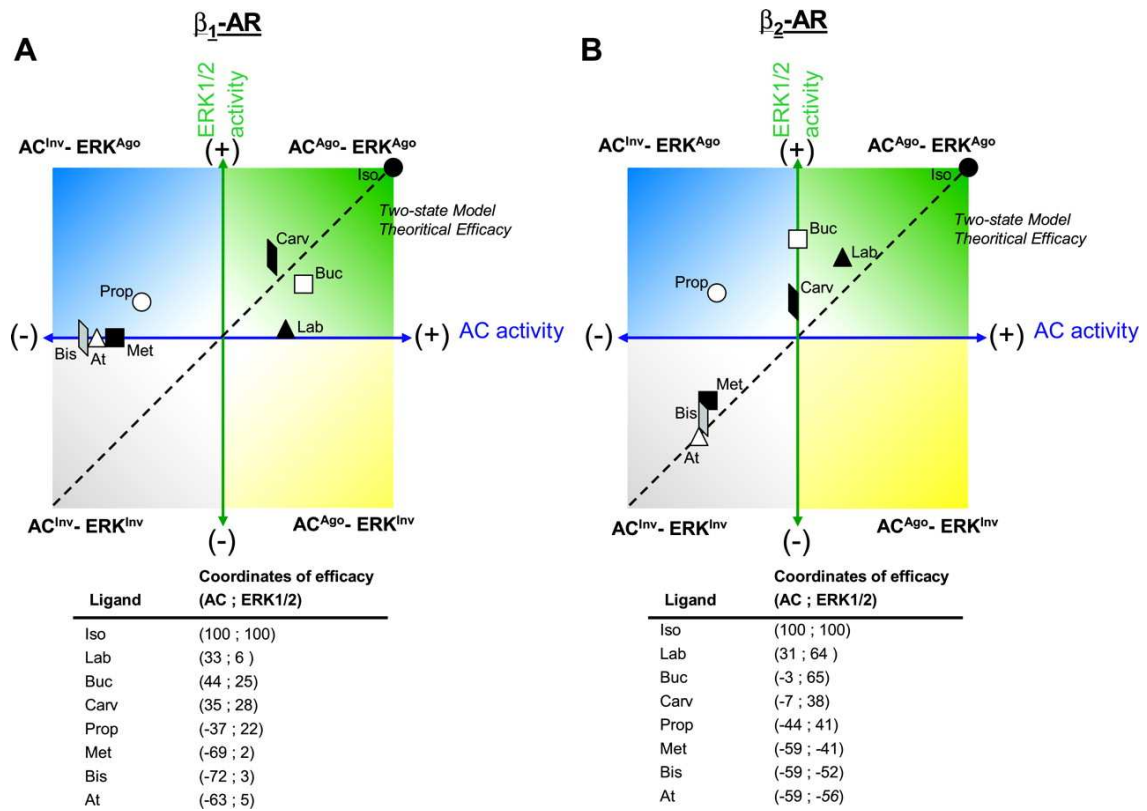


FIGURE 1.11 – Représentation cartésienne des profils d’efficacité des ligands testés sur le R7TM  $\beta_2$ -adrénergique [Galandrin and Bouvier, 2006]. Les axes représentent la voie AC en abscisse et la voie MAPK ERK1/2 en ordonnée. A/ Profil d’efficacité des ligands sur le récepteur  $\beta_1$ -adrénergique. B/ Profil d’efficacité des ligands sur le récepteur  $\beta_2$ -adrénergique. Les coordonnées sont obtenues à partir des tableaux figurant en dessous des graphiques. On retrouve les caractéristiques du propranolol (Prop), agoniste inverse au niveau de la production d’AC et agoniste partiel au niveau d’une production de ERK pour le récepteur  $\beta_2$ -adrénergique.

Abréviations : Inv : agoniste inverse, Ago : agoniste.

Noms des ligands : Iso : isoproterenol, Lab : labetalol, Buc : bucindolol, Carv : carvedilol, Prop : DL-propranolol, Bis : bisoprolol, Met : metoprolol, At : atenolol.

$\beta$ -arrestine en décrivent la signalisation du ligand du SII ([sarcosine1,Ile4,Ile8]Ang II) uniquement dépendante des  $\beta$ -arrestines. C’est d’ailleurs avec ce ligand qu’a été produit le phosphoprotéome de la voie  $\beta$ -arrestine [Xiao et al., 2010].

Dans la suite de l’introduction, nous présenterons avec plus de détails la biologie des récepteurs FSH et angiotensine 1A, deux R7TMs sur lesquels notre travail s’est concentré.

## 1.2 Le récepteur de la FSH

Le RFSH est la cible de l'hormone folliculo-stimulante (FSH), fortement impliquée dans la régulation de la gamétogénèse mâle et femelle chez les mammifères. Chez l'homme, la recherche de mode de contraception, l'augmentation des problèmes de fertilité, ainsi que diverses mutations génétiques du RFSH ou de la FSH gardent des zones d'ombre ou n'ont pas encore de traitements adaptés. Au niveau des animaux de rente, les éleveurs font massivement appel à l'insémination artificielle qui implique l'utilisation de eCG comme agoniste des RFSH et RLH, ce qui est une source potentielle de problèmes sanitaires et n'est pas toujours efficace. Un préalable pour apporter des réponses à ces différents problèmes est de comprendre comment agit la FSH au niveau de ses cellules cibles.

**Physiologie** La fonction de reproduction chez les mammifères est contrôlée par les gonadotrophines hypophysaires : la LH et la FSH. Ces gonadotrophines sont synthétisées et sécrétées par l'hypophyse antérieure en réponse à une stimulation pulsatile de GnRH et vont intervenir dans la maturation des gamètes. Chez la femelle, le gamète est l'ovule, il est produit par l'ovaire. Les ovules proviennent d'ovogonies qui prolifèrent durant le début de la vie embryonnaire. Cette multiplication des ovogonies cesse puis ces cellules germinales s'entourent de cellules de granulosa pour former des follicules. La femelle présente ainsi un nombre maximal de cellules germinales à la naissance. Cette quantité va diminuer au cours de la vie : les follicules peuvent dégénérer ou produire une ovulation à partir de la puberté. L'ovulation est l'expulsion d'un ovule fertile ; prêt à être fécondé. Au sein de ce système, la FSH aide à la maturation des follicules ovariens en induisant leur croissance. Pour ce faire, la FSH cible ses récepteurs situés en majorité à la surface des cellules de granulosa et induit la prolifération et la différenciation de ces dernières (Figure 1.12). La FSH induit aussi la formation de l'antrum au sein des follicules et la maturation des ovocytes. Une fois l'ovocyte mature, la LH dont la quantité augmente transitoirement, provoque l'ovulation et la lutéinisation.

Chez le mâle, le gamète est le spermatozoïde, il est contenu dans les tubes séminifères des testicules. Contrairement aux cellules germinales femelles, les spermatogonies, dont sont issus les spermatozoïdes, ne vont pas s'arrêter de proliférer au cours de la vie. Par contre, les cellules de Sertoli, cible majoritaire de la FSH, cessent de se multiplier avant la puberté. Ces cellules agissent sous l'action de la FSH comme des cellules nourricières des spermatogonies et aident à la maturation en spermatozoïdes. Un troisième type de cellule joue un rôle dans la spermatogénèse, ce sont les cellules de Leydig. Elles sont situées autour des tubes séminifères et en réponse à la LH, sont responsables de 95% de la production de testostérone, une hormone dont dépend principalement la spermatogénèse adulte (Figure 1.12).

**Pathologies liées à la FSH** Des anomalies dans la structure de la FSH ou de son récepteur (RFSH) conduisent à des problèmes de fertilité. Ainsi, une mutation de la sous-unité  $\beta$  de la FSH a été décrite pour les deux sexes. Tous étaient stériles, cependant, les conséquences chez les hommes sont moins sévères, malgré leur azoospermie, ils ont une puberté normale [Layman et al., 2002]. D'autre part, diverses mutations du RFSH,

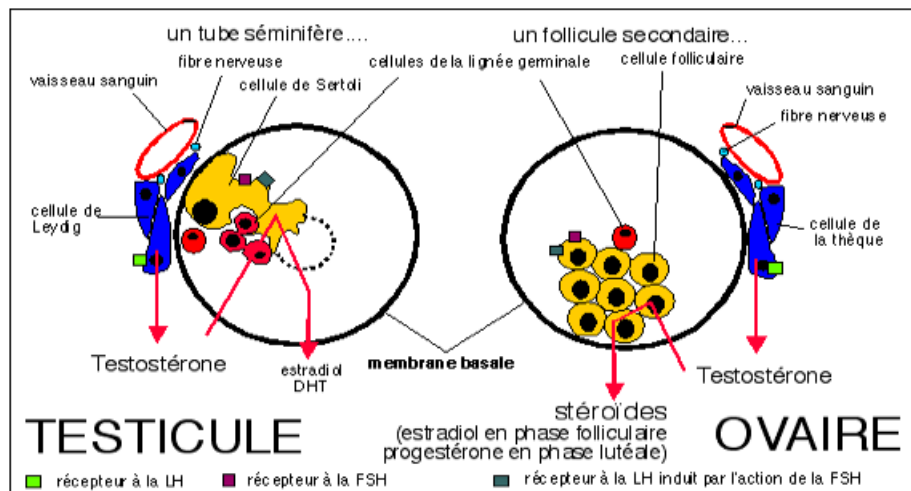


FIGURE 1.12 – Représentation schématique de coupes dans le testicule et l’ovaire (<http://pst.chez-alice.fr/svtium/vie.htm>). Il y a analogie entre les cellules de Leydig bordant les tubes séminifères et les cellules de thèque des follicules, ces deux types cellulaires produisent des stéroïdes en réponse à la LH pour soutenir la gamétogenèse. De même les cellules internes : les cellules de Sertoli et les cellules de granulosa sont les cibles de la FSH et nourrissent les futurs gamètes.

comme par exemple Ala189Val, conduisent à une hypofertilité chez les hommes et un arrêt de la croissance folliculaire chez les femmes empêchant l’ovulation [Tapanainen et al., 1997; Aittomäki et al., 1995].

**Structure** Le RFSH fait partie de la famille 1c des R7TMs selon le classement de [Bockaert and Pin, 1999]. Chez l’humain, le RFSH a été cartographié sur le chromosome 2 et contient 695 amino-acides pour 78kDa. Il est muni d’un domaine extra-cellulaire proéminent et fixe la FSH entre l’arginine 265 et la sérine 296 [Liu et al., 1994; Fan and Hendrickson, 2005] (Figure 1.13). Cependant l’intégralité du domaine N-terminal est nécessaire au bon fonctionnement du RFSH. Il est composé de domaines riches en leucines qui lui confèrent une structure en fer à cheval. Sept domaines trans-membranaires suivent le domaine extra-cellulaire. On notera l’importance du motif d’acides aminés ERW présent au sein de la deuxième boucle intracellulaire, il permet la liaison du récepteur à la protéine G [Timossi et al., 2002]. Enfin, le RFSH est constitué d’un domaine C-terminal intracellulaire, les multiples sérines et thréonines forment 14 sites de phosphorylations (Figure 1.13).

Lorsque la FSH vient se fixer au sein du domaine extra-cellulaire, elle entraîne un changement de conformation, départ de la signalisation.

### 1.2.1 Réseaux de signalisation de la FSH

Suite à la liaison de la FSH, les passages transmembranaires II et IV bougent l’un par rapport à l’autre et libèrent le site d’interaction du RFSH avec la protéine G.



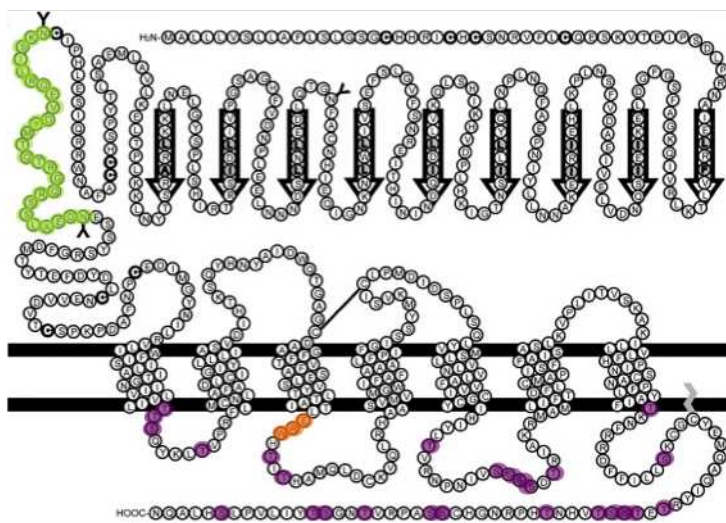


FIGURE 1.13 – Structure du récepteur FSH humain [Bogerd et al., 2005]. Il est composé d’un domaine extra-cellulaire important comportant des séquences riches en leucine induisant une structure en fer à cheval. Ces domaines riches en leucines sont indiqués par des flèches noires. Le site de liaison de la FSH est indiqué en vert et correspond à la suite d’acides aminés [Arg265 .. Ser296]. Suit les 7 domaines trans-membranaires ainsi que la partie C-terminale au sein du cytoplasme. La protéine G vient se fixer sur le motif ERW au départ de la seconde boucle intracellulaire (en orange). De nombreux sites de phosphorylation sont présents au sein de la queue C-terminale et des boucles intracellulaires (en violet), ce sont les sérines et les thréonines. Enfin, le RFSH subit des transformations post-traditionnelles comme la palmitoylation du C terminal représenté par un zigzag gris et permettant un point d’ancrage dans la membrane.

### 1.2.1.1 Voie $G_{\alpha s}$ /AC/AMPC/PKA

Une partie de la signalisation découle du couplage à la protéine G hétérotrimérique contenant la sous-unité  $\alpha s$  [Dattatreya et al., 1987]. Cette sous-unité  $G_{\alpha s}$  va échanger son GDP contre un GTP et activer l’adénylate cyclase (AC). L’AC convertit ensuite l’ATP en AMPc, principal second messager de la FSH. L’AMPc est régulé négativement par les PDEs et peut à son tour activer la PKA.

**PKA régulateur des MAP kinases** La PKA induit l’activation de la cascade MAPK Raf/MEK/ERK dans les cellules de granulosa [Cottom et al., 2003], dans les cellules de Sertoli [Crépieux et al., 2002] et dans les cellules HEK293 sur-exprimant le RFSH [Kara et al., 2006]. Dans les cellules HEK293, en inhibant PKA à l’aide de H89, [Kara et al., 2006] ont montré que les ERKs dépendantes de PKA, *GpERK*, étaient activés de manière rapide et transitoire (Figure 1.14). Il est supposé que ces *GpERK* formés au sein du cytoplasme migrent au sein du noyau. Dans les cellules de granulosa, les *GpERK* inhibent l’expression de StAR (Steroidogenic Acute Regulatory protein), une enzyme prenant part à la formation de progestérone [Seger et al., 2001]. Dans la granulosa comme dans

les cellules de Sertoli issus de rats prépubères, les ERK sont nécessaires à la réponse proliférative induite par la FSH [Crépieux et al., 2001; Kayampilly and Menon, 2004].

Outre les ERKs, la PKA active p38. Dans les cellules primaires de granulosa, la PKA a été montrée comme étant nécessaire et suffisante pour phosphoryler la protéine kinase p38 sur la théonine 180 et la tyrosine 182 [Maizels et al., 1998]. En effet, un inhibiteur de PKA (H89) annule totalement la formation de Pp38{*Thr*180, *Tyr*182} tandis que sans stimulation FSH, la forskoline (activateur de l'AMPc en aval de PKA) rétablit le signal Pp38. Cette Pp38 va ensuite phosphoryler Heat Shock Protein 27 (HSP-27). Cette dernière phosphoprotéine est responsable d'une réorganisation des filaments d'actine au sein de la cellule de granulosa immature, ce qui préviendrait ces cellules de granulosa en maturation de l'apoptose. Enfin, comme ERK, p38 inhibe StAR et la production de stéroïdes qui en découle.

**PKA impliquée dans l'activation de facteurs de transcriptions** La PKA active aussi le facteur de transcription CREB et son régulateur CREM, tous deux sont essentiels au bon déroulement de la spermatogenèse [Scobey et al., 2001]. L'expérience le montrant a consisté à exprimer un mutant de CREB non phosphorylable au sein de cellules de Sertoli ré-injectées au sein des tubes séminifères de rats. Cela a conduit à une altération de la spermatogenèse, les cellules germinales sont entrées en apoptose avant d'être matures. Au sein des cellules de granulosa et après stimulation FSH, CREB va aussi soutenir la gamétogenèse en stimulant l'expression du récepteur LH en vue du pic ovulatoire et en favorisant l'expression de l'aromatase capable de transformer la testostérone produite pas les cellules de thèque en oestradiol [Carlone and Richards, 1997]. La PKA induit une multitude d'autres facteurs de transcription, parmi lesquels : NF- $\kappa$ B (nuclear factor-kappa B), SF-1 (steroidogenic factor 1), USF (upstream stimulating factor), NF-1, HIF-1 $\alpha$  (hypoxia-inductible factor 1  $\alpha$ ) et C/EBP $\beta$  (CCAAT-box/enhancer-binding protein  $\beta$ ) [Lécureuil et al., 2007]. Plus généralement, la majorité des gènes cibles de la FSH n'ont pas d'éléments de réponse à l'AMPc et la PKA active une majorité des gènes en réponse à la FSH mais pas tous, d'autre voies de signalisation impliquant entre autre PKB sont essentielles [Salvador et al., 2001; Zeleznik et al., 2003; Escamilla-Hernandez et al., 2008].

**PKA régule la traduction** D'autre part et plus récemment, la PKA a été impliquée dans la régulation de l'initiation de la traduction. En effet, PKA va phosphoryler p70S6K sur les sites {*Thr*389, *Ser*421/424} au sein des cellules de Sertoli et sur le site {*Thr*389} au sein de cellules de granulosa. La Pp70S6K entraîne ensuite l'activation par phosphorylation des protéines rpS6 impliquées dans des complexes d'initiation de la traduction [Alam et al., 2004; Lécureuil et al., 2005].

### 1.2.1.2 Voie Gas/PI3K/PKB

D'autres voies de signalisation existent en aval de l'AMPc indépendamment de la PKA. Au sein des cellules OSE, [Choi et al., 2009] ont montré que la FSH favorisait la croissance cellulaire. Ils ont découvert que cette régulation intervenait via les voies de signalisation AMPc/Epac/PI3K et AMPc/Epac/ERK1/2. La PI3K (Phosphoinositide 3-kinase), aussi

activée directement par la sous-unité  $G_{\alpha s}$  va jouer un rôle sur la phosphorylation de p70S6K et ce de façon complémentaire à PKA. PI3K convertit le PIP2 en PIP3, ce dernier active en s'y complexant la PDK (phosphoinositide-dependent kinase). La PDK phosphoryle ensuite PKB (protein kinases B) qui active Rheb (Ras homolog enriched in brain) et par conséquent mTOR (Target of Rapamycin) [Stokoe et al., 1997; Gonzalez-Robayna et al., 2000; Alam et al., 2004]. mTOR agit à son tour sur différents facteurs de transcription tel HIF-1 [Alam et al., 2004], mais aussi sur la phosphorylation des sites  $\{Thr389\}$  et  $\{Ser421/424\}$  de la p70S6K. Cette régulation de p70S6K a un impact sur l'activation de la protéine ribosomale S6 et sur EIF-4B qui régulent la traduction. Dans la partie résultats de cette thèse, nous présenterons une étude et une modélisation dynamique des différentes formes phosphorylées de la p70S6K dans les cellules de Sertoli primaires [Musnier et al., 2009].

### 1.2.1.3 Autres voies dépendantes de $G_{\alpha s}$ et $G_{\alpha i}$

**Voie  $G_{\alpha s}$ /Epac/Rap** Toujours en aval d'Epac, au sein de cellules de granulosa, [Wayne et al., 2007] ont observé, qu'après stimulation par la FSH, Epac induit l'activation de Rap qui déclenche la cascade MAPK p38 cette fois indépendamment de la PKA.

**Voie  $G_{\alpha i}$**  Au sein des cellules de Sertoli, [Loganzo and Fletcher, 1992] ont montré que la voie  $G_{\alpha s}$ , activée par la FSH, régule négativement la formation des mRNAs codant pour les sous-unités  $G_{\alpha i-1}$  et  $G_{\alpha i-2}$  et positivement les mRNA codant pour la sous-unités  $G_{\alpha i-3}$ . Or, [Eskola et al., 1998] montrent qu'au sein des cellules de Sertoli, l'AMPc augmente de façon plus significative lorsque  $G_{\alpha i}$  est inhibée en présence de FSH. Plus précisément, les protéines G contenant une sous-unité  $\alpha i-2$  inhibent la voie  $G_{\alpha s}$ /AC. On a alors une boucle de rétroaction positive de la voie  $G_{\alpha s}$ /AC sur elle même.

Comme pour les autres R7TMs, les différentes voies déclenchées par les protéines G ont longtemps été considérées comme l'unique signalisation du RFSH. Mais, petit à petit, au sein des gonades, d'autres effecteurs ont montré leur importance comme les  $\beta$ -arrestines.

### 1.2.1.4 Signalisation indépendante des protéines G

**Phosphorylation du récepteur par les GRKs** Les  $\beta$ -arrestines sont majoritairement recrutées au récepteur lorsqu'il est phosphorylé. Le récepteur FSH comporte quatorze sites de phosphorylation (surlignés Figure 1.13). Soit d'innombrables combinaisons de phosphorylations possibles ( $2^{14}$ ) et autant de modulations du signal. Par la suite, nous allons considérer deux combinaisons de phosphorylation  $\{P1\}$  et  $\{P2\}$ . La première combinaison est le résultat des GRKs 2 tandis que la seconde est le résultat des GRKs 5 et 6 (Figure 1.15). Remarquons que cette différenciation s'est effectuée en fonction de la similitude des effets produits par les deux couples de GRKs [Kara et al., 2006].

Une fois phosphorylé le récepteur FSH recrute les  $\beta$ -arrestines 1 et 2 [Marion et al., 2002; Marion et al., 2006]. Ces protéines d'échafaudages établissent un complexe stable avec le récepteur. Le RFSH est donc un R7TM de classe B. Puis, ce complexe est internalisé par un processus dépendant des clathrines.

**Induction des MAP kinases ERK par la voie des  $\beta$ -arrestines** Les récepteurs phosphorylés par les GRKs 5 et 6 induisent une signalisation ERK. Les ERK phosphorylés provenant de la voie GRK56/ $\beta$ -arrestines (*bpERK*) apparaissent plus lentement. Une deuxième différence avec la voie G se situe au niveau de leur localisation sub-cellulaire. Alors que les *GpERK* sont transférés au sein du noyau, les *bpERK* restent séquestrées dans le cytoplasme.

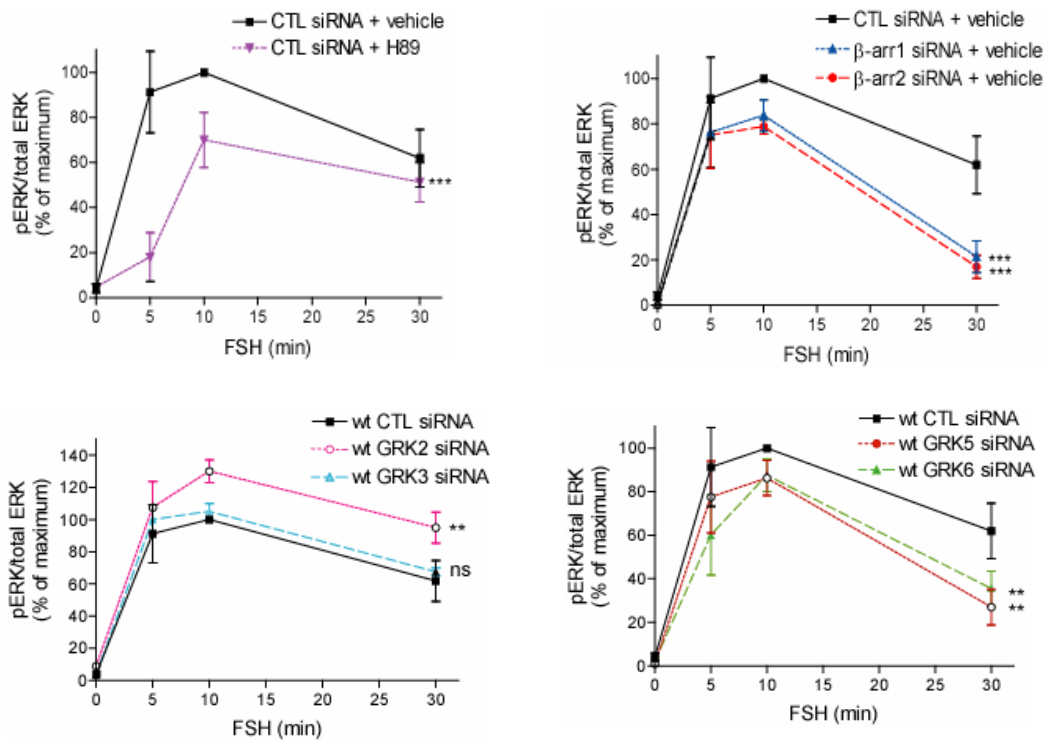


FIGURE 1.14 – L’activation du RFSH induit une augmentation des ERK phosphorylés par deux voies de signalisation distinctes [Kara et al., 2006]. A/ et B/ En inhibant sélectivement la PKA et les  $\beta$ -arrestines, on observe la part *GpERK* et la part *bpERK* au sein de signalisation ERK. C/ et D/ On observe l’implication opposée des GRK2 et GRK5 et 6 dans la signalisation ERK.

### 1.2.1.5 Désensibilisation du signal FSH

La cible classique de la désensibilisation du RFSH est l’AMPc. Dans un premier temps, le couplage du récepteur à la protéine *G $\alpha$ s* est inhibé afin de diminuer la formation d’AMPc. Cet état résulte des phosphorylations du récepteur par les GRKs ainsi que du recrutement des  $\beta$ -arrestines [Troispoux et al., 1999; Marion et al., 2002]. Les  $\beta$ -arrestines vont participer à l’internalisation du récepteur. Une fois au sein du cytoplasme ce récepteur n’induit

plus d'AMPc, il peut engendrer d'autres voies de signalisation comme les *bpERK* décrit ci-dessus, être recyclé ou dégradé. En parallèle, la PDE augmente en réponse à la FSH et inhibe directement l'AMPc [Conti et al., 1982; Levallet et al., 2007].

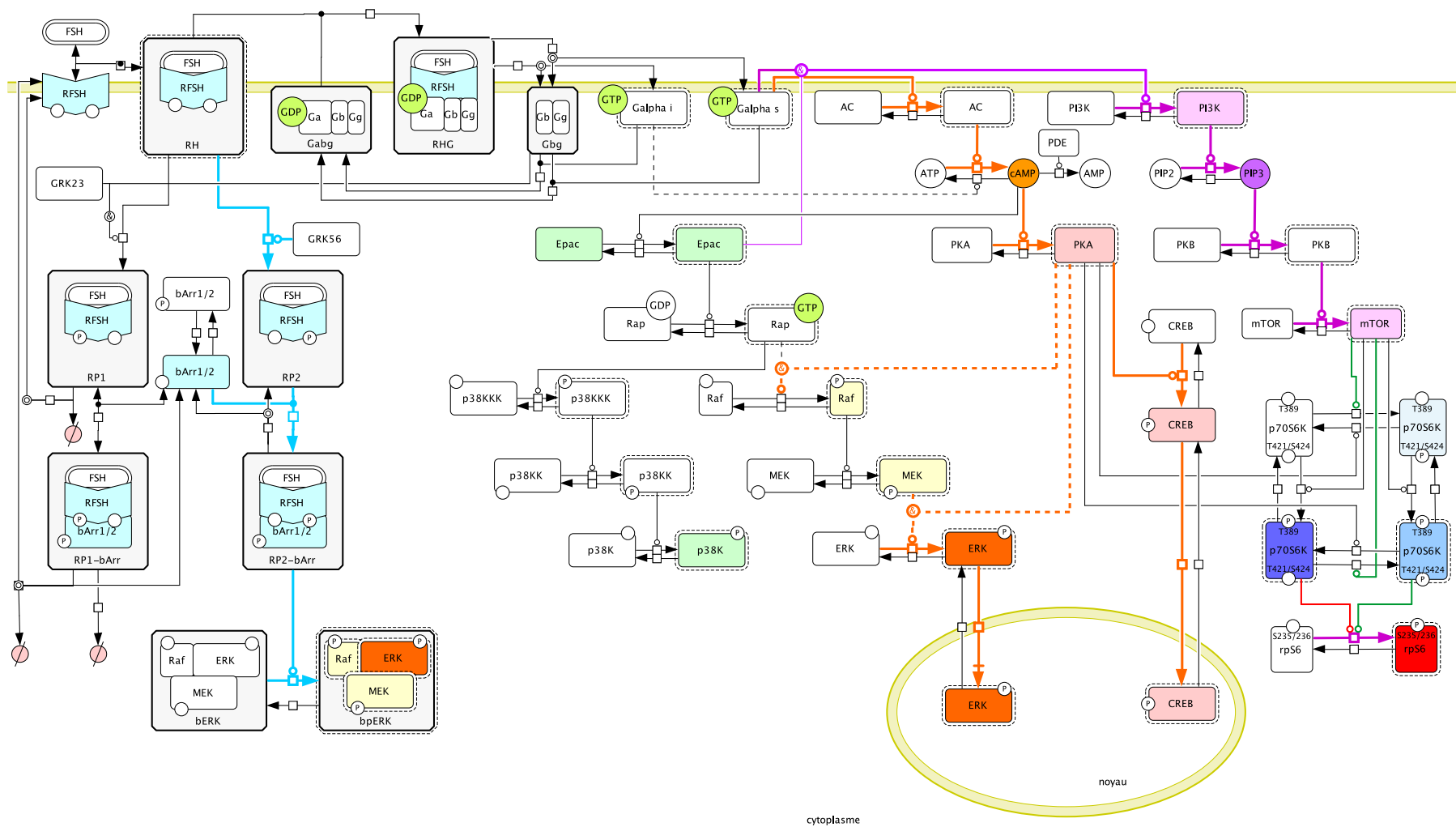


FIGURE 1.15 – Réseau de signalisation de la FSH comprenant les voies dépendantes des protéines G et des  $\beta$ -arrestines, les régulations de p70S6K qui seront détaillées dans un modèle de signalisation au sein de cellules de Sertoli à deux stades de développement (en vert, les catalyses présentes en prolifération et en rouge celles présentes au stade différencié)

### 1.3 Le récepteur de type 1A de l'angiotensine

Le deuxième R7TM que nous avons étudié est le récepteur cible de l'angiotensine II. C'est un R7TM modèle dont le réseau de signalisation est déjà bien étudié. Des similitudes au sein des réseaux de signalisation FSH et angiotensine existent, même si les fonctions physiologiques qu'ils régulent sont vraiment différentes.

**Physiologie** L'angiotensine II est un puissant vasoconstricteur Elle provoque la diminution du diamètre des vaisseaux et est à l'origine d'une activation des muscles lisses de l'organisme. Ces muscles lisses tapissent entre autres les artérioles mais aussi l'utérus. Il existe cinq types de récepteurs à l'angiotensine II : AT1, AT2, AT3, AT4, AT5. Les deux premiers sont les plus fréquemment étudiés et l'AT1 est considéré comme le principal responsable des effets de l'hormone. L'AT2 est un R7TM qui a des effets antagonistes (vasodilatation), il est très actif chez le fœtus puis son importance s'inverse avec l'AT1 chez l'adulte. L'AT1 est aussi un R7TM présent dans le rein, le foie, la glande surrénale, mais aussi au niveau du système cardiovasculaire et du cerveau. L'angiotensine II induit une élévation de la pression artérielle en ciblant le système cardiovasculaire mais aussi en prenant part à la régulation du volume sanguin. Pour cela, l'angiotensine II est très liée à l'absorption ou relargage de liquide et d'ions dans le sang à partir des reins. L'angiotensine peut entraîner une sensation de soif au niveau du cerveau, toujours pour augmenter indirectement le volume sanguin, ce qui a pour conséquence une augmentation de la pression artérielle (Figure 1.16).

**Pathologies liées à la signalisation angiotensine II** La principale pathologie liée à l'angiotensine est l'hypertension. Ce problème est très répandu et très largement traité, il fait partie des 5 grands thèmes de recherche des entreprises pharmaceutiques. De nombreux médicaments antihypertenseurs existent. Certains inhibent la rénine qui convertit l'angiotensinogène en angiotensine II (Figure 1.16). D'autres médicaments, abrégés ICE, bloquent l'enzyme de conversion de l'angiotensine (ACE). Enfin, un troisième mécanisme propose des agents actifs agissant comme antagonistes du récepteur à l'angiotensine II (abrégée ARA2) c'est notamment le cas du valsartan et du losartan.

Les ICEs et ARA2 inhibent l'activité angiotensine II et empêchent le rétrocontrôle négatif de la signalisation angiotensine sur la sécrétion de rénine. La rénine s'accumule et induit l'expression des récepteurs AT2 responsables de vasodilatation. Pour cet ensemble de raisons, le RAT<sub>1A</sub> est l'objet de nombreux travaux notamment concernant ses mécanismes de signalisation cellulaire. Il constitue également un R7TM modèle pour l'étude de la signalisation  $\beta$ -arrestine. C'est à ce titre que nous avons abordé l'étude de ce récepteur.

**Structure du RAT<sub>1A</sub>** Tout d'abord, comme le RFSH, le RAT1 fait partie de la famille 1 des R7TMs. Il appartient au sous-groupe 1a des récepteurs de type rhodopsine. Chez l'humain, le RAT<sub>1</sub> a été cartographié sur le chromosome 3 et contient 359 amino-acides pour 41kDa. Au cours de cette thèse nous avons travaillé sur des cellules HEK 293 exprimant le RAT<sub>1A</sub> de rat. Comme indiqué à la figure 1.17, le RAT<sub>1A</sub> présente un ectodomaine beaucoup plus court que le RFSH. Il possède un motif DRY (Asp, Arg, Tyr) à la fin du

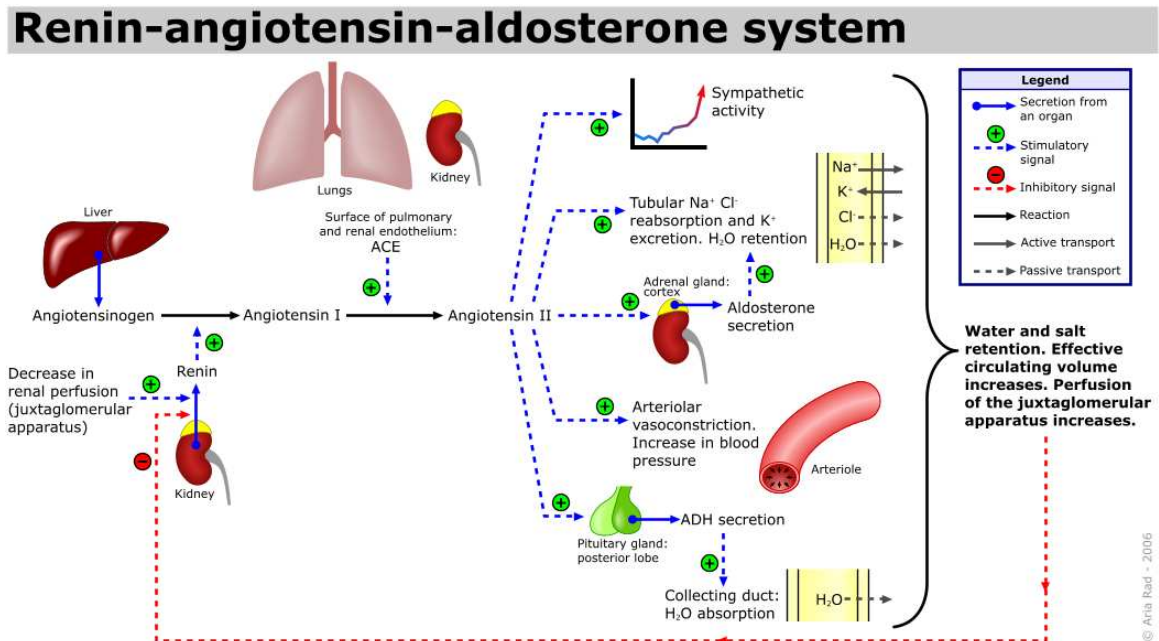


FIGURE 1.16 – Illustration issue d'un article wikipédia. L'angiotensine est issue d'une conversion d'angiotensinogène. Cette protéine inactive est produite par le foie puis transformée par la rénine et l'enzyme de conversion de l'angiotensine (l'ACE). Lorsque la pression baisse, la rénine stimule la conversion d'angiotensine. L'angiotensine II va se fixer sur ses R7TMs et favorise l'élévation de la pression artérielle par les différents mécanismes décrits ci-dessus.

domaine transmembranaire III fonctionnellement comparable au motif ERW (Glu, Arg, Try) du RFSH. Son domaine C-terminal contient 13 sérines et théonines potentiellement phosphorylables.

### 1.3.1 Signalisation de l'AT<sub>1A</sub>

Comme pour la plupart des R7TMs, il a d'abord été identifié que le RAT<sub>1A</sub> se couple aux protéines G. Cependant, alors que le RFSH activé se couple préférentiellement aux protéines G<sub>αs</sub>, le RAT<sub>1A</sub> une fois activé recrute principalement les protéines G<sub>αq</sub>/11 [Shirai et al., 1995]. Un parallèle peut s'effectuer entre les réseaux FSH et angiotensine II, les cinétiques d'activation des ERKs sont comparables (Figures 1.8, 1.14 et 1.20). Cependant, au niveau de la voie de signalisation ERK dépendantes des protéines G le récepteur angiotensine admet comme initiateur de la cascade ERK la voie G<sub>αq</sub>/PLC/PKC.

#### 1.3.1.1 Signalisation dépendante des protéines G

**Voie G<sub>αq</sub>/PLC/Ca** La principale voie d'action du AT<sub>1A</sub> passe par le couplage du récepteur aux protéines G<sub>αq</sub> [De Lean et al., 1980; Samama et al., 1993; Milligan and Kostenis, 2006]. L'activation de la protéine G après couplage du récepteur à l'angioten-



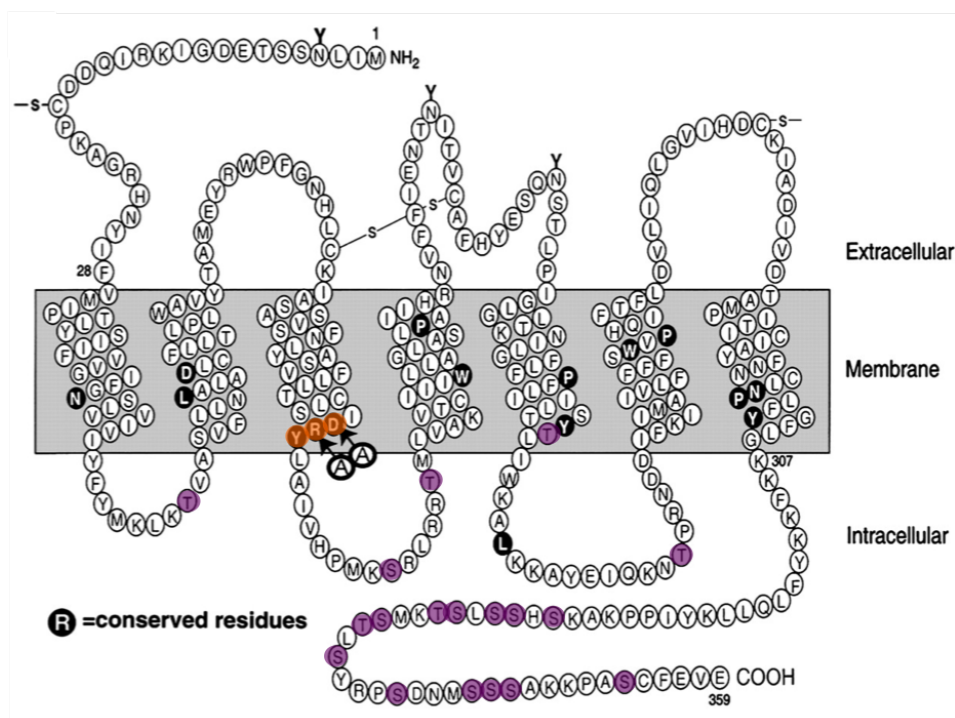


FIGURE 1.17 – Structure du récepteur  $AT_{1A}$  (rat) [Wei et al., 2003]. Contrairement au RFSH, le  $RAT_{1A}$  admet un N-terminal de petite taille. Des ponts disulfures sont formés entre les résidus extra-cellulaire Cys18 et Cys274 ainsi qu’entre Cys101 et Cys180. Enfin, comme pour le RFSH, le motif DRY est surligné en orange et sa mutation en AAY donne un récepteur mutant incapable de recruter les protéines G.

sine II induit un recrutement immédiat des PLC- $\beta$ 1, PLA<sub>2</sub> (Phospholipase A 2) et PLD (Phospholipase D) [Mehta and Griendling, 2007]. Les sous-unités  $\beta\gamma$  des protéines G sont effectrices de ces activations de phospholipases [Ushio-Fukai et al., 1998; Ushio-Fukai et al., 1999].

La PLC va induire une augmentation d’IP3 et de DAG en quelques secondes. L’IP3 va se complexer à son récepteur situé à la membrane du réticulum endoplasmique et induire une augmentation de calcium dans le cytoplasme. Le calcium va aller se lier à la calmodulin et activer les MLCK (MLC, myosin light chain kinase). Puis, les MLCK vont induire la phosphorylation des chaînes légères de myosine (MLC, myosin light chain). Ces MLC vont alors consommer un ATP pour se courber et interagir avec les filaments d’actine pour provoquer la contraction des cellules musculaires lisses vasculaire (CMLV) responsable de l’effet vasoconstricteur [Horowitz et al., 1996]. La régulation de ce mouvement passe par la déphosphorylation de MLC par des MLCP (myosin light chain phosphatase). Lorsque ces MLCP sont inhibées par la Rho kinase, la contraction est prolongée [Jin et al., 2006]. Il a même été montrée que sans Rho, le calcium ne pouvait pas agir sur les filaments d’actine [Uehata et al., 1997].

La PLD va induire de l’acide phosphatidique (PA) à partir de phosphatidylcholine (PC). Ce PA va soutenir la réponse DAG [Mehta and Griendling, 2007].

**Signalisation MAPK ERK1/2** En parallèle de l'IP3, le DAG active PKC. La PKC induit l'activation d'une protéine initiatrice de la cascade ERK, Ras, par un échange GDP/GTP. Puis cette protéine va phosphoryler une MAPKKK : Raf. Cette kinase va à son tour phosphoryler la MAPKK de la cascade : MEK1/2. Enfin, cette dernière kinase active les MAPK ERK1/2. L'activation de ces MAPK ERK conduit à une augmentation de l'expression des gènes *c-fos*. En parallèle, la MAPK JNK va induire *c-jun*. En se complexant, *c-fos* et *c-jun* forment l'hétérodimère AP1 qui est un facteur de transcription. En réponse à l'angiotensine II, les MAPK vont alors induire la différenciation, migration et adhésion cellulaire en se complexant à des promoteurs de gènes. Nous appellerons les ERKs phosphorylés au sein de la voie G : *GpERK*. Ces *GpERK* ont une cinétique d'activation rapide et brève et sont transportés au sein du noyau (Figures 1.8 et 1.20) [Ahn et al., 2004a].

Enfin, la PKC phosphoryle la p47 phagocyte oxidase (p47phox) ce qui a pour effet l'activation du complexe enzymatique NADPH oxydase qui catalyse le transfert d'électron du NADPH et induit de multiples radicaux libres oxydant (ROS) [Brown and Griendling, 2009] et (Figures 1.18, 1.19 et 1.21). Comme illustré (Figure 1.19), la NADPH oxydase est capable d'activer les cascades MAPK p38 et JNK.

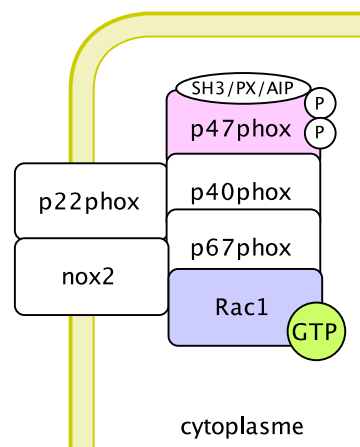


FIGURE 1.18 – La NADPH oxydase est un complexe enzymatique formé de 6 sous-unités : p22phox et Nox 2 qui sont transmembranaires ainsi que p40phox, p47phox, p67phox et Rac cytoplasmique. En mauve, la p47phox a particulièrement été étudiée, elle comporte un domaine d'homologie phox, deux domaines d'homologie src 3 (SH3) et une région auto-inhibitrice (AIR) ainsi que plusieurs sites de phosphorylation entre les Ser303 et Ser379 [El-Benna et al., 2009]. La NADPH oxydase est activée par l'angiotensine de manière directe, rapide et brève en aval de la PKC qui agit sur p47phox. En bleu, Rac1 qui est une cible de la transactivation du récepteur EGF par l'angiotensine et permet l'activation de NADPH oxydase de façon plus lente et soutenue [Brown and Griendling, 2009] et (Figure 1.21)

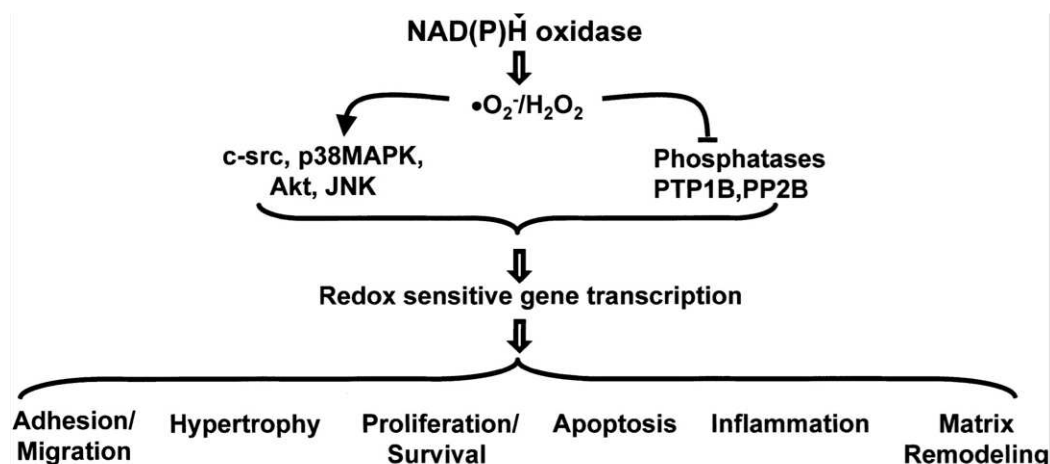


FIGURE 1.19 – La NADPH induit la production de ROS qui conduit à l’activation de kinases sensibles à une oxydoréduction mais aussi à l’inactivation de phosphatases spécifiques. Il en résulte une modulation de l’expression génique. L’impact de l’oxydase NADPH se ressent dans l’adhérence cellulaire, la migration des monocytes/macrophages, dans la prolifération et la survie de plusieurs types de cellules vasculaires, l’apoptose, l’inflammation et le remodelage de la matrice extra-cellulaire. [Griendling et al., 2000; Lassègue and Griendling, 2010]

**Signalisation MAPK p38 et JNK** Ces cascades sont soutenues par l’activation via le  $RAT_{1A}$  de la protéine  $G\alpha_{11/12}$  qui va induire l’activation de la kinase Rho qui a son tour active la petite protéine G Rac. La petite protéine Rac est impliquée dans l’initiation des cascades p38 et JNK. Une fois active, Rac va activer les MAPKKK TAK (TGF-beta activated protein kinase) et TAO (thousand-and-one amino acids) pour la cascade p38 et MEKK1-3 pour la cascade JNK. Ces kinases vont à leur tour phosphoryler la MAPKK de la cascade. Pour la cascade p38, il s’agit de MEK3 et MEK6 et pour la cascade JNK de MEK4. Enfin, ces dernières kinases activent les MAPK JNK et p38 [Touyz and Schiffrin, 2000].

Comme au sein du réseau FSH, la Pp38 induit une phosphorylation de la chaperone HSP-27 qui va préserver et stabiliser les protéines intracellulaires.

**Signalisation dépendante des tyrosines kinases de la famille src** Le  $RAT_{1A}$  induit aussi l’activation de Src, une tyrosine kinase impliquée dans un large éventail de phosphorylations et activations de protéines intracellulaires. Src est en fait activée par le complexe  $G\beta\gamma$  et de manière dépendantes de ROS. Src active Ras, FAK, JAK/STAT, et  $PLC\gamma$ . Ce deuxième mécanisme d’activation de la PLC a été identifié récemment. Après interaction avec la  $PLC\beta$  dépendante de l’activation des protéines G, le récepteur se complexe à la  $PLC\gamma$ , ce processus permet d’activer la cascade MAPK Raf/MEK/ERK et de faire durer la vasoconstriction initiée par  $PLC\beta_1$  [Mehta and Griendling, 2007]. Src va induire la cascade MAPK JNK via FAK et soutenir l’expression de c-jun et la formation d’AP1 qui en découle. FAK va aussi mobiliser des protéines du cytosquelette telles p130Cas, Pyk2 qui

en association avec PI3K vont activer p70S6K, impliqué, comme au sein du réseau FSH, dans la synthèse protéique. Src va aussi être activateur de la PKD induisant la croissance cellulaire et l'adhésion des cellules nécessaire lors de lésions.

**Signalisation JAK/STAT** La signalisation JAK/STAT est impliquée dans la division cellulaire. La queue C-terminale du RAT<sub>1A</sub> a été identifiée comme site de liaison à JAK2. En se complexant au récepteur, JAK2 est phosphorylé et activé. [Frank et al., 2002] ont montré que pour ce recrutement, l'activation au préalable de PKC par la voie dépendante des protéine G était nécessaire. À travers, cette voie de signalisation, l'angiotensine induit croissance, migration et remodelage cellulaire.

**Transactivation** Enfin, il n'est pas à négliger que les R7TMs peuvent transactiver d'autres récepteurs et induire des réseaux de signalisation qui ne leur sont pas propre. Le plus fréquent est l'activation du récepteur EGF, c'est le cas au sein du réseau angiotensine et cela se produit via src. En découle une vaste quantité d'effecteurs, en particulier la cascade PI3K/PDK1/PKB et la voie PI3K/Rac1/NADPH oxydase. Le RAT<sub>1A</sub> transactive d'autre récepteurs tyrosine kinase comme le récepteur PDGF (Platelet-derived growth factor) ainsi que le récepteur insuline [Yin et al., 2003].

#### 1.3.1.2 Voies $\beta$ -arrestines

**Phosphorylation du récepteur par les GRKs** La phosphorylation du récepteur par les GRKs facilite le recrutement des  $\beta$ -arrestines. Alors que le RFSH comporte 14 sites de phosphorylation, le RAT<sub>1A</sub> en compte 13 (surlignés Figure 1.17). Pour ces deux modèles, les sites sont répartis sur la queue C-terminale et les boucles intracellulaires. Par la suite nous considérerons deux combinaisons ou ensembles de combinaisons de sites phosphorylés :  $\{P1\}$  et  $\{P2\}$ . La première combinaison est le résultat de l'action des GRKs 2 et 3 tandis que la seconde est induite par les GRKs 5 et 6 (Figure 1.21). Comme pour le RFSH, cette différenciation s'effectue sur base de ressemblance des effets produits par les deux couples de GRKs [Kim et al., 2005].

**Désensibilisation** Comme au sein du réseau FSH, les  $\beta$ -arrestines 1 et 2 désensibilisent le récepteur (Figure 1.21). Il faut noter que les DGKs et les PDEs font partie de l'interaction des  $\beta$ -arrestines 1/2 induit pas le RAT<sub>1A</sub> ??, Tableau 1.2. Tout porte donc à croire que la double désensibilisation énoncés pour les récepteurs  $\beta$ 2-adrénergique et M1 muscarinique en introduction générale (section 1.1) pourraient être généralisable au RAT<sub>1A</sub>. Les  $\beta$ -arrestines qui, complexées au récepteur, bloquent les interactions avec les protéines G, activeraient en parallèle la dégradation des seconds messagers issus de la voie G.

**Signalisation  $\beta$ -arrestines** Outre le double rôle supposé des  $\beta$ -arrestines dans la désensibilisation, et comme pour le RFSH, le RAT<sub>1A</sub> fait appel aux  $\beta$ -arrestines pour une signalisation indépendante des protéines G. Le recrutement des  $\beta$ -arrestines au RAT<sub>1A</sub> est facilité par la phosphorylation du récepteur par les GRKs. Ce complexe induit ensuite une multitude de mécanismes de signalisation [Lefkowitz and Shenoy, 2005]. Comme le RFSH,

le  $RAT_{1A}$  se complexe de manière stable aux  $\beta$ -arrestines (récepteur de classe B). Une étude menée sur un récepteur dénué des ses sites de phosphorylation ( $RAT_{1A} \Delta 324$ ) montre que le récepteur devient de classe A [DeWire et al., 2007]. En revanche, le rôle des isoformes de la  $\beta$ -arrestine dans la signalisation ERK est différent au sein de la signalisation angiotensine. Alors que les  $\beta$ -arrestines 1 prennent part à la signalisation ERK au sein du réseau FSH, les  $\beta$ -arrestines 1 sont cantonnées à leur rôle de désensibilisation pour l'angiotensine (Tableau 1.1). Les  $\beta$ -arrestines 1 ont ainsi un rôle antagoniste aux  $\beta$ -arrestines 2 [Ahn et al., 2004b]. Les  $\beta$ -arrestines 2, elles, comme au sein du réseau FSH, se complexent au récepteur et aux éléments de la cascade MAPK. De cette voie, il résulte des ERK phosphorylées dans le cytoplasme (*bpERK*). Ces *bpERK* ont une cinétique d'activation plus lente que les *GpERK* mais soutenue dans le temps (Figure 1.20) [Ahn et al., 2004a].

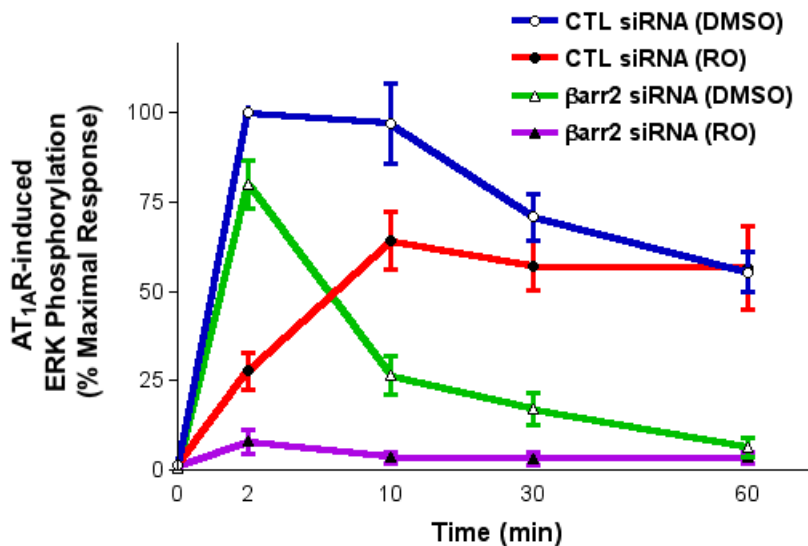


FIGURE 1.20 – Cinétiques ERK. En bleu la cinétique ERK phosphorylés en condition contrôle, en rouge lorsque la voie G est inhibée, en vert, lorsque les  $\beta$ -arrestines 2 sont inhibées et en violet lorsque la voie G et la voie  $\beta$ -arrestines sont inhibées. On remarque que la somme des signaux indépendants des protéines G et indépendants des  $\beta$ -arrestines donne la résultante contrôle. On dira donc souvent par approximation que la composante verte illustre les *GpERK* tandis que la rouge illustre les *bpERK*.

Enfin, cela a été énoncé au sein de la section générale (section 1.1), la signalisation dépendante des  $\beta$ -arrestines a récemment fait l'objet de deux études à haut débit basées sur le récepteur angiotensine [Xiao et al., 2007; Xiao et al., 2010]. Au sein du tableau 1.2, certains acteurs de l'interactome des  $\beta$ -arrestines [Xiao et al., 2007] ont été recensés. En particulier les protéines qui participent déjà à la signalisation dépendante des protéines G (Ras, Rho, ERK), les protéines de régulation de cette première voie (PDE, DGK) mais aussi les espèces impliquées dans d'autres fonctions tel le cycle cellulaire (CDC, CDK, Wee1) ou la traduction (eIF 2).

### 1.3. LE RÉCEPTEUR DE TYPE 1A DE L'ANGIOTENSINE

Protéines d'échafaudage :	$\beta$ -arrestin 1 ; $\beta$ arrestin 2 ; 14-3-3 $\beta/\alpha$ ; $\gamma$ ;
Kinases :	DGK ; PI3K ; MEKK1 ; Serine/threonine kinase 38 like p160ROCK, c-Yes ; ERK-2 ; Casein kinases ; CaM kinase II ; CDC2L5 ; CDK3 ; CDK4 ; CDK7 ; eIF2 A kinase 4 ; Wee1A kinase ; TAK1 ;
Phosphatases :	PP2C ; Myosin phosphatase-targeting subunit 1
Liées au trafficking	Clathrin ; dynamin
Protéines G et les petites protéines G :	Ras ; Rho
Autres protéines de signalisation :	Calmodulin ; I $\kappa$ -B kinase $\alpha$ ; TAK1-binding protein 1 ; PDE4D protein ;

TABLE 1.2 – Exemples d'interactants avec les  $\beta$ -arrestines au sein de la transduction du signal du ligand SII via le RAT<sub>1A</sub>. En violet les interactants spécifiques de  $\beta$ -arrestine 1 et en vert les interactants spécifiques de  $\beta$ -arrestine 2 [Xiao et al., 2007].

#### 1.3.2 Conclusion

Chaque R7TM est initiateur d'un large réseau de signalisation. Comme pour nos deux récepteurs d'intérêt, RFSH et RAT<sub>1A</sub>, l'étude de ces réseaux est devenu un travail de titan et l'intégration de l'ensemble des connaissances une tâche impossible à l'aide des seuls outils expérimentaux. La biologie systémique est une tentative pour apporter des solutions à cette situation. Dans cet esprit et avec pour but d'intégrer les données présentées dans cette introduction, nous avons construit le réseau de signalisation en aval du RFSH (Figure 1.15) et du RAT<sub>1A</sub> (Figure 1.21) dans le formalisme CellDesigner [Kitano et al., 2005; Novère et al., 2009].

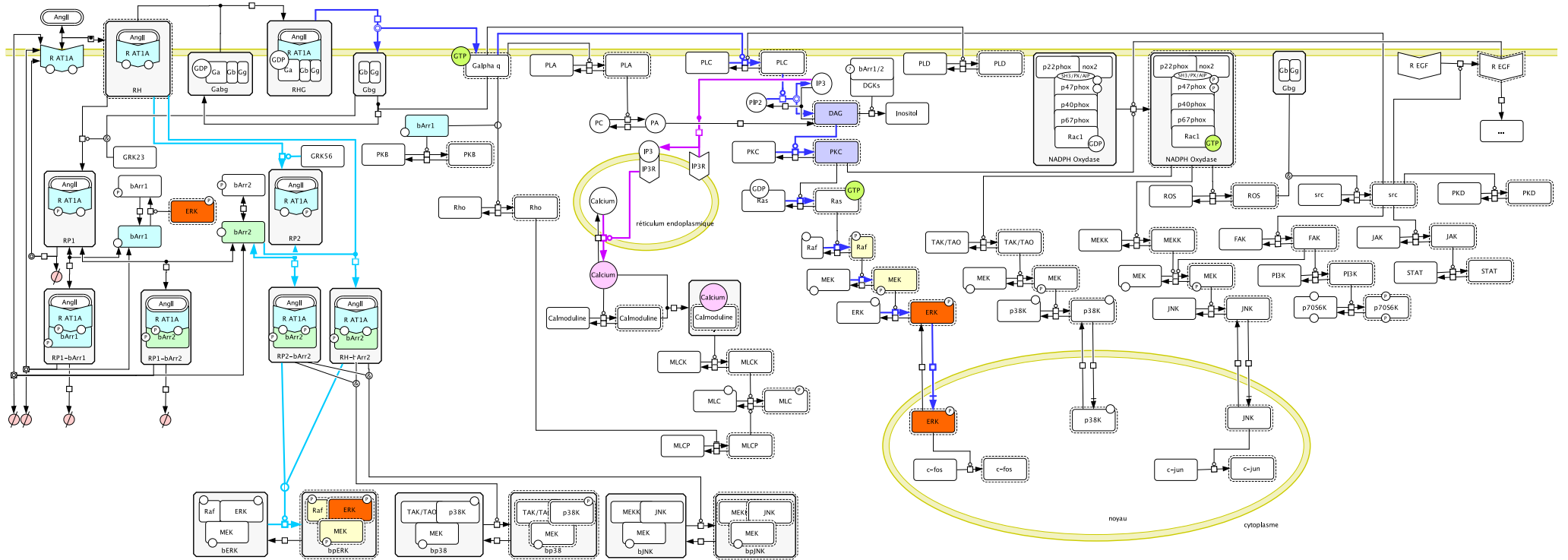


FIGURE 1.21 – Tentative d'interprétation des connaissances disponibles sur le réseau de signalisation de l'angiotensine II comprenant les différentes voies dépendantes des protéines G et de  $\beta$ -arrestines. En couleur les espèces et les voies de signalisation intégrées au sein du modèle dynamique du réseau angiotensine (Chapitre 3)

## 1.4 Vers une biologie systémique des R7TMs

Les R7TMs contrôlent toutes les grandes fonctions physiologiques et sont la cible d'une grande partie des médicaments. Notre perception de la signalisation des R7TMs s'est considérablement complexifiée puisqu'il est maintenant admis qu'ils activent de larges réseaux de signalisation capables d'intégrer les flux d'information en des réponses biologiques appropriées, comme détaillé en Introduction pour la FSH et l'angiotensine II. Il devient essentiel de pouvoir observer un maximum d'éléments du réseau de signalisation et ce, sous plusieurs conditions expérimentales. Outre la complexité numérique des observations, la signalisation évolue en fonction du temps et diffère en fonction du type cellulaire ainsi que du stade de développement de la cellule. Le besoin d'expériences biologiques s'est nettement accru, et des méthodes à haut débit se sont développées pour répondre à ces demandes.

En particulier, les méthodes sur puces qui permettent une miniaturisation et une multiplication des points expérimentaux. Les puces à protéines permettent une étude à très haut débit des différentes protéines du réseau. Certaines vont suivre l'évolution d'une protéine au cours du temps, dans différentes conditions tandis que d'autres, appelées puces à anticorps, vont détailler les différentes molécules d'un même extrait cellulaire. Deux avancées majeures de ces méthodes sont :

- la possibilité d'effectuer beaucoup de mesures et de les répéter aisément (jusqu' à 12 000 spots par puce) ce qui améliore la rapidité d'obtention des résultats ainsi que leur fiabilité
- la quantification des protéines, une information précieuse (voire chapitre 2 et l'article en Annexe).

Au vu de ces quantités de données en plusieurs dimensions (molécules, temps, perturbations, type cellulaire...), le stockage des expériences, leur traitement, leur partage et leur visualisation nous a semblé un réel défi à relever. Quant à la compréhension et la prédiction des réponses cellulaires, cela ne peut plus être intégré par une simple analyse. Les études *in vivo* des mécanismes de signalisation intracellulaire observent une partie des molécules. L'expression d'une voie de signalisation est souvent interprétée à tort comme la réponse lorsque les autres voies sont inhibées. Le besoin d'une vision globale du réseau est crucial. Pour cela, les approches *in silico* de modélisation permettent de regarder le système en intégrant plusieurs dimensions. Cependant, chaque méthode apporte avantages et inconvénients. Les méthodes basées sur le graphe de régulation (ensemble des influences positives et négatives entre les espèces) et les méthodes basées sur les simulations booléennes d'un modèle peuvent intégrer de très larges réseaux et en obtenir des propriétés temporelles mais n'offrent pas les simulations dynamiques souhaitées pour nos réseaux de signalisation. En revanche, les approches basées sur des Petri nets qui offrent une vision graphique et discrète autant que les modèles ODEs et les méthodes stochastiques donnent l'évolution dynamique des différentes espèces. Cependant, ces trois méthodes nécessitent la connaissance de paramètres pour simuler le modèle. Ces derniers étant en majorité inconnus, des démarches de réduction et d'optimisation doivent être couplées aux approches de modélisation dynamiques.

Cette rigueur et ces formalismes apportent en compréhension mais aussi en remise en cause des paradigmes biologiques impliquant un retour aux expériences pour plus de



précisions. En effet, les modèles intègrent et demandent de préciser les compétitions, les recyclages, les éléments intermédiaires non observés et les quantités de réactifs. Cette compréhension des mécanismes de signalisation contribue à accélérer l'émergence de nouvelles classes de médicaments telle une classe d'agonistes (ou d'antagonistes) sélectifs de voies de signalisation présentant moins d'effets indésirables.

### **1.4.1 Vers une biologie systémique de la signalisation des récepteurs couplés aux protéines G : Défis et attentes**

Ce travail a fait l'objet d'une publication dans les *Comptes Rendus Biologies*. Dans cette revue, nous avons tout d'abord listé et comparé les méthodes expérimentales actuelles, naissantes et futures. Puis, nous présentons les différentes possibilités techniques et conceptuelles permettant la mise en place d'une approche systémique de la signalisation des R7TMs. Enfin, les principaux défis auxquels nous restons confrontés sont également discutés.



Review / Revue

## Towards a systems biology approach of G protein-coupled receptor signalling: Challenges and expectations

Domitille Heitzler<sup>a,b,c,d</sup>, Pascale Crépieux<sup>a,b,c,d</sup>, Anne Poupon<sup>a,b,c,d</sup>, Frédérique Clément<sup>e</sup>, François Fages<sup>f</sup>, Eric Reiter<sup>a,b,c,d,\*</sup>

<sup>a</sup> BIOS group, INRA, UMR85, Unité physiologie de la reproduction et des comportements, 37380 Nouzilly, France

<sup>b</sup> CNRS, UMR6175, 37380 Nouzilly, France

<sup>c</sup> Université François-Rabelais, 37041 Tours, France

<sup>d</sup> Haras Nationaux, 37380 Nouzilly, France

<sup>e</sup> Sisyphe team, INRIA Paris-Rocquencourt, 78153 Le Chesnay, France

<sup>f</sup> Contraintes team, INRIA Paris-Rocquencourt, 78153 Le Chesnay, France

Available online 14 October 2009

Presented by Michel Thellier

### Abstract

G protein-coupled receptors (GPCRs) control all the main physiological functions and are targeted by more than 50% of therapeutics. Our perception of GPCRs signalling has grown increasingly complex since it is now accepted that they activate large signalling networks which are integrating the information fluxes into appropriate biological responses. These concepts lead the way to the development of pathway-selective agonists (or antagonists) with fewer side effects. Systems biology approaches focused on GPCR-mediated signalling would help dealing with the huge complexity of these mechanisms therefore speeding-up the discovery of new drug classes. In this review, we present the various technical and conceptual possibilities allowing a systems approach of GPCR-mediated signalling. The main remaining limitations are also discussed. **To cite this article: D. Heitzler et al., C. R. Biologies 332 (2009).**

© 2009 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

### Résumé

**Vers une biologie systémique de la signalisation des récepteurs couplés aux protéines G : Défis et attentes.** Les récepteurs couplés aux protéines G (RCPGs) contrôlent toutes les grandes fonctions physiologiques et sont la cible de plus de 50% des médicaments. Notre perception de la signalisation des RCPGs s'est considérablement complexifiée puisqu'il est maintenant admis qu'ils activent de larges réseaux de signalisation capables d'intégrer les flux d'information en des réponses biologiques appropriées. Ces concepts ouvrent la voie au développement d'agonistes (ou d'antagonistes) sélectifs de voies de signalisation qui présenteraient moins d'effets indésirables. Une démarche de biologie systémique appliquée à la signalisation des RCPGs aiderait à appréhender la complexité des mécanismes de signalisation et accélérerait ainsi l'émergence de nouvelles classes de médicaments. Dans cette revue, nous présentons les différentes possibilités techniques et conceptuelles permettant la mise en place d'une approche systémique de la signalisation des RCPGs. Les principaux défis auxquels nous restons confrontés sont également discutés. **Pour citer cet article : D. Heitzler et al., C. R. Biologies 332 (2009).**

\* Corresponding author. Address for correspondence: INRA, UMR85, Unité physiologie de la reproduction et des comportements, 37380 Nouzilly, France.

E-mail address: [eric.reiter@tours.inra.fr](mailto:eric.reiter@tours.inra.fr) (E. Reiter).

© 2009 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

**Keywords:** GPCR; Signalling; Pharmacology; Systems biology; Bioinformatics; Modelling

**Mots-clés :** RCPG ; Signalisation ; Pharmacologie ; Biologie systémique ; Bioinformatique ; Modélisation

## 1. Introduction

An impressive amount of detailed information has been gathered over the past decades on how external stimuli activate plasma membrane receptors, how they translate to the activation of linear downstream signalling cascades and eventually affect cell fate. Recently, the advent of highly sensitive proteomic methods has produced maps of protein interactions and led to the reconstruction of biochemical networks [1]. As a consequence, it is now widely accepted that signalling pathways are organized as coordinated communication networks in which multi-protein complexes process and integrate the signal fluxes. Now, the challenge in cell signalling is to understand the behaviour of these intertwined communication networks in order to decipher the cellular language [2]. G protein-coupled receptors (GPCR) represent the largest class of membrane receptors. They are capable of binding a wide diversity of molecules that regulate most physiological processes and are involved in a plethora of diseases. Noteworthy, GPCRs have long been preferential targets of therapeutic research and development and they currently account for up to 50% of marketed drugs [3].

## 2. The growing complexity of GPCR-induced signalling

Classically, upon ligand binding, GPCRs undergo a conformational change that leads to heterotrimeric G protein recruitment and activation, followed by the generation of diffusible second messengers such as cAMP (cyclic Adenosine Mono-Phosphate), calcium or phosphoinositides. However, it is increasingly recognized that GPCRs trigger multiple signalling pathways which lead to the formation of signalling networks [4] (Fig. 1). For instance, some GPCRs have the ability to couple to multiple G protein subtypes [5] and many GPCRs directly interact with non-G protein signalling effectors through specific protein–protein interaction domains [6]. But quite remarkably, outside of heterotrimeric G proteins, only two protein families are able to specifically interact with the majority of GPCRs in their activated conformation: G protein-coupled receptor kinases (GRKs) and  $\beta$ -arrestins [7]. Historically, GRKs and  $\beta$ -arrestins have been associated with the

desensitization and internalization/recycling of most GPCRs [8]. However, recently, GPCRs have also been demonstrated to elicit signals, independently of heterotrimeric G protein coupling, through interaction with  $\beta$ -arrestins 1 and 2. Indeed,  $\beta$ -arrestins have been shown to act as multifunctional scaffolds and activators for a growing number of signalling proteins including ERK, p38, JNK, I- $\kappa$ B, Akt and RhoA [7,9–13]. Moreover, a recent proteomic study has reported as many as 337 protein interactions involving  $\beta$ -arrestins [14], strongly suggesting that they play a central role in the ability of GPCRs to activate very complex signalling networks. In addition, GRKs have also been reported to elicit signalling responses on their own right through protein/protein interactions [13]. Indeed, GRKs interact with a variety of proteins involved in signalling and trafficking such as G $\alpha$ q, G $\beta$  $\gamma$ , PI3K $\gamma$ , clathrin, GIT (G protein-coupled receptor kinase-interacting protein) and caveolin [15]. Phosphorylation of Raf kinase inhibitor protein (RKIP) by PKC displaces it from Raf and increases its association with GRK2 [16]. In addition, the physical interaction between GRK2 and Akt leads to the inhibition of Akt activity [17]. Finally, GRK2 and MEK1 have been found in the same multimolecular complex and this interaction is correlated with an inhibition of MEK activity [18].

Adding to this complexity is the fact that GPCR-induced signals can be spatially and temporally encoded. Signalling networks actively modulate the transmitted signals: negative feedback allows pathways to adapt or desensitize to persistent stimuli whereas cross inhibition is used to avoid crosstalk between pathways [19,20]. In addition to transmit qualitative information (e.g. the presence or absence of a stimulus), signalling pathways must also convey quantitative information about the strength of the stimulus (i.e. ligand concentration). It has been recently shown that signalling pathways can take advantage of their non-linear nature to convert stimulus intensity into signal duration [21]. Modulation of signal duration increases the range of stimulus concentrations for which dose-dependent responses are possible as dose-dependent responses are still possible after apparent saturation of the receptors. Another well documented example of spatial and temporal encoding in GPCR-induced sig-

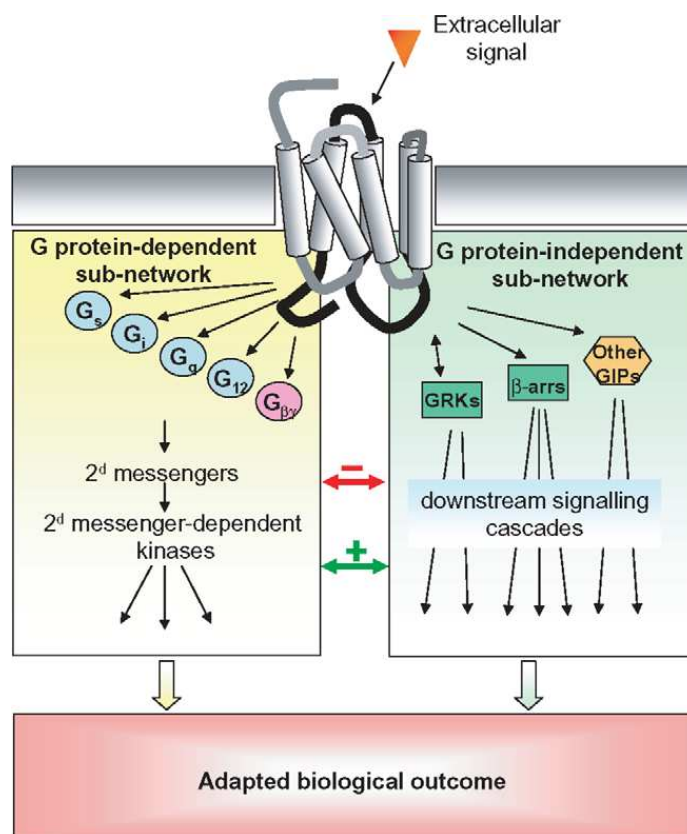


Fig. 1. GPCR-induced signalling networks. Schematic representation of the different transduction mechanisms leading to the activation of signalling pathways upon activation of a GPCR and how they define a complex signalling network. Cross-activations and cross-inhibitions between the G protein-dependent and -independent sub-networks are depicted as green and red arrows respectively.  $G_s$ ,  $G_i$ ,  $G_q$ ,  $G_{12}$ : different alpha subunits of heterotrimeric G protein;  $G\beta\gamma$ : G beta and gamma subunits of heterotrimeric G proteins; other GIPs: other GPCR interacting proteins;  $\beta$ -arrest:  $\beta$ -arrestin 1 and 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

nalling pathways is the dual activation mechanism of ERK by G protein and  $\beta$ -arrestins [22–26]. G protein-mediated ERK activation is rapid, transient and translocates to the nucleus. In contrast, the ERK activated *via*  $\beta$ -arrestins are slower in onset ( $\sim 5$ – $10$  min to reach maximum), very persistent ( $t_{1/2} > 1$  hour) and are sequestered in the cytosol. Such spatial and temporal differences in GPCR-induced signals substantially increase the complexity of signalling systems, hence their processing power. Collectively, these features highlight the importance of considering the dynamic properties of signalling pathways when characterizing their behaviour.

### 3. Pathway-selective ligands for GPCR: A new era in drug discovery?

Interestingly, this emerging conceptual framework opens new research avenues for the development of therapeutics [27]. There is increasing evidence that some GPCR targeting drugs can selectively modulate a subset of the signalling events triggered by the full

agonist. These effects have been given various names including “stimulus-trafficking”, “biased agonism”, “collateral efficacy” or “functional selectivity” [28]. Moreover, several GPCR ligands have already been reported to selectively activate or inhibit  $\beta$ -arrestin signalling [29–31]. Consideration of these new concepts might lead to the development of therapeutics with more selective actions, hence less side-effects. For instance, one drug, carvedilol, a  $\beta$  adrenergic receptor antagonist, has proven particularly effective in the treatment of heart failure. Interestingly, of 16 clinically relevant  $\beta$  adrenergic receptor antagonists, carvedilol displays a unique ability to stimulate  $\beta$ -arrestin-mediated signalling while preventing receptor coupling to  $G_s$  [31].

### 4. What added value can systems biology provide?

GPCRs’ complex signalling mechanisms probably lead to context-adapted cellular responses relying on emerging system-level properties that cannot be predicted from the individual components of the induced

networks. Therefore, it would be of paramount interest to provide a conceptual framework for deciphering and possibly predicting how an extra-cellular signal that activates a GPCR translates into a given biological or pathological response. This “global” level of analysis of GPCRs’ biology is in its infancy. Pioneer work carried out in yeast has recently shown the value of systems biology for elucidating complex signalling mechanisms triggered by GPCRs [32,33]. Aspects of cell signalling and the mechanisms (i.e. feedback and feed-forward regulations) that regulate pathway activity triggered by GPCRs have been studied in yeast, and nicely illustrate how mathematical modelling can be used to understand the logic of various pathway architectures. In mammals, a system-level grasp of GPCR-mediated signalling networks would be a significant asset to rationalize and speed-up the discovery of new “pathway-selective” drugs. Indeed, the rate of new drug discovery using standard approaches, based essentially on heterotrimeric G protein-dependent activities, such as second messenger accumulation, has been slowing down despite increased investments by the pharmaceutical industry.

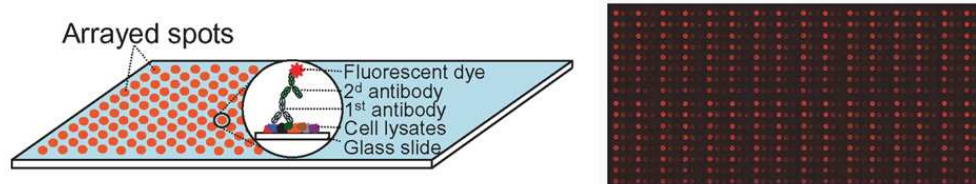
It has been proposed that computational modelling offers a powerful tool for examining GPCR pathways [34]. Such models can be used to better understand hypothesized mechanisms, run virtual (*in silico*) experiments, interpret data, suggest new drug targets, motivate experiments, and offer new explanations for observed phenomena. In the remaining part of the present review, we will identify and discuss the different challenges that the scientists in the field will be confronted with in their efforts to establish a systems biology approach to GPCR signalling.

Understanding the whole cell as an integrated system has been a goal for almost a century [35]. As in other fields of cell biology, deciphering GPCR signalling by at least descriptive, at most analytical methods has occupied the last 4 decades. Consequently, an enormous amount of data is published every year on some GPCR signalling aspects. However, these studies are highly heterogeneous in terms of the nature of the GPCR studied, the signalling pathway studied, the cellular system used, etc. As a consequence, it has been tremendously difficult to transform this huge amount of information into general concepts. An important effort of standardization and data sharing between laboratories in the field is needed. In addition, publicly accessible databases gathering and distributing standardized raw data related to GPCR signalling, especially dynamic data, would undoubtedly stimulate the modelling of GPCR-induced signalling networks.

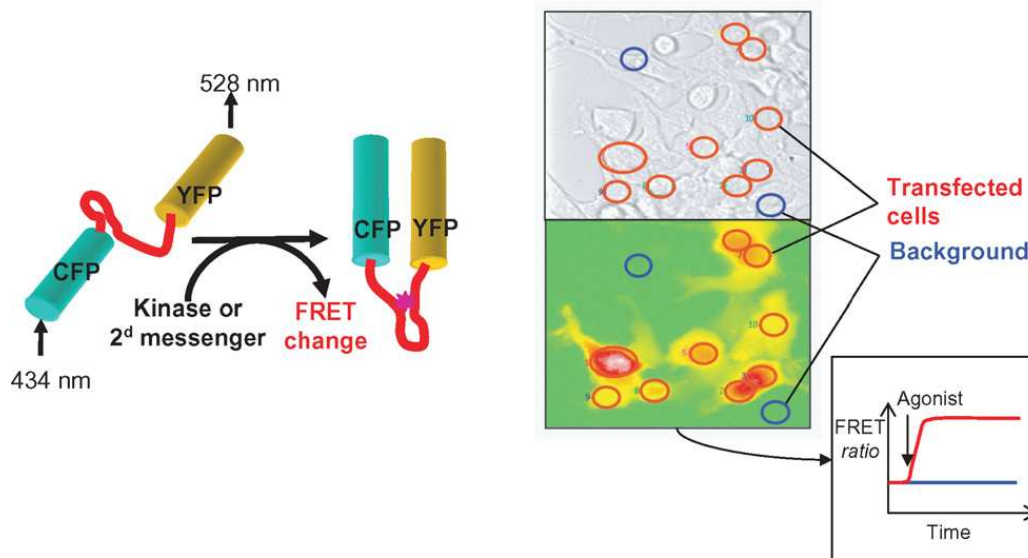
## 5. Challenges in high throughput generation of signalling data

More recently, feeding system-level analyses with relevant high quality biological data have become possible thanks to new experimental techniques that allow large scale accumulation of unbiased signalling data (Fig. 2). Speeding up the production of new data sets and enhancing their quality is not only essential to feed model-building but also to allow testing of key model findings. Signalling events are often propagated within the cell by post-translational modifications involving protein–protein interactions and enzymatic activities. Noteworthy, reversible protein phosphorylation is centrally involved in signal transmission within cells. The comprehensive and quantitative analysis of the protein phosphorylation patterns in different cellular backgrounds is therefore critical to reach a system level analysis of cell signalling. Lately, high-throughput has spread to molecular biology and biochemistry, giving access to most items of information necessary for the comprehension of organisms’ behaviour [36]. In particular, breakthroughs have been achieved in the isolation of phosphorylated peptides from complex samples, as well as in their analysis by mass spectrometry coupled to computational methods. Using such approaches, thousands of phosphopeptides and phosphorylation sites can now be identified in a single sample [37–39]. However, despite their unparalleled analytical power, mass spectrometry-based approaches present a static snapshot of cellular events; they do not allow the acquisition of dynamical phosphorylation data at large-scale and with high throughput. This limitation remains a major hurdle towards the development of powerful systems biology approaches in the field of cellular signalling. Over the last decade, the possibility to analyze the proteome using protein microarray-based methods has emerged into proteomics research, diagnostics, and drug discovery. In particular, automated spotting of concentrated and complex protein extracts permits their analysis with phosphospecific antibodies. This method, referred to as Reverse-Phase Protein Array (RPPA), uses very small quantities of biological material, which allows a wide sample collection with a high number of different antibodies to be screened [40–43]. Since it has a very high throughput, detailed kinetic experiments can be systematically carried out and analyzed. Moreover, the simultaneous quantification of thousands of samples achieved with RPPA drastically reduces data heterogeneity and variability which traditionally hamper modelling. Therefore, RPPA potentially represents a very attractive approach to cap-

### A. Reverse phase protein array



### B. FRET biosensors



### C. Transfected cell array

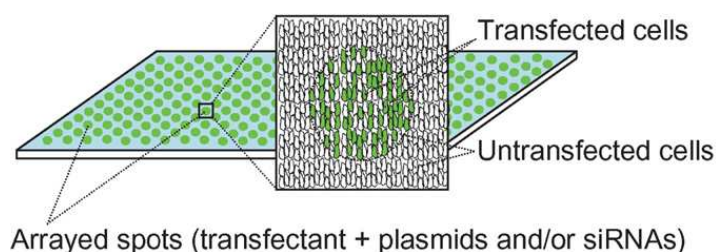


Fig. 2. Schematic representation of three high-throughput methods particularly adapted to the model-building process in the context of GPCR-induced signalling. A. Reverse phase protein array (RPPA) relies on the ordered micro-arranging of cellular lysates. Specific signals are detected using a primary antibody-fluorescently-labelled secondary antibody revelation system. This technology is easy to automate and when coupled with the use of a repertoire of phospho-specific antibodies raised against signalling intermediates, it allows the analysis of signalling networks at an unprecedented throughput. B. Fluorescence resonance energy transfer (FRET) biosensors are genetically-encoded fusion proteins linking two fluorescent proteins with a peptide sensitive to either phosphorylation by a specific kinase or second messenger binding. Once altered by a signalling event, the linker undergoes a conformational change that modifies the relative positions of the two fluorescent proteins hence their resonance. Cells transfected with the sensor plasmid are analyzed using a microscope. The emission wavelength of the first fluorescent protein (generally CFP) corresponds to the excitation wavelength of the second fluorescent protein (generally YFP). Therefore, FRET can be quantified in living cells with exquisite time-resolution by using an excitation wavelength specific of the first fluorescent protein while recording the emission at the wavelength specific of the second fluorescent protein. C. Transfected cell array (TCA) consists of parallelized micro-spots containing a transfection reagent, a plasmid and/or a siRNA. Each spot is different and a very large number of them can be arrayed on one slide allowing large-scale screening. Cells are seeded directly on the slide so that they attach and grow on it while being specifically transfected by the genetic material underneath. The effect of each transfection can be analyzed for each spot individually based on morphology *criteria*, immunocytochemistry or reporter-gene activity.

ture the subtle and highly dynamic nature of phosphorylation cascades out of very large sample collections.

Recently developed imaging approaches that use fluorescent sensors of signalling activities combine unmatched time and spatial resolution. Genetically en-

coded fluorescence resonance energy transfer (FRET)-based reporters have been used in living cells to monitor the spatiotemporal patterns of diffusible second messengers, kinase activities and GPCR activation [44–46]. Protein–protein interactions can also be analyzed in real time in living cells using either FRET or BRET (bioluminescence resonance energy transfer) [46–48]. When used in multi-well plate format, both FRET and BRET-based methods ensure the production of huge amounts of high content dynamic data which are very well suited to feed systems biology approaches.

If the ability to measure signalling outputs in a dynamic way is a key-point, so is the aptitude to control the inputs (i.e. agonist stimulation) both in time and concentrations. Chemical signalling in/between cells is crucial as it determines the cellular response, and is characterized by various time-scales ranging from a few milliseconds to several minutes [44]. However, with conventional experiments, mimicking *in vivo* conditions remains a challenge. The implementation of microfluidic devices might help suppress the limitations found with conventional approaches and allows multiplexed analysis with various stimulation patterns to be performed. Indeed, microfluidics offers the possibility of not only tightly controlling and modulating cell culture conditions but also of applying well-defined (in intensity, space and time) chemical stimulation (receptor ligands, inhibitors, etc.) [49,50]. This should notably facilitate an iterative dialog between computational modelling and “wet lab” experiments (i.e. prediction *versus* validation).

An important aspect when trying to decipher and model signalling networks is the ability to specifically apply perturbations and to measure the interactions of signalling pathways considered in the recent past as isolated entities (Fig. 3). In addition to classical approaches (e.g. kinase inhibitors, dominant negative constructs, etc.), interfering RNAs offer the unique opportunity to easily and specifically achieve gene knock-downs. Moreover, genome-wide siRNA screening are now available either in multi-well liquid phase or in transfected cell array format [51–53]. Large-scale siRNA screening may very well be an experimental breakthrough facilitating the edification of highly complex signalling networks.

## 6. Challenges in bioinformatics

Knowledge and data management is instrumental to systems biology. Indeed, high-throughput approaches generate huge quantities of heterogeneous data that cannot be handled with classical labbooks, or even flat files.

Therefore, computational methods, standards and tools must be developed and used to tackle this problem. Beyond the handling problem, the design, validation and refinement of high fidelity models require the use of all available experimental and non-experimental data. To reach this objective, data produced by experimental biology and bioinformatics have to be accessible, in compatible formats, and easy to correlate. Experimental data management systems, called Laboratory Information Management Systems (LIMS) should allow biologists to capture all the data relative to an experiment (conditions, protocol, results). So far, the availability of a LIMS to manage all the pieces of information required has been the major stumbling-block of the design of models in systems biology.

Once the data and knowledge are available and organized, the first modelling step consists in the construction of a detailed representation of the studied system. For intracellular signalling networks, the reaction/interaction graph is commonly used. Its construction is achieved manually, in a hypothesis-driven manner by the expert, using prior knowledge of the system and chosen experimental data. Sophisticated data analysis and visualization tools, such as clustering, help in this task. However, when the system is very large, choices have to be made by the expert, and the resulting model is biased, incomplete and often fails to reveal emerging features from the studied network. The development of methods to automatically infer influence graphs directly from the ensemble of data would allow the production of unbiased, thus potentially very innovative models from large data sets. It is certainly a promising research area but to date, only a few and limited attempts have been reported [54].

## 7. Challenges in mathematical modelling

Currently, relatively detailed mechanistic models of GPCR-induced signalling network can be formalized using the systems biology Markup Language (SBML) for representing the elementary interactions [55]. In this initial formalization step, signalling networks can be either directly written in SBML or drawn using CellDesigner which provides an intuitive SBML-based graphical modelling environment [56,57]. Of note is the Biochemical Abstract Machine BIOCHAM [58] which can be used to compute the influence graph between molecular species from the reaction/interaction graph [59], formalize the (observed) biological properties of the system by temporal logic formulae, and automatically verify their satisfiability by model-checking algorithms. From the influence graph, temporal logics and model-

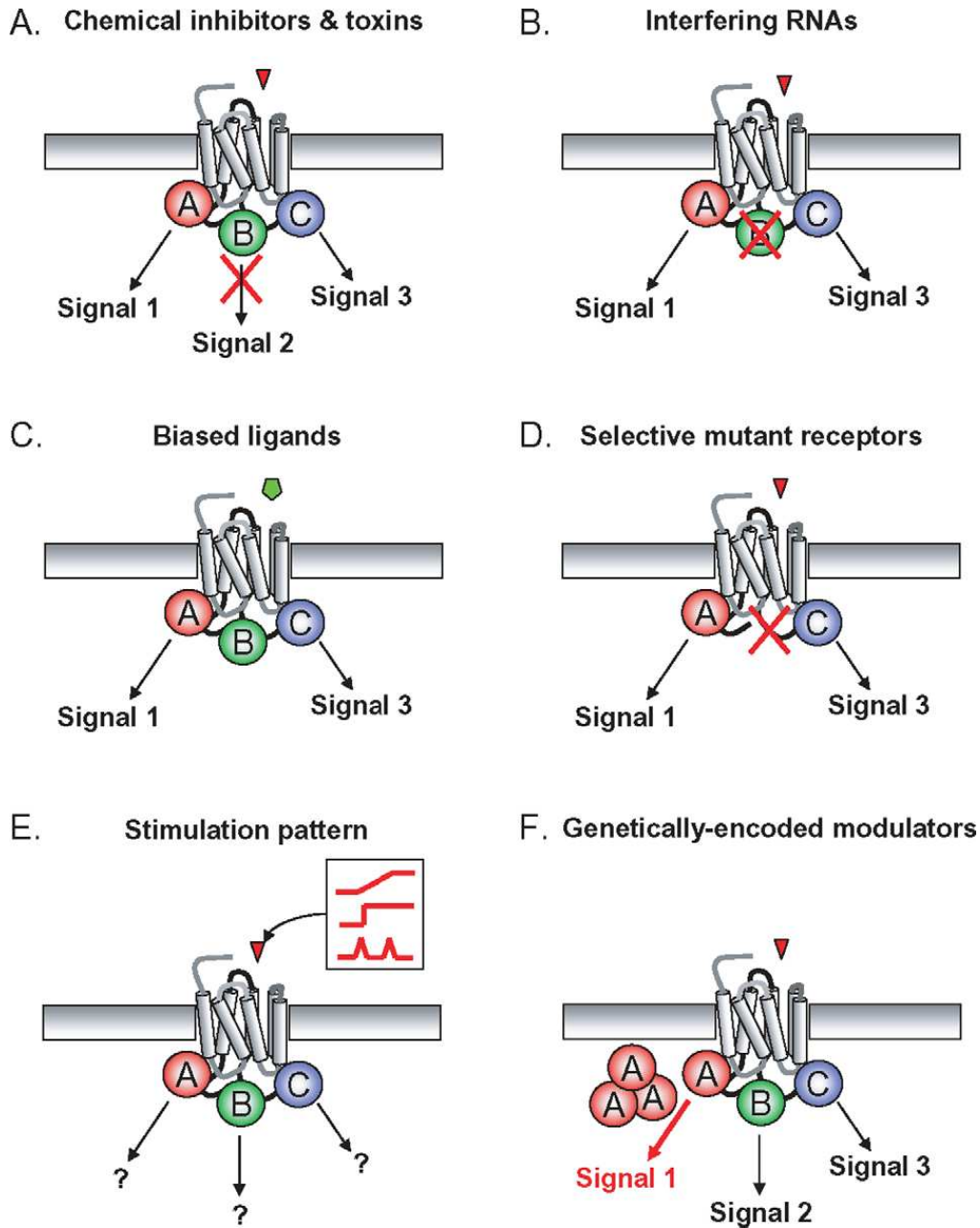


Fig. 3. GPCR-induced signalling systems can be specifically perturbed using different strategies: A. chemical inhibitors and toxins can be used to specifically block some of the proximal transduction mechanisms as well as some of the nodes of the downstream network; B. interfering RNA provide the unique possibility to knock-down any single gene putatively involved in the signalling network; C. when available, the use of signalling pathway-selective “biased ligands” provides a very interesting way to imbalance the proximal transduction mechanism allowing to measure the consequences on downstream network; D. alternatively, signalling pathway-selective mutant GPCR can be engineered; E. an interesting and potentially powerful way to favour the much needed iterative interactions between model-based predictions and experimental validations consists in applying tightly controlled agonist stimulation patterns to cells. Of course, this approach implies the use of adapted microfluidic devices; F. finally, genetically-encoded modulators can also be serviceable for instance to modify reaction stoichiometry by overexpressing a protein. Dominant negative constructs can also block a specific reaction or interaction.

checking algorithms have proven useful to express biological properties of complex biochemical systems and automatically verify if they are appropriate. This approach has allowed the analysis of reachability and temporal logic properties of signalling systems under various conditions [60].

Considering the highly dynamic nature both in time and space of GPCR-induced signalling pathways, a dynamic modelling approach is interesting and should yield more predictive power than static approaches. The aim of a dynamic model is to reproduce time-course experiments and provide intrinsic concentrations



for molecular species that are experimentally unreachable. There are several dynamical modelling approaches such as ordinary differential equations (ODEs, population view), Petri nets (discrete and independent mechanisms), and pi-calculus (stochastic approach) [61,62]. For all those methods, kinetics parameters (activation rates, probability of transition) are needed to simulate the changes in molecule concentrations over time. Frequently, too many parameters are unknown and the initial large reaction/interaction graphs have to be reduced. Stability analysis of the model (study of steady states, bifurcation diagrams, etc.) can help to restrict the parameter space to biologically relevant areas and reveal important differences in reaction speeds therefore providing an accurate way to reduce the model. Despite its usefulness, stability analysis remains too limited to ensure sufficient model reduction on its own. In addition, the qualitative properties of the influence graph can help develop a reduced dynamic model amenable to numerical simulations and parameter optimization with respect to quantitative data [63]. This model reduction approach remains however mostly empirical and even if the main graphical properties of the network (i.e. positive and negative circuits in the influence graph, reachability properties in the reaction graph) are preserved, the simplification of the structure of the network may suppress some delays and may limit the repertoire of dynamic behaviours permitted by the original model.

Even though stability analysis can lead to the determination of some unknown kinetic parameters in function of the others, the number of remaining unknown parameters is often large. Thus, another central difficulty concerns the non-linear optimization techniques that are needed to infer the unknown kinetic parameter values from the experimental data obtained under various conditions. Several techniques of parameter learning by data/property fitting can be used, such as gradient-based methods, Monte Carlo methods, and the Covariance Matrix Adaptation Evolution Strategy CMAES [64]. The latter two approaches minimize the error by repeated random sampling. Once the unknown parameters have been optimized, simulations can be performed and provide all component quantities over time. Simulations can be performed with different stimulation patterns or with *in silico* perturbations (modifications of total protein amounts, suppression of pathways, modulation of kinetics parameters, etc.). Thanks to these perturbations, the robustness of the system can be appreciated. Indeed, robustness is classically defined as the error between the perturbed and the initial simulations. Interestingly, the formalization of the expected dynamic properties of the system in temporal logic with numerical constraints also

makes it possible to quantify the robustness of the system with respect to some important properties [65]. This opens the way for integrating robustness criteria in the process of building the model.

## 8. Necessity of an iterative dialog between experimentation and modelling

The value of a computational model also needs to be assessed by its ability to fit all the available experimental observations made in either control or perturbed conditions. Next, it can be used to make *in silico* predictions. Simulations present the decisive advantage of providing the temporal evolution of all the constitutive molecular species involved, including those that are experimentally unreachable. In addition, it is easy to systematically perturb the signalling system *in silico*, for instance by suppressing one molecule or one reaction; the predicted effects can then be verified experimentally. Interesting predictions often result from this process. Another way to evaluate the predictive power of a signalling network model is to modify the agonist input to the system. Typically, different agonist concentrations or patterns (i.e. positive or negative gradients, pulsatile mode, etc.) can be applied *in silico*.

After initial experimental data gathering, a model is designed, parameterized and simulated. Predictions are made based on this initial model, and validation experiments are undertaken. In this initial phase, it often happens that predictions are not experimentally validated. In this case, the model has to be modified accordingly, and submitted to experimental validation again. This iterative process between modelling and experimentation is referred to as model refinement. According to this workflow, any prediction leads to either hypothesis validation or model refinement (Fig. 4). Thanks to this virtuous circle, the decoding of the intimate functioning of signalling networks' should move forward. Understanding how signalling pathways encode and transmit quantitative information about the external environment not only deepens our understanding of these systems, but will also lead to major achievements both in agriculture and medicine: models will point to molecular targets to be optimized in normal animal physiology and will help restore the proper function of pathways that have become deregulated in disease.

## 9. Conclusion

The acquisition of reliable data has long been the primary challenge in the field of GPCR signalling. Now,

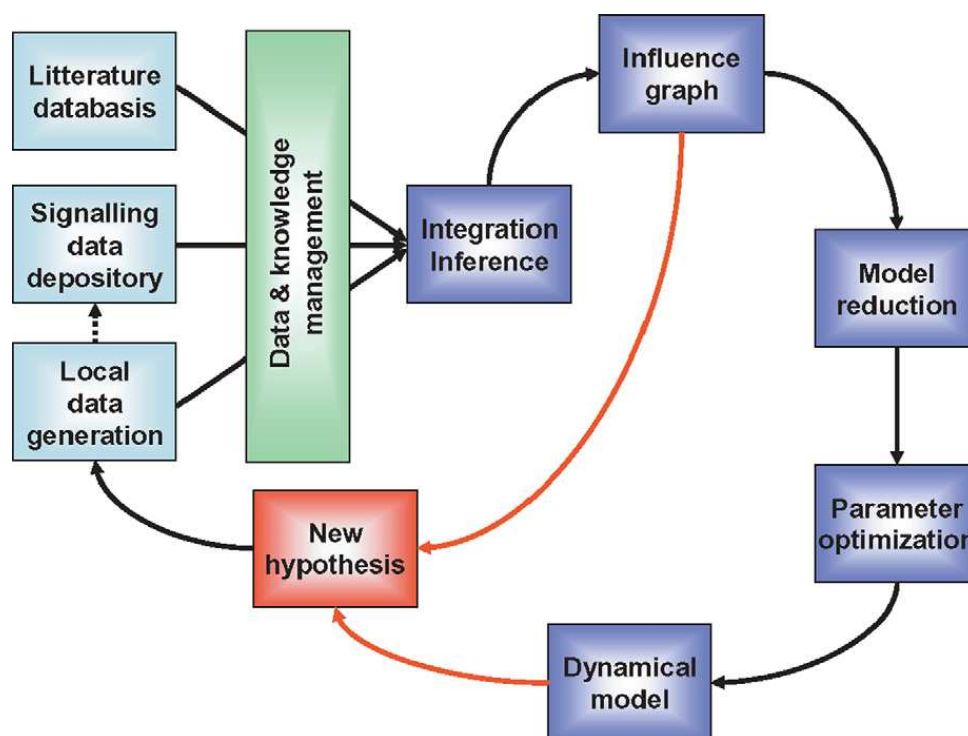


Fig. 4. Proposed workflow to model signalling systems. Data are produced in the laboratory, using low and high throughput methods. Already existing relevant and carefully standardized data can be extracted from publicly accessible databases (note that, once exploited, locally produced data have the vocation to be deposited in such databases). Literature mining is also essential. Data and knowledge from these three sources are combined, organized, analyzed and made available for modelling using various bioinformatics tools integrated into a comprehensive LIMS (laboratory information management system). Influence graph is constructed and its qualitative analysis can lead to new hypotheses. In parallel, this initial model is reduced and parameters are optimized. The dynamical model can be analyzed and its predictive power evaluated. An iterative process then is engaged between experimentation and modelling which leads to novel findings and model refinement.

with the advent of high throughput technologies potentially able to generate an unprecedented amount of dynamic signalling data, and to take into account the growing complexity of GPCR signalling, the challenges are being re-centred towards more theoretical and conceptual aspects. State of the art bioinformatics and mathematics can already help better handle the complexity associated with the large signalling networks and with the tremendous development of systems biology, rapid improvements can be expected.

### Acknowledgements

The authors thank members of the BIOS team for their advices and support. D.H. was funded by a fellowship from the INRA (ASC). This work was supported by the INRA AIP AgroBI and by the large scale project REGATE (REGulation of the GonAdoTropE axis).

### References

- [1] C. Preisinger, A. von Kriegsheim, D. Matallanas, W. Kolch, Proteomics and phosphoproteomics for the mapping of cellular signalling networks, *Proteomics* 8 (2008) 4402–4415.
- [2] D.B. Kell, Theodor Bucher Lecture: Metabolomics, modelling and machine learning in systems biology – Towards an understanding of the languages of cells, *FEBS J.* 273 (2006) 873–894. Delivered on 3 July 2005 at the 30th FEBS Congress and the 9th IUBMB Conference in Budapest.
- [3] P. Ma, R. Zimmel, Value of novelty? *Nat. Rev. Drug Discov.* 1 (2002) 571–572.
- [4] L.M. Luttrell, Transmembrane signaling by G protein-coupled receptors, *Methods Mol. Biol.* 332 (2006) 3–49.
- [5] R.J. Lefkowitz, K.L. Pierce, L.M. Luttrell, Dancing with different partners: Protein kinase a phosphorylation of seven membrane-spanning receptors regulates their G protein-coupling specificity, *Mol. Pharmacol.* 62 (2002) 971–974.
- [6] J. Bockaert, A. Dumuis, L. Fagni, P. Marin, GPCR-GIP networks: A first step in the discovery of new therapeutic drugs? *Curr. Opin. Drug Discov. Devel.* 7 (2004) 649–657.
- [7] R.J. Lefkowitz, S.K. Shenoy, Transduction of receptor signals by beta-arrestins, *Science* 308 (2005) 512–517.
- [8] R.J. Lefkowitz, G protein-coupled receptors. III. New roles for receptor kinases and beta-arrestins in receptor signaling and desensitization, *J. Biol. Chem.* 273 (1998) 18677–18680.
- [9] D.S. Witherow, T.R. Garrison, W.E. Miller, R.J. Lefkowitz, Beta-arrestin inhibits NF-kappaB activity by means of its interaction with the NF-kappaB inhibitor IkappaBalpha, *Proc. Natl. Acad. Sci. USA* 101 (2004) 8603–8607.
- [10] W.G. Barnes, E. Reiter, J.D. Violin, X.R. Ren, G. Milligan, R.J. Lefkowitz, {beta}-Arrestin 1 and G{alpha}q/11 coordinately ac-

- tivate RhoA and stress fiber formation following receptor stimulation, *J. Biol. Chem.* 280 (2005) 8041–8050.
- [11] J.M. Beaulieu, T.D. Sotnikova, S. Marion, R.J. Lefkowitz, R.R. Gainetdinov, M.G. Caron, An Akt/beta-arrestin 2/PP2A signaling complex mediates dopaminergic neurotransmission and behavior, *Cell* 122 (2005) 261–273.
- [12] D.L. Hunton, W.G. Barnes, J. Kim, X.R. Ren, J.D. Violin, E. Reiter, G. Milligan, D.D. Patel, R.J. Lefkowitz, Beta-arrestin 2-dependent angiotensin II type 1A receptor-mediated pathway of chemotaxis, *Mol. Pharmacol.* 67 (2005) 1229–1236.
- [13] E. Reiter, R.J. Lefkowitz, GRKs and beta-arrestins: Roles in receptor silencing, trafficking and signaling, *Trends Endocrinol. Metab.* 17 (2006) 159–165.
- [14] K. Xiao, D.B. McClatchy, A.K. Shukla, Y. Zhao, M. Chen, S.K. Shenoy, J.R. Yates 3rd, R.J. Lefkowitz, Functional specialization of beta-arrestin interactions revealed by proteomic analysis, *Proc. Natl. Acad. Sci. USA* 104 (2007) 12011–12016.
- [15] P. Penela, C. Ribas, F. Mayor Jr., Mechanisms of regulation of the expression and function of G protein-coupled receptor kinases, *Cell. Signal.* 15 (2003) 973–981.
- [16] K. Lorenz, M.J. Lohse, U. Quitterer, Protein kinase C switches the Raf kinase inhibitor from Raf-1 to GRK-2, *Nature* 426 (2003) 574–579.
- [17] S. Liu, R.T. Premont, C.D. Kontos, S. Zhu, D.C. Rockey, A crucial role for GRK2 in regulation of endothelial cell nitric oxide synthase function in portal hypertension, *Nat. Med.* 11 (2005) 952–958.
- [18] M.C. Jimenez-Sainz, C. Murga, A. Kavelaars, M. Jurado-Pueyo, B.F. Kraastak, C.J. Heijnen, F. Mayor Jr., A.M. Araga, G protein-coupled receptor kinase 2 negatively regulates chemokine signaling at a level downstream from G protein subunits, *Mol. Biol. Cell* 17 (2006) 25–31.
- [19] N.L. Komarova, X. Zou, Q. Nie, L. Bardwell, A theoretical framework for specificity in cell signaling, *Mol. Syst. Biol.* 1 (2005) 2005.0023.
- [20] M.N. McClean, A. Mody, J.R. Broach, S. Ramanathan, Cross-talk and decision making in MAP kinase pathways, *Nat. Genet.* 39 (2007) 409–414.
- [21] M. Behar, N. Hao, H.G. Dohlman, T.C. Elston, Dose-to-duration encoding and signaling beyond saturation in intracellular signaling networks, *PLoS Comput. Biol.* 4 (2008) e1000197.
- [22] X.R. Ren, E. Reiter, S. Ahn, J. Kim, W. Chen, R.J. Lefkowitz, Different G protein-coupled receptor kinases govern G protein and {beta}-arrestin-mediated signaling of V2 vasopressin receptor, *Proc. Natl. Acad. Sci. USA* 102 (2005) 1448–1453.
- [23] S.K. Shenoy, M.T. Drake, C.D. Nelson, D.A. Houtz, K. Xiao, S. Madabushi, E. Reiter, R.T. Premont, O. Lichtarge, R.J. Lefkowitz, Beta-arrestin-dependent, G protein-independent ERK1/2 activation by the beta2 adrenergic receptor, *J. Biol. Chem.* 281 (2006) 1261–1273.
- [24] S. Ahn, S.K. Shenoy, H. Wei, R.J. Lefkowitz, Differential kinetic and spatial patterns of beta-arrestin and G protein-mediated ERK activation by the angiotensin II receptor, *J. Biol. Chem.* 279 (2004) 35518–35525.
- [25] H. Wei, S. Ahn, S.K. Shenoy, S.S. Karnik, L. Hunyady, L.M. Luttrell, R.J. Lefkowitz, Independent beta-arrestin 2 and G protein-mediated pathways for angiotensin II activation of extracellular signal-regulated kinases 1 and 2, *Proc. Natl. Acad. Sci. USA* 100 (2003) 10782–10787.
- [26] E. Kara, P. Crepieux, C. Gauthier, N. Martinat, V. Piketty, F. Guillou, E. Reiter, A phosphorylation cluster of five serine and threonine residues in the C-terminus of the follicle-stimulating hormone receptor is important for desensitization but not for beta-arrestin-mediated ERK activation, *Mol. Endocrinol.* 20 (2006) 3014–3026.
- [27] J.D. Violin, R.J. Lefkowitz, Beta-arrestin-biased ligands at seven-transmembrane receptors, *Trends Pharmacol. Sci.* 28 (2007) 416–422.
- [28] T. Kenakin, Collateral efficacy in drug discovery: Taking advantage of the good (allosteric) nature of 7TM receptors, *Trends Pharmacol. Sci.* 28 (2007) 407–415.
- [29] D. Gesty-Palmer, M. Chen, E. Reiter, S. Ahn, C.D. Nelson, S. Wang, A.E. Eckhardt, C.L. Cowan, R.F. Spurney, L.M. Luttrell, R.J. Lefkowitz, Distinct beta-arrestin- and G protein-dependent pathways for parathyroid hormone receptor-stimulated ERK1/2 activation, *J. Biol. Chem.* 281 (2006) 10856–10864.
- [30] M.T. Drake, J.D. Violin, E.J. Whalen, J.W. Wisler, S.K. Shenoy, R.J. Lefkowitz, Beta-arrestin-biased agonism at the beta2-adrenergic receptor, *J. Biol. Chem.* 283 (2008) 5669–5676.
- [31] J.W. Wisler, S.M. DeWire, E.J. Whalen, J.D. Violin, M.T. Drake, S. Ahn, S.K. Shenoy, R.J. Lefkowitz, A unique mechanism of beta-blocker action: Carvedilol stimulates beta-arrestin signaling, *Proc. Natl. Acad. Sci. USA* 104 (2007) 16657–16662.
- [32] N. Hao, M. Behar, T.C. Elston, H.G. Dohlman, Systems biology analysis of G protein and MAP kinase signaling in yeast, *Oncogene* 26 (2007) 3254–3266.
- [33] N. Hao, M. Behar, S.C. Parnell, M.P. Torres, C.H. Borchers, T.C. Elston, H.G. Dohlman, A systems-biology analysis of feedback inhibition in the Sho1 osmotic-stress-response pathway, *Curr. Biol.* 17 (2007) 659–667.
- [34] J.J. Linderman, Modeling of G-protein-coupled receptor signaling pathways, *J. Biol. Chem.* 284 (2009) 5427–5431.
- [35] L. Von Bertalanffy, The theory of open systems in physics and biology, *Science* 111 (1950) 23–29.
- [36] U. Stelzl, U. Worm, M. Lalowski, C. Haenig, F.H. Brembeck, H. Goehler, M. Stroedicke, M. Zenkner, A. Schoenherr, S. Koepfen, J. Timm, S. Mintzlaff, C. Abraham, N. Bock, S. Kietzmann, A. Goedde, E. Toksoz, A. Droegge, S. Krobitsch, B. Korn, W. Birchmeier, H. Lehrach, E.E. Wanker, A human protein–protein interaction network: A resource for annotating the proteome, *Cell* 122 (2005) 957–968.
- [37] R. Linding, L.J. Jensen, G.J. Ostheimer, M.A. van Vugt, C. Jorgensen, I.M. Miron, F. Diella, K. Colwill, L. Taylor, K. Elder, P. Metalnikov, V. Nguyen, A. Pasculescu, J. Jin, J.G. Park, L.D. Samson, J.R. Woodgett, R.B. Russell, P. Bork, M.B. Yaffe, T. Pawson, Systematic discovery of in vivo phosphorylation networks, *Cell* 129 (2007) 1415–1426.
- [38] E.S. Witze, W.M. Old, K.A. Resing, N.G. Ahn, Mapping protein post-translational modifications with mass spectrometry, *Nat. Methods* 4 (2007) 798–806.
- [39] J.V. Olsen, B. Blagoev, F. Gnäd, B. Macek, C. Kumar, P. Mortensen, M. Mann, Global, in vivo, and site-specific phosphorylation dynamics in signaling networks, *Cell* 127 (2006) 635–648.
- [40] B. Spurrier, S. Ramalingam, S. Nishizuka, Reverse-phase protein lysate microarrays for cell signaling analysis, *Nat. Protoc.* 3 (2008) 1796–1808.
- [41] S.M. Chan, J. Ermann, L. Su, C.G. Fathman, P.J. Utz, Protein microarrays for multiplex analysis of signal transduction pathways, *Nat. Med.* 10 (2004) 1390–1396.
- [42] D.S. Gembitsky, K. Lawlor, A. Jacovina, M. Yaneva, P. Tempst, A prototype antibody microarray platform to monitor changes in protein tyrosine phosphorylation, *Mol. Cell. Proteomics* 3 (2004) 1102–1118.
- [43] K.M. Sheehan, V.S. Calvert, E.W. Kay, Y. Lu, D. Fishman, V. Espina, J. Aquino, R. Speer, R. Araujo, G.B. Mills, L.A. Liotta,

- E.F. Petricoin 3rd, J.D. Wulfkuhle, Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma, *Mol. Cell. Proteomics* 4 (2005) 346–355.
- [44] M.J. Lohse, V.O. Nikolaev, P. Hein, C. Hoffmann, J.P. Vilaradaga, M. Bunemann, Optical techniques to analyze real-time activation and signaling of G-protein-coupled receptors, *Trends Pharmacol. Sci.* 29 (2008) 159–165.
- [45] M.D. Allen, L.M. DiPilato, B. Ananthanarayanan, R.H. Newman, Q. Ni, J. Zhang, Dynamic visualization of signaling activities in living cells, *Sci. Signal.* 1 (2008) pt6.
- [46] B.N. Giepmans, S.R. Adams, M.H. Ellisman, R.Y. Tsien, The fluorescent toolbox for assessing protein location and function, *Science* 312 (2006) 217–224.
- [47] P.G. Charest, S. Terrillon, M. Bouvier, Monitoring agonist-promoted conformational changes of beta-arrestin in living cells by intramolecular BRET, *EMBO Rep.* 6 (2005) 334–340.
- [48] C. Gales, R.V. Rebois, M. Hogue, P. Trieu, A. Breit, T.E. Hebert, M. Bouvier, Real-time monitoring of receptor and G-protein interactions in living cells, *Nat. Methods* 2 (2005) 177–184.
- [49] F. Azizi, C.H. Mastrangelo, Generation of dynamic chemical signals with pulse code modulators, *Lab Chip* 8 (2008) 907–912.
- [50] P.J. Hung, P.J. Lee, P. Sabounchi, R. Lin, L.P. Lee, Continuous perfusion microfluidic cell culture array for high-throughput cell-based assays, *Biotechnol. Bioeng.* 89 (2005) 1–8.
- [51] H. Erfle, B. Neumann, P. Rogers, J. Bulkescher, J. Ellenberg, R. Pepperkok, Work flow for multiplexing siRNA assays by solid-phase reverse transfection in multiwell plates, *J. Biomol. Screen* 13 (2008) 575–580.
- [52] V. Starkuviene, R. Pepperkok, H. Erfle, Transfected cell microarrays: An efficient tool for high-throughput functional analysis, *Expert Rev. Proteomics* 4 (2007) 479–489.
- [53] D. Castel, M.A. Debily, A. Pitaval, X. Gidrol, Cell microarray for functional exploration of genomes, *Methods Mol. Biol.* 381 (2007) 375–384.
- [54] S. Nelander, W. Wang, B. Nilsson, Q.B. She, C. Pratilas, N. Rosen, P. Gennemark, C. Sander, Models from experiments: Combinatorial drug perturbations of cancer cells, *Mol. Syst. Biol.* 4 (2008) 216.
- [55] M. Hucka, A. Finney, H.M. Sauro, H. Bolouri, J.C. Doyle, H. Kitano, A.P. Arkin, B.J. Bornstein, D. Bray, A. Cornish-Bowden, A.A. Cuellar, S. Dronov, E.D. Gilles, M. Ginkel, V. Gor Goryanin II, W.J. Hedley, T.C. Hodgman, J.H. Hofmeyr, P.J. Hunter, N.S. Juty, J.L. Kasberger, A. Kremling, U. Kummer, N. Le Novere, L.M. Loew, D. Lucio, P. Mendes, E. Minch, E.D. Mjolsness, Y. Nakayama, M.R. Nelson, P.F. Nielsen, T. Sakurada, J.C. Schaff, B.E. Shapiro, T.S. Shimizu, H.D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, J. Wang, The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models, *Bioinformatics* 19 (2003) 524–531.
- [56] H. Kitano, A. Funahashi, Y. Matsuoka, K. Oda, Using process diagrams for the graphical representation of biological networks, *Nat. Biotechnol.* 23 (2005) 961–966.
- [57] A. Funahashi, Y. Matsuoka, A. Jouraku, M. Morohashi, N. Kikuchi, H. Kitano, CellDesigner 3.5: A versatile modeling tool for biochemical networks, in: *Proceedings of the IEEE*, 2008, pp. 1254–1265.
- [58] L. Calzone, F. Fages, S. Soliman, BIOCHAM: An environment for modeling biological systems and formalizing experimental knowledge, *Bioinformatics* 22 (2006) 1805–1807.
- [59] F. Fages, S. Soliman, From reaction models to influence graphs and back: A theorem, in: *Formal Methods in Systems Biology*, in: LNBI, vol. 5054, Springer-Verlag, 2008, pp. 90–102.
- [60] F. Fages, S. Soliman, Formal cell biology in BIOCHAM, in: *International School on Formal Methods for the Design of Computer, Communication and Software Systems: Computational Systems Biology*, Bertinoro, Italy, in: LNCS, vol. 5016, Springer-Verlag, 2008, pp. 54–80.
- [61] D. Gilbert, H. Fuss, X. Gu, R. Orton, S. Robinson, V. Vyshe-mirsky, M.J. Kurth, C.S. Downes, W. Dubitzky, Computational methodologies for modelling, analysis and simulation of signalling networks, *Brief Bioinform.* 7 (2006) 339–353.
- [62] R.J. Orton, O.E. Sturm, V. Vyshe-mirsky, M. Calder, D.R. Gilbert, W. Kolch, Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway, *Biochem. J.* 392 (2005) 249–261.
- [63] A. Rizk, G. Batt, F. Fages, S. Soliman, On a continuous degree of satisfaction of temporal logic formulae with applications to systems biology, in: *Computational Methods in Systems Biology CMSB'08*, in: LNBI, vol. 5307, Springer-Verlag, 2008, pp. 251–268.
- [64] N. Hansen, A. Ostermeier, Completely derandomized self-adaptation in evolution strategies, *Evol. Comput.* 9 (2001) 159–195.
- [65] A. Rizk, G. Batt, F. Fages, S. Soliman, A general computational method for robustness analysis with applications to synthetic gene networks, in: *17th Annual International Conference on Intelligent Systems for Molecular Biology, Bioinformatics 25 (2009)* 169–178.

### 1.4.2 Conclusion

La modélisation de la signalisation des R7TMs est un vaste sujet qui regroupe plusieurs domaines de recherche. Nous allons nous concentrer par la suite sur les modèles dynamiques. Plus précisément, nous avons choisi le formalisme ODE. Ces modèles mathématiques sont largement étudiés tandis que leur définition et leur simulation sont réalisables facilement car déjà implémentées dans différents logiciels et outils de programmation. Enfin, les simulations dynamiques d'un modèle ODE fournissent l'évolution en moyenne, quantitative et continue des différentes espèces du réseau de signalisation au cours du temps. Ceci peut être identifié aux concentrations des espèces du réseau et directement comparé aux données expérimentales comme nous le souhaitons pour nos systèmes de signalisation intracellulaire induit par l'activation des R7TMs.

## 1.5 Modèles dynamiques

Comme nous l'avons détaillé précédemment, la transduction du signal au sein d'un réseau de signalisation est un processus dynamique. Ces mécanismes de signalisation peuvent être considérés comme une suite de réactions biochimiques. La traduction de ces mécanismes biochimiques en équations différentielles ordinaires (ODE) permet de suivre la concentration de chaque espèce en fonction du temps. Ce formalisme ODE, de type « biochimique », fait correspondre chaque espèce à une variable, chaque réaction à une loi et chaque constante cinétique (constante d'affinité, vitesse d'interaction, vitesse de translocation) à un paramètre (Tableau 1.3). Les équations représentent les vitesses d'évolution des espèces à partir d'un point de départ (conditions initiales).

au sein du réseau biologique	au sein du modèle	notations
espèces	variables	$X(t) = (x_1(t), x_2(t), \dots, x_{NEQ}(t))$
réactions	lois	
constantes cinétiques	paramètres cinétiques	$K = (k_1, k_2, \dots, k_{NCIN})$
quantités initiales	conditions initiales	$CI = (c_1, c_2, \dots, c_{NEQ})$
quantités totales	quantité totales	$QT = (q_1, q_2, \dots, q_{NQT})$
paramètres	paramètres de l'ODE	$P = (K, CI, QT) = (p_1, p_2, \dots, p_{NPR})$
normes	paramètre de normes	$Norm = (norm_1, norm_2, \dots, norm_{NOR})$

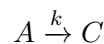
TABLE 1.3 – Correspondance entre les réseaux de signalisations et le formalisme ODE en notant  $NEQ$  le nombre de variables au sein du système,  $NCIN$  le nombre de paramètres cinétiques,  $NQT$  le nombre de quantités totales,  $NPR$  le nombre de paramètre de l'ODE et  $NOR$  le nombre de normes associées au modèle. L'utilisation des quantités totales ainsi que la définition des paramètres de normes sera explicité plus loin (sections 1.6 et 1.7).

Nous noterons le système ODE équivalent à l'ensemble des vitesses des variables  $f$  muni de ses paramètres  $P$

$$\frac{dX(t)}{dt} = f(t, X(t), P) \quad (1.1)$$

D'autres modélisations ODE existent où des paramètres dynamiques externes (ou entrée dynamiques) sont introduits au sein des équations. On parlera dans ce cas de modélisation « en module » car certaines quantités agrégées ou non décrites sont exprimées par des fonctions externes. C'est le cas de [Vidal et al., 2009] qui intègre à l'expression de la quantité de LH dans le sang, une fonction dépendante du temps représentant l'effet de la photopériode sur la sécrétion de LH. Ce type de modèle est aussi très fréquent pour décrire les phénomènes oscillants tel la décharge calcique [Dupont et al., 2007; Maltsev and Lakatta, 2010] et la pulsativité du GnRH [Csercsik et al., 2010]. Par la suite, nous nous concentrerons sur les modèles ODE de type « biochimique » (décrits par un système de type (1.1)). Ces modèles sont centrés sur les espèces du réseau de signalisation et détaillent l'ensemble des interactions sans dépendre de fonctions externes.

Par exemple, pour une réaction simple dont les constituants inter-agissent proportionnellement à leur quantité (loi d'action de masse) :



La « traduction » en équation différentielle ordinaire (ODE) se fait en décrivant les vitesses d'évolution des espèces  $A$  et  $C$  telles que :

$$\frac{da(t)}{dt} = -k * a(t) \quad (1.2)$$

$$\frac{dc(t)}{dt} = k * a(t) \quad (1.3)$$

$$a(0) = a_0 \quad (1.4)$$

$$c(0) = c_0 \quad (1.5)$$

avec

- $X(t) = (a(t), c(t))$  les variables
- $k$  le paramètre cinétique
- $CI = (a_0, c_0)$  les conditions initiales

Une fois cette traduction effectuée, nous chercherons, à partir des vitesses de formation/disparition de chaque espèce (1.2 et 1.3) et des quantités initiales (1.4 et 1.5), les quantités au cours du temps  $a(t)$  et  $c(t)$ . Cette démarche correspond à l'intégration numérique du système d'équations différentielles (voire section 2.2 pour une présentation détaillée). Le résultat, appelé simulation du modèle, offre une vision globale de l'ensemble des espèces évoluant dynamiquement. Il peut y avoir plus d'espèces décrites au sein du modèle qu'observées expérimentalement, le modèle prédit alors les évolutions de ces espèces non observées. On notera que les simulations d'un modèle sont fonction de la structure du modèle (ensemble des interactions), des lois régissant chaque interaction, des valeurs de paramètres cinétiques et des conditions initiales. Le travail de modélisation est d'établir une structure et les lois associées ainsi que de déterminer les paramètres cinétiques, le tout en confrontant le modèle aux données expérimentales. Le modèle et les simulations qui en découlent vont permettre la compréhension du fonctionnement dynamique d'un réseau

de signalisation. Enfin, à partir d'un modèle dynamique qualifié, de nouvelles expériences peuvent être réalisées *in silico* et ce, aisément puisqu'il suffit de simuler le modèle en modifiant les paramètres associés à la perturbation voulue. Cette possibilité est évidemment très intéressante pour disséquer le fonctionnement de réseaux complexes et aboutir à des hypothèses qui pourront être testés expérimentalement.

### 1.5.1 Modèles ODE - État de l'art

Un des premiers modèles dynamiques à base d'équations différentielles ordinaires (ODE) « biochimiques », modélisait l'activation du R7TM  $\beta$ 2-adrénergique et sa liaison à la protéine G [De Lean et al., 1980]. Comme tout modèle, il représentait une petite partie de la réalité et a été souvent remis en question et amélioré en fonction des avancées de la biologie. C'est ainsi que [Samama et al., 1993] décrit un « Extended Ternary Complex » en remarquant qu'une mutation du récepteur  $\beta$ 2-adrénergique conduit à l'activation agoniste-indépendante de l'adénylcyclase alors que le modèle de [De Lean et al., 1980], exige l'agoniste pour former un complexe actif, (Figure 1.22). Enfin, de multiples améliorations ont donné lieu à de nouvelles extensions comme [Bridge et al., 2010] qui tient compte de la possibilité de concurrence entre agoniste et antagoniste pour les récepteurs. Pour plus d'informations, [Woodroffe et al., 2010] récapitule l'ensemble des modèles issus du « Ternary complex » et propose une nouvelle extension. Pour ces modèles, la connaissance des quantités initiales et l'étude des états stables comparés aux données permettent de déterminer les paramètres cinétiques.

Les modèles de signalisation intracellulaire se développent tout d'abord pour l'activation du récepteur EGF (Epidermal growth factor receptor) qui appartient à la famille des Récepteurs à activité Tyrosine Kinase (RTK). Les connaissances liées à l'EGF sont plus vastes, ainsi [Kholodenko et al., 1999] dispose pour son modèle des quantités totales d'hormones, de récepteurs, de PLC (phosphoinositide-specific phospholipase C) phosphorylée, de Shc (Src homology and collagen domain protein) phosphorylée, de Grb2 (growth factor receptor-binding protein 2) complexées au récepteur et à la Shc. À ces quantités totales sont associées des lois de conservation (voir paragraphe « Réduction de modèle »), ce qui diminue la taille du modèle. Pour déterminer les paramètres cinétiques, [Kholodenko et al., 1999] s'aide des états stables et des expériences comme [De Lean et al., 1980] et [Samama et al., 1993], auxquels il rajoute des lois sur ces paramètres dues à la présence de cycles au sein du modèle. Des modèles plus larges apparaissent, toujours sur la signalisation EGF avec [Sasagawa et al., 2005; Birtwistle et al., 2007], mais contrairement aux premiers modèles, les données utilisées ne sont pas homogènes. Elles proviennent d'expériences réalisées au sein de cellules de différents types (PC12, hépatocyte, A431 ou HeLa). D'une part, on sait que la transduction du signal est intimement liée à la cellule cible : la variation des espèces et de leurs quantités au sein d'une cellule module le signal. En effet, l'angiotensine II ciblant des cellules musculaires lisses induit une contraction de celles-ci tandis que l'angiotensine II ciblant des fibroblastes induit une prolifération de ceux-ci. [Zou et al., 1998] montre que cette différence de réponse à l'angiotensine II trouve sa source dans une transduction du signal spécifique à chaque type cellulaire. D'autre part, lorsque les expériences sont réalisées sur des lignées de cellules immortelles embryonnaires (HEK293) ou tumorales (PC12), on peut émettre des réserves quant à leur pertinence pour

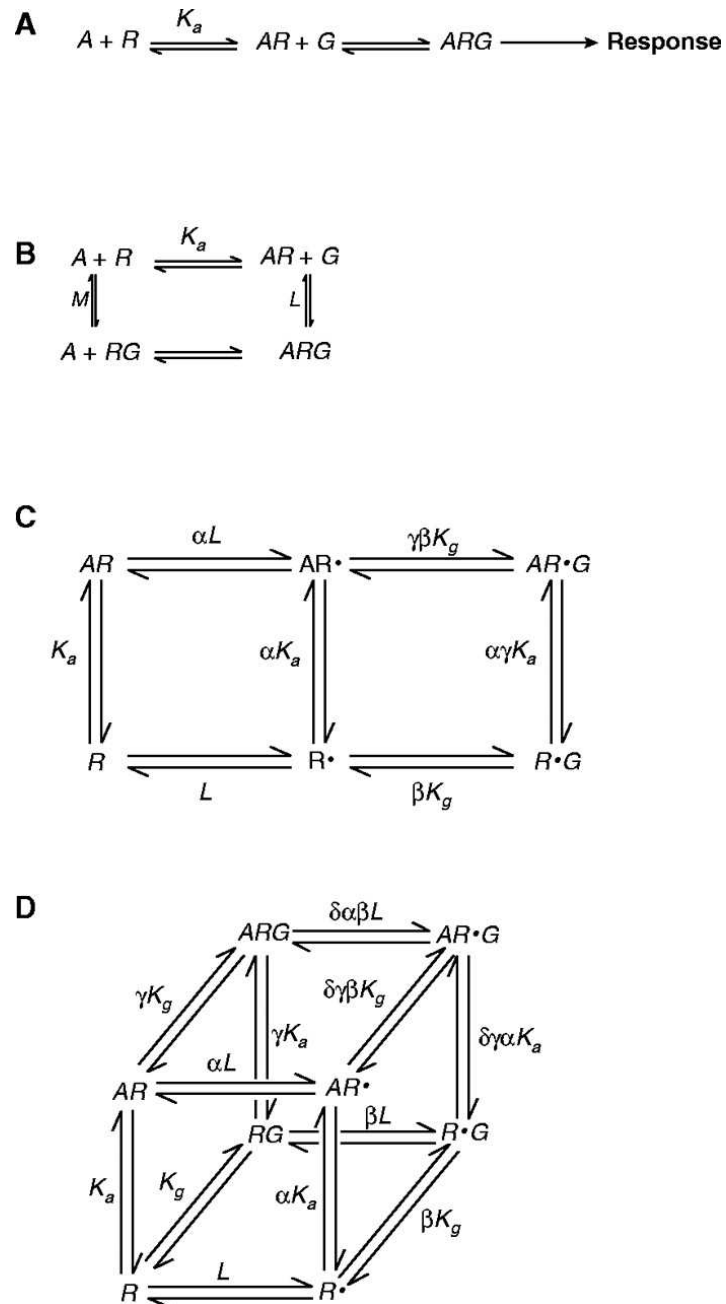


FIGURE 1.22 – Evolution des modèles décrivant le couplage des R7TMs aux protéines G [Limbird, 2004]. A. Les premiers modèles de signalisation par les R7TM sont tels que : un récepteur est activé par un agoniste. Une fois ce complexe actif formé, il y a recrutement de la protéine G qui active les voies de signalisation. B. Le modèle « Ternary Complex », [De Lean et al., 1980] décrit une relation entre l'affinité des récepteurs pour les agonistes et affinité du récepteur pour G. C. L'extension du « Ternary Complex », [Samama et al., 1993], intègre une signalisation spontanée du récepteur, sans complexation préalable à un agoniste. D. Le modèle « Cubic Ternary Complex », [Weiss et al., 1996], ajoute la possibilité de former des complexes agoniste, récepteur et protéine G (ARG) inactifs.



l'étude de la cellule normale.

De manière différente, [Bhalla and Iyengar, 1999] construisent des sous-réseaux indépendants, ils déterminent certains paramètres à partir de réactions entre espèces purifiées, ajustent les autres par rapport aux données, puis couplent les différentes parties pour reconstruire le modèle global. Leur premier modèle porte également sur la signalisation du récepteur EGF. Cependant, tout comme pour les expériences provenant de différents types cellulaires, un problème d'homogénéité se révèle : les réactions entre espèces purifiées ne sont pas identiques aux réactions entre espèces au sein du réseau. Cette hétérogénéité est due à l'absence des compétitions, des régulations et des synthèses au sein des réactions entre espèce purifiée mais aussi aux conditions de diffusion qui sont différentes *in vitro* et au sein d'une cellule.

Une partie du réseau, la cascade MAPK après activation du récepteur EGF, a été beaucoup modélisée. Une revue, [Orton et al., 2005], recense et compare les modèles réalisés jusqu'en 2005.

L'évolution vers des modèles de plus en plus large s'est heurtée à l'impossibilité de déterminer toutes les constantes. Ces constantes ou paramètres, nécessaires à l'intégration numérique de l'équation différentielle, sont les vitesses de formation des complexes, les taux de synthèse ou de dégradation, les quantités totales et initiales des espèces. Ces paramètres, dans la majorité des cas, ne sont pas accessibles expérimentalement.

Face à ce constat, les modélisateurs ont utilisé des méthodes de recherche de paramètres pour réaliser leurs modèles. Ainsi, l'équipe de Kholodenko dans [Nakakuki et al., 2010] utilise un algorithme génétique pour déterminer les paramètres manquants. [Nelander et al., 2008] associe à la recherche des paramètres, une recherche de la structure du modèle en privilégiant la simplicité du réseau. Sa fonction objectif (voir paragraphe « Ajustement Paramétrique ») contient la distance entre simulations et données ainsi qu'un poids du réseau en fonction des liaisons qu'il contient. Plus intéressant, la partie de la fonction objectif « distance entre les simulations et les données » contient une observation de sept variables en 23 conditions : non-stimulé, stimulé en condition contrôle et 21 combinaisons de perturbations avec stimulation. Ces combinaisons miment des couples de traitements médicamenteux et/ou d'altérations génétiques. Toutefois, malgré homogénéité des expériences (toutes réalisées sur des cellules de cancer du sein (MCF7)), un seul point de temps est pris en compte alors que le modèle est dynamique. De plus, les prédictions du modèle sur les conditions perturbées sont à relativiser car elles sont utilisées dans la procédure d'ajustement du modèle.

Sur un sujet totalement différent, la régulation du cycle cellulaire chez les levures de bière, les perturbations du système ont aussi été utilisées pour paramétrer un modèle de signalisation. C'est le travail remarquable de la thèse de Laurence Calzone [Chen et al., 2004] qui étudient les propriétés phénotypiques de 131 souches génétiquement modifiées. A l'issue de cette étude, ils proposent un modèle des mécanismes moléculaires de régulation du cycle cellulaire des levures de bière expliquant 120 des 131 souches mutées. Les modèles dynamiques de cette envergure sont très rares, ils nécessitent énormément de données qualitatives, quantitatives et/ou cinétiques homogènes.

Les modèles qui traitent de la signalisation cellulaire liée à l'activation des R7TMs sont peu nombreux car la mise en œuvre des expériences sur les R7TMs est plus délicate

(difficultés de transfection des R7TM, de marquage) que pour les RTKs.

On signalera le modèle de [Clément et al., 2001] qui a examiné le récepteur FSH couplé à la protéine G comme une espèce unique qui peut être activée par une hormone. Ce complexe active à son tour l'adénylcyclase pour produire l'AMPc. Ce modèle a été développé pour un horizon temporel long et focalisé sur l'évolution de la réponse cellulaire (croissance folliculaire) en fonction de l'historique d'exposition à FSH.

Pour un tout autre R7TM, [Bhalla and Iyengar, 1999] établissent des modèles portant sur le récepteur au glutamate au sein de neurones. Comme le premier modèle de l'article [Bhalla and Iyengar, 1999], portant sur la signalisation EGF, le deuxième est construit à partir de sous-réseaux indépendants. Ces sous-réseaux portent sur quatre kinases clés : la PKA, la PKC, la CaMKII et la cascade MAPK. Le comportement de ce réseau a été analysé pour déterminer les connections spécifiques entre les sous-réseaux permettant à l'activité CaMKII d'atteindre un état stable après l'arrêt des effets du calcium induit par le glutamate. Ceci est réalisé lorsqu'il y a une boucle de rétro-action positive de la PKC sur l'AC. À ce moment-là, la PKA, en aval de l'AC a un effet négatif sur un inhibiteur de la CaMKII, ce qui permet à la CaMKII de se stabiliser en un état non nul. Ce modèle a ensuite été ré-introduit dans divers réseaux de régulation. Un premier l'élargit par l'ajout du sous-réseau de régulation de l'IP3 [Mishra and Bhalla, 2002]. Un second qui cherche à déterminer les mécanismes moléculaires liés à l'apprentissage et intègre au réseau du glutamate les interactions avec les récepteurs EGF et NMDA (N-méthyl-D-aspartate), ainsi que la cascade MAPK/ERK suite à une stimulation électrique des neurones [Ajay and Bhalla, 2004]. Enfin, récemment, toujours autour des neurones et du récepteur au glutamate, le laboratoire de Bhalla a mis au point un large réseau intégrant la synthèse protéique dans les dendrites afin de prendre en compte les modifications synaptiques qui permettent la mémoire à long terme. Ce modèle relie sept sous-réseaux de régulation. En particulier, ce modèle montre que mTOR est un intermédiaire clé des récepteurs au glutamate et au facteur de croissance BDNF (brain-derived neurotrophic factor) pour la synthèse protéique [Jain and Bhalla, 2009]. Cette modélisation ODE en sous-réseaux juxtaposés nécessite énormément de données pour chaque sous-réseau. Comme spécifié dans les articles cités précédemment, une grande partie des paramètres a été estimée à partir de réactions entre espèces purifiées, puis des mesures sur certaines espèces ont permis d'établir chaque sous-modèle. De même que pour la signalisation EGF, on remarque qu'au sein d'un système cellulaire les réactions entre espèces purifiées ne sont pas toujours représentatives des réalités comprenant compétitions et recyclages. Enfin, la modélisation par modules reste très délicate, les comportements obtenus pour chaque sous-réseaux séparément peuvent être modifiés une fois qu'ils sont reliés entre eux au sein du réseau principal.

Contrairement au modèle précédent, un troisième R7TM a vu son réseau de signalisation modélisé en un système ODE sans module, c'est le récepteur  $\beta$ 2-adrénergique. Le modèle réalisé par [Violin et al., 2008], prend en compte le récepteur qui, une fois activé, induit la formation d'AC. L'AC active à son tour l'AMPc qui induit une activation de la PKA. Au sein du modèle, la PKA désensibilise le récepteur en le phosphorylant et soutient l'activation de PDEs qui dégradent l'AMPc. Le récepteur peut aussi être phosphorylé par les GRKs et se complexer aux  $\beta$ -arrestines pour être désensibilisé. Tout d'abord, au vu de leur temps d'observation, 20 minutes, ils considèrent constantes les quantités totales des espèces au sein de leur modèle et en déduisent cinq lois de conservation. Puis, se basant

sur les données FRET portant sur l'AMPC, [Violin et al., 2008] ajustent leur modèle à l'aide d'une méthode de Monte Carlo (voire paragraphe 1.7 et les Algorithmes 2 et 3 au chapitre 2).

Enfin, un modèle angiotensine existe et prend en compte la signalisation ERK dépendante des protéines G et des  $\beta$ -arrestines [Csercsik et al., 2008]. Ce modèle est construit par étapes et les paramètres sont recherchés par une méthode d'optimisation à chacune des étapes. Les espèces observées sur lesquelles se base l'ajustement paramétrique sont les ERKs en condition contrôle et sous les conditions perturbées d'inhibition des  $\beta$ -arrestines 2 et d'inhibition de la PKC. Cependant, les différents types de GRKs ne sont pas pris en compte, la sensibilité des paramètres est très importante et le modèle est arbitrairement normé à 1 ce qui ne permet pas de comparaisons avec des paramètres de la littérature et introduit notre travail qui prend en compte ces différents points ainsi que des validations sur des données publiées supplémentaires et des prédictions originales (section 3.2).

## 1.6 Réduction de modèle

Précédemment, nous avons vu que pour pallier un manque de connaissances sur les constantes du réseaux de signalisation, des méthodes de recherche de paramètres avaient été ajoutées à la réalisation des modèles. Une deuxième possibilité qui peut être complémentaire de la première est de réduire les modèles. De plus, certaines réductions permettent de ne pas détailler l'ensemble de la structure du réseau (ou ensemble des réactions). C'est un point intéressant quand de grandes incertitudes persistent à ce niveau. Parfois, les méthodes de réductions sont couplées au méthodes d'ajustement paramétrique. En effet, si le nombre de paramètres inconnus à rechercher est trop important, l'ajustement paramétrique est difficile à réaliser et parfois impossible. C'est pourquoi, une méthode de réduction pourra être appliquée avant l'ajustement des paramètres. On se rappellera que les paramètres et la structure du modèle sont nécessaires à l'intégration numérique de l'ODE.

Une première méthode de réduction, très utilisée, se base sur l'établissement de lois de conservation sur une espèce. Établir une loi de conservation revient à supposer constante une espèce du modèle. Le système est ainsi clos pour cette espèce, soit sa synthèse compense sa dégradation, soit ces deux réactions sont absentes du modèle. L'espèce peut être sous plusieurs formes : phosphorylée, au sein d'un complexe etc... Avoir une loi de conservation sur cette espèce est équivalent à que considérer la somme de toutes ses formes est constante. Lorsqu'une loi de conservation est établie, une équation du système ODE peut être enlevée contre l'ajout d'un paramètre représentant la quantité totale de l'espèce. Selon les conditions, il peut être plus avantageux d'enlever une équation et d'ajouter un paramètre que de garder le modèle initial. La forme de l'espèce dont l'équation de vitesse a été supprimée peut être retrouvée en soustrayant à la quantité totale de l'espèce, la somme des autres formes de l'espèce. Cette réduction est considérée exacte et réversible, elle ne change rien à la dynamique du modèle. Ainsi, au sein de l'article, [Musnier et al., 2009], nous avons noté  $RT$  la quantité totale de récepteurs supposée constante au cours du temps d'observation,  $R$  le récepteur FSH et  $RH$  le récepteur FSH complexé à la FSH.

La loi de conservation s'écrit :

$$RT = R + RH$$

Et, il suffit de prendre en compte la dynamique de  $RH$ , celle de  $R$  peut se déduire telle que :

$$R = RT - RH$$

Le système ODE (1.1) peut ainsi être ré-écrit après élimination d'une variable et ajout de la loi de conservations ainsi que du paramètres de quantité totale,  $QT$ , correspondant.

Une seconde réduction de modèle que nous avons utilisée pour notre modèle d'activation des ERK par l'angiotensine est la somme des variables qui découlent d'une suite de réactions linéaires. Ce type de réduction rend inaccessibles les états intermédiaires de la variable, donc est irréversible. L'exemple du modèle angiotensine est illustré ci-dessous (Figure 1.23).

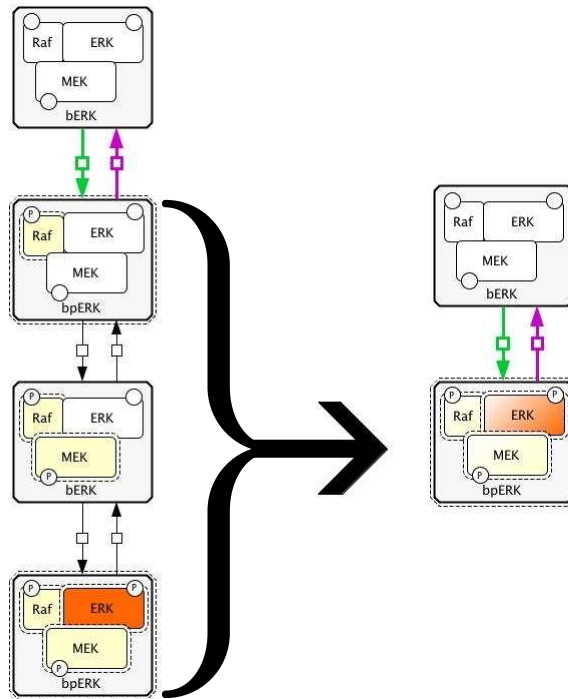


FIGURE 1.23 – Réduction qui consiste à faire la somme des variables qui découlent d'une suite de réactions linéaires.

Pour des modèles présentant des espèces ayant une multitude d'états possibles (phosphorylée sur différents sites, au sein d'un complexe etc.), une réduction en « macro » est proposée par [Borisov et al., 2008]. Il regroupe en une variable les différents complexes ou états auxquels participe une forme de l'espèce. Ce regroupement est possible si la forme de l'espèce est indépendante des autres formes. Par exemple, le récepteur  $AT_{1A}$  comprend 13 sites de phosphorylation, la combinaison  $\{P1\}$  et la combinaison  $\{P2\}$  apparaissent indépendamment l'une de l'autre et n'interagissent pas (voire modèle angiotensine section 3.2). Selon [Borisov et al., 2008],  $RP1$  et  $RP2$  sont deux « macro » qui contiennent l'ensemble des espèces lié à  $RP1$  et lié  $RP2$ . De même, au sein du modèle

FSH, les  $Pp70S6K\{T389\}$ ,  $Pp70S6K\{T421,S424\}$  et  $Pp70S6K\{T389,T421,S424\}$  sont des « macro » qui regroupe l'ensemble des  $Pp70S6K$  étant phosphorylée sur le site  $\{T389\}$ , sur les sites  $\{T421,S424\}$  et sur les sites  $\{T389,T421,S424\}$  (voire modèle FSH section 3.1). Cette réduction se focalise sur un état d'une espèce et fait la somme des différentes variables auxquelles cette espèce participe. Cette réduction n'est pas réversible et modifie le comportement dynamique du modèle.

De manière similaire, et proposant une généralisation des deux réductions précédentes, il existe une méthode graphique qui consiste à fusionner des nœuds (ou variables) pour former des variables agrégées. A la différence de la méthode précédente, ces variables agrégées ne se limitent pas aux différentes formes d'une même espèce. Cette méthode ne nécessite aucune connaissance préalable du réseau et est basée sur un formalisme de théorie des graphes avec nœud de fusion et suppression d'opérations. Bien souvent, il existe déjà dans la littérature plusieurs modèles réduits d'un même réseau (cascade MAPK). C'est pourquoi, [Gay et al., 2010] a élaboré un algorithme qui compare deux modèles : un détaillé et un réduit pour déterminer s'il existe une réduction par cette méthode graphique entre les deux et, le cas échéant,, cet algorithme donne les étapes de la réduction.

Il existe aussi des réductions liées à l'étude très précise des dynamiques du système ODE. Contrairement à la méthode précédente, ces méthodes requièrent la connaissance de toutes les constantes cinétiques afin de déterminer quels ensembles de molécules ont une dynamique lente par rapport aux autres. Le système ainsi découpé peut être réduit en considérant constante, à leur point d'équilibre, les variables aux dynamiques lentes [Glendinning, 1994; Kaper and Kaper, 2006]. Toujours au niveau des réductions de modèles paramétrés, [Gorban et al., 2010] listent les méthodes classiques actuelles d'états stationnaires et proposent, pour les modèles ODE linéaire, le calcul d'un modèle réduit asymptotiquement « dominant » basé sur l'étude des valeurs propres et de leurs vecteurs propres associés au modèle.

Pour les réductions visant à simplifier une chaîne de réaction en ligne, pour les réductions graphiques et les réductions liées aux dynamiques lentes, les résultats dynamiques sont affectés et les réductions sont irréversibles : on ne peut pas retrouver les états détaillés à partir du modèle réduit. Cependant, pour les modèles ODE, nous observons une simulation du modèle qui est le résultat d'une intégration numérique. Cette intégration numérique étant une approximation de la solution, cela produit une erreur au niveau des simulations (Voire chapitre 2, Figure 2.5). Lorsque l'erreur générée par l'intégration numérique est de même grandeur que l'erreur produite par la réduction, on ne distingue pas les simulations des modèles détaillé et réduit. Dans le même sens, les données sur lesquelles le modèle est ajusté comportent un bruit, cela est fréquent et est représenté par les écart-types. Alors, l'important est de vérifier que les erreurs dues à l'intégration numérique et aux réductions soient d'un ordre inférieur aux données expérimentales modélisées.

Enfin, et pour conclure, les modèles comportent en premier lieu les variables liées aux quantités observées expérimentalement.

**Définition 1.6.0.1.** Une *observable* est une fonction dépendante des variables du modèle qui est directement comparable à une quantité mesurées expérimentalement. L'ensemble des observables d'un modèle sera noté par la suite  $Obs(X)$ .

Par exemple, l'observable du modèle angiotensine à comparer aux données ERK phos-

phorylées est la fonction somme des variables  $GpERK$  et  $bpERK$ . Pour le DAG et la PKC du même modèle, les fonctions sont plus complexes et détaillées au sein de la section 2.1 puis utilisées dans l'article angiotensine section 3.2. Ces observables sont mises en parallèle de leurs valeurs expérimentales pour guider l'ajustement paramétrique du modèle dynamique. Les variables qui composent chaque observable sont importantes. Supprimer une de ces variables reviendrait à ne plus pouvoir calculer une observable, donc ne plus considérer les données expérimentales correspondantes, donc diminuer l'apport d'informations sur le modèle. Au delà de ces observables, l'erreur commise ne peut pas être quantifiée. C'est pourquoi les modèles réduits se concentrent souvent sur les variables liées aux observables. L'important à retenir au niveau des réductions non réversibles est que la précision ou le manque de précision qui en découlent doivent être adaptés au modèle biologique que l'on veut étudier.

## 1.7 Ajustement Paramétrique - État de l'art

Les ajustements paramétriques peuvent être utilisés indépendamment des réductions, ou après celles-ci, pour déterminer les paramètres inconnus. En effet, les systèmes dynamiques nécessitent la connaissance de l'ensemble des paramètres de l'ODE comprenant les constantes de réaction, les conditions initiales et des quantités totales ( $P = (K, CI, QT)$ , Tableau 1.3) pour réaliser l'intégration numérique du modèle, qui donne l'évolution des différentes variables au cours du temps. Ces évolutions, aussi appelées simulation du modèle, représentent les quantités prédites par le modèle. Le but du modélisateur est d'avoir une correspondance entre ces simulations et les données biologiques.

Le plus souvent, les données sont relatives, exprimées en pourcentage d'un maximum. Pour les rendre quantitatives, on leur joint une norme ( $Norm$ ) qui fait correspondre les pourcentages à une grandeur. Pour le modèle angiotensine, la grandeur du modèle est le  $\mu mol.L^{-1}$  et nous avons trois normes associées aux quantités de  $DAG$ ,  $PKC$  et  $pERK$  (Figure 1.24). Ces constantes ( $Norm = (norm_1, norm_2, \dots, norm_{NOR})$ ) peuvent aussi faire partie des quantités inconnues. Elles ne sont pas nécessaires à l'intégration numérique de l'ODE mais indispensables à la comparaison des simulations aux données.

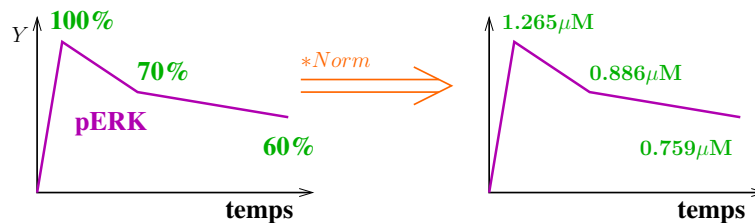


FIGURE 1.24 – Pour faire correspondre les données  $Y$  fonction du maximum de réponse (100%) à une concentration molaire, on multiplie ces données par un paramètre de norme,  $Norm$ .

Pour confronter les simulations du modèle aux données expérimentales, joindre une norme aux observables simulées ( $Obs(X, Norm)$ ) plutôt qu'aux données est équivalent et est utilisé au sein de la fonction objectif (définition ci-dessous).

**Définition 1.7.0.2.** On nommera *fonction objectif*, la distance entre les quantités observées expérimentalement,  $Y$ , et leurs observables correspondantes,  $Obs(X, Norm)$ . Cette fonction dépend des valeurs de paramètres  $P$  qui permettent de calculer les simulations  $X$  par intégration numérique de l'ODE et des valeurs de normes ( $Norm$ ) associées aux observables. Elle sera notée  $f_{obj}$  telle que :

$$f_{obj}(P, Norm) = d(Y, Obs(X, Norm)) \quad (1.6)$$

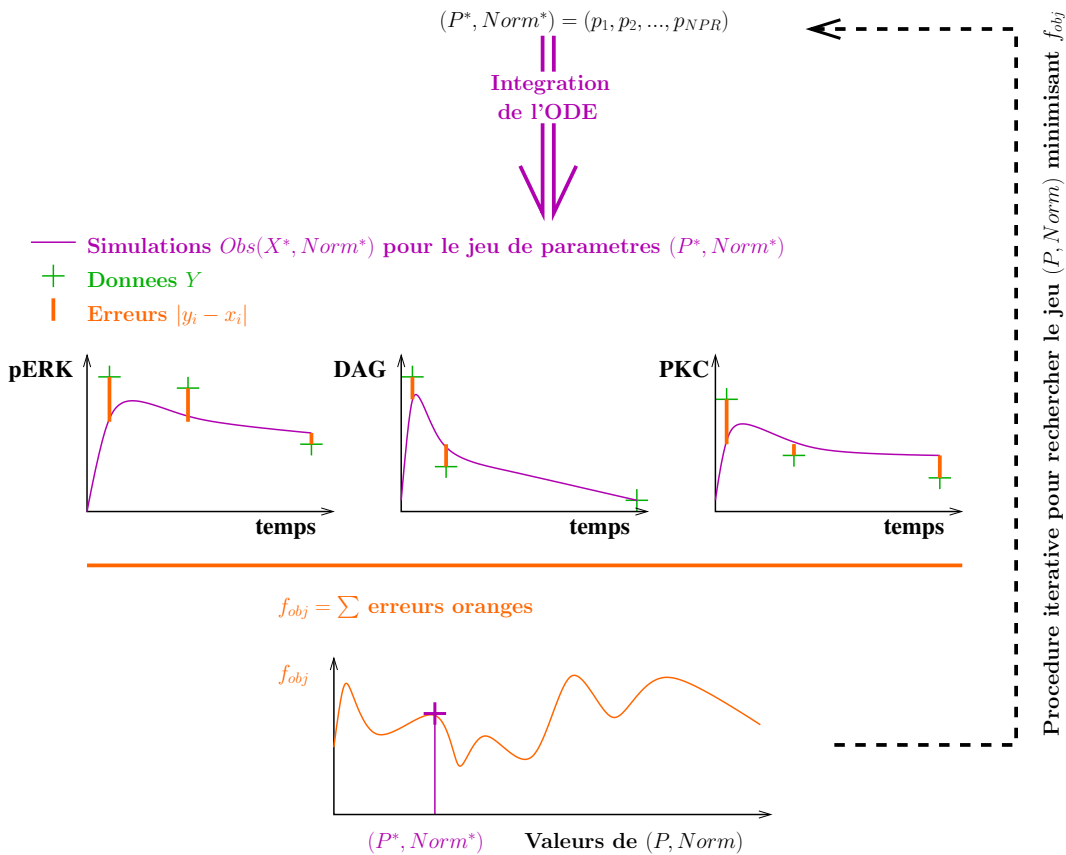


FIGURE 1.25 – La fonction objectif va représenter la somme des distances entre les données et la simulation. Selon les valeurs des paramètres  $(P^*, Norm^*)$ , les simulations sont plus ou moins proches des données et la valeur de la fonction objectif sera plus ou moins petite. L’ajustement paramétrique consiste à rechercher  $(P, Norm)$ .

Le but de l’ajustement paramétrique est d’obtenir une valeur de tous les paramètres inconnus parmi  $P = (K, CI, QT)$  et  $Norm$  afin de minimiser la distance entre les simulations et les données ou encore de minimiser la fonction objectif (Figure 1.25). Par la suite on parlera de jeu de paramètres pour désigner l’ensemble des constantes inconnues et par abus de langage, on désignera ce jeu  $P$ .

La distance  $d$  de la fonction objectif peut être différente selon les logiciels et la nature des éléments composant la fonction objectif. Si  $Y$  et  $Obs(X, Norm)$  sont des vecteurs de dimensions finies ( $Y = (y_1, \dots, y_n)$  et  $Obs(X, Norm) = (x_1, \dots, x_n)$ ), alors  $d$  peut être une

distance usuelle (Figure 1.25). C'est le cas au sein de Copasi [Hoops et al., 2006] et GLSDC [Kimura and Konagaya, 2003] où  $d(Y, Obs(X, Norm)) = \sum_{i=1}^n (y_i - x_i)^2$ . Pour le modèle FSH p70S6K avec Scilab (section 3.1) et au sein de notre implémentation *HYPE* (section 3.3), la distance utilisée est  $d(Y, Obs(X, Norm)) = \sum_{i=1}^n w_i |y_i - x_i|$  où  $w_i$  est un poids qui permet aux erreurs d'être comparables entre observables : les jeux comprenant plus de points de temps ainsi que ceux ayant des valeurs plus fortes ne seront pas privilégiés. Pour le logiciel Biocham, la distance est différente car  $Y$  et  $Obs(X, Norm)$  ne sont pas des vecteurs mais des propriétés temporelles et une trace, simulation du modèle (voire paragraphe 2.2.3 et [Rizk et al., 2008]).

Une fois le problème bien posé : paramètres inconnus listés et fonction objectif établie, il existe plusieurs stratégies pour faire évoluer les paramètres en vue de minimiser la fonction objectif. Regroupées sous le nom de *méthodes d'optimisation*, nous allons en détailler plusieurs ci-dessous, entre autre celles qui seront utilisées au sein des résultats.

### 1.7.1 Méthodes déterministes

Tout d'abord, les méthodes déterministes, elles sont pré-établies et ne s'adaptent pas au cours de la recherche du jeu de paramètre idéal. Souvent, elles utilisent la structure du modèle et leurs champs de recherche est limité au voisinage du jeu de paramètre initial.

C'est le cas des descentes de gradient qui calculent les pentes de la fonction objectif par rapport aux paramètres et fait évoluer le jeu de paramètres dans le sens de la descente pour la fonction objectif (Figure 1.26 et Algorithm 1). D'un point de vue mathématique, la pente d'une fonction correspond à la dérivée de cette dernière. En dimension supérieur, la dérivée devient un vecteur : le gradient de la fonction. Chaque élément de ce vecteur correspond à la dérivée partielle de la fonction objectif selon l'un des paramètres du jeu.

Le gradient  $\nabla_P f_{obj}$  indique la direction dans laquelle les paramètres induisent une augmentation maximale de la fonction objectif, donc  $-\nabla_P f_{obj}(P^0)$  la plus forte descente. Dans cette direction, on avance d'un pas  $\alpha$  pour obtenir le jeux de paramètres suivant :

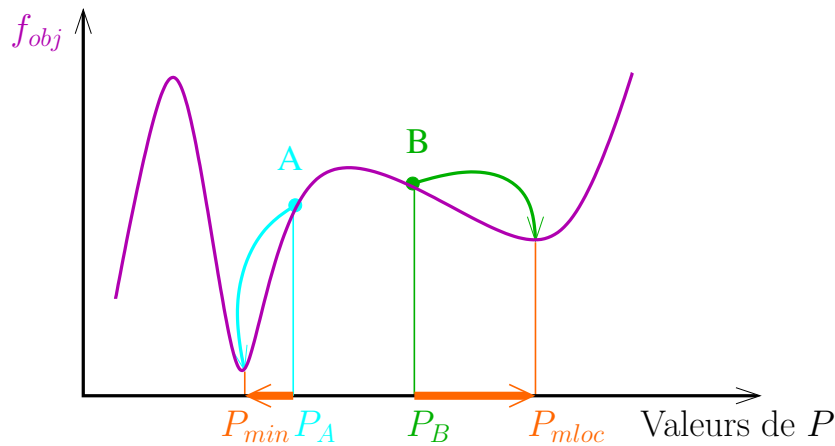
$$P^1 = P^0 - \alpha * \nabla_P f_{obj}(P^0)$$

L'algorithme repart de ce nouveau jeu de paramètres  $P^1$  et le fait évoluer linéairement selon une nouvelle pente,  $-\nabla_P f_{obj}(P^1)$ , pour améliorer le minimum trouvé.

Ces méthodes (Descente de Gradient, Gradient conjugué, Newton, ...) dépendent beaucoup du point de départ et du changement de pente. De plus, le calcul de la pente nécessite que la fonction objectif ait de bonnes propriétés : dérivable une fois dans toutes les directions pour le gradient conjugué et dérivable deux fois pour la méthode de Newton. D'autres formules de descente ont été développées afin d'approcher la précision de la méthode de Newton. Ce sont les méthodes de Quasi-Newton qui, à l'aide du gradient, approximent les dérivées de second ordre comme les méthodes BFGS [Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970].

De telles méthodes de recherche impliquent des calculs de gradients de la fonction objectif par rapport aux paramètres. Pour le cas des fonctions objectifs dont on ne connaît rien (BBO pour Black Box Optimisation), le calcul du gradient peut être remplacé par





→ Evolution du jeu de paramètre

FIGURE 1.26 – En violet les valeurs d’une fonction objectif en fonction d’un paramètre. Pour rechercher le minimum de la fonction objectif, la méthode du gradient conjugué descend le long de la plus grande pente et s’arrête dans le premier minimum rencontré (le pas d’évolution du paramètre peut faire sauter des optima locaux). Pour un paramètre initial  $P_A$ , la méthode trouvera le minimum global  $P_{min}$  tandis qu’en partant de  $P_B$  la méthode s’arrêtera dans un minimum local  $P_{mloc}$ .

des différences finies. Le gradient de  $f_{obj}$  en  $P$  est approché par l’évaluation suivante :

$$\frac{\partial f_{obj}(P)}{\partial p_j} \approx \frac{f_{obj}(P + \epsilon * e_j) - f_{obj}(P)}{\epsilon} \quad (1.7)$$

le long de la  $j$ -ème composante,  $e_j$  est le vecteur unité et  $\epsilon$  est un nombre réel positif. Afin d’assurer une certaine précision dans l’approximation du gradient,  $\epsilon$  doit être petit par rapport aux composantes de  $P$ .

D’autres méthodes évitent le calcul du gradient comme les « Pattern Search ». A chaque itération, les directions sont pré-établies et l’algorithme descend successivement suivant chacune avant de recalculer une série de directions pour l’itération suivante. Cette méthode peut être très rapide comparé aux méthodes comportant un calcul du gradient [Hooke and Jeeves, 1961].

Un autre type de méthodes déterministes appelé « smoothing technique » revient à lisser localement la fonction objectif. Puis, dans le voisinage lissé appelé « Trust region », un minimum est recherché. Un lissage peut être fait à chaque itération mais la fonction optimisée sera de plus en plus différente de la fonction objectif de départ. C’est pourquoi deux contrôles sont effectués sur le jeu de paramètre minimisant localement la fonction lissée : il doit être au sein de la « Trust region » et il doit améliorer la vraie fonction objectif (Figure 1.27).

On retiendra de ces méthodes qu’elles sont très usuelles, cependant, elles nécessitent souvent des propriétés de dérivabilité de la fonction objectif ce qui engendrent des calculs

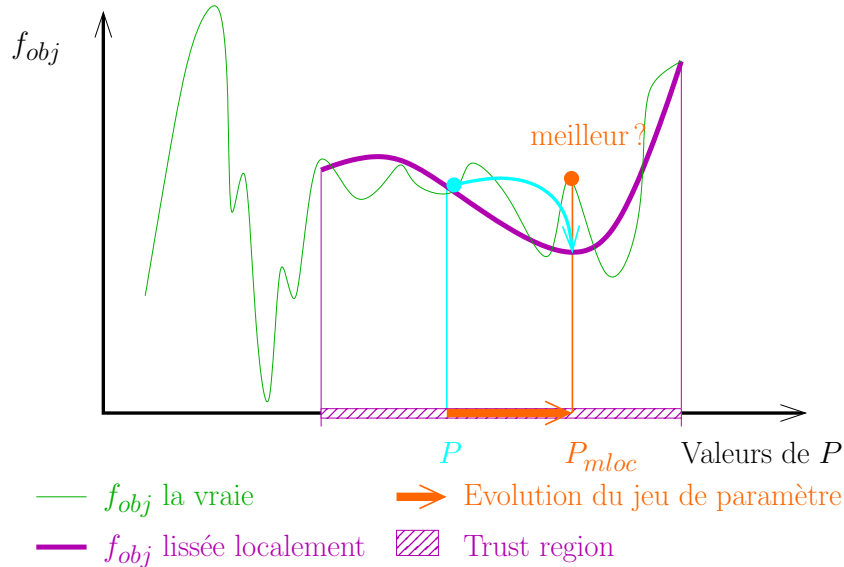


FIGURE 1.27 – Recherche d’un jeu de paramètre qui minimise la fonction objectif par une méthode de « smoothing ». Tout d’abord, la fonction objectif initiale (en vert) est lissée localement (en violet). Puis, le jeu  $P$  est modifié pour minimiser cette fonction objectif lissée. Enfin, on vérifie que le jeu  $P_{mloc}$  trouvé appartienne à la « Trust region » et soit meilleur pour la vraie fonction objectif.

très lourds. D’autre part, la recherche est pré-établie (méthodes déterministes) ce qui ne laisse pas la possibilité à l’optimisation de s’adapter à la fonction objectif. Enfin, elles explorent peu l’espace des paramètres et donnent souvent une solution locale. Pour conclure et en remarquant que leurs résultats dépendent du jeu de départ, une amélioration simple existe et consiste à essayer plusieurs jeux de départ choisis de manière stochastique. Une implémentation a été réalisée sous le nom de BFGS-restart.

## 1.7.2 Méthodes stochastiques globales et locales

Contrairement aux méthodes déterministes, les méthodes stochastiques sont basées sur un échantillonnage aléatoire de l’espace des paramètres ce qui permet une exploration plus large (Figure 1.28).

La première méthode, utilise un algorithme de Monte Carlo. Cet algorithme a été inventé par Nicholas Metropolis, et publiée en 1949 dans un article co-écrit avec Stanislas Ulam [Metropolis and Ulam, 1949]. Elle consiste à créer une population de jeux de paramètres candidats de façon indépendante en utilisant plusieurs fois une distribution de probabilités uniforme fixe. Puis, le meilleur jeu trouvé, celui donnant la plus petite valeur de fonction objectif, est gardé. Ce processus est répété, le meilleur jeu de paramètre, des populations actuelles et précédentes, est enregistré. Il a été prouvé que cette méthode converge en théorie vers le minimum avec une probabilité de 1 si l’optimum fait partie du domaine de définition de la distribution (si aucune région de l’espace des paramètres est ignorée) [Solis and Wets, 1981]. Dans la pratique, le temps de convergence est très grand et

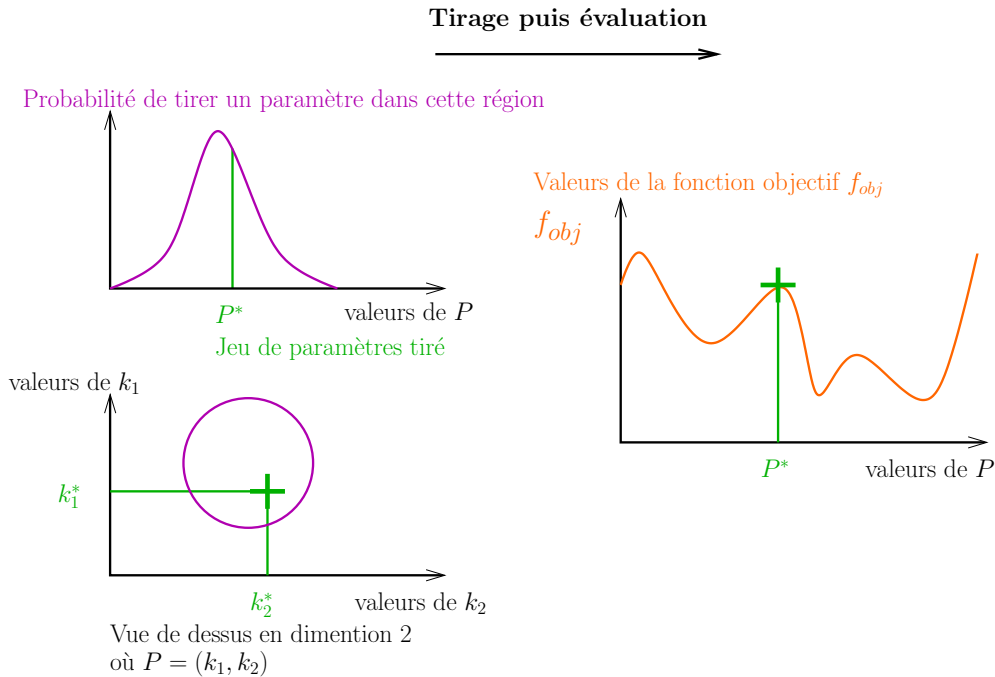


FIGURE 1.28 – Tirage d’un jeu de paramètre  $P^*$  selon une distribution puis évaluation de  $f_{obj}(P^*)$ .

augmente exponentiellement en fonction de la dimension spatiale de recherche (en fonction du nombre de paramètres inconnus).

Un inconvénient de cette méthode est que la recherche par tirage uniforme n’avantage aucune région de l’espace des paramètres. Pour introduire un échantillonnage préférentiel, la population de jeux de paramètres candidats sera générée selon une densité normale par exemple. C’est le cas de « Important Sampling », « Important Sampling Resampling », « Population Monte Carlo » et « Population Monte Carlo à noyaux de transition ».

De plus, ces méthodes font évoluer une population de jeux de paramètres. La fonction objectif est évaluée pour l’ensemble des jeux de la population et la meilleure solution est gardée [Cappé et al., 2004; Douc et al., 2005].

Le but de ces méthodes est d’estimer  $P$  tel qu’il ait la meilleure vraisemblance au vu des données. Soit  $\pi(\xi)$  la loi *a priori* et  $\pi(\xi|Y)$  la loi *a posteriori* qui désignent les emplacements où sont *a priori* et réellement (*a posteriori*) les jeux de paramètres solution. Un estimateur du jeu de paramètre  $P$  est l’espérance de  $\xi$  suivant la loi *a posteriori*.

Avec pour base le théorème de Bayes, on va chercher cette estimation de  $P$  ou espérance *a posteriori* en multipliant la distribution *a priori*, par la vraisemblance, et en effectuant une normalisation.

$$\begin{aligned} E_{\pi(\xi|Y)}(\xi) &= \int_{\Theta} \xi \cdot \pi(\xi|Y) \cdot d\xi \\ &= \frac{\int_{\Theta} \xi \cdot L(\xi|Y) \cdot \pi(\xi) \cdot d\xi}{\int_{\Theta} L(\xi|Y) \pi(\xi) \cdot d\xi} \end{aligned}$$

**Important Sampling** Il peut s'avérer très difficile voire impossible de générer un échantillon suivant la loi  $\pi(\xi|Y)$ . L'idée clé d'« Important Sampling » est de générer l'échantillon selon une densité d'importance (proposal distribution) :  $g$ . L'espérance ci-dessus peut alors se réécrire :

$$E_{\pi(\xi|Y)}(\xi) = \frac{\int_{\Theta} \xi \cdot L(\xi|Y) \cdot \frac{\pi(\xi)}{g(\xi)} \cdot g(\xi) \cdot d\xi}{\int_{\Theta} L(\xi|Y) \frac{\pi(\xi)}{g(\xi)} \cdot g(\xi) \cdot d\xi}$$

On appellera  $w(\xi) = \frac{\pi(\xi)}{g(\xi)}$  le poids d'importance.

Soient  $(\xi_j)_{j=1..Npop}$  variables aléatoires indépendantes suivant la loi  $g$ , d'après la loi forte des grands nombres (LFGN), on a :

$$E_{\pi(\xi|Y)}^{IS}(\xi) = \lim_{Npop \rightarrow \infty} \sum_{j=1}^{Npop} \xi_j \omega_j$$

avec

$$\omega_j = \frac{L(\xi_j|Y) \cdot \frac{\pi(\xi_j)}{g(\xi_j)}}{\sum_{j=1}^{Npop} L(\xi_j|y_{0..p4}) \cdot \frac{\pi(\xi_j)}{g(\xi_j)}}$$

(LFGN : Pour  $(Xn)_n$ , suite de variables aléatoires indépendantes qui suivent la même loi de probabilité et sont intégrables i.e.  $E(|Xn|) < \infty$ ,  $\frac{1}{n} \sum Xn$  converge vers  $E(X)$  « presque sûrement »)

Après avoir choisi une loi *a priori*  $\pi$  et une loi d'importance  $g$ ,  $E_{\pi(\xi|Y)}^{IS}(\xi)$  peut être calculée. Les choix de  $\pi$  et  $g$  offrent une large diversité :  $\pi = g$ ,  $\pi$  gaussienne,  $g$  loi gaussienne,  $g$  loi de Student ...

Cette deuxième méthode, Important Sampling Resampling (Algorithme 2), consiste à ré-échantillonner un échantillon  $(\xi_1, ..\xi_{Npop})$  obtenu par Important Sampling selon une loi multinomiale  $\mathcal{M}(\omega_1, .., \omega_{Npop})$ . On obtient un nouvel échantillon :  $(\tilde{\xi}_1, ..\tilde{\xi}_{Npop})$

$$E_{\pi(\xi|Y)}^{ISR}(\xi) = \frac{1}{Npop} \cdot \sum_{j=1}^{Npop} \tilde{\xi}_j$$

**Population Monte Carlo** Cette troisième méthode d’ajustement paramétrique est aussi un raffinement d’Important Sampling. Population Monte Carlo appelle itérativement Important Sampling Resampling. A chaque itération  $t$ , l’échantillon  $\xi^{t+1}$  est calculé à partir de l’échantillon  $\xi^t$ .

Enfin une dernière variante propose d’utiliser plusieurs lois d’importance pour la génération de chaque  $\xi_j^t$  en pondérant les résultats, elle est appelée Population Monte Carlo avec noyaux de transition (Algorithme 3). Cela permet en faisant varier l’écart type des distributions de raffiner la recherche en se déplaçant « vite » dans les zones où les paramètres induisent une grande erreur et en effectuant une recherche plus précise et locale dans les zones où les paramètres donnent de meilleurs résultats. Ainsi, on effectue une recherche à la fois globale et locale des paramètres minimisant la fonction objectif. Le principal défaut de ces méthodes est le besoin d’*a priori* sur les paramètres à rechercher.

Les méthodes précédentes utilisent des distributions « fixes » au cours de la recherche, d’où sont tirées des populations de jeux de paramètres. Au contraire, la méthode du « Recuit Simulé » génère un seul jeu de paramètres à partir d’une proposition de distribution qui est modulée au cours des itérations. Le jeu de paramètres tiré à l’itération  $n$  suivant la distribution est gardé s’il est meilleur que le jeu  $n - 1$  ou s’il est moins bon avec une probabilité qui évolue. Cette dernière probabilité est fonction d’une « température » qui diminue au cours de la recherche. Pour une distribution normale, cela correspond à faire diminuer le rayon de recherche des jeux de paramètres (Figure 1.29). Ainsi, la recherche plus globale au départ devient progressivement locale, on dira de cet algorithme qu’il est « adaptatif ». Mais notons bien que la « température » (rayon de recherche) est pré-établie dès le départ, l’algorithme n’est pas auto-adaptatif. Une méthode de recuit simulé très utilisée est l’algorithme ASA (Adaptative Simulated Annealing) [Ingber and Ingber, 1989; Ingber, 1993]. Cette méthode comporte des lois de proposition différentes pour chaque paramètre (dans chaque direction) au sein d’un jeu ainsi qu’une « température » qui évolue au fil des itérations.

Les méthodes présentées ci-dessus impliquent le choix des distributions qui doivent couvrir le voisinage du jeu de paramètre qui minimise  $f_{obj}$ . Pour éviter ces problèmes d’*a priori*, les méthodes évolutionnaires sont de bons outils. Dans un premier temps nous allons voir deux méthodes évolutionnaires inspirées de la biologie puis un algorithme récent : CMAES (Covariance Matrix Adaptation Evolution Strategy)

### 1.7.3 Méthodes évolutionnaires inspirées de la biologie

Selon la théorie de l’évolution de Darwin, au sein d’une espèce, des modifications du génome apparaissent au fil des générations. Si le nouvel individu est adapté à son milieu, il survit et par la même occasion transmet ses gènes à la génération suivante. Par contre, l’individu inadapté à l’environnement ne survit pas et entraîne avec lui la perte son génome inadapté. S’inspirant de cette théorie de l’évolution, les méthodes évolutionnaires font évoluer une population d’individus (ou jeux de paramètres). Si le nouvel individu permet

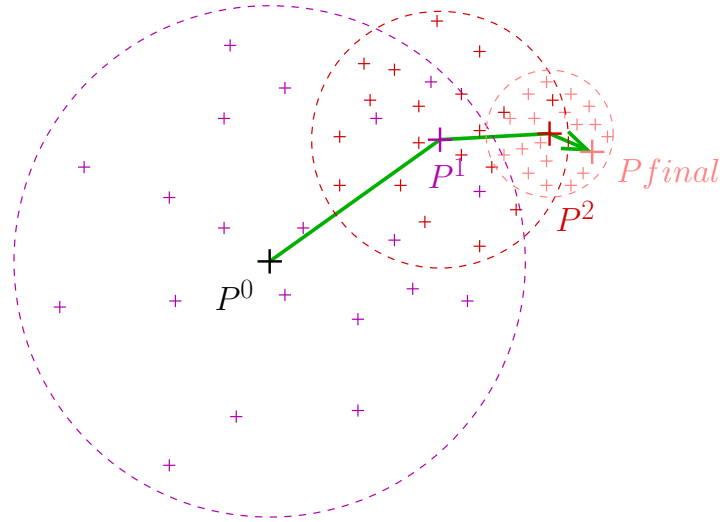


FIGURE 1.29 – Pour rechercher le jeu de paramètre qui minimise la fonction objectif par une méthode de « Recuit Simulé », les jeux de paramètres candidats ( $P^1$ ,  $P^2$ ,  $P^{final}$ ) sont choisis au sein d'une distribution dont le rayon de recherche diminue, il est fonction d'une « température ».

une amélioration de la minimisation de la fonction objectif, il est gardé, sinon éliminé.

Sur ce principe, la méthode « Particle Swarm » mime le comportement d'un essaim de particules qui « volent » sur le paysage de la fonction objectif et échangent des informations sur les meilleures solutions qu'elles ont « vues » [Clerc and Kennedy, 2002]. L'essaim des particules correspond à une population de jeux de paramètres. La méthode consiste à faire évoluer ces jeux de l'essaim en se rapprochant des jeux les plus adaptés. Plus précisément, à chaque génération ou itération  $n$ , chaque jeu de paramètres  $P_i^n$  évolue en fonction :

- du meilleur jeu gardé en mémoire au fil des itérations :  $Pm_i$
- du meilleur jeu de l'itération précédente  $Pg_i^{n-1}$
- du meilleur « voisin »  $Pn_i^n$

Le jeu  $P_i^{n+1}$  à l'itération  $n + 1$  est une somme du jeu  $P_i^n$  de l'itération  $n$  ainsi qu'une vitesse  $v_i^{n+1}$  telle que :

$$v_i^{n+1} \leftarrow wv_i^n + c1r1 \cdot (Pg_i^{n-1} - P_i^n) + c2r2 \cdot (Pm_i - P_i^n) + c3r3 \cdot (Pn_i^n - P_i^n) \quad (1.8)$$

$$P_i^{n+1} \leftarrow P_i^n + v_i^{n+1} \quad (1.9)$$

avec  $w, c1, c2, c3$  des constantes,  $r1, r2$  et  $r3$  des nombres aléatoires.

Une autre méthode évolutionnaire inspirée de la biologie est l'algorithme génétique. C'est même la plus connue et la plus ancienne des méthodes évolutionnaires. Holland développa l'algorithme génétique pour simuler et comprendre comment les organismes s'adaptent à leur environnement naturel [Holland, 1975]. Mais, ce n'est que bien plus tard que ces algorithmes sont repris et utilisés en optimisation, en particulier par [Goldberg, 1989]. Cette méthode consiste à considérer des chromosomes monobrin et leur appliquer des croisements et des mutations. Chaque chromosome représente un jeu de paramètres. Les

paramètres sont les mêmes d'un chromosome à l'autre, seules les valeurs de ces paramètres peuvent être différentes. Cela est représenté par  $k$  et  $k'$  figure 1.31. Tout d'abord deux définitions :

**Définition 1.7.3.1.** Un *croisement*, en biologie, est un échange de gènes homologues entre deux chromosomes d'une même paire. L'opérateur de croisement au sein de l'algorithme génétique va échanger entre deux jeux, des paramètres homologues pouvant avoir des valeurs différentes (par exemple  $p_8$  et  $p'_8$  figure 1.31). Le *croisement barycentrique* appliqué à une paire de parents ( $P_1^n, P_2^n$ ) donne une nouvelle paire telle que :

$$P_1^n = (p'_1, p'_2, \dots, p'_{bar}, p_{bar+1}, \dots, p_{NPR}) \quad (1.10)$$

$$P_2^n = (p_1, p_2, \dots, p_{bar}, p'_{bar+1}, \dots, p'_{NPR}) \quad (1.11)$$

où *bar* l'emplacement du pivot barycentrique

**Définition 1.7.3.2.** En biologie, une *mutation* correspondent à une modification de gène. Adapté à un jeu de paramètres cela revient à modifier légèrement certains paramètres. Dans le cas gaussien, le nouveau jeu  $P^{n+1}$  est piochée autour de  $P^n$  suivant une distribution (Figure 1.30) tel que :

$$P^{n+1} = P^n + \sigma \mathcal{N}(0, C) \quad (1.12)$$

où  $\sigma$  est l'écart type de la mutation et  $\mathcal{N}(0, C)$  un tirage de vecteur gaussien centré de matrice de covariance  $C$  (définie ci-dessous).

**Définition 1.7.3.3.** La *matrice de covariance* est une mesure de dispersion autour d'une moyenne (Figure 1.30). Elle est utilisée pour définir et délimité la distribution gaussienne,  $\mathcal{N}(P^n, C)$  autour de  $P^n$ , où vont être pioché les jeux de paramètres  $P^{n+1}$ .

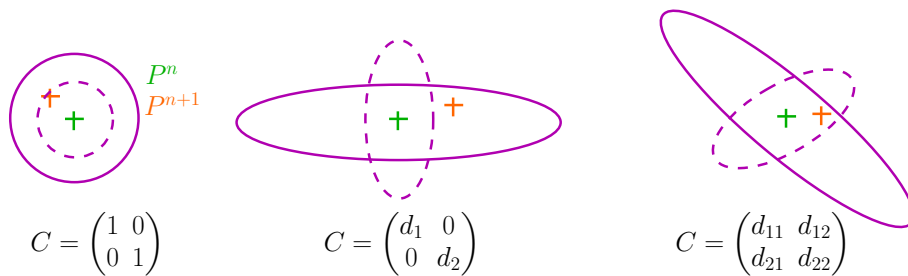


FIGURE 1.30 – La matrice de covariance  $C$  détermine la forme de la gaussienne où est piochée le nouveau jeu de paramètres  $P^{n+1}$  (orange) à partir du jeu  $P^n$  (vert) lors de la mutation. Géométriquement, la matrice de covariance peut être représentée par ses ellipses d'iso-densité (en violet).

Si la mutation est petite, elle permet une exploration du voisinage du jeu de paramètres  $P^n$ , appelée « exploitation ». En revanche, l'opérateur de croisement apporte de la diversité au sein des jeux de paramètre et permet l'exploration de l'espace de recherche.

L'algorithme génétique va utiliser ces opérateurs pour faire évoluer plusieurs jeux de paramètres. Au départ, il y a  $N_{par}$  chromosomes monobris « parents » (ou jeux de paramètres initiaux) choisis aléatoirement. Puis, parmi cet ensemble de « parents » une paire est piochée. Les opérations, mutations et croisements, sont effectuées sur la paire pour former une nouvelle paire de chromosomes dit « enfants ». L'étape de pioche d'une paire de « parents » auxquels on applique des opérations « génétiques » est répétée autant de fois que nécessaire pour obtenir  $N_{enf}$  « enfants ». Puis, la fonction objectif est évaluée pour chacun des nouveaux chromosomes « enfants » créée. Les meilleurs chromosomes parmi les « parents » et « enfants » sont gardés pour former la génération de « parents » suivante et recommencer le processus de création d'« enfants » (Figure 1.31 et Algorithme 4).

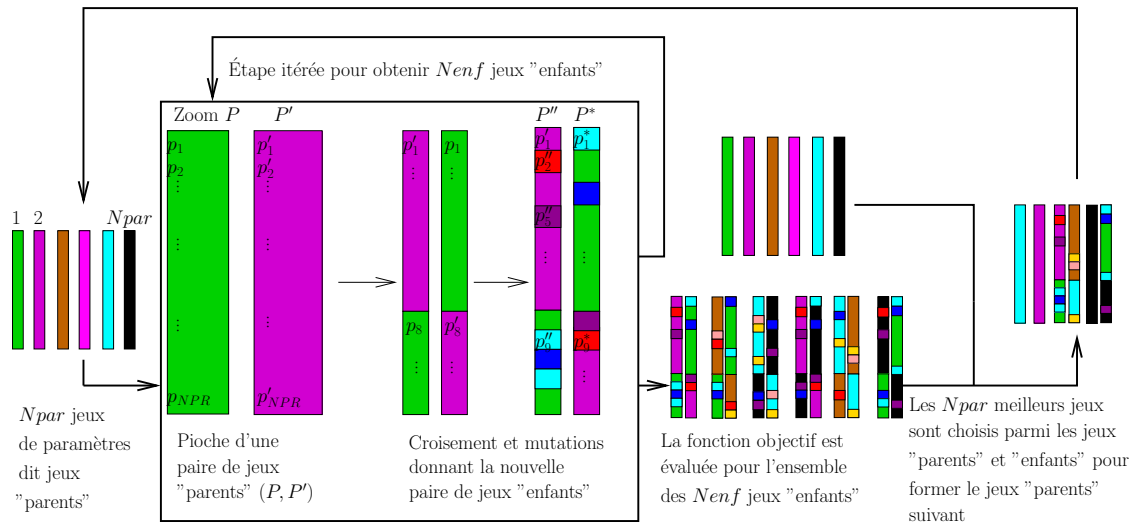


FIGURE 1.31 – Illustration de l'algorithme génétique à l'aide de chromosomes monobris

Lorsqu'on veut effectuer une phase d'exploration (ou recherche globale), on a besoin de fortes mutations alors qu'en phase d'exploitation (ou recherche locale), de petites mutations sont plus adaptées. Pour rendre adaptatif l'algorithme génétique, on peut modifier l'amplitude des mutations ( $\sigma$ ). Intuitivement, une règle assez simple consiste à repérer quand on ne trouve plus d'« enfants » meilleurs que les « parents », ce qui amène à penser que les mutations sont trop fortes et éloignent les enfants des solutions. Dans ce genre de cas, une diminution de l'amplitude des mutations peut être envisagée afin d'exploiter le voisinage des parents. Inversement, lorsque les enfants sont en grande partie meilleurs que les parents, il n'y a aucune raison de rester dans un voisinage proche des solutions déjà trouvées, les mutations peuvent être élargies pour tendre à une exploration de l'espace. De manière similaire, une grande quantité de croisements permet d'explorer l'espace des paramètres tandis qu'un petit nombre favorise l'exploitation des voisinages des solutions pré-existantes. Cependant, un algorithme génétique comporte toujours des croisements, c'est une de ses caractéristiques propre. Cela permet de toujours apporter au fil de la recherche des jeux de paramètres originaux, pas forcément proche des jeux pré-existants.

Pour l'opérateur de mutation, une première adaptation simple est la règle des 1/5.



Dans ce cas, la matrice de covariance  $C$  est l'identité, l'opérateur de mutation est le même pour tous les paramètres d'un jeu. Lorsqu'au bout de  $T$  générations, on a plus d'un enfant sur cinq de gardés, l'écart type régissant la taille de la mutation est multiplié par 1.22, dans le cas contraire, il est multiplié par 0.83 [Auger, 2004].

#### 1.7.4 Stratégie d'Évolution et Adaptation de la Matrice de Covariance

SA-ES pour Self Adaptative Evolution Strategy se focalise aussi sur l'évolution de l'opérateur de mutation. Cette fois-ci, les paramètres de mutation vont évoluer par mutation. L'idée est que la sélection des meilleurs individus va aussi être la sélection des bons opérateurs de mutation. Les jeux qui auront un opérateur de mutation inadapté vont être éliminés à l'itération suivante car ils vont donner de mauvais paramètres. On nomme auto-adaptatif ce type d'algorithme car la sélection se fait d'elle-même sans intervention déterministe. Il existe plusieurs SA-ES, je vous présenterai ci-dessous les modèles isotrope et non-isotrope pour des matrices de covariance identité et diagonale. Le cas d'une matrice de covariance pleine est plus compliqué mais offre la possibilité d'avoir des corrélations entre les différents écarts types des différents paramètres d'un même jeu [Auger, 2004] (Figure 1.30).

*SA-ES isotrope :*

l'écart type est muté par multiplication avec une loi log-normale :

$$\sigma^{n+1} = \sigma^n \exp(\tau \mathcal{N}(0, 1)) \quad (1.13)$$

puis utilisé pour muter le parent :

$$P^{n+1} = P^n + \sigma^{n+1} \mathcal{N}(0, I) \quad (1.14)$$

où  $\tau$  est un paramètre fixé par l'utilisateur et  $\mathcal{N}(0, I)$ , un tirage gaussien.

*SA-ES non-isotrope :*

l'écart type est un vecteur  $(\sigma_1^n, \dots, \sigma_{NPR}^n)$ , chaque paramètre au sein du jeu a un écart type spécifique.

Et, pour chaque coordonnée ou paramètre au sein du jeu, on a

$$\sigma_i^{n+1} = \sigma_i^n \exp(\tau \mathcal{N}(0, 1)) \exp(\tau' \mathcal{N}_i(0, 1)) \quad (1.15)$$

$$P_i^{n+1} = P_i^n + \sigma_i^{n+1} \mathcal{N}'_i(0, 1) \quad (1.16)$$

où  $\tau$  et  $\tau'$  deux paramètres fixés par l'utilisateur,  $\mathcal{N}(0, 1)$ , des tirages indépendants de loi normale centrée réduite.

Ces méthodes sont plus riches que la méthode 1/5, cependant elles nécessitent des populations « enfants » assez larges et donc un coût en temps de calcul non négligeable. L'inconvénient de ces techniques est que les mutations des opérateurs de mutation peuvent compenser la mutation elle-même [Auger, 2004].

C'est pourquoi un nouveau type d'ajustement de l'opérateur de mutation est apparu utilisant le tirage aléatoire utilisé pour muter le paramètre. Ces approches sont appelées « dérandomisées » car on enlève les tirages utilisés pour muter  $\sigma$  pour ne garder qu'un tirage commun à sigma et à la mutation des paramètres. La méthode CSA (Cumulative Step-Size Adaptation) utilise ce principe et simplifie la population « parent ». Elle ne se base que sur un élément « parent » pour créer une population « enfant », puis calcule une somme pondérée de ses enfants en guise de parent suivant. Les poids sont inversement proportionnels à la valeur de l'erreur commise avec le jeu. Enfin, le trait qui donne son nom à la méthode est la prise en compte du « chemin » parcouru entre les parents au fil des itérations. On entend par chemin parcouru  $c_n = P_{n+1} - P_n$  (représenté en vert Figure 1.29). Si ce chemin est « grand » et toujours dans la même direction, c'est que le même chemin aurait pu être suivi plus rapidement, alors l'écart type est augmenté. Dans le cas contraire il est diminué pour une recherche plus locale, une exploitation du voisinage du parent [Ostermeier et al., 1994].

Alors que CSA se base sur une matrice de covariance diagonale, CMA (pour Covariance Matrix Adaptation) utilise une matrice de covariance pleine. Une actualisation de la matrice dans sa globalité est ajoutée pour CMA par rapport à CSA où seuls les différents écarts types sont modifiés. L'actualisation de la covariance CMA comprend une somme de la matrice de covariance précédente, du produit scalaire des deux derniers chemins parcourus ainsi qu'une prise en compte du chemin parcouru par les  $\mu$  meilleurs jeux de paramètres [Hansen N, 2003]. Faisant partie des stratégies d'évolution, cette méthode sera aussi appelée : CMAES (Algorithme 5). C'est un algorithme très performant qui est utilisé pour les recherches de paramètres au sein de l'environnement Biocham [Rizk et al., 2010]. Cet algorithme a soulevé de nombreuses questions et motivé la troisième partie des résultats qui suivent.



# Objectifs scientifiques

Comme nous l'avons vu, les R7TMs ont une importance indéniable. Ils sont au cœur de larges réseaux de signalisation ce qui permet une multitude de modulations selon le type cellulaire, les quantités des espèces du réseau etc... La complexité de ces réseaux est telle que l'utilisation d'approches mathématiques et informatiques est devenue nécessaire.

À l'INRA, au sein de l'équipe BIOS, la signalisation FSH est abordée en observant individuellement quelques éléments du réseau. En particulier, l'AMPC, la PKA, le PIP3, la p70S6K, la rpS6, les  $\beta$ -arrestines et les formes phosphorylées de la cascade MAPK Raf/MEK/ERK. Afin de démêler les comportements de chaque espèce au sein du réseau, des études du système en conditions normales et perturbées sont réalisées. Les perturbations peuvent être des mutations du récepteur, des inhibitions enzymatiques ou des inactivations d'ARN messagers. En parallèle, le récepteur de l'angiotensine (RAT<sub>1A</sub>) est d'avantage étudié, nous avons accès à des jeux de données plus abondants dans le cadre d'une collaboration entre RJ Lefkowitz (Duke University) et plusieurs expériences publiées permettaient de valider le modèle. En particulier, le RAT<sub>1A</sub> est le R7TM modèle pour l'étude de la voie arrestine.

D'un autre côté, à l'INRIA, au sein de l'équipe Contraintes, un logiciel : Biocham a été conçu pour étudier les réseaux de signalisation. Cet environnement encore naissant au début de ma thèse (sa création date de 2002) couple un formalisme de description des réseaux simple et précis à des outils numériques et informatiques. En parallèle, un grand programme de modélisation multi-échelle de l'axe hypothalamo-hypophyso-gonadique est en cours de réalisation et rassemble plusieurs équipes de l'INRIA et de l'INRA. A terme nos travaux sur la signalisation FSH pourraient alimenter le niveau moléculaire de ce modèle.

Dans ce contexte nous avons voulu relever plusieurs défis :

- modéliser les réseaux de signalisation induits par deux R7TMs et comprendre comment ils induisent des réponses biologiques distinctes (*eg* prolifération *vs* différenciation)
- réaliser un modèle original à l'aide du logiciel Biocham et permettre de faire progresser le logiciel afin de l'adapter aux contraintes biologiques.
- apporter notre contribution à la compréhension et à l'utilisation de méthodes d'optimisation paramétrique.

En effet, les modèles dynamiques nécessitent la recherche préalable des paramètres inconnus pour être simulés. L'ajustement paramétrique par apprentissage à partir des données biologiques s'est révélé comme un enjeu majeur pour atteindre nos deux autres objectifs.



Deuxième partie

Matériels et Méthodes



## Chapitre 2

# Bases expérimentales et mathématiques

Tout d'abord, je vous propose de décrire les principes de bases des expériences biologiques utilisées au cours de cette thèse ainsi que les bases mathématiques sous-jacente à la modélisation ODE, les algorithmes d'optimisations qui ont servit aux résultats et une brève présentation des propriétés temporelles liées au logiciel Biocham.

### 2.1 Expériences biologiques et Interprétations

#### 2.1.1 Dosage ELISA

Cette méthode permet de mesurer la concentration d'un produit à un temps de stimulation. En particulier, pour le modèle p70S6K, nous avons utilisés ce type d'expérience pour mesurer les variations d'AMPc. Le principe de la méthode est de mettre en compétition l'AMPc à doser avec de l'AMPc marqué. On observe ensuite la quantité d'AMPc marqué, plus elle est importante, plus le signal d'AMPc cellulaire est faible (Figure 2.1). La quantification est ensuite rendue possible en comparant les intensités obtenues à une gamme d'expériences réalisées avec de l'AMPc purifiée en doses connues. Par opposition, la méthode FRET va observer au sein des cellules et en temps réel l'évolution d'une espèce.

#### 2.1.2 FRET : d'un quotient de fluorescence à une variable

**Description de la méthode biologique :** Le FRET ou Förster Resonance Energy Transfert est une méthode expérimentale qui permet d'observer des évènements de signalisation en temps réel sur des cellules vivantes. Plus précisément, le FRET permet de mesurer un transfert d'énergie entre deux fluorophores. Plusieurs senseurs existent qui permettent de mesurer différents signaux (*eg* : activité PKC, PKA, DAG, AMPc, ...). Nous avons eu recours aux senseurs PKC et DAG pour générer les données ayant servi à l'ajustement du modèle angiotensine. L'interprétation en terme de variable et de concentration de ce type de données pose des problèmes au modélisateur. C'est pourquoi, nous détaillons ci-dessous la façon dont nous avons résolu ces problèmes en prenant l'exemple



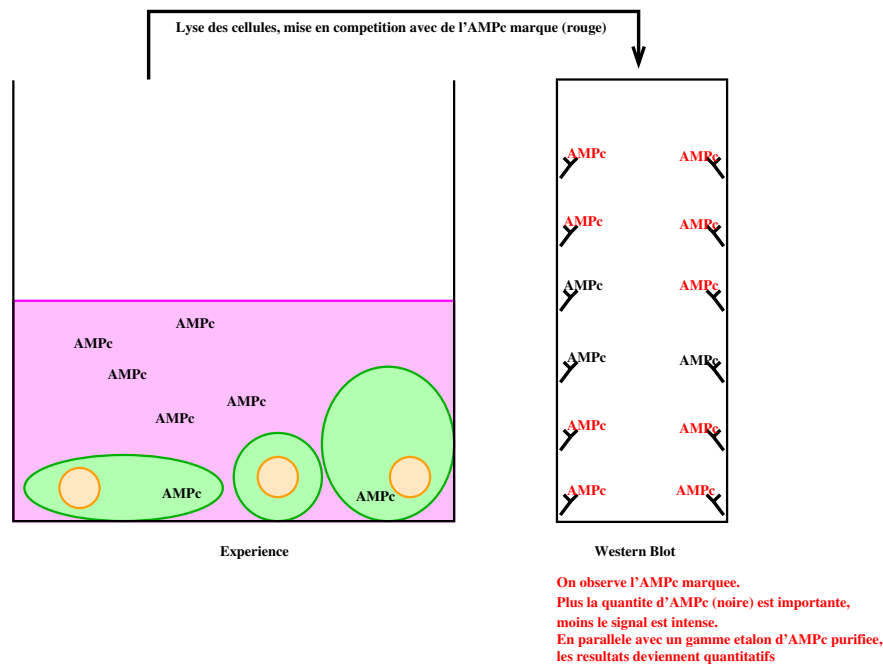


FIGURE 2.1 – Méthode de dosage permettant de mesurer la concentration d'une espèce au sein d'une expérience.

de données AMPc obtenues au sein du laboratoire sur des cellules HEK293 stimulées par de la FSH. Dans ce cas, nous voulons observer la production d'AMPc. Or l'AMPc n'est pas une protéine donc on ne peut pas lui greffer des senseurs FRET. Cependant, l'AMPc libre active Epac. Deux protéines fluorescentes sont donc greffées de part et d'autre de l'Epac tel que CFP-Epac-YFP. Puis, cette construction est transfectée au sein des cellules. Enfin, sous le microscope, la CFP, excitée à la longueur d'onde 425 nm, émet à son tour à  $470 \pm 20 \text{ nm}$  [Ponsioen et al., 2004]. Cette nouvelle longueur d'onde fait partie du spectre d'excitabilité de la YFP qui émet à son tour une lumière à  $530 \pm 25 \text{ nm}$ . Ces deux intensités lumineuses sont enregistrées : 470nm et 530nm (Figure 2.2). Le quotient de ces intensités représente la résonance entre la CFP et la YPF qui est elle-même dépendante de la distance entre les deux fluorochromes. L'AMPc en se liant au senseur induit un changement de conformation qui provoque une modification de la distance entre les deux fluorochromes et donc s'accompagne d'un changement du signal FRET.

**Analyse des données :** Les données brutes sont des relevés d'intensité en lumière bleu et jaune (longueurs d'ondes 470nm et 530nm). Pour observer une variation de transfert d'énergie, on regarde le quotient duquel on déduit le bruit (quotient bleu/jaune d'une zone sans cellule) puis on norme à 1 (Figure 2.3).

**Interprétation en terme de variables et quantités molaires :** La ligne basale vaut 1 alors qu'en réalité, il existe une concentration initiale d'AMPc :  $[AMPc(t = 0)]$ .

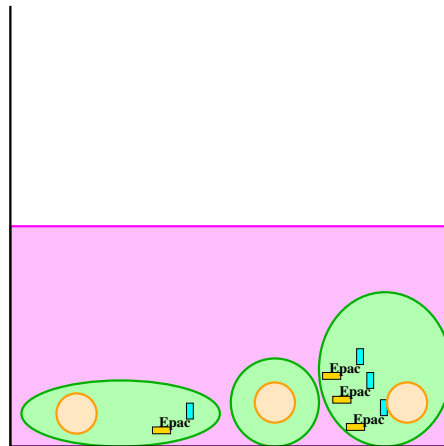


FIGURE 2.2 – La méthode FRET mesure une quantité au sein de cellules *in vivo*.

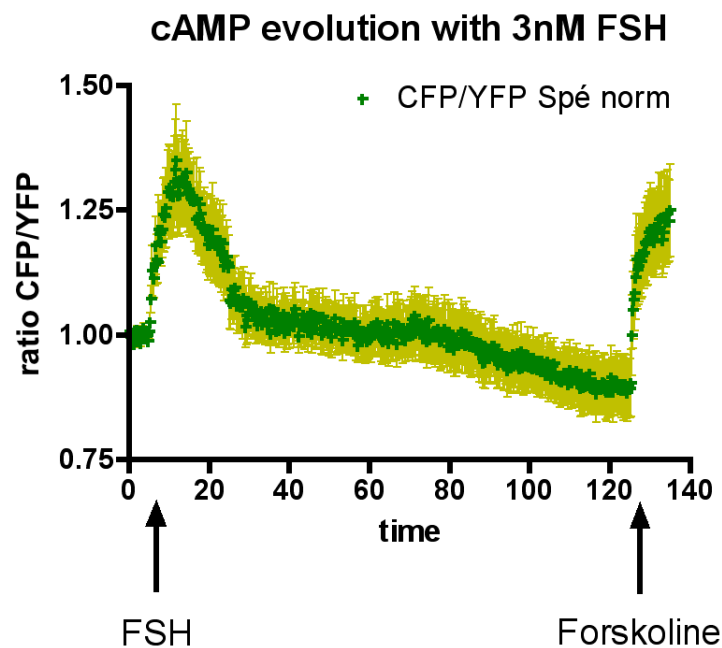


FIGURE 2.3 – Cinétique Epac FRET suite à l'injection de 3nM de FSH sur des cellules HEK293.

De plus, en observant le quotient des intensités, on observe une variation relative et non quantitative, il faut donc ajouter une norme :  $NormAMPc$ .

L'observable,  $Obs(AMPc, NormAMPc)$ , qui peut être comparée aux données FRET,  $Y_{AMPc}$ , est telle que :

$$Obs(AMPc, NormAMPc) = \frac{[AMPc(t)] - [AMPc(t=0)]}{NormAMPc} + 1 \quad (2.1)$$

Pour le DAG et la PKC au sein du modèle angiotensine cela donne :

$$Obs(DAG, NormDAG) = \frac{[DAG(t)] - [DAG(t=0)]}{NormDAG} + 1 \quad (2.2)$$

$$Obs(PKC, NormPKC) = \frac{[PKC(t)] - [PKC(t=0)]}{NormPKC} + 1 \quad (2.3)$$

Alors que la précision en temps est impressionnante : mesure toutes les 15 secondes, la grande inconnue est au niveau des quantités absolues, de la valeur des normes  $NormAMPc$ ,  $NormDAG$ ,  $NormPKC$ . Ces paramètres doivent donc être recherchés lors du processus d'optimisation.

### 2.1.3 Des Blots aux Puces : une miniaturisation des procédés

Le même problème de quantités relatives existe pour les expériences de « Western blot » où tout est fonction d'un 100% habituellement fixé au maximum de la réponse contrôle. Dans notre équipe l'utilisation d'une approche de puce à protéine en phase inverse a permis d'effectuer des cinétiques plus fines et des répétitions plus nombreuses. Et en plus, les puces permettent la mise en parallèle de données d'une gamme étalon et donc une quantification précise des observations (Figure 2.4).

Les extraits protéiques sont spottés, chaque spot correspond à une condition expérimentale (capacité maximale de l'ordre de 12 000 spots par puce). On analyse ensuite la puce avec un anticorps sélectif donné (eg pERK ou pAkt) ce qui permet de mesurer l'évolution de la phosphorylation d'une protéine donnée (ERK ou Akt) dans de très nombreuses conditions expérimentales (eg différents points de cinétique ou de dose réponse, différents modèles cellulaires, différents stades,...). D'autre part, cette méthode consomme très peu de matériel expérimental, il est donc possible de réaliser de multiples répliques de la même lame et de les analyser avec des anticorps différents.

La méthode et sa validation sont présentées en détail dans l'article :

A highly sensitive near-infrared fluorescent detection method to  
analyze signalling pathways by reverse-phase protein array

L. Dupuy, C. Gauthier, G. Durand, A Musnier, D Heitzler, A Herledan, V. Sakanyan, P.  
Crépieux and E. Reiter

*Proteomics*, 2009, 9, 5446-5454

qui est en annexe A.

## 2.2 Bases mathématiques et algorithmiques

### 2.2.1 Résolution de l'équation différentielle

Les réseaux de signalisation énoncés dans la première partie de l'introduction représentent des processus dynamiques. Le formalisme ODE utilisé pour modéliser ces systèmes fait cor-

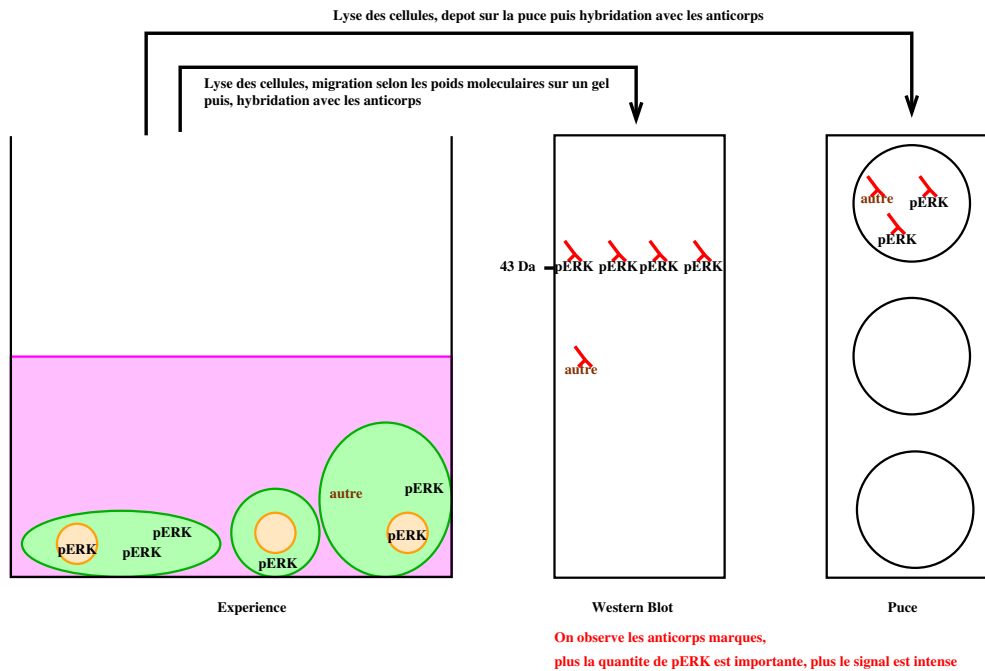


FIGURE 2.4 – La méthode de « Western Blot » mesure la quantité de pERK piégée par un anticorps. Cette mesure est réalisée après une répartition sur un gel en fonction du poids moléculaire. Ainsi, le « Western blot » ne mesure pas le non spécifique. De manière assez similaire, la mesure sur puce repère les pERK muni d’un anti-corps mais aussi un peu de non spécifique. toutefois, lorsque l’anti-corps est très spécifique, les résultats sont identiques.

respondre chaque molécule à une variable, chaque réaction à une loi et chaque constante cinétique à un paramètre (Tableau 1.3). Pour obtenir les concentrations des espèces au cours du temps, une intégration numérique du modèle ODE doit être réalisée.

*Remarque.* Il existe *une unique* solution à notre système pour un jeu de paramètres donné, mais à part pour certains cas particuliers (linéaire...) on ne sait pas la calculer. On effectue donc *une* intégration numérique, c’est-à-dire *une* approximation de la solution.

### 2.2.1.1 La démarche de base

Reprenons l’ODE (1.1) énoncée en introduction, et détaillons une intégration numérique de type Euler en dimension 1 :

$$\frac{dx(t)}{dt} = f(t, x(t), p)$$

Pour intégrer le modèle, on part de la valeur initiale :  $x_0$  puis on cherche la valeur suivante :  $x_1$  comme une approximation de  $x(t_0 + h)$  où  $h$  est un petit pas en avant sur l’axe du temps.

Cela revient à considérer la forme discrète de l'équation différentielle :

$$\frac{\Delta x(t)}{\Delta t} = f(t, x(t), p)$$

Pour un pas de temps  $\Delta t = h$  :

$$\begin{aligned} \Rightarrow \quad \frac{x(t+h)-x(t)}{h} &= f(t, x(t), p) \\ \Leftrightarrow \quad x(t+h) &= x(t) + h \cdot f(t, x(t), p) \end{aligned}$$

En partant de  $x_0$  et ainsi de suite jusqu'au temps désiré ( $t = t_0 + n \cdot h$ ).

A titre d'exemple, imaginons-nous quittant Tours à 8h00 (position initiale :  $x(t_0)$  au temps  $t_0 = 8h00$ ) roulant vers le nord à une vitesse de  $60km.h^{-1}$  ( $f(t, x(t), p) = 60km.h^{-1}$ ).

L'intégration numérique pourrait être imagée de la sorte :

- à  $t_1 = t_0 + 5 = 8h05mns$   
On se situe à  $x(t_0) + h \cdot f(t_0, x(t_0), p)$   
Soit Tours + 5mns \* vitesse = Tours + 5km = Tours Nord
- à  $t_2 = t_1 + 5 = 8h10mn$   
On se situe à  $x(t_1) + h \cdot f(t_1, x(t_1), p)$   
Soit Tours Nord + 5mns · vitesse = Tours Nord + 5km = Parçay Meslay
- à  $t_3 = t_2 + 5 = 8h15mns$   
On se situe à  $x(t_2) + h \cdot f(t_2, x(t_2), p)$   
Soit Parçay Meslay + 5mns · vitesse = Parçay Meslay + 5km = Monnaie
- à  $t_4 = t_3 + 5 = 8h20mns$   
On se situe à  $x(t_3) + h \cdot f(t_3, x(t_3), p)$   
Soit Monnaie + 5mns · vitesse = Monnaie + 5km = l'INRA de Nouzilly

Cette intégration numérique dite d'Euler n'est exacte que pour les modèles linéaires (vitesse constante). Lorsque le modèle est non-linéaire, l'estimation de l'évolution de  $x(t+h)$  par rapport à  $x(t)$  va devoir être affinée car la solution n'est pas exactement  $h * f(t, x(t), p)$  (Figure 2.5).

Pour déterminer  $x(t+h)$  en fonction de  $x(t)$ , le développement de Taylor (2.4) offre une bonne précision lorsque le pas  $h$  est de taille raisonnable et l'ordre  $n$  est assez grand. Cependant, il nécessite le calcul des dérivées de  $f$  jusqu'à l'ordre  $n$ . Cela exige des propriétés de dérivabilité et implique d'important calculs souvent long. C'est pourquoi la méthode de Runge Kutta retrouvant la précision du développement de Taylor sans dériver la fonction  $f$  est très intéressante.

$$\begin{aligned} x(t+h) &= x(t) + h \cdot f(t, x(t), p) + \frac{h^2}{2!} \cdot f'(t, x(t), p) + \dots \\ &\quad + \frac{h^n}{n!} \cdot f^{[n]}(t, x(t), p) + o(h^n) \end{aligned} \tag{2.4}$$

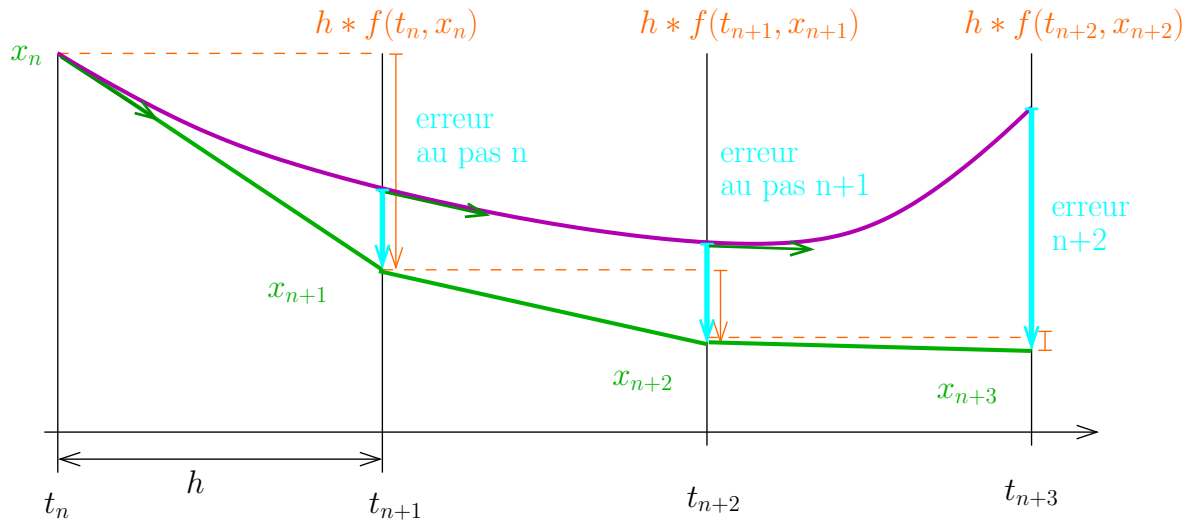


FIGURE 2.5 – En violet la vraie solution de l’ODE, en vert son approximation par la méthode d’Euler. A chaque pas, on part du point précédent, par exemple  $x_n$ , on approxime la fonction (en violet) par un tronçon linéaire (en vert) pour déterminer le point suivant  $x_{n+1} = x_n + h * f(t_n, x_n, p)$ . Les erreurs commises à chaque pas (en bleu) s’accumulent, les changements de pente accentuent l’erreur.

### 2.2.1.2 Runge Kutta

Les méthodes de Runge Kutta vont retrouver le développement de Taylor (2.4) à l’aide de  $f$  évalué en une combinaison de points bien choisis. Ainsi, pour obtenir le développement de Taylor à l’ordre 4 avec un pas de temps  $h$ , on évalue  $f$  quatre fois :

$$\begin{aligned}
 rk1 &= hf(t_n, x_n, p) \\
 rk2 &= hf(t_n + h/2, x_n + rk1/2, p) \\
 rk3 &= hf(t_n + h/2, x_n + rk2/2, p) \\
 rk4 &= hf(t_n + h, x_n + rk3, p)
 \end{aligned}$$

puis, on ajoute la combinaison linéaire comme ci-dessous à  $x_n$  pour obtenir  $x_{n+1}$

$$x_{n+1} = x_n + \frac{1}{6}rk1 + \frac{1}{3}rk2 + \frac{1}{3}rk3 + \frac{1}{6}rk4 + o(h^4) \quad (2.5)$$

La méthode de Runge Kutta offre une grande précision de  $x_{n+1}$  en fonction de  $x_n$  en utilisant seulement la fonction  $f$  et non comme dans la formule de Taylor l’ensemble des dérivées de la fonction  $f$  jusqu’à l’ordre 4 [Press et al., 1992].

### 2.2.1.3 Méthodes explicite, implicite et semi-implicite

Au sein des méthodes d'intégration numérique, il existe deux grands types de méthodes : *explicites* et *implicites*.

**Définition 2.2.1.1.** On dit qu'une méthode est *explicite* lorsque  $x_{n+1}$  s'exprime en fonction de  $x_n$ , on peut alors calculer aisément cette nouvelle valeur. Dans le cas contraire, la méthode est dite *implicite*.

Par exemple, pour l'équation linéaire suivante :

$$\frac{dx(t)}{dt} = f(t, x(t), p) = -k \cdot x(t)$$

( $k$  étant une constante  $> 0$ )

le schema Euler implicite, s'écrit

$$\begin{aligned} x_{n+1} &= x_n + h \cdot f(t_n, x_{n+1}, p) \\ \Leftrightarrow x_{n+1} &= x_n + h \cdot (-k \cdot x_{n+1}) \\ \Leftrightarrow x_{n+1} \cdot (1 + h \cdot k) &= x_n \\ \Leftrightarrow x_{n+1} &= x_n / (1 + h \cdot k) \end{aligned} \tag{2.6}$$

Les méthodes implicites engendrent plus de calculs mais offre une meilleure précision. Ce résultat se généralise dans le cas d'un système d'équations linéaires à plusieurs variables. Le coût supplémentaire pour utiliser une méthode implicite est alors l'inversion d'une matrice à chaque pas. Si les équations ne sont pas linéaires, on doit procéder à une linéarisation locale. Ainsi, partant de l'équation générale vectorielle (1.1), sur laquelle nous avons basée nos modèles, la formule d'intégration implicite analogue à (2.6) devient :

$$X_{n+1} = X_n + h[I - h \cdot \frac{\partial f}{\partial X}(t, X_n, P)]^{-1} \cdot f(t, X_n, P) \tag{2.7}$$

où  $X$ ,  $P$  et  $f$  désignent des vecteurs,  $I$  la matrice unité et  $\partial f / \partial X$  la matrice jacobienne. (La matrice jacobienne d'une fonction vectorielle  $f$  par rapport à un vecteur  $X$  est la matrice de ses dérivées partielles par rapport aux composantes de  $X$ ). Cette résolution de l'équation implicite par linéarisation est appelée méthode *semi-implicite*. Il n'est pas totalement garanti qu'elle soit stable, mais elle l'est en général car localement le comportement est semblable à celui des systèmes linéaires.

### 2.2.1.4 Pas d'intégration auto-adaptatif

Enfin, un autre levier pour améliorer la précision de l'intégration numérique en limitant les calculs est d'adapter la largeur du pas  $h$ . Lorsque les variations des espèces au sein de l'ODE sont modérées, un pas de taille moyenne suffit tandis qu'en cas de grandes variations des espèces, un petit pas est nécessaire afin de garder une bonne précision.

Je vais vous présenter deux techniques. La première consiste à effectuer successivement deux pas d'intégration différents :

- une fois avec un pas égal à  $2h$  ; résultat :  $x_1$
- une deuxième fois en faisant deux pas égaux à  $h$  ; résultat :  $x_2$

La différence

$$\Delta = x_2 - x_1$$

est une estimation de l'erreur de troncature commise. Dans le cas des méthodes de Runge Kutta, il suffit de bien choisir deux combinaisons de coefficients en facteur des  $rki$  pour obtenir  $x_1$  et  $x_2$ .

D'une autre manière, en désirant une erreur égale ou inférieure à  $\Delta_0$ , on peut utiliser la formule :

$$h' = h|\Delta_0/\Delta|^\alpha$$

où  $\alpha$  est un exposant positif (par exemple  $\alpha = 0.2$ ).

Elle permet d'obtenir une valeur du pas  $h'$  qui donnera directement une erreur de l'ordre de  $\Delta_0$ . Ainsi, si  $\Delta < \Delta_0$  (pas d'intégration réussi), le pas pourra être augmenté ( $h' > h$ ) pour l'étape suivante. Au contraire, si  $\Delta > \Delta_0$ , l'intégration numérique est erronée et doit être recommencée avec un pas plus petit ( $h' < h$ ).

### 2.2.1.5 Rosenbrock

La méthode de Rosenbrock utilise la linéarisation semi-implicite et dans le même esprit que la méthode de Runge Kutta pousse les calculs à un ordre supérieur. De plus, cette méthode utilise un pas d'intégration auto-adaptatif [Rosenbrock, 1963]. Il en résulte une méthode stable devant une grande partie des résolutions d'ODE non-linéaire ainsi qu'une bonne précision en un temps de calcul réduit par rapport à Runge Kutta. Cette méthode est utilisée au sein du logiciel Biocham ainsi que dans l'environnement HYPE qui ont servi aux résultats (sections 3.2 et 3.3)

## 2.2.2 Méthodes de recherches de paramètres

La résolution de l'équation différentielle donne une simulation du modèle pour un jeu de paramètres donné. Or, une partie des paramètres de l'ODE nécessaires à l'intégration numérique est inconnue au départ. Nous allons donc chercher ces paramètres manquants,  $P$ , tels qu'ils minimisent la fonction objectif (Définition 1.7.0.2). Au cours de la thèse, j'ai testé différentes méthodes d'optimisation que je vais expliciter ci-dessous. En particulier, celles qui ont servi à l'élaboration des résultats : la méthode du gradient conjugué pour le modèle FSH (section 3.1), l'algorithme CMAES pour le modèle angiotensine (section 3.2) et l'algorithme génétique au sein de HYPE (section 3.3).

### 2.2.2.1 Gradient Conjugué

Tout d'abord, la méthode du gradient conjugué qui minimise une fonction objectif en se laissant glisser suivant la pente vers un minimum, comme l'illustre la Figure (1.26). Plus précisément, pour  $f_{obj}$  une fonction objectif et  $P$  les paramètres à rechercher, la méthode du gradient conjugué le modifie le jeu de départ suivant la direction du gradient pour



trouver une première amélioration  $P^1$  du jeu de départ. Puis, l'étape suivante consiste à se déplacer le long d'une direction conjuguée en partant de  $P^1$ . Le processus peut être réitéré. L'algorithme s'arrête quand il trouve un minimum, si aucun progrès n'est fait après une réinitialisation de la direction ou quand un critère de tolérance est atteint. Cette méthode est très simple mais elle ne permet pas l'exploration d'un grand espace de paramètres car son défaut est de s'arrêter au sein de minima locaux rencontrés (Figure 1.26 et Algorithme 1).

---

**Algorithme 1** : Algorithme du Gradient Conjugué

---

**pour chaque itération faire**  
 Calculer  $\Delta P^n = -\nabla_P f_{obj}(P^n)$  la direction de plus forte pente  
 Calculer  $\beta^n = \frac{\Delta P^{nT} \Delta P^n}{\Delta P^{n-1T} \Delta P^{n-1}}$ , formule de Fletcher-Reeves.  
 Mettre à jour la direction conjuguée :  $\Lambda P^n = \Delta P^n + \beta^n \Lambda P^{n-1}$   
 Chercher  $\alpha^n$  pour avoir le minimum de  $f_{obj}(P^n + \alpha^n \Lambda P^n)$   
 Mettre à jour la position :  $P^{n+1} = P^n + \alpha^n \Lambda P^n$   
**fin**

---

**2.2.2.2 Important Sampling et Population Monte Carlo**

Contrairement au gradient conjugué, Important Sampling et Population Monte Carlo sont des méthodes stochastiques. L'algorithme propose  $N_{pop}$  jeux de paramètres générés suivant une distribution de probabilité à chaque itération de la recherche. La fonction objectif est évaluée pour l'ensemble des possibilités et la meilleure solution est gardée [Cappé et al., 2004; Douc et al., 2005].

---

**Algorithme 2** : Important Sampling Resampling

---

**pour chaque jeu de paramètres  $j = 1..N_{pop}$  faire**  
 Générer  $\xi_j$   
 Calculer sa fonction d'importance  $g(\xi_j)$   
 Evaluer la fonction objectif  
 Calculer la fonction de vraisemblance  $L(\xi_j|Y)$   
 Calculer le prior  $\pi(\xi_j)$   
 Calculer les poids d'importance  $w_{isj}$ .  
 Normaliser les poids :  $w_{nisj} = \frac{w_{isj}}{\sum w_{isj}}$   
 Piocher avec remise une nouvelle population  $pop_{isr}$  :  $(\tilde{\xi}_{j=1..N_{pop}})$  parmi les  $\xi_j$   
 selon une  $\mathcal{M}(\omega_1, \dots, \omega_{N_{pop}})$

**fin**

Une estimation de  $P$  est

$$E_{\pi(\xi|Y)}^{ISR}(\xi) = \frac{1}{N_{pop}} \cdot \sum_{j=1}^{N_{pop}} \tilde{\xi}_j$$


---

**Population Monte Carlo avec noyaux de transition** Une variante de la méthode de population monte Carlo qui propose plusieurs écart types pour les distributions de recherche de paramètres. Cela permet de faire varier le déplacement du jeu de paramètre selon son erreur. La recherche est à la fois globale et locale avec plusieurs écart types.

---

**Algorithme 3** : Population Monte Carlo avec noyaux de transition

---

```

pour chaque itération  $t = 1..Niter$  faire
  pour chaque fonction d'importance  $d = 1..Nd$  faire
    Important Sampling Resampling
    pour chaque  $j = 1..Npop$  faire
      | Garder le meilleur jeu de paramètres parmi les  $d$  générés
    fin
  fin
   $\xi_j^{t+1}$  est constitué pour l'itération suivante en piochant parmi les jeux selon la
  loi  $\mathcal{M}(\omega_1^{t+1}, \dots, \omega_{Npop}^{t+1})$ 
fin

```

---

### 2.2.2.3 Algorithme génétique

Les méthodes évolutionnaires permettent d'éviter les problèmes liés au caractère déterministe des fonctions ou aux besoins d' *a priori*. Parmi ces méthodes, la plus connue est l'algorithme génétique (Algorithme 4) détaillé en introduction (section 1.7, Figure 1.31).

### 2.2.2.4 CMAES : Covariance Matrix Adaptation Evolution Strategy

Contrairement à l'algorithme génétique, CMAES propose l'évolution d'une population, en ne gardant à chaque itération qu'une moyenne pondérée des  $\mu$  meilleurs jeux, puis en leur ajoutant  $\sigma y_i$  avec  $y_i \sim N(0, C)$ . Pour cette méthode évolutionnaire, tout réside dans l'évolution de la matrice de covariance  $C$  ainsi que dans la taille des mutations, représentée par  $\sigma$ . Comme on peut le montre l'algorithme 5, ces évolutions dépendent d'un grand nombre de paramètres, ainsi que du chemin parcouru jusqu'à la nouvelle moyenne pour la matrice de covariance. Pour l'évolution de  $\sigma$ , un calcul de variance effective est pris en compte.

---

**Algorithme 4** : Algorithme génétique utilisé au sein du programme HYPE

---

```

Données : Population "parents" de  $Npar$  jeux de paramètres
/* Evaluation de la fonction objectif sur les jeux « parents » */
pour chaque jeu « parents » faire
    Résolution de l'ODE;
    Calcul de l'erreur;
fin
Tri des jeux « parents » initiaux suivant leur valeur de fonction objectif
/* Boucle de génération des jeux « enfants » et mise à jour des jeux
« parents » */
pour ( $k = 0; k < Niter; k ++$ ) faire
    /* Création des Enfants deux par deux */
    pour ( $i = 0; i < \frac{Nenf}{2}; i ++$ ) faire
        Pioche de deux jeux « parents » proportionnellement à leur erreur ;
        Croisement ;
        Mutation ;
    fin Pour
    /* Évaluation de la fonction objectif sur les jeux « enfants » */
    pour chaque Enfant faire
        Résolution de l'ODE;
        Calcul de l'erreur ;
    fin
    Tri des jeux « enfants » + « parents ».
    Prendre les  $Npar$  meilleurs pour former la population de jeux
    « parents » suivante.
fin Pour

```

---

---

**Algorithme 5** : Algorithme CMAES utilisé par Biocham et le programme HYPE

---

Initialisation et valeurs des coefficients

$$m \in \mathbb{R}^n, \sigma \in \mathbb{R}^+, C = I,$$

$$p_c = 0, p_\sigma = 0, c_c \approx 4/n, c_\sigma \approx 4/n, c_1 \approx 2/n^2, c_\mu \approx \mu_w/n^2, c_1 + c_\mu \leq 1, \\ d_\sigma \approx 1 + \sqrt{\mu_w/n}, \lambda \text{ et } w_i, i = 1, \dots, \mu_w \text{ tel que } \mu_w \approx 0.3\lambda$$

**tant que** *un critère d'arrêt n'est pas atteint* **faire**

    Générer une population de jeux « enfants »

**pour chaque** ( $i = 0; i < N_{enf}, i++$ ) **faire**

$$| \quad PR_i = m + \sigma y_i \text{ où } y_i \sim N(0, C)$$

**fin**

    Mettre à jour la moyenne  $m$  en ajoutant à la moyenne précédente une somme pondérée des jeux centrés obtenus :

$$m = m + \sigma y_w \text{ où } y_w = \sum_{i=1}^{\mu} w_i y_{i:\lambda}$$

    Mettre à jour la matrice de covariance  $C$  en utilisant :

        le chemin parcouru jusqu'à la nouvelle moyenne  $m$

$$p_c = (1 - c_c)p_c + 1_{\{\|p_\sigma\| < 1.5\sqrt{n}\}} \sqrt{1 - (1 - c_c)^2} \sqrt{\mu_w} y_w$$

$$C = (1 - c_1 - c_\mu)C + c_1 p_c p_c^T + c_\mu \sum_{i=1}^{\mu} w_i y_{i:\lambda} y_{i:\lambda}^T$$

    Mettre à jour le coefficient  $\sigma$  en utilisant :

        des indications sur la variance effective :

$$p_\sigma = (1 - c_\sigma)p_\sigma + \sqrt{1 - (1 - c_\sigma)^2} \sqrt{\mu_w} C^{-1/2} y_w$$

$$\sigma = \sigma \cdot \exp\left(\frac{c_\sigma}{d_\sigma} \left(\frac{\|p_\sigma\|}{E\|N(0, I)\|} - 1\right)\right)$$

**fin Tq**

---

### 2.2.3 Logiciels et environnements de programmation disponibles

Pour conclure cette partie, je dirai quelques mots sur l'accessibilité de ces méthodes ou sur leur mode d'utilisation pour réaliser des modèles dynamiques de réseau de signalisation. Une possibilité est de trouver un logiciel ou un environnement de programmation qui les contient sous forme de fonctions pré-programmées. C'est généralement le cas pour les fonctions de résolution numériques des ODEs. En effet, Scilab [Gomez et al., 1998], Biocham [Chabrier-Rivier et al., 2003; Calzone et al., 2006; Fages and Soliman, 2008a], CellDesigner [Kitano et al., 2005], Copasi [Hoops et al., 2006] et bien d'autres logiciels proposent l'intégration numérique des ODEs. En fournissant l'ODE et l'ensemble de ses paramètres à ces logiciels et environnements de programmation, ils fournissent les simulations désirées. En revanche, concernant les méthodes de recherche de paramètres, seuls certains logiciels et environnements sont dotés de méthodes, quasi Newton et le gradient conjugué pour Scilab, CMAES pour Biocham [Hansen N, 2003; Rizk et al., 2010], algorithme génétique pour GLSDC [Kimura and Konagaya, 2003; Nakakuki et al., 2010] et 13 méthodes variées pour Copasi [Hoops et al., 2006]. Ci-dessous, nous détaillons le logiciel Biocham qui a servi à l'élaboration du modèle angiotensine (section 3.2). Contrairement aux autres logiciels, il est basé sur un formalisme un peu particulier : la logique temporelle.

#### 2.2.3.1 Biocham et la logique temporelle

La machine abstraite biochimique (Biocham) propose de décrire un modèle selon un formalisme de règles de réaction avec expressions cinétiques. De ce haut niveau d'abstraction, Biocham peut calculer l'ensemble des influences positives et négatives entre espèces (le graph d'influence) mais aussi dériver un système d'ODE [Fages and Soliman, 2008b]. Comme CellDesigner et Copasi, Biocham propose une écriture du modèle dans le formalisme de description des réseaux de signalisation standard : le *SBML* pour Systems Biology Markup Language. Mais, au delà de ces particularités génériques, Biocham est doté d'une possibilité d'interroger le modèle sur des propriétés booléennes (Computation Tree Logic, CTL), des propriétés temporelles (Linear Time Logic, LTL) et des propriétés temporelles avec degré de liberté (quantifier free LTL(R) formulae, QFLTL) [Calzone et al., 2006; Rizk et al., 2008; Rizk et al., 2010].

Plus précisément,

- les propriétés CTL sont utilisés pour décrire les états du modèle, par rapport à l'absence ou la présence de composés moléculaires. Ces propriétés booléennes ne nécessitent pas la connaissance des cinétiques du modèle. Un exemple :

$E_i(EF(bpERK))$ .

signifie qu'il existe un état initial à partir duquel il existe une suite de règles permettant d'obtenir *bpERK*. On peut poser la question, le logiciel Biocham vérifiera si cela est vrai ou faux.

- les propriétés LTL décrivent des comportements temporels quantitatifs. Elles sont basées sur la simulation dynamique du modèle et qualifie les concentrations et leurs dérivés. Un exemple :

$G((5=<Time=<35) \rightarrow (1.055266378=<0bs\_DAG=<1.198806693))$ .

exprime le fait qu'entre 5mns et 35mns, l'observable DAG doit être comprise entre  $c_1 = 1.055266378$  et  $c_2 = 1.198806693$ . Le logiciel Biocham peut vérifier si cela est vrai ou faux.

- les propriétés QFLTL ajoutent aux propriétés LTL un degré de liberté et permettent non pas une vérification de la propriété mais un calcul des valeurs qui rendent la propriété vraie. Ainsi, on peut ré-exprimer la propriété ci-dessus :

$G((5 \leq \text{Time} \leq 35) \rightarrow (v_1 \leq \text{Obs\_DAG} \leq v_2))$ .

et Biocham calcule le  $v_1$  et le  $v_2$  correspondant à la dernière simulation. Par exemple, *Obs\_DAG* peut être comprise entre 1.255 et 1.278.

Pour l'optimisation paramétrique, la fonction objectif d'un modèle peut être décrite à l'aide de propriétés LTL (appelée propriétés LTL « objectif » par la suite). Puis, des variables  $(v_1, v_2, \dots)$  sont associées à chaque constante  $(c_1, c_2, \dots)$  régissant ces propriétés LTL « objectif » ce qui produit des propriétés QFLTL « objectif ». Pour chaque jeu de paramètres candidat  $P$ , le modèle est simulé (intégration numérique) puis l'erreur est évaluée avec la « distance QFLTL » (Figure 1.25). Cette distance consiste au calcul du domaine de validité des QFLTL « objectif » pour la simulation (calcul des  $(v_1, v_2, \dots)$  valable pour la simulation) puis, à la recherche du degré de violation de ces propriétés *i.e.* la distance entre ces valeurs et les constantes initiales  $(c_1, c_2, \dots)$ . C'est ce degré de violation des propriétés temporelles (ou somme des degrés pour l'ensemble des propriétés temporelles objectif) qui servira d'erreur, de valeur à la fonction objectif pour guider l'algorithme d'ajustement vers un jeu solution [Rizk et al., 2008].

Avec ces propriétés des études de robustesses sont aussi possibles [Rizk et al., 2009] mais cela n'a pas été développé pour les modèles étudiés au cours de cette thèse. Plus particulièrement, nous verrons dans la partie qui suit, un modèle FSH réalisé avec Scilab, le modèle angiotensine utilisant Biocham puis le développement d'un environnement de recherche de paramètres, « HYPE » (HYbrid Parameter Estimation).



Troisième partie

Résultats





## Chapitre 3

# Résultats et articles

### 3.1 Modélisation dynamique du réseau de signalisation du RFSH

L'un des constats réalisés par les biologistes en signalisation était que la complexité des réseaux ne permettait plus leur analyse et leur compréhension par une étude « manuelle » des données expérimentales. Au chapitre 1, nous avons vu les différentes méthodes mathématiques et informatiques qui pouvaient être appliquées aux problèmes de signalisation. Pour l'évolution dynamique des concentrations intracellulaires des différentes espèces du réseau, nous avons choisi le formalisme ODE. Nous avons détaillé les modèles existants et constaté que, concernant les R7TMs, les modèles sont rares et ne permettent pas de répondre aux questions biologiques actuelles. Par ailleurs, la prise en compte au sein des modèles des différentes conditions expérimentales est balbutiante tandis que l'évolution d'un réseau en fonction des différents stades de développement n'a jamais été modélisé.

#### 3.1.1 La p70S6K une kinase clé pour l'activation de la traduction des protéines par la FSH

Pour le réseau FSH, le mécanisme de traduction est en partie régulé par la p70S6K qui régule le facteur d'initiation rpS6. Différentes formes phosphorylées modulant l'activation de la p70S6K ont été observées en cellules primaires de Sertoli en prolifération et différenciées. Pour comprendre l'incidence de la signalisation p70S6K suivant le stade de développement, nous avons réalisé un modèle d'activation de la rpS6 via la p70S6K après stimulation par la FSH pour les deux stades de développement. Plus précisément, au sein des cellules primaires de Sertoli, l'activation et l'activité enzymatique de la p70S6K est régulée par deux voies de signalisation indépendantes, PI3K/mTOR et PKA, et ces régulations diffèrent en fonction du stades de développement cellulaire :

- en prolifération (cellules issues de rats de 5 jours *post partum* (5dpp)) et en différenciation (cellules issues de rats de 19 jours *post partum* (19dpp)) :
  - \* la FSH active la voie PKA qui induit une dephosphorylation des sites {T421, S424} de la p70S6K.

- en prolifération (5dpp) :
  - \* la FSH active la voie PI3K/mTOR qui induit une phosphorylation sur le site  $\{T389\}$  de la p70S6K.
  - \* la forme de la p70S6K qui active rpS6 est la Pp70S6K $\{T389, S421/424\}$ .
  
- en différenciation (19 dpp).
  - \* la voie PI3K/mTOR est constitutivement activée ainsi la phosphorylation du site  $\{T389\}$  de la p70S6K.
  - \* la forme de la p70S6K qui active rpS6 est la Pp70S6K $\{T389\}$ .

Ainsi, dans les cellules en prolifération stimulées par la FSH les deux voies de signalisation (PI3K et PKA) ont des rôles concurrents et opposés sur les sites de phosphorylation  $\{T389\}$  et  $\{T421, S424\}$  de la P70S6K. Il en résulte que le PI3K semble avoir un effet positif sur l'activation de la rpS6 à travers p70S6K tandis que le PKA a un effet négatif. Au sein des cellules en différenciation, p70S6K est déjà phosphorylée sur le site  $\{T389\}$ , la voie PI3K est indépendante de la stimulation FSH. Mais, la forme active de p70S6K ayant changé, cette fois-ci, la PKA exerce un rôle positif dans l'activation de la rpS6 par p70S6K.

### 3.1.2 L'apport d'un modèle dynamique

Au delà de ces observations basées sur les expériences, nous avons montré qu'un seul modèle dynamique du réseau de régulation de la p70S6K par la FSH suffit à représenter les deux stades de développement. Dans l'article qui suit nous détaillons comment, en conservant la structures et en ne modifiant que certaines quantités initiales et catalyses, le modèle explique les deux stades. Enfin, le modèle a permis de prédire les cinétiques des quatre formes de p70S6K sachant que l'on a observé que des sommes de formes (Figure 3.1). Le modèle prédit que l'apparente stabilité de Pp70S6K $\{T389\}$  observée en différenciation serait due à une compensation de l'augmentation de la forme Pp70S6K $\{T389\}$  par une diminution des Pp70S6K $\{T389, T421, S424\}$  ce qui n'est pas le cas en prolifération.

Ce travail a fait l'objet de la publication suivante, détaillant les expériences biologiques ainsi que le modèle unique de régulation de la p70S6K au cours du développement. Le modèle comporte 7 équations différentielles ordinaires et 24 paramètres ajustés en fonction des données.

### 3.1. MODÉLISATION DYNAMIQUE DU RÉSEAU DE SIGNALISATION DU RFSH

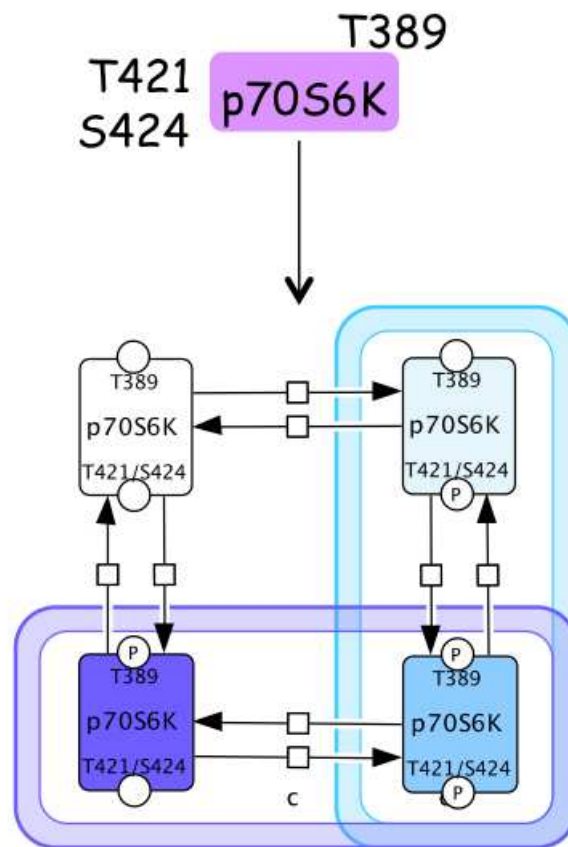


FIGURE 3.1 – La p70S6K comporte plusieurs sites de phosphorylation, nous nous sommes focalisés sur deux sites :  $\{T389\}$  et  $\{T421, S424\}$ . Les p70S6K observées avec l'anticorps  $\{T389\}$  sont une somme des formes  $p70S6K\{T389\}$  et  $p70S6K\{T389, T421, S424\}$ , de même pour les données générées avec l'anticorps  $\{T421, S424\}$ .

# Developmental regulation of p70 S6 kinase by a G protein-coupled receptor dynamically modeled in primary cells

Astrid Musnier · Domitille Heitzler · Thomas Boulo ·  
Sophie Tesseraud · Guillaume Durand · Charlotte Lécureuil ·  
Hervé Guillou · Anne Poupon · Eric Reiter · Pascale Crépieux

Received: 8 April 2009 / Revised: 9 July 2009 / Accepted: 12 August 2009  
© Birkhäuser Verlag, Basel/Switzerland 2009

**Abstract** The mechanisms whereby G protein-coupled receptors (GPCR) activate signalling pathways involved in mRNA translation are ill-defined, in contrast to tyrosine kinase receptors (TKR). We compared a GPCR and a TKR, both endogenously expressed, for their ability to mediate phosphorylation of 70-kDa ribosomal S6 kinase p70S6K in primary rat Sertoli cells at two developmental stages. In proliferating cells stimulated with follicle-stimulating hormone (FSH), active p70S6K was phosphorylated on T389 and T421/S424, through cAMP-dependent kinase (PKA) and phosphatidylinositol-3 kinase (PI3K) antagonizing actions. In FSH-stimulated differentiating cells,

active p70S6K was phosphorylated solely on T389, PKA and PI3K independently enhancing its activity. At both developmental stages, insulin-induced p70S6K regulation was consistent with reported data. Therefore, TKR and GPCR trigger distinct p70S6K active conformations. p70S6K developmental regulation was formalized in a dynamic mathematical model fitting the data, which led to experimentally inaccessible predictions on p70S6K phosphorylation rate.

**Keywords** p70S6K · G protein-coupled receptor · Dynamic model · Differentiation

---

A. Musnier · D. Heitzler · T. Boulo · G. Durand · A. Poupon ·  
E. Reiter · P. Crépieux (✉)  
BIOS Group, INRA, UMR85,  
Unité Physiologie de la Reproduction et des Comportements,  
37380 Nouzilly, France  
e-mail: Pascale.Crepieux@tours.inra.fr

A. Musnier · D. Heitzler · T. Boulo · G. Durand · A. Poupon ·  
E. Reiter · P. Crépieux  
CNRS, UMR6175, 37380 Nouzilly, France

A. Musnier · D. Heitzler · T. Boulo · G. Durand · A. Poupon ·  
E. Reiter · P. Crépieux  
Université François Rabelais, 37041 Tours, France

A. Musnier · D. Heitzler · T. Boulo · G. Durand · A. Poupon ·  
E. Reiter · P. Crépieux  
Haras Nationaux, 37380 Nouzilly, France

S. Tesseraud  
INRA UR83 de Recherches Avicoles, 37380 Nouzilly, France

C. Lécureuil · H. Guillou  
The Inositide Laboratory, The Babraham Institute,  
Cambridge CB2 4AT, UK

## Introduction

The 70-kDa ribosomal S6 kinase (p70S6K) is a serine/threonine kinase critically involved in cell growth control in vivo [1, 2]. In vitro, serum-induced entry into S phase of fibroblasts is impaired by anti-p70S6K antibodies [3], indicating a role in cell cycle control. Among other substrates, p70S6K phosphorylates ribosomal S6 protein (rpS6), a component of the 40S eukaryotic ribosomal subunit [4] which plays a role in the ribosome cohesion and in its recruitment to the m7GTP structure of the mRNAs [5]. p70S6K also phosphorylates eukaryotic initiation factor 4B (eIF4B) which enhances the helicase activity of the eukaryotic initiation factor 4A (eIF4A) [6, 7]. Both are part of the pre-initiation translational complex, and their activation induced by p70S6K is thought to enhance translation initiation [8].

In response to growth factors, p70S6K gets phosphorylated [9]. One reference model postulates that the phosphorylation of p70S6K induced by serum or growth factor addition proceeds stepwise [10, 11]. Phosphorylation

of the T421/S424 site located in the autoinhibitory domain relieves the intramolecular inhibition exerted on T389, located in the linker region. T389 is the important residue which, when phosphorylated, serves as an anchoring point to PDK1 (3-phosphoinositide-dependent kinase) whose substrate, among others, is the T229 located in p70S6K activation loop [12, 13]. All those inducible phosphorylation events depend on the phosphoinositide-3 kinase (PI3K)/mammalian target of rapamycin (mTOR) pathway and, accordingly, are inhibited by the PI3K inhibitor LY294002 and the mTOR inhibitor rapamycin [12]. The complexity of the PI3K/mTOR pathway is incremented with feedback loops and cross-talks with other pathways, such as the MAP kinase-dependent signaling, which are mainly integrated at the level of TSC2 [14, 15].

It is commonly admitted that p70S6K is conformationally active only once those three sites at least have been phosphorylated. However, data suggesting that this particular sequence of p70S6K phosphorylation may be functionally uncoupled from p70S6K enzymatic activity now lead to stem from this view. For example, mutations in the carboxyterminal end of p70S6K indicate that these phosphorylation sites are not involved in controlling enzyme activity in mitogen-stimulated 3T3 fibroblasts [16]. Even more, in mitotic cells, paclitaxel, a chemotherapeutic agent which destabilizes microtubules, has been shown to stimulate phosphorylation of p70S6K on T421/S424, concurrently with catalytic inactivation of the enzyme [17], indicating that phosphorylation of this region of p70S6K may even be linked to enzyme inactivation. In addition, in response to G protein-coupled receptors (GPCR), agonists recognizing the thyroid-stimulating hormone receptor [18], the prostaglandin F2 $\alpha$  receptor [19], the endothelin receptor [20], or the  $\alpha$ 1-adrenergic receptor [21] respectively, epistatic elements such as PI3K may not be activated [18, 19], or may even be inhibited [21], despite activation of p70S6K.

Our own results in differentiating primary Sertoli cells stimulated by follicle-stimulating hormone (FSH) also contradict the commonly admitted model of p70S6K regulation [22]. FSH importance in mammalian reproduction has long been acknowledged and has been confirmed both in male and in female, by knocking out the FSH receptor (FSH-R) [23] or the specific FSH $\beta$  subunit [24]. At the cellular level, FSH controls the proliferation and differentiation of somatic cells of the gonads: granulosa cells of the ovarian follicle and Sertoli cells of the seminiferous tubules [25, 26]. Those stages are separate during the testis development: Sertoli cells proliferate perinatally and proceed through their differentiation program until the end of puberty.

The FSH-R is a GPCR whose coupling to heterotrimeric protein Gs and to scaffolding proteins, namely  $\beta$ -arrestins

[27], accounts for most FSH cellular responses. Gs coupling leads to protein kinase A (PKA) activation, which phosphorylates several nuclear as well as cytosolic targets. In primary differentiating Sertoli cells, our group has previously shown that FSH enhances p70S6K catalytic activity via PKA activation, which is also responsible for the dephosphorylation of the T421/S424 residues. PI3K was not involved in the regulation of this site [22], and was not stimulated by FSH in these cells. Nevertheless, PI3K dynamically maintained a constitutive level of T389 phosphorylation, required for further p70S6K activation in response to FSH. Of note, insulin stimulation of these cells led to inducible phosphorylation of both the T421/S424 and T389 residues, consistently with the current model of stepwise phosphorylation of p70S6K in response to growth factors as mentioned above [10].

All these observations outline the complexity of p70S6K regulation and address the following question: is the particular activation of p70S6K that we have observed in differentiating Sertoli cells, a hallmark of GPCRs such as the FSH-R, or is it dependent on the cell developmental stage, i.e., does it vary in proliferating versus terminally differentiating Sertoli cells? To address this question, we sought to analyze and to dynamically model the p70S6K activation pathways in proliferating primary Sertoli cells isolated from 5-day-old rats and stimulated with FSH or with insulin.

## Materials and methods

### Materials

Human FSH was purchased from Merck Serono (Geneva, Switzerland) and insulin from Sigma Chemical (St Louis, MO). Porcine FSH was purified by Dr Jean Closset (Université de Liège, Belgium). LY294002 and rapamycin were purchased from Cell Signaling Technology (Beverly, MA) and Myr-PKI from Calbiochem (San Diego, CA). Antibodies raised against phospho-p70S6K [T389], phospho-p70S6K [T421/S424], phospho-rpS6 [S235/Ser236], phospho-mTOR [S2448], phospho-Foxo1 [S256], rpS6, and Foxo1 were purchased from Cell Signaling Technology. Antibodies to ERK1, 2, to mTOR and to p70S6K were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Horseradish-coupled peroxidase antibody was purchased from GE Healthcare UK (Buckinghamshire, UK). Cholera toxin (CTX) was from Sigma Chemical.

### Primary Sertoli cell culture

Animals were treated following the current ethical guidelines of the European Community. Sertoli cells were

isolated from testes of 5- or 19-day-old Wistar rats and seeded in DMEM (Invitrogen Life Technologies, Carlsbad, CA) complemented (DMEMc) with penicillin (100 IU/ml), streptomycin (100 µg/ml), glutamine (2 mM), retinol (50 ng/ml), vitamin E (200 ng/ml), and human transferrin (5 µg/ml), all purchased from Sigma Chemical, as reported previously [28]. At 24–48 h after purification, cells were exposed to the PI3K inhibitor LY294002 (50 µM, 10 min) or to the mTOR inhibitor rapamycin (10 nM, 60 min), or to the PKA inhibitor Myr-PKI (50 µM, 60 min), or were left untreated before being exposed to human FSH (100 ng/ml) or to human insulin (100 ng/ml) from 5 to 60 min, or to cholera toxin for 60 min. To minimize non-specific effects, dose-responses of the various pharmacological inhibitors were carried out. The experiments were performed with the minimal efficient concentrations of pharmacological inhibitors that did not affect cell viability, determined by a MTT assay, as described [29]. The results presented herein show the response at 60 min for the p70S6K activity and hormone stimulation alone and at 30 min for hormone stimulation following pharmacological treatment.

#### p70S6K activity assay

The activity of the enzyme was measured by quantifying the incorporation of  $\gamma$ -<sup>33</sup>P-ATP on 11 amino-acids of S6 sequence (<sup>229</sup>AKRRRLSSLRA<sup>239</sup>) according to the manufacturer's instructions (Upstate Cell Signaling, Lake Placid, NY), as described previously [22].

#### Western blot analysis

Cells were lysed in p70S6K activity lysis buffer as described above, and 20 µg of proteins were separated by SDS-PAGE and transferred to nitrocellulose membrane (Whatman, Dassel, Germany). Membranes were saturated with TBST-BSA (Tris Buffer Saline with 0.1% Tween 20, 3% Bovine Serum Albumin) and were probed with rabbit primary antibodies (1:1000): anti-phospho-p70S6K1 [T421/S424], anti phospho-p70S6K1 [T389], anti-phospho-rpS6 [S235/Ser236] or anti-phospho-mTOR [S2448]. Membranes were also probed to detect the total corresponding protein, and re-probed with an anti-ERK1, 2 antibody to monitor gel loading. Western blots were revealed with Supersignal West Pico Chemiluminescent Substrate or Supersignal Femto Chemiluminescent Substrate from Pierce Biotechnology (Rockford, IL). Films were scanned and the optical density of the signals was measured with ImageMaster 1D Elite version 4 software (Amersham Biosciences).

#### Phosphatidyl-inositide (3, 4, 5) trisphosphate assay by protein-lipid overlay

To estimate PI3K activity, PIP3 was quantified in Sertoli cells as described recently [30]. Briefly, PtdIns-rich fractions were extracted with chloroform/methanol, selectively isolated on neomycin beads, eluted with methanol–chloroform–HCl. Purified PtdIns were spotted onto a nitrocellulose membrane, blotted overnight at 4°C with the PH domain of the general receptor for phosphoinositide 1 (GRP1) fused to GFP. Membranes were then probed with a rabbit anti-GFP polyclonal antibody coupled to HRP, the chemiluminescence was detected with the ECL reagent (Amersham) and acquired with the Image Reader LAS-1000 charge-coupled device (CCD) camera (Fujifilm). Densitometry was measured with the AIDA software with local background subtraction.

#### cAMP quantification by ELISA

One million Sertoli cells was seeded in 24-well plates and stimulated with increasing doses of FSH ranging from 10 to 400 ng/ml, during 15 min, or with 100 ng/ml FSH in a time-course ranging from 5 to 120 min. The experiments have been performed in the presence of 100 µM phosphodiesterase inhibitor IBMX (isobutylmethylxanthine; Sigma Chemicals). Results were expressed in percent of the response to 10 µM of the adenylate cyclase activator Forskolin (Sigma Chemicals) for 15 min in the dose-response experiments or for 60 min for the time-course experiments. cAMP was measured directly in the cell media using a competition assay, according to the manufacturer's instructions (Biomedical Technologies, Stoughton, MA).

#### Binding assays and Scatchard analysis

One million Sertoli cells stimulated with porcine FSH was plated in poly-lysine-coated 24-well plates. Binding assays were performed the next day, essentially as described in [27]. Both the number of cells and the global protein content were found as equivalent in the two cell types, at the end of the experiment. The affinity constant was calculated from the binding data using Scatchard analysis.

#### Mathematical model

A single model of the regulation of p70S6K was built and parameterized in differentiating and proliferating cells. The differences that were observed experimentally in some reactions between the two stages were expressed by considering two reactions, one being catalyzed, and one being

spontaneous. In the cell stage where the considered reaction was catalyzed, the parameter corresponding to the spontaneous reaction was null and, conversely, in the cell stage where the reaction was spontaneous, the parameter corresponding to the catalyzed reaction was null. The model can be represented by a non-linear system of 13 ODEs (ordinary differential equations, see below) with 24 unknown parameters (Table 1), corresponding to kinetic rates and total amounts of the different molecules involved.

Parameters were estimated as follows: the amount of FSH was considered as constant throughout the 60 min of experiment, regardless its capture by the FSH-R. In addition, we considered that no FSH receptor was lost during

the experiment. Consequently, the total quantity of receptors ( $R_T$ ), which is the quantity of free receptors ( $R$ ) plus the quantity of receptors complexed with the hormone ( $RH$ ), was considered as constant (see conservation law CL1). Similarly, the total amount of nucleoside  $A_T$  (ATP and cAMP), of phosphatidyl-inositol  $P_T$  (PIP2 and PIP3), of p70S6K  $S_T$  (non-phosphorylated p70S6K and the three different phosphorylated forms) and rpS6: rpS6<sub>T</sub> (phosphorylated PrpS6 and non-phosphorylated rpS6) were considered as constant, as further supported by our experimental data for p70S6K and rpS6. This can be formalized by the following 5 conservation laws, which leads to removing 5 ODEs:

**Table 1** Parameters, constants, and variables in the dynamic model of p70S6K regulation by FSH

Definition	Name in ODEs	Value in proliferating cells (5 dpp)	Value in differentiating cells (19 dpp)
Receptor binding association rate	$k_1$	0.01 $\mu\text{g}/\text{nmol}/\text{s}$	
Activation of cAMP PKA pathway	$k_2$	0.7 $\mu\text{g}/\text{nmol}/\text{s}$	
Activation of PI3K PIP mTOR pathway	$k_3$	5 $\mu\text{g}/\text{nmol}/\text{s}$	0.25 $\mu\text{g}/\text{nmol}/\text{s}$
Desactivation of cAMP PKA pathway	$k_4$	0.025 $\mu\text{g}/\text{nmol}/\text{s}$	
Phosphorylation rate (S6KP1 $\rightarrow$ S6KP1P2)	$k_5$	0.1 $\mu\text{g}/\text{nmol}/\text{s}$	0.00001 $\mu\text{g}/\text{nmol}/\text{s}$
Phosphorylation rate (S6KP2 $\rightarrow$ S6KP1P2)	$k_6$	0.001 $\mu\text{g}/\text{nmol}/\text{s}$	
Phosphorylation rate (S6K $\rightarrow$ S6KP1)	$k_7$	0.018 $\mu\text{g}/\text{nmol}/\text{s}$	0.00001 $\mu\text{g}/\text{nmol}/\text{s}$
Phosphorylation rate (S6K $\rightarrow$ S6KP2)	$k_8$	0.0001 $\mu\text{g}/\text{nmol}/\text{s}$	
Dephosphorylation rate (S6KP1P2 $\rightarrow$ S6KP1)	$k_9$	0.1 $\mu\text{g}/\text{nmol}/\text{s}$	
Dephosphorylation rate (S6KP2 $\rightarrow$ S6K)	$k_{10}$	7 $\mu\text{g}/\text{nmol}/\text{s}$	
Dephosphorylation rate (S6KP1P2 $\rightarrow$ S6KP2)	$k_{11}$	0.02/s	
Dephosphorylation rate (S6KP1 $\rightarrow$ S6K)	$k_{12}$	0.000001/s	
Phosphorylation rate (rpS6 $\rightarrow$ PrpS6)	$k_{13}$	4 $\mu\text{g}/\text{nmol}/\text{s}$	
Total amount of FSH	$k_{14}$	5.58 nmol/ $\mu\text{g}$	
Total amount of receptor	$k_{15}$	0.0558 nmol/ $\mu\text{g}$	0.0744 nmol/ $\mu\text{g}$
Total amount of ATP	$k_{16}$	1 nmol/ $\mu\text{g}$	4.37 nmol/ $\mu\text{g}$
Total amount of PIP	$k_{17}$	0.5 nmol/ $\mu\text{g}$	0.75 nmol/ $\mu\text{g}$
Total amount of p70S6K	$k_{18}$	0.4 nmol/ $\mu\text{g}$	4 nmol/ $\mu\text{g}$
Total amount of rpS6	$k_{19}$	0.8 nmol/ $\mu\text{g}$	
1% cAMP value	$k_{20}$	0.00336 ( $10^6$ cells)/ $\mu\text{g}$	0.00263 ( $10^6$ cells)/ $\mu\text{g}$
1% insulin value	$k_{21}$	0.0005 ( $10^6$ cells)/ $\mu\text{g}$	
Dephosphorylation rate (PrpS6 $\rightarrow$ rpS6)	$k_{22}$	0.22/s	0.18/s
Receptor affinity constant	$K_d_{RH}$	10.64664 nmol/ $\mu\text{g}$	6.88344 nmol/ $\mu\text{g}$
PIP2 $\leftrightarrow$ PIP3 reaction constant ( $k_{23}/k_3$ )	$K_d_{PIP}$	0.1407 nmol/ $\mu\text{g}$	13.420304 nmol/ $\mu\text{g}$
Dephosphorylation rate (S6KP2 $\rightarrow$ S6K) at equilibrium	$k_{10}/k_8$	70,000	
Phosphorylation rate (S6KP2 $\rightarrow$ S6KP1P2) at equilibrium	$k_6/k_{11}$	0.05	
Dephosphorylation rate (S6KP1P2 $\rightarrow$ S6KP1) at equilibrium	$k_9/k_5$	1	10,000
Phosphorylation rate (S6K $\rightarrow$ S6KP1) at equilibrium	$k_7/k_{12}$	18,000	10

Common parameters to both stages are indicated for 5 dpp only. At both stages, for insulin  $k_9$  and  $k_{10} = 0.00001$ , hence  $k_9/k_5 = 0.0001$  and  $k_{10}/k_8 = 0.1$

p70S6K[T389] is noted by S6KP1, p70S6K[T421/S424] is noted S6KP2 and p70S6K[T389/T421/S424] is noted S6KP1P2

$\text{s}^{-1}$  corresponds to the time needed by S6KP1P2 to be converted to S6KP2 in 1  $\mu\text{g}$  protein



$$R_T = R + RH \quad (\text{CL1})$$

$$A_T = \text{ATP} + \text{cAMP} \quad (\text{CL2})$$

$$P_T = \text{PIP2} + \text{PIP3} \quad (\text{CL3})$$

$$\begin{aligned} S_T &= \text{p70S6K} + \text{Pp70S6K}[\text{T389}] \\ &+ \text{Pp70S6K}[\text{T421/S424}] \\ &+ \text{Pp70S6K}[\text{T389/T421/S424}] \end{aligned} \quad (\text{CL4})$$

$$\text{rpS6}_T = \text{rpS6} + \text{PrpS6}[\text{S235/236}] \quad (\text{CL5})$$

The quantities of R, ATP, PIP2, p70S6K and rpS6 can then be computed from the differential equations given below. For the sake of clarity, null quantities were removed from the equations when relevant, leading to two expressions of the same ODE at the two stages: parameters corresponding to a unique reaction with different regulations at the two stages (catalyzed or spontaneous) are noted p (for proliferation stage) and d (for differentiation stage).

For both stages:

$$dRH/dt = k1 \cdot R \cdot H_T - k1 \cdot K_{dRH} \cdot RH \quad (\text{ODE1})$$

$$dcAMP/dt = k2 \cdot \text{ATP} \cdot RH - k4 \cdot \text{cAMP} \quad (\text{ODE2})$$

For the proliferation stage:

$$d\text{PIP3}/dt = k3p \cdot \text{PIP2} \cdot RH - k3p \cdot K_{d\text{PIP}} \cdot \text{PIP3} \quad (\text{ODE3p})$$

$$\begin{aligned} d\text{Pp70S6K}[\text{T389}]/dt &= k7p \cdot \text{PIP3} \cdot \text{p70S6K} \\ &+ k9 \cdot \text{cAMP} \cdot \text{Pp70S6K}[\text{T389/T421/S424}] \\ &- k5p \cdot \text{PIP3} \cdot \text{Pp70S6K}[\text{T389}] - k12 \cdot \text{Pp70S6K}[\text{T389}] \end{aligned} \quad (\text{ODE4p})$$

$$\begin{aligned} d\text{Pp70S6K}[\text{T421/S424}]/dt &= k8p \cdot \text{PIP3} \cdot \text{p70S6K} + k11 \cdot \text{Pp70S6K}[\text{T389/T421/S424}] \\ &- k6 \cdot \text{PIP3} \cdot \text{Pp70S6K}[\text{T421/S424}] \\ &- k10 \cdot \text{cAMP} \cdot \text{Pp70S6K}[\text{T421/S424}] \end{aligned} \quad (\text{ODE5p})$$

$$\begin{aligned} d\text{Pp70S6K}[\text{T389/T421/S424}]/dt &= k5p \cdot \text{PIP3} \cdot \text{Pp70S6K}[\text{T389}] \\ &+ k6 \cdot \text{PIP3} \cdot \text{Pp70S6K}[\text{T421/S424}] \\ &- k9 \cdot \text{cAMP} \cdot \text{Pp70S6K}[\text{T389/T421/S424}] \\ &- k11 \cdot \text{Pp70S6K}[\text{T389/T421/S424}] \end{aligned} \quad (\text{ODE6p})$$

$$\begin{aligned} d\text{rpS6}[\text{S235/236}]/dt &= k13p \cdot \text{rpS6} \cdot \text{Pp70S6K}[\text{T389/T421/S424}] \\ &- k22 \cdot \text{PrpS6}[\text{S235/236}] \end{aligned} \quad (\text{ODE7p})$$

For the differentiation stage:

$$d\text{PIP3}/dt = k3d \cdot \text{PIP2} - k3d \cdot K_{d\text{PIP}} \cdot \text{PIP3} \quad (\text{ODE3d})$$

$$\begin{aligned} d\text{Pp70S6K}[\text{T389}]/dt &= k7d \cdot \text{PIP3} \cdot \text{p70S6K} \\ &+ k9 \cdot \text{cAMP} \cdot \text{Pp70S6K}[\text{T389/T421/S424}] \\ &- k5d \cdot \text{PIP3} \cdot \text{Pp70S6K}[\text{T389}] - k12 \cdot \text{Pp70S6K}[\text{T389}] \end{aligned} \quad (\text{ODE4d})$$

$$\begin{aligned} d\text{Pp70S6K}[\text{T421/S424}]/dt &= k8d \cdot \text{p70S6K} + k11 \cdot \text{Pp70S6K}[\text{T389/T421/S424}] \\ &- k6 \cdot \text{PIP3} \cdot \text{Pp70S6K}[\text{T421/S424}] \\ &- k10 \cdot \text{cAMP} \cdot \text{Pp70S6K}[\text{T421/S424}] \end{aligned} \quad (\text{ODE5d})$$

$$\begin{aligned} d\text{Pp70S6K}[\text{T389/T421/S424}]/dt &= k5d \cdot \text{Pp70S6K}[\text{T389}] \\ &+ k6 \cdot \text{PIP3} \cdot \text{Pp70S6K}[\text{T421/S424}] \\ &- k9 \cdot \text{cAMP} \cdot \text{Pp70S6K}[\text{T389/T421/S424}] \\ &- k11 \cdot \text{Pp70S6K}[\text{T389/T421/S424}] \end{aligned} \quad (\text{ODE6d})$$

$$\begin{aligned} d\text{rpS6}[\text{S235/236}]/dt &= k13d \cdot \text{rpS6} \cdot \text{Pp70S6K}[\text{T389}] \\ &- k22 \cdot \text{PrpS6}[\text{S235/236}] \end{aligned} \quad (\text{ODE7d})$$

#### Data and units

Quantities were expressed in nmol/ $\mu\text{g}$  of proteins, as deduced from the experimental data presented herein, or as % of the response to insulin, arbitrarily fixed at  $5 \times 10^{-7}$  mol/g in proliferating cells and at  $10^{-6}$  mol/g in differentiating cells, based on a previous PIP3 quantification in CHO cells [31]. The quantities measured experimentally were PIP3, cAMP, and PrpS6, which correspond to a single node in the model, p70S6K phosphorylated on Thr389, and p70S6K phosphorylated on T421/S424. p70S6K phosphorylated on Thr389 is in fact the sum of Pp70S6K[T389] plus Pp70S6K[T389/T421/S424], while the amount of p70S6K phosphorylated on T421/S424 is the sum of Pp70S6K[T421/S424] and Pp70S6K[T389/T421/S424].

#### Program, software, fitting algorithm

The ODE model was simulated with the SCILAB software (<http://www.scilab.org>). Parameter values were adjusted by analyzing the data back and forth with the conjugate gradient method implemented in datafit intern Scilab procedure. The minimized normalized error is the distance between the data and the corresponding quantity simulated in the model.

The p70S6K developmental regulation by FSH has been graphically represented with the CellDesigner graphic interface (<http://celldesigner.org>).

## Statistics

Measurements were performed in duplicates or triplicates and results were expressed as mean values  $\pm$  SEM of at least 3 independent experiments. Statistical analysis of the biological data was performed using one-way ANOVA (Bonferroni's multiple comparison tests) to compare samples or two-way ANOVA to compare entire curves (GraphPad PRISM Software, San Diego, CA). Values with  $P < 0.05$  were considered as significantly different.

## Results

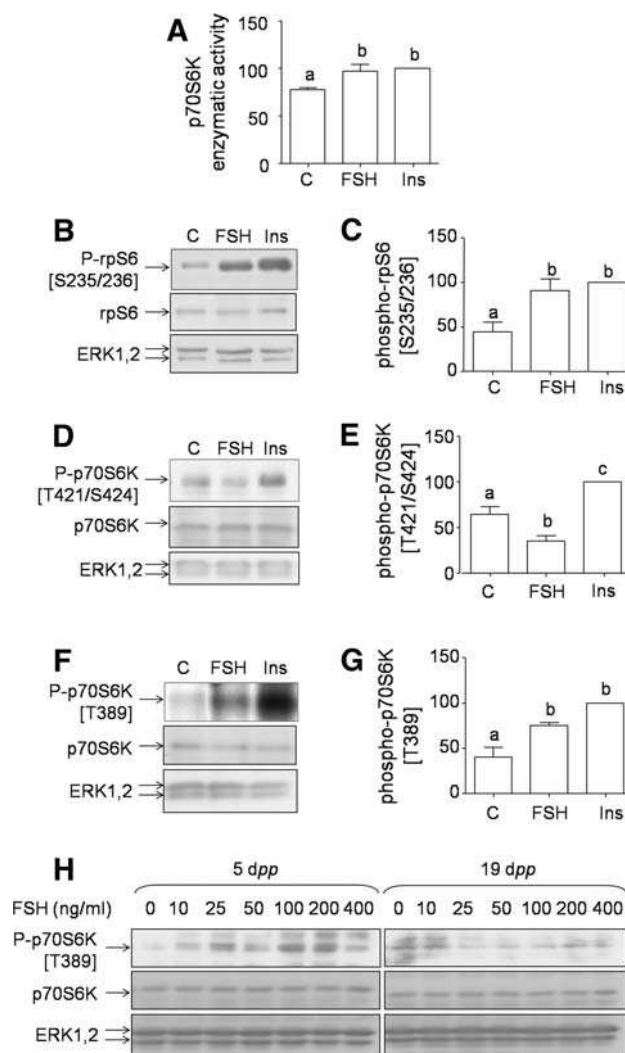
FSH or insulin stimulates p70 S6 kinase enzymatic activity and alters its phosphorylation pattern in primary Sertoli cells from 5 day-old rats

We first sought to determine whether FSH stimulated the activity of p70S6K in proliferating Sertoli cells from 5 day-old rats, as it does in differentiating Sertoli cells from 19 day-old rats [22]. After 1 h of exposure to FSH, p70S6K enzymatic activity increased slightly but significantly (Fig. 1a). Consistently, the phosphorylation of rpS6 also increased after 60 min of FSH stimulation, as shown by western blot analysis (Fig. 1b, c), confirming that p70S6K was activated by FSH. Likewise, both p70S6K activity and rpS6 phosphorylation were stimulated by insulin.

Then, we analyzed p70S6K pattern of phosphorylation on T389 and T421/S424, both sites supposedly required for p70S6K activation [10]. The T421/S424 site was dephosphorylated following FSH stimulation, whereas it was phosphorylated in response to insulin (Fig. 1d, e). In contrast to differentiating Sertoli cells, in proliferating Sertoli cells, T389 phosphorylation was enhanced by FSH (Fig. 1f, g) in a dose-responsive manner (Fig. 1h). However, the level of T389 phosphorylation in response to FSH was lower than the one observed in response to insulin (Fig. 1g).

The hormone-bound FSH-R generates PIP3 and cAMP with a different efficacy in 5 and 19 day-old rat primary Sertoli cells

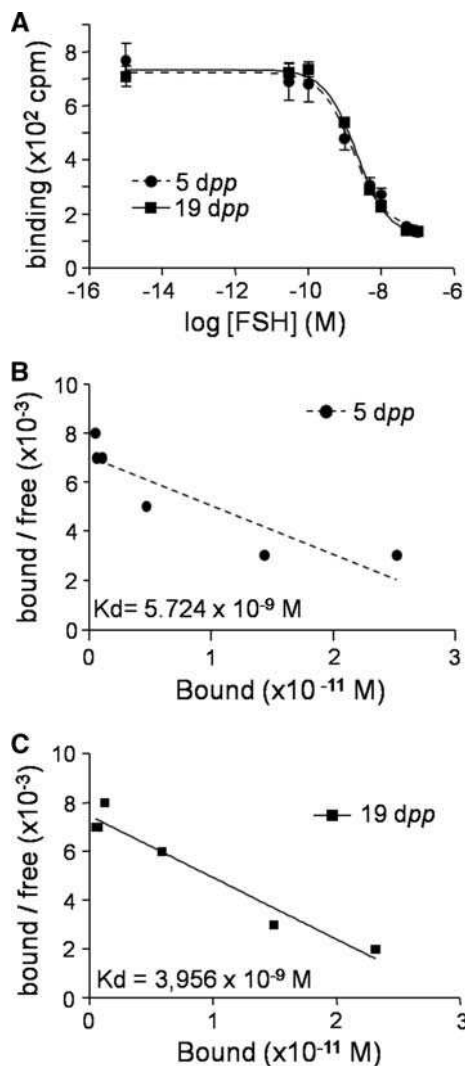
One characteristic feature of GPCRs when compared to tyrosine kinase receptors is their ability to couple to G proteins. Our previous data suggest that the FSH-R differentially couples to both Gs and Gi in proliferating Sertoli cells and only to Gs in differentiating Sertoli cells [32]. Therefore, we hypothesized that the FSH-R activation might induce the production of second messengers such as cAMP and PIP3 with different efficacy in proliferating and



**Fig. 1** p70S6K is activated by FSH and by insulin in 5-day-old rat primary Sertoli cells. Sertoli cells were treated for 60 min with 100 ng/ml of FSH or insulin (Ins) or were left untreated C. **a** Mean  $\pm$  SEM of p70S6K activity obtained at least from 3 independent experiments each performed in triplicates. **b** Representative western blot of rpS6 phosphorylation on S235/236, reprobated with an anti-total-rpS6 antibody. **c** Mean  $\pm$  SEM of rpS6 phosphorylation obtained from more than 3 independent experiments each performed in duplicates. **d**, **f** Representative western blots of p70S6K phosphorylation on T421/S424 and T389 residues, respectively, reprobated with an anti-total-p70S6K antibody. **h** Representative western blot of p70S6K phosphorylation on T389 induced by increasing doses of FSH in proliferating Sertoli cells (5 dpp) and in differentiating Sertoli cells (19 dpp) reprobated with an anti-total-p70S6K antibody. In **b**, **d**, **f**, **h**, gel loading was monitored by reprobating the membranes with an anti-ERK1, 2 antibody. **e**, **g** Means  $\pm$  SEM of phospho-T421/S424 and phospho-T389 signals respectively, obtained from 3 independent experiments each performed in duplicates. All the results were expressed as percent of the response to insulin. Shared *superscripts* indicate no significant difference, while different *superscripts* indicate significant differences, at the  $P < 0.05$  level

in differentiating Sertoli cells, to account for the developmental differences in signaling to p70S6K. To compare cAMP and PIP3 production at the two developmental

stages, we first had to make sure that second messenger production was measured at similar receptor density in both cell types. By binding assays performed with  $^{125}\text{I}$ -pFSH, the displacement curves obtained with proliferating as well as differentiating Sertoli cells matched (Fig. 2a). The  $K_d$  values were in similar range, i.e.,  $5.7 \times 10^{-9}$  M and  $4 \times 10^{-9}$  M in Sertoli cells from 5- and 19-day-old rats, respectively, as inferred from Scatchard plot analysis (Fig. 2b, c). These results show that the affinity of FSH for its receptor and the receptor density do not vary significantly in both developmental stages ( $0.04$  vs  $0.03 \times 10^{-9}$  M, respectively), in agreement with previous

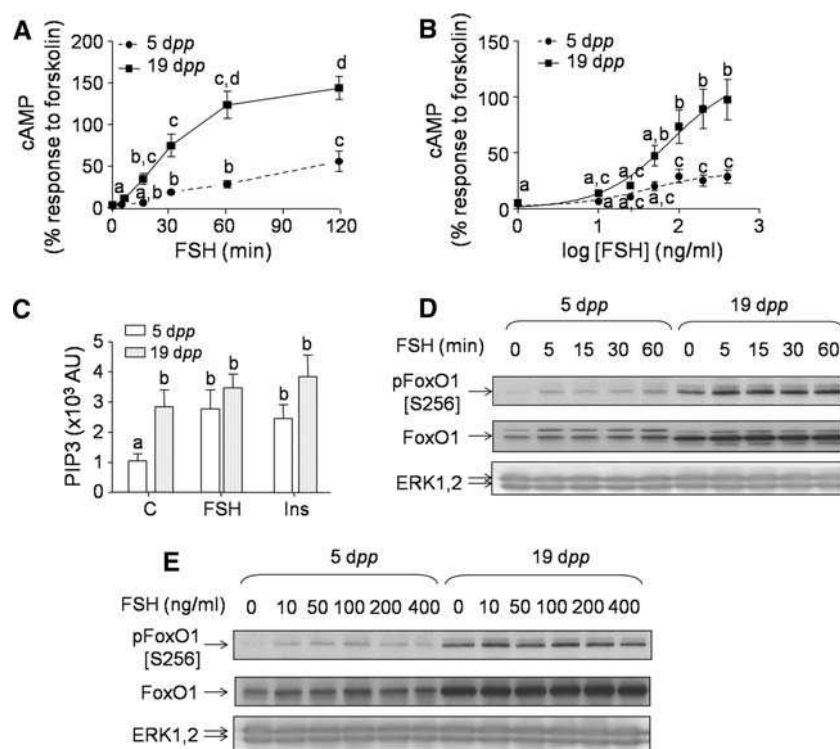


**Fig. 2** Comparison of the binding properties of the FSH-R in proliferating and in differentiating Sertoli cells. Binding parameters were estimated in proliferating (*circles*) and differentiating (*squares*) Sertoli cells incubated for 1 h at 37°C in the presence of  $^{125}\text{I}$ -pFSH and increasing concentrations of unlabeled pFSH ( $10^{-14}$  to  $10^{-7}$  M). Triplicates were done for each concentration of pFSH. **a** The displacement curve represents the mean of three different experiments  $\pm$  SEM. **b,c** The affinity constant was calculated from the binding data using Scatchard analysis

data obtained from testis homogenates [33]. Therefore, second messenger production can be directly compared.

We previously reported that only in differentiating Sertoli cells did FSH cause a time-dependent increase in cAMP levels, whereas in proliferating cells, cAMP was constitutively elevated, similarly to the cAMP maximum levels reached in FSH-stimulated Sertoli cells from pre-pubertal rats [32]. Since phosphodiesterases are expressed at much higher levels in differentiating Sertoli cells [34], a phosphodiesterase inhibitor, IBMX, was added to cell cultures 15 min prior to FSH stimulation, in order to compare cAMP production in both cellular backgrounds more accurately. Clearly, the efficacy of FSH to produce cAMP was much higher in differentiating than in proliferating Sertoli cells, as can be observed in time-course (Fig. 3a) as well as in FSH dose-response experiments (Fig. 3b). Forskolin stimulation led to comparable cAMP production in both cell types ( $4.5 \pm 0.53$  and  $3.79 \pm 0.19$  pmoles/ $10^6$  cells in proliferating and in differentiating Sertoli cells, respectively; data not shown). Hence, it was considered as an internal control, so that cAMP production could be directly compared at both stages.

Then, PIP3 production was directly quantified by a protein-lipid overlay. In 5-day-old rat cells, PIP3 was significantly increased by FSH stimulation whereas it was not in 19-day-old rat cells (Fig. 3c). This observation suggests an FSH-dependent activation of PI3K in proliferating cells and is consistent with PI3K constitutive activation in differentiating cells. To estimate the efficacy of FSH to produce PIP3, we first assayed whether phosphorylation of the phosphoinositide-dependent transcription factor Foxo1 would correlate to our data. Time-course stimulation with FSH indicated that Foxo1 gets inducibly phosphorylated in proliferating Sertoli cells, whereas it was constitutively phosphorylated in differentiating Sertoli cells (Fig. 3d). Since this result was concordant with the PIP3 direct quantification, the efficacy of FSH to stimulate PIP3-dependent events was next estimated by assaying Foxo1 phosphorylation in dose-response experiments. As expected, we found that increasing doses of FSH nicely enhanced Foxo1 phosphorylation in proliferating cells, but not in differentiating cells where Foxo1 phosphorylation was constitutively high, and could not be further enhanced by FSH (Fig. 3e). Although insulin stimulation for 15 min did not lead to a significant increase in PIP3 production, it enhanced Foxo1 phosphorylation after 1 h. Therefore, PI3K activity is constitutive in differentiating cells, and still can be activated further at both stages by insulin, but only in proliferating Sertoli cells by FSH. In conclusion, the efficacy of the FSH-R coupling to cAMP was more efficient in differentiating Sertoli cells, whereas PIP3-generating mechanisms were more potent in proliferating Sertoli cells stimulated by FSH.



**Fig. 3** Developmental regulation of FSH efficacy to induce second messenger production. **a** Mean  $\pm$  SEM of cAMP production obtained from 3 independent experiments each performed in triplicate, as a function of FSH stimulation time, in proliferating or in differentiating Sertoli cells as indicated. **b** Mean  $\pm$  SEM of cAMP production obtained from 3 independent experiments each performed in triplicate, as a function of FSH concentration, in proliferating or in differentiating Sertoli cells as indicated. Results are expressed as % of the forskolin-induced cAMP response, taken as an internal control. **c** Sertoli cells isolated from 5 or 19 day-old rats were stimulated with FSH for 15 min or were left untreated. Results shown are the mean values  $\pm$  SEM of PIP3 production, expressed in arbitrary units (AU)

obtained from more than 3 independent fat ligand blot experiments. **d** Representative western blot of Foxo1 phosphorylation on S256 as a function of FSH stimulation time, in proliferating (5 dpp) or in differentiating Sertoli cells (19 dpp). **e** Representative western blot of Foxo1 phosphorylation as a function of FSH concentration, in proliferating (5 dpp) or in differentiating Sertoli cells (19 dpp) reprobated with an anti-total-Foxo1 antibody. In **d** and **e**, gel loading was monitored by reprobating the membranes with an anti-ERK1, 2 antibody. Shared *superscripts* indicate no significant difference, while different *superscripts* indicate significant differences, at the  $P < 0.05$  level

#### Dynamic mathematical simulation of p70S6K regulation by FSH in control conditions

To gain a clearer view of the active forms of p70S6K in response to FSH in the development of Sertoli cells, we achieved dynamic mathematical modelling. All the molecular species studied here were listed as follows: the FSH receptor, the hormone, the hormone-receptor complex, PI3K, activated PI3K, PIP2, PIP3, mTOR, activated mTOR, ATP, cAMP, PKA, the cAMP-PKA complex, rpS6, PrpS6, p70S6K, Pp70S6K[T389], Pp70S6K[T421/S424], and Pp70S6K[T389/T421/S424]. Table 1 lists all the parameters determined in this study, in proliferating and in differentiating Sertoli cells. Quantification of the molecular species and Kd values have been determined as explained in “Materials and Methods”. Parameters such as  $k_3$  and  $k_5$  vary, as expected from modification of the model structure: for  $k_3$ , PIP3 is no longer increased in differentiating cells. As for  $k_5$ , the p70S6K regulation is uncoupled

to the PIP3-dependent component in differentiated Sertoli cells.

In our model, we have four different forms for p70S6K: the non-phosphorylated form, two mono-phosphorylated forms (either T389 or T421/S424) and the form with both sites phosphorylated. Consequently, there are 12 possible transitions. Among those, four were eliminated: the double phosphorylation, direct and reverse, and the transitions between the two mono-phosphorylated forms, because dephosphorylation on the T421/S424 site requires FSH stimulation, whereas the dephosphorylation on the T389 site is spontaneous. Consequently, none of these four transitions can be made in a single step, and are in fact the sum of two of the retained transitions.

Since this study focuses on p70S6K regulation nodes, parts of the model were simplified. Based on previous data obtained in differentiating Sertoli cells where inhibiting PI3K or mTOR led to similar consequences on p70S6K regulation [22], the model was reduced by aggregating the

PI3K/PIP/mTOR cascade in a single reaction, considered as a black box: transformation of PIP2 (inactive) in PIP3 (active). Again, on the basis of our previous report [22], the same was done for the cAMP/PKA pathway, also considered as one reaction: the formation of cAMP from ATP.

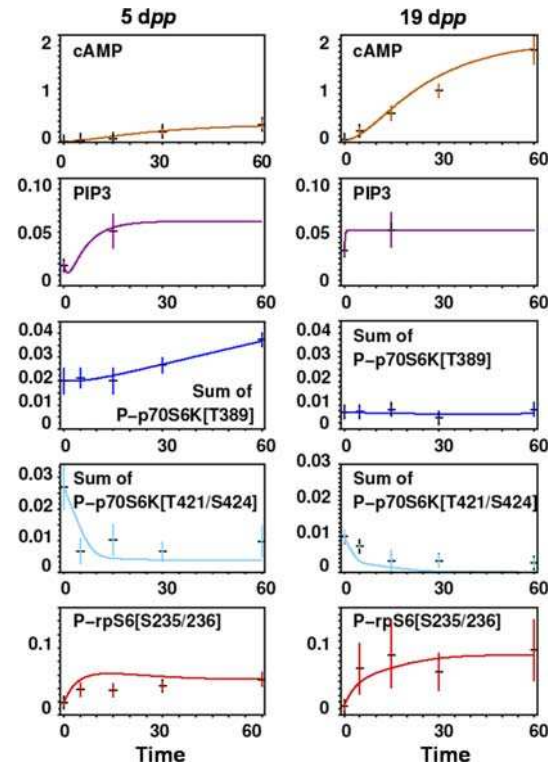
From experimental data, regulations of four of the reactions were differently regulated in both stages. For example, in differentiating cells, PIP3 reached its steady state in the absence of hormone, and the reaction could be considered as spontaneous. Thereof, the regulation of the conversion rate of PIP2 to PIP3 upon FSH binding was not taken into account for this stage, and only the spontaneous reaction was considered. In contrast, FSH binding to the FSH-R enhanced PIP3 level in proliferating cells and the model indicates that this PIP3 amount was produced more slowly due to low hormone-receptor complex formation. At this cell stage, the parameter representing the spontaneous conversion rate of PIP2 to PIP3 was considered as null.

We obtained two dynamic simulations of the model based on optimized parameters, which fitted the biological data obtained at both stages (Fig. 4). Of note, all but one kinetic parameters corresponding to reactions having the same regulation in both stages are equal (Table 1). This shows that the differences in regulations, formalized as differences in parameters, in initial conditions and total amounts are sufficient to explain the observed differences in p70S6K regulation.

The model provided the intrinsic concentrations of Pp70S6K[T389], Pp70S6K[T421/S424], and Pp70S6K[T389/T421/S424], which are not experimentally observable (Fig. 5a, b). At both stages, Pp70S6K[T421/S424] was initially the major form, to be transformed into p70S6K, Pp70S6K[T389/T421/S424] and indirectly into Pp70S6K[T389]. Whereas the clearance of Pp70S6K[T421/S424] was fast, the evolution of Pp70S6K[T389/T421/S424] and Pp70S6K[T389] remained slower and sustained. When considering Pp70S6K[T389/T421/S424], it remained constant only in proliferating cells, where an increase in Pp70S6K[T389] compensated the decrease in Pp70S6K[T421/S424].

FSH modulates PI3K and mTOR-dependent phosphorylation of p70S6K T389 and T421/S424 residues

To unravel the relationships linking the PI3K/mTOR pathway to p70S6K activation, we analyzed the effect of their inhibition on the p70S6K phosphorylation pattern. LY294002, a PI3K inhibitor, and rapamycin, an mTOR inhibitor, abolished the basal phosphorylation of the T421/S424 site (Fig. 6a, b). Hence, the FSH-induced dephosphorylation was not visible any more. Similarly,

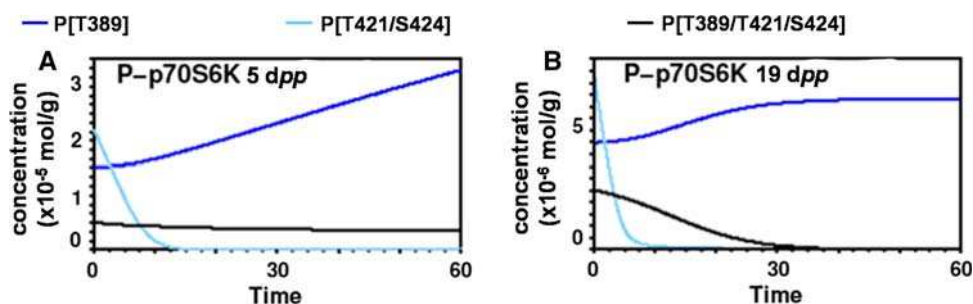


**Fig. 4** Simulation of FSH-induced p70S6K regulation. Computational simulations (plain lines) compared with experimental data (black crosses  $\pm$  SEM) of FSH-R-mediated signaling in proliferating (left) and differentiating (right) Sertoli cells. The concentration of molecules in mmol/g is expressed as a function of time in minutes. “cAMP” stands for the aggregate cAMP-PKA module, “PIP3” stands for the aggregate PI3K/PIP3/mTOR module

both LY294002 and rapamycin considerably reduced basal and FSH-induced T389 phosphorylation (Fig. 6c, d). Therefore, the PI3K-mTOR pathway seems to be involved in the basal phosphorylation of both sites and in the FSH-inducible phosphorylation of T389. As expected, LY294002 and rapamycin inhibited insulin-stimulated phosphorylation of both T421/S424 and T389 sites.

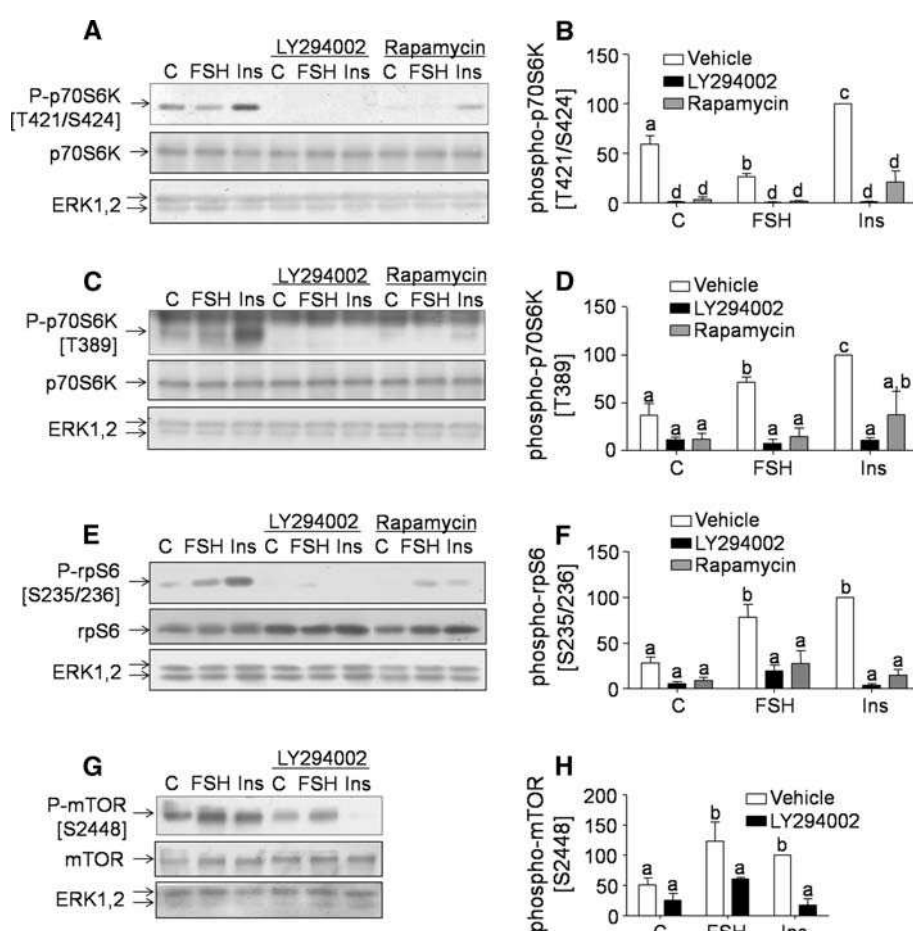
To define whether the PI3K-mTOR pathway impacted on p70S6K activity, we measured rpS6 phosphorylation following cell pre-treatment with LY294002 or with rapamycin. Those inhibitors strongly decreased the phosphorylation of rpS6, which reveals inhibition of p70S6K activity, in response to FSH, and even more to insulin (Fig. 6e, f).

Consistent with the involvement of PI3K and mTOR in FSH-dependent signaling, FSH stimulated mTOR phosphorylation, in a PI3K-dependent manner (Fig. 6g, h). Surprisingly, mTOR was more potently activated by FSH-dependent signaling than by insulin-dependent signaling. However, mTOR phosphorylation was more sensitive to PI3K inhibitors in cells stimulated by insulin than following FSH stimulation, suggesting that other upstream events,



**Fig. 5** Intrinsic Pp70S6K quantities predicted by computational simulations, respectively, Pp70S6K phosphorylated on [T389] only, Pp70S6K phosphorylated on [T421/S424] only, and Pp70S6K

phosphorylated on [T389/T421/S424], as indicated. Simulations in **a** proliferating and **b** differentiating Sertoli cells are shown



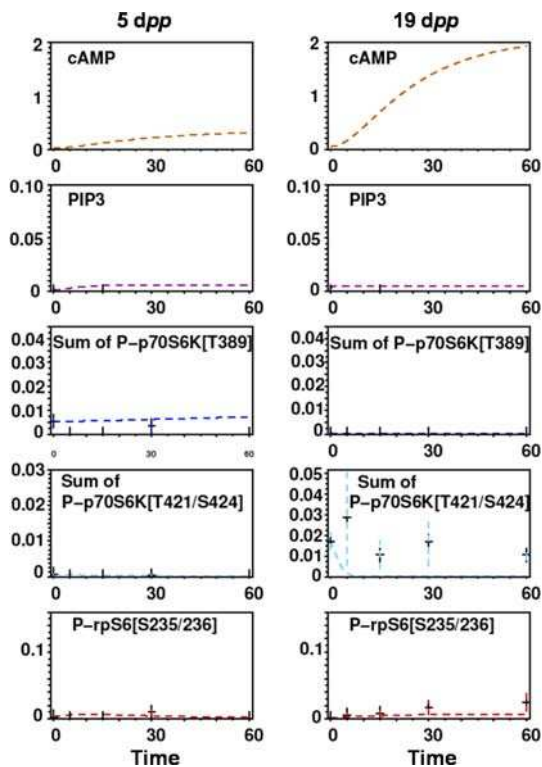
**Fig. 6** FSH enhances PI3K and mTOR-dependent phosphorylation of p70S6K T389 and decreases PI3K and mTOR-dependent tonic phosphorylation of T421/S424. Sertoli cells were pretreated with the PI3K inhibitor LY294002 (50 mM, 10 min) or with the mTOR inhibitor rapamycin (10 nM, 60 min) or were left untreated before being stimulated for 30 min with FSH or insulin. **a,c** Representative western blots of p70S6K phosphorylation on T421/S424 and T389 residues, respectively, reprobred with an anti-total-p70S6K antibody. **b,d** Mean  $\pm$  SEM of phospho-T421/S424 and phospho-T389 signals respectively, each obtained from 3 independent experiments performed in duplicates. **e** Representative western blot of rpS6

phosphorylation on S235/236 reprobred with an anti-total-rpS6 antibody. **f** Mean  $\pm$  SEM of phospho-rpS6 signals obtained from 3 independent experiments each performed in duplicates. **g** Representative western blot of mTOR phosphorylation on S2448 reprobred with an anti-total-mTOR antibody. In **a, c, e, and g**, gel loading was monitored by reprobred the membranes with an anti-ERK1, 2 antibody. **h** Mean  $\pm$  SEM of phospho-mTOR signals obtained from 3 independent experiments each performed in duplicates. Results are expressed as percent of the response to insulin. Shared *superscripts* indicate no significant difference, while different *superscripts* indicate significant differences, at the  $P < 0.05$  level

unrelated to PI3K, are required for FSH to stimulate mTOR.

We kept the same sets of parameters as for the control conditions to simulate pharmacological perturbations, except for the total amounts of those molecules that were affected by the treatment, namely PI3K and mTOR, which were not represented in our models. In both cases, the consequence of the inhibition of these molecules was a drastic decrease of the available amount of the active form, equivalent to a drastic decrease in the amount of the aggregate variable PIP3, which was modeled by reducing 10-fold the initial and total PIP3 amounts at both stages (Table 1 and Fig. 7). The simulation confirmed that in active PI3K-depleted conditions, Pp70S6K[T389] decreased at both stages, and that Pp70S6K[T421/S424] also decreased in proliferating cells (Fig. 8a, b, to be compared with the simulation in control conditions in Fig. 5a, b). In proliferating cells, loss of Pp70S6K[T389/T421/S424] was concomitant to the loss of rpS6 phosphorylation, while in differentiating cells, only the loss of Pp70S6K[T389] was seen.

Importantly, the fact that our model still fits the observations in perturbed conditions (Fig. 7) without modifying



**Fig. 7** Simulation of the PI3K/PIP3 component of FSH-induced p70S6K regulation. Computational simulations (plain lines) compared with experimental data (black crosses  $\pm$  SEM) of FSH-R-mediated signaling in the presence of PI3K inhibitor, in proliferating (left) and differentiating (right) Sertoli cells. The concentrations of molecules are expressed as a function of time, as above

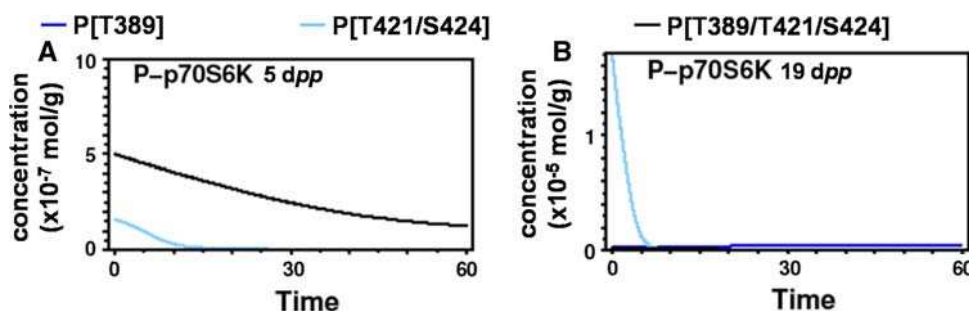
the parameters that were obtained in normal conditions (Table 1), further confirms its validity.

p70S6K activation is constitutively restrained by PKA in proliferating Sertoli cells

PKA stimulates p70S6K activity by dephosphorylating the T421/S424 site in differentiating Sertoli cells [22]. Since FSH stimulation also led to dephosphorylation of these residues in proliferating Sertoli cells, we sought to define the role of PKA in these cells. The enzymatic activity of p70S6K was measured in proliferating Sertoli cells pre-treated with the PKA highly specific inhibitor, Myr-PKI [35]. Contrasting with differentiating cells, at this stage of development, inhibition of PKA over-stimulated p70S6K activity, which indicates that PKA would have an inhibitory effect on p70S6K (Fig. 9a). These data were further confirmed by western blot analysis, which showed a great increase in rpS6 phosphorylation in similar experimental conditions (Fig. 9b, c).

Since PKA antagonistically acted on p70S6K activity depending on Sertoli cell developmental stage, it was interesting to explore the impact of PKA inhibition on p70S6K phosphorylation pattern in proliferating cells. By western blot analysis, we observed that cell exposure to increasing doses of Myr-PKI led to a gradual raise in phosphorylation of the T421/S424 site (Fig. 9d). This increase was observed in basal as well as in FSH- or insulin-induced conditions (Fig. 9d, e). Similar results were obtained when using a structurally unrelated PKA inhibitor, C-MIQ (data not shown). Conversely, activating G $\alpha$ s, known to stimulate PKA, by increasing doses of cholera toxin mimicked FSH-induced dephosphorylation of the T421/S424 site (Fig. 9f). In contrast, the T389 site appeared to be insensitive to regulation by PKA (Fig. 9g, h).

As above, PKA, which is the target of Myr-PKI, is not represented in the simplified model. The direct consequence of PKA impairment is a drastic decrease of the available amount of active PKA, equivalent to a drastic decrease of the aggregate variable cAMP. This was modeled in both cell types by dividing initial and total amounts of cAMP 50-fold (Table 1). The simulation in PKA-depleted conditions revealed a large amount of Pp70S6K[T421/S424] at both stages (Fig. 11a, b, to be compared with the simulation in control conditions in Fig. 5a, b), leading to the hypothesis that cAMP indirectly catalyses the dephosphorylation of Pp70S6K[T421/S424]. In addition, in proliferating cells, Pp70S6K[T389] was restrained by PKA depletion. Therefore, at both stages, the PKA pathway indirectly catalyses the transition of Pp70S6K[T389/T421/S424] to Pp70S6K[T389]. The major difference between both stages is that this transition is



**Fig. 8** Intrinsic Pp70S6K quantities predicted by computational simulations, in PI3K-perturbed conditions. As above, the respective levels of Pp70S6K phosphorylated on [T389] only, Pp70S6K

phosphorylated on [T421/S424] only and Pp70S6K phosphorylated on [T389/T421/S424], are compared. Simulations in **a** proliferating and **b** differentiating Sertoli cells are shown

inhibitory of rpS6 phosphorylation in proliferating cells, and stimulatory in differentiating cells, as can be inferred from rpS6 phosphorylation level. This is due to the fact that the active form of p70S6K is different at the two stages. We can conclude from these simulations (Fig. 10) that the difference in initial conditions and total amount of the aggregate variable cAMP, figuring the inhibition of PKA, is sufficient to explain the evolution of the phosphorylated forms of p70S6K levels in FSH-stimulated conditions (Fig. 11).

#### Graphic formalization of p70S6K regulation by FSH

To formalize the differences in p70S6K regulation by FSH accurately, we graphically represented the signaling mechanisms engaged at both developmental stages by using the CellDesigner graphic interface (Fig. 12). It should be noted that in these diagrams, the PKA and mTOR pathways are figured, although they have been aggregated in the model used for simulation. When focusing on p70S6K phosphorylations, it appeared clearly that in proliferating Sertoli cells, only p70S6K phosphorylated on both T421/S424 and T389 was active, regardless the hormone, i.e. FSH (Fig. 12a) or insulin (Fig. 12b). Therefore, we propose that the major difference between FSH versus insulin-induced p70S6K activation is that p70S6K is refrained by the PKA signaling component emanating from FSH stimulation. In contrast, in differentiating Sertoli cells, only p70S6K phosphorylated on T389 is the active form, and PKA relieves the inhibition mediated by constitutive phosphorylation of T421/S424. Insulin-induced phosphorylation of p70S6K was similar at both stages.

#### Discussion

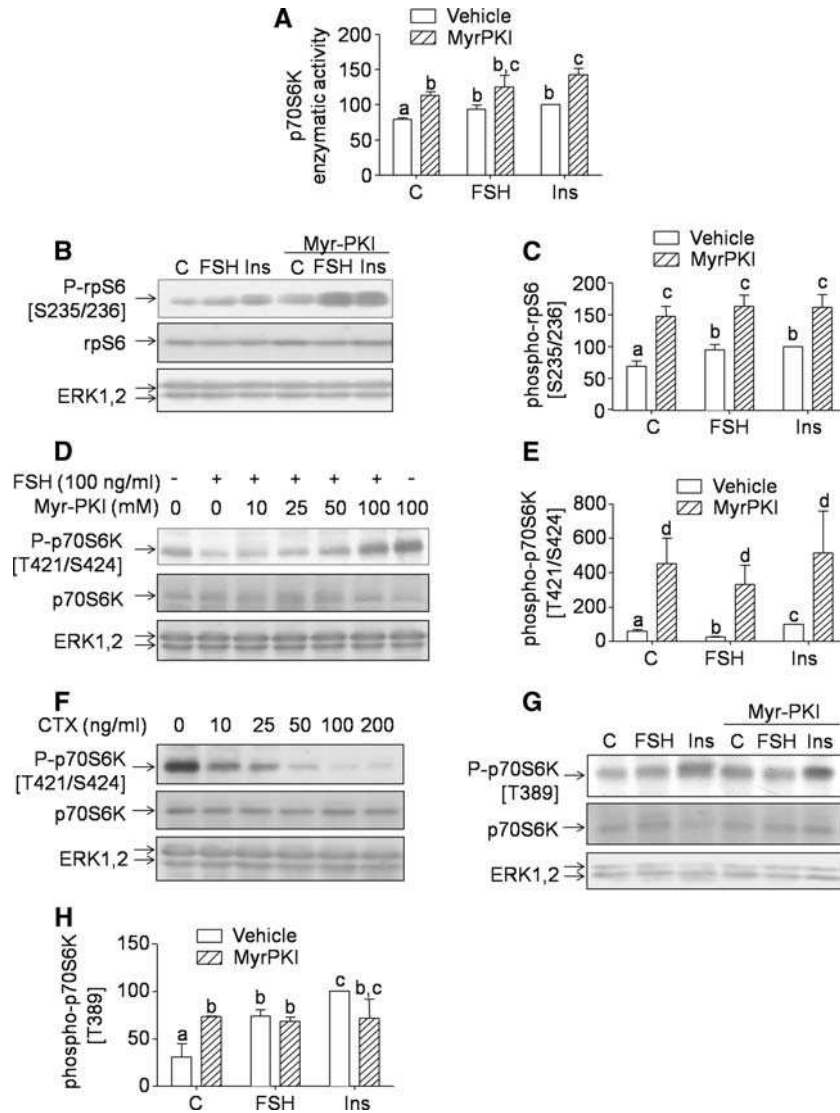
Here, we demonstrate that FSH-induced p70S6K activation pattern varies as a function of the Sertoli cell

developmental stage. In proliferating Sertoli cells isolated from 5-day-old rats, FSH weakly activates p70S6K, and induces dephosphorylation of its T421/S424 site and phosphorylation of its T389 site. FSH-induced T421/S424 dephosphorylation relies on PKA activation, and is detrimental to p70S6K enzymatic activity. The PI3K/mTOR pathway counteracts this inhibitory effect, and is also amenable for FSH-induced T389 phosphorylation. In striking contrast, in differentiating Sertoli cells stimulated by FSH, PKA-induced dephosphorylation of the T421/S424 site is required for its enzymatic activity, and p70S6K is not inducibly phosphorylated on T389 [22]. Tonic activation of the PI3K pathway is important for FSH to activate p70S6K in these cells. Resting concentrations of PIP3 have been shown previously to contribute permissively to carbachol-induced S6K1 activation mediated by the Gq-coupled M3 muscarinic receptor [36].

Importantly, in proliferating as well as in differentiating Sertoli cells treated with insulin, p70S6K pattern of phosphorylation was similar and was consistent with the current model, describing the enzyme as getting fully activated through sequential phosphorylation on multiple sites, including T421/S424 and T389 [10]. Therefore, our results clearly indicate that various patterns of phosphorylation are able to confer an active conformation to p70S6K. More precisely, extracellular signal-induced phosphorylation of the T421/S424 is supposed to relieve p70S6K auto-inhibition. This happens to be the case in proliferating Sertoli cells, but not in differentiating Sertoli cells, which suggests additional mechanisms to support this auto-inhibitory role. Furthermore, one important advance provided by our data is that p70S6K pattern of phosphorylation could be dictated by the class of receptor engaged, i.e., via a tyrosine-kinase receptor for the insulin receptor *versus* a GPCR for the FSH-R.

Here, we proposed a dynamical model of FSH-regulated p70S6K in primary Sertoli cells. The structure of the model was hypothesized from the experiment, and the parameters





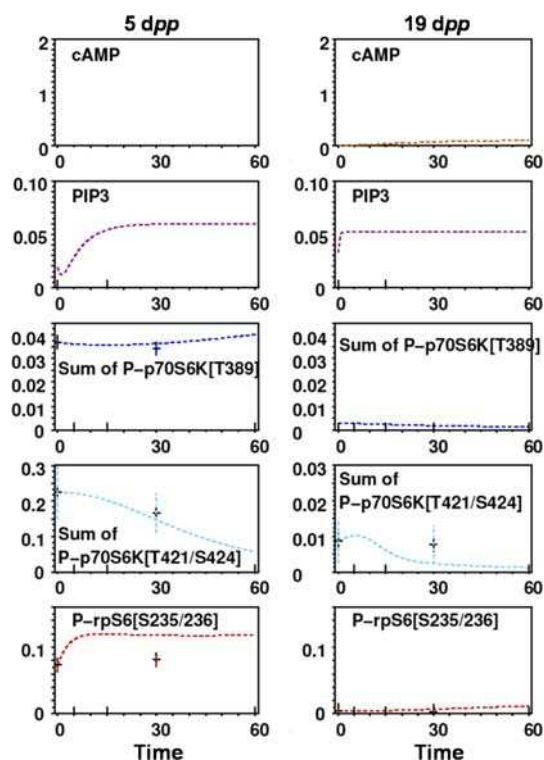
**Fig. 9** FSH-induced p70S6K activation includes a PKA-dependent inhibitory component. Sertoli cells were pretreated with the PKA inhibitor Myr-PKI (50 nM, 1 h) or left untreated before stimulation with FSH or insulin (each at 100 ng/ml) for 1 h to measure p70S6K activity or for 30 min for western blot assay. **a** Mean  $\pm$  SEM of p70S6K activity obtained from 3 independent experiments each performed in triplicates. **b** Representative western blot of rpS6 phosphorylation reprobbed with an anti-total-rpS6 antibody. **c** Mean  $\pm$  SEM of phospho-rpS6 signals obtained from more than 3 independent experiments each performed in duplicates. **d, f** Representative western blots of p70S6K phosphorylation on T421/S424 in

response to FSH, in cells incubated with increasing doses of Myr-PKI or CTX, respectively. The membrane was reprobbed with an anti-total-p70S6K antibody. **g** Representative western blot of T389 phosphorylation in response to FSH, reprobbed with an anti-total-p70S6K antibody. In **b, d, f, and g**, gel loading was monitored by reprobbed the membranes with an anti-ERK1, 2 antibody. **e, h** Mean  $\pm$  SEM of phospho-T421/S424 and T389 signals obtained from 3 independent experiments each performed in duplicates. Results are expressed as percent of the response to insulin. Shared *superscripts* indicate no significant difference, while different *superscripts* indicate significant differences, at the  $P < 0.05$  level

where optimized to fit data in normal conditions. Although we cannot rule out that this is not the unique solution, we found it satisfactory to obtain a set of parameters which fit the data, given the number of readouts, i.e., cAMP, PIP3, sum of Pp70S6K[T389], sum of Pp70S6K[T421/S424], PrpS6[S235/236],  $K_{d_{RH}}$ , FSH-R, FSH. Therefore, the probability that our parameters are close to reality is enhanced. Our model is further supported by the fact that, although it was obtained by fitting the control conditions,

all the parameters, i.e., kinetic rates and total amounts, still fitted the data in pharmacologically perturbed conditions.

Considering the phosphorylated p70S6K isoforms, our model predicts that, in differentiating cells, FSH stimulation increases the level of Pp70S6K[T389], which is ultimately the active form, at the expense of Pp70S6K [T389/T421/S424]. However, this cannot be observed experimentally, since western blot analyses done with a unique anti-phospho-antibody do not discriminate between

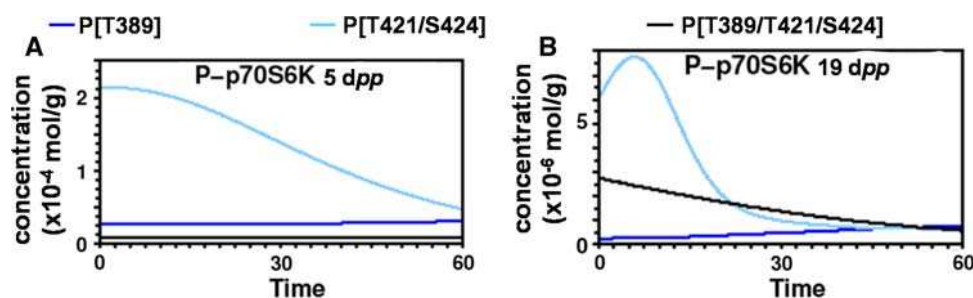


**Fig. 10** Simulation of the cAMP/PKA component of FSH-induced p70S6K regulation. Computational simulations (*plain lines*) compared with experimental data (*black crosses*  $\pm$  SEM) of FSH-R-mediated signaling in the presence of PKA inhibitor, in proliferating (*left*) and differentiating (*right*) Sertoli cells. The concentrations of molecules are expressed as a function of time, as above

Pp70S6K[T389] and Pp70S6K[T389/T421/S424]. Therefore, although the PIP3 level was not significantly increased following FSH stimulation, we propose that there was still an increase in the phosphorylation of the PIP3-sensitive site, T389. In contrast, in proliferating cells, as a result from conflicting interactions of the PKA- and the PI3K-dependent component, Pp70S6K[T389/T421/S424] is the active form following FSH stimulation. When compared to the admitted model of p70S6K

regulation by insulin [10], we admit that phosphorylation on the [T389] residue participates in p70S6K activation. Nevertheless, our results indicate that phosphorylation of the [T421/S424] is not a pre-requisite. Therefore, inhibition relief, if any, should be achieved by other aminoacid residues of p70S6K. This view is further supported by the fact that no fit was found unless a supplemental degree of freedom,  $k_7$ , was added. Presumably, an unidentified effector catalyzes the conversion of p70S6K to p70S6K[T389] in proliferating cells, or, more likely, p70S6K gets phosphorylated at other sites than the two studied here.

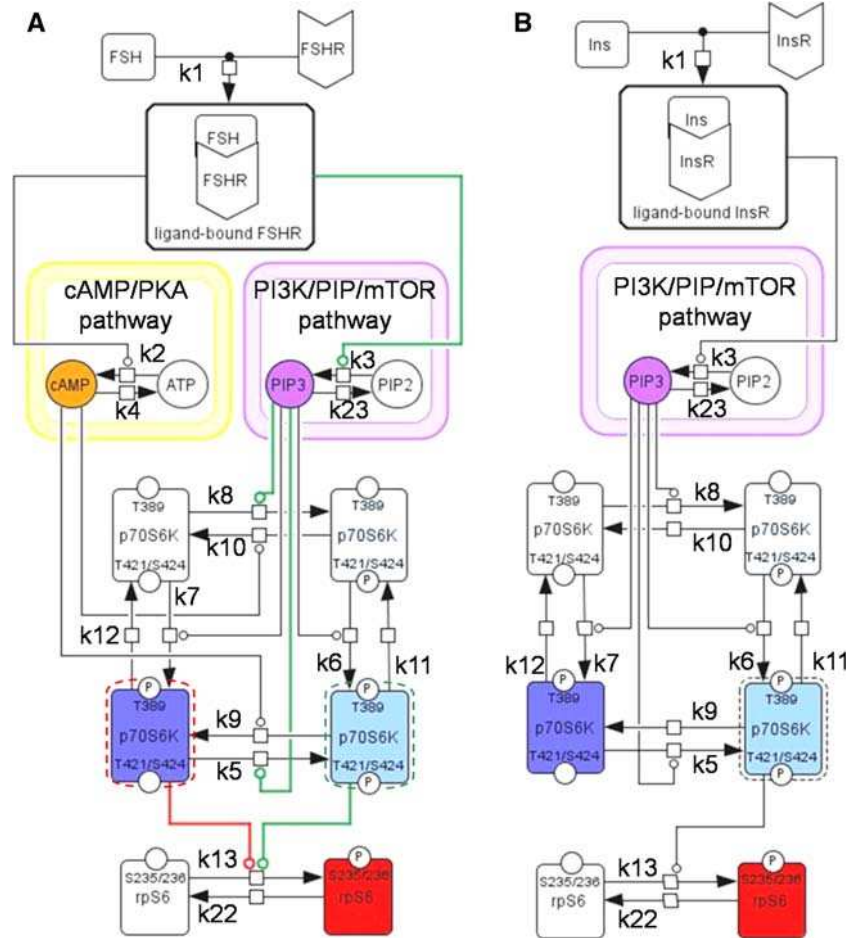
Sertoli cells can be considered as a particular cell type. In the neonate, they are differentiating cells which retain some mitotic potential, whereas in the pre-pubertal animal, they lose their mitotic ability and their terminal differentiation is finally acquired in the adult. Interestingly, a single hormone, FSH, is a key regulator of these developmental processes, hence its signaling can be analyzed in two different and physiologically relevant backgrounds. It is now clear that the signaling pathways activated by FSH at each stage of Sertoli cell post-natal development are different, or, at least, differentially modulated to lead to those various cellular responses. For example, in 5-day-old rat proliferating Sertoli cells, the MAP kinases ERK1 and ERK2 are phosphorylated by FSH [32]. ERK phosphorylation induces a cyclin D1 increase which is directly related to the proliferation state of the cells. In contrast, FSH totally switches off the ERK pathway in 19-day-old rat differentiating Sertoli cells. Likewise, we hypothesized that the signaling pathways activated by FSH upstream of the p70S6K might be different at each of those developmental stages and would result in differential phosphorylation patterns. And actually, the signaling pathways induced by FSH to phosphorylate p70S6K seem to be modulated during the post-natal development of Sertoli cells, leading to different phosphorylation patterns, with differential integration of PI3K- and PKA-dependent signaling mechanisms. In addition, T421/S424 dephosphorylation seems



**Fig. 11** Intrinsic Pp70S6K quantities predicted by computational simulations, in PKA-perturbed conditions. As above, the respective levels of Pp70S6K phosphorylated on [T389] only, Pp70S6K

phosphorylated on [T421/S424] only and Pp70S6K phosphorylated on [T389/T421/S424], are compared. Simulations in **a** proliferating and **b** differentiating Sertoli cells are shown

**Fig. 12** Graphic representation of the developmental regulation of p70S6K by FSH and by insulin. **a** Regulation of p70S6K by FSH has been superimposed to highlight the differences between both stages. The reactions specifically induced by FSH in proliferating cells are in *green*, while those specifically induced by FSH in differentiating cells are in *red*, and the common ones are in *black*. The active form of p70S6K is surrounded by *green* and *red dots*, respectively. **b** Insulin induces a similar pattern of p70S6K phosphorylation at both stages. The molecular species experimentally measured in this study and mined for the simulations are in color, while the others are in *black* and *white*. All the reactions are equilibrium, but for clarity, they are represented with simple arrows, unless necessary. Reaction parameters, as determined by data fitting, are indicated. Activating interactions are figured with a line ended in a circle



to have different consequences on p70S6K activity in the two stages: activating in differentiating cells and inhibitory in proliferating cells. Therefore, the cellular background appears to be instrumental for p70S6K phosphorylation and activity, in agreement with the notion that the developmental stage of cells dictates the way the hormonal messages are integrated. p70S6K is also activated by FSH in granulosa cells of the ovary, in a PKA-dependent manner, but only phosphorylation on the T389 site has been analyzed in this study [37].

One possibility to explain the developmental regulation of p70S6K by second messenger-dependent kinases was that second messenger production in response to FSH would be developmentally regulated, depending on FSH-R coupling to various G proteins. For this reason, we analyzed the efficacy of FSH to induce PIP3 and cAMP downstream. The cAMP basal level was much higher in proliferating cells than in differentiating cells [32] but the FSH-induced raise in production was much more important in differentiating cells. Conversely, basal PIP3 production was lower in proliferating cells and significantly increased by FSH, when compared to differentiating cells. These observations could explain that the main pathway required

for FSH to stimulate p70S6K in proliferating cells is the PI3K/mTOR pathway, leading to the inducible phosphorylation of p70S6K T389. This is consistent with the proliferating state of the cells and to the well-known role of this pathway in cell growth [38]. In differentiating cells, the main pathway activated by FSH is the PKA pathway, required for the expression of specific markers, such as transferrin [39]. Again, these findings indicate that a single ligand can stimulate diverse intracellular signaling pathways with differential efficacy. The FSH-R has the ability to couple to both Gs and Gi, depending on the developmental stage [32] or the cell type [40]. The nature of the G proteins that the receptor couples to could be critical for signaling to p70S6K activation, since GPCR coupled to Gq, for example, such as the M3-muscarinic [36], PGF2 $\alpha$  [19],  $\alpha$ 1-adrenergic receptor [41], stimulate p70S6K in a PI3K-independent manner, unlike Gi-coupled receptors, such as the LPA [42] or the  $\mu$ -opioid [43] receptors. Alternatively, a yeast two-hybrid screen has revealed the direct interaction of the FSH-R with a PI3K adaptor, APPL1 [44], in a signalosome including Foxo1 [45]. Unfortunately, the expression pattern of APPL1 in the development of Sertoli cells is not yet available.

We cannot yet appreciate the biological consequences of the differences in p70S6K regulation in Sertoli cell development and in response to a GPCR or to a tyrosine kinase receptor. Since p70S6K has 8 potential phosphorylation sites [46], more studies are needed to better define the p70S6K pattern of phosphorylation in response to each of those receptors and the outcomes on the biological role of p70S6K. And in fact, modeling the entire combinatorial complexity of p70S6K regulation sites in developing Sertoli cells stimulated by FSH is an interesting question to address in future work. In particular, p70S6K is known to be implicated in mRNA translation, by phosphorylating rpS6 [47], and eIF4B [7]. Whether the different active isoforms of p70S6K that we have studied here equally regulate their target proteins within the translational machinery remains to be defined. GPCRs have recently been linked to the control of translation by the discovery that the GPCR scaffolding proteins  $\beta$ -arrestins are involved in protein synthesis by positively regulating Mnk [48]. Hence, whether the FSH-R and the insulin receptor act on the translational machinery in different manners in Sertoli cell development also needs to be addressed in the future.

**Acknowledgments** The authors thank the technical assistance of the rat breeders Jean-Claude Braguer and Claude Cahier. Drs Florian Guillou, Laurence Dupuy in our laboratory, and Drs François Fages and Frédérique Clément (INRIA, Rocquencourt), are acknowledged for critical reading of the manuscript. We are also indebted to Drs Len Stephens and Phil Hawkins for the PIP3 quantification and to Dr Anne Segonds-Pichon for statistical analysis (Babraham Institute, Cambridge, UK). A.M. was funded by a fellowship from the Region Centre and from the Institut National de la Recherche Agronomique (INRA), D.H. was funded by a fellowship from the INRA (ASC), G.D. by a fellowship from the Ministère de la Recherche et de la Technologie, C.L. was granted by the Fondation pour le Recherche Médicale, and H.G. was a BBSRC fellow. This work was supported by the INRA AgroBI AIP, by the Centre National de la Recherche Scientifique, by Université de Tours and by the AE INRIA/INRA REGATE.

## References

- Ohanna M, Sobering AK, Lapointe T, Lorenzo L, Praud C, Petroulakis E, Sonenberg N, Kelly PA, Sotiropoulos A, Pende M (2005) Atrophy of S6K1(-/-) skeletal muscle cells reveals distinct mTOR effectors for cell cycle and size control. *Nat Cell Biol* 7:286–294
- Pende M, Kozma SC, Jaquet M, Oorschot V, Burcelin R, Le Marchand-Brustel Y, Klumperman J, Thorens B, Thomas G (2000) Hypoinsulinaemia, glucose intolerance and diminished beta-cell size in S6K1-deficient mice. *Nature* 408:994–997
- Lane HA, Fernandez A, Lamb NJ, Thomas G (1993) p70S6K function is essential for G1 progression. *Nature* 363:170–172
- Blenis J, Erikson RL (1985) Regulation of a ribosomal protein S6 kinase activity by the Rous sarcoma virus transforming protein, serum, or phorbol ester. *Proc Natl Acad Sci USA* 82:7621–7625
- Stewart MJ, Thomas G (1994) Mitogenesis and protein synthesis: a role for ribosomal protein S6 phosphorylation? *Bioessays* 16:809–815
- Rogers GW Jr, Richter NJ, Merrick WC (1999) Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. *J Biol Chem* 274:12236–12244
- Raught B, Peiretti F, Gingras AC, Livingstone M, Shahbazian D, Mayeur GL, Polakiewicz RD, Sonenberg N, Hershey JW (2004) Phosphorylation of eucaryotic translation initiation factor 4B Ser422 is modulated by S6 kinases. *EMBO J* 23:1761–1769
- Holz MK, Ballif BA, Gygi SP, Blenis J (2005) mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. *Cell* 123:569–580
- Burnett PE, Barrow RK, Cohen NA, Snyder SH, Sabatini DM (1998) RAFT1 phosphorylation of the translational regulators p70 S6 kinase and 4E-BP1. *Proc Natl Acad Sci USA* 95:1432–1437
- Dennis PB, Pullen N, Pearson RB, Kozma SC, Thomas G (1998) Phosphorylation sites in the autoinhibitory domain participate in p70(s6k) activation loop phosphorylation. *J Biol Chem* 273:14845–14852
- Pullen N, Thomas G (1997) The modular phosphorylation and activation of p70S6K. *FEBS Lett* 410:78–82
- Dennis PB, Pullen N, Kozma SC, Thomas G (1996) The principal rapamycin-sensitive p70(s6 k) phosphorylation sites, T-229 and T-389, are differentially regulated by rapamycin-insensitive kinase kinases. *Mol Cell Biol* 16:6242–6251
- Pullen N, Dennis PB, Andjelkovic M, Dufner A, Kozma SC, Hemmings BA, Thomas G (1998) Phosphorylation and activation of p70S6K by PDK1. *Science* 279:707–710
- Ma L, Chen Z, Erdjument-Bromage H, Tempst P, Pandolfi PP (2005) Phosphorylation and functional inactivation of TSC2 by Erk implications for tuberous sclerosis and cancer pathogenesis. *Cell* 121:179–193
- Proud CG (2007) Signalling to translation: how signal transduction pathways control the protein synthetic machinery. *Biochem J* 403:217–234
- Edelmann HM, Kuhne C, Petrutsch C, Ballou LM (1996) Cell cycle regulation of p70 S6 kinase and p42/p44 mitogen-activated protein kinases in Swiss mouse 3T3 fibroblasts. *J Biol Chem* 271:963–971
- Le XF, Hittelman WN, Liu J, McWatters A, Li C, Mills GB, Bast RC Jr (2003) Paclitaxel induces inactivation of p70 S6 kinase and phosphorylation of Thr421 and Ser424 via multiple signaling pathways in mitosis. *Oncogene* 22:484–497
- Suh JM, Song JH, Kim DW, Kim H, Chung HK, Hwang JH, Kim JM, Hwang ES, Chung J, Han JH, Cho BY, Ro HK, Shong M (2003) Regulation of the phosphatidylinositol 3-kinase, Akt/protein kinase B, FRAP/mammalian target of rapamycin, and ribosomal S6 kinase 1 signaling pathways by thyroid-stimulating hormone (TSH) and stimulating type TSH receptor antibodies in the thyroid gland. *J Biol Chem* 278:21960–21971
- Arvaisis EW, Romanelli A, Hou X, Davis JS (2006) AKT-independent phosphorylation of TSC2 and activation of mTOR and ribosomal protein S6 kinase signaling by prostaglandin F2alpha. *J Biol Chem* 281:26904–26913
- Moschella PC, Rao VU, McDermott PJ, Kuppuswamy D (2007) Regulation of mTOR and S6K1 activation by the nPKC isoforms, PKCepsilon and PKCdelta, in adult cardiac muscle cells. *J Mol Cell Cardiol* 43:754–766
- Ballou LM, Cross ME, Huang S, McReynolds EM, Zhang BX, Lin RZ (2000) Differential regulation of the phosphatidylinositol 3-kinase/Akt and p70 S6 kinase pathways by the alpha(1A)-adrenergic receptor in rat-1 fibroblasts. *J Biol Chem* 275:4803–4809
- Lecureuil C, Tesseraud S, Kara E, Martinat N, Sow A, Fontaine I, Gauthier C, Reiter E, Guillou F, Crepieux P (2005) Follicle-stimulating hormone activates p70 ribosomal protein S6 kinase

- by protein kinase A-mediated dephosphorylation of Thr 421/Ser 424 in primary Sertoli cells. *Mol Endocrinol* 19:1812–1820
23. Dierich A, Sairam MR, Monaco L, Fimia GM, Gansmuller A, LeMeur M, Sassone-Corsi P (1998) Impairing follicle-stimulating hormone (FSH) signaling in vivo: targeted disruption of the FSH receptor leads to aberrant gametogenesis and hormonal imbalance. *Proc Natl Acad Sci USA* 95:13612–13617
  24. Kumar TR, Wang Y, Lu N, Matzuk MM (1997) Follicle stimulating hormone is required for ovarian follicle maturation but not male fertility. *Nat Genet* 15:201–204
  25. Sairam MR, Krishnamurthy H (2001) The role of follicle-stimulating hormone in spermatogenesis: lessons from knockout animal models. *Arch Med Res* 32:601–608
  26. Wayne CM, Fan HY, Cheng X, Richards JS (2007) Follicle-stimulating hormone induces multiple signaling cascades: evidence that activation of Rous sarcoma oncogene, RAS, and the epidermal growth factor receptor are critical for granulosa cell differentiation. *Mol Endocrinol* 21:1940–1957
  27. Kara E, Crepieux P, Gauthier C, Martinat N, Piketty V, Guillou F, Reiter E (2006) A phosphorylation cluster of five serine and threonine residues in the C-terminus of the follicle-stimulating hormone receptor is important for desensitization but not for beta-arrestin-mediated ERK activation. *Mol Endocrinol* 20:3014–3026
  28. Guillou F, Martinat N, Combarous Y (1986) Study of the superactivity of equine follicle-stimulating hormone in in vitro stimulation of rat Sertoli cells. *Biochem Biophys Acta* 887:196–203
  29. Bernard L, Martinat N, Lécureuil C, Crépieux P, Reiter E, Tilloy-Ellul A, Chevalier S, Guillou F (2007) Dichlorodiphenyltrichloroethane impairs follicle-stimulating hormone receptor-mediated signaling in rat Sertoli cells. *Reprod Toxicol* 23:158–164
  30. Guillou H, Lecureuil C, Anderson KE, Suire S, Ferguson GJ, Ellson CD, Gray A, Divecha N, Hawkins PT, Stephens LR (2007) Use of the GRP1 PH domain as a tool to measure the relative levels of PtdIns(3, 4, 5)P3 through a protein-lipid overlay approach. *J Lipid Res* 48:726–732
  31. Nishio Y, Nagata S, Umeda M, Shirai R, Yokogawa T, Ihara S, Fukui Y (2000) Quantification of phosphatidylinositol 3, 4, 5-trisphosphate by liposome lysis assay with specific monoclonal antibodies. *Anal Biochem* 285:270–273
  32. Crépieux P, Marion S, Martinat N, Fafeur V, Vern YL, Kerboeuf D, Guillou F, Reiter E (2001) The ERK-dependent signalling is stage-specifically modulated by FSH, during primary Sertoli cell maturation. *Oncogene* 20:4696–4709
  33. Bortolussi M, Zanchetta R, Belvedere P, Colombo L (1990) Sertoli and Leydig cell numbers and gonadotropin receptors in rat testis from birth to puberty. *Cell Tissue Res* 260:185–191
  34. Monn E, Desautel M, Christiansen RO (1972) Highly specific testicular adenosine-3', 5'-monophosphate phosphodiesterase associated with sexual maturation. *Endocrinology* 91:716–720
  35. Murray AJ (2008) Pharmacological PKA inhibition: all may not be what it seems. *Sci Signal* 1, re4
  36. Tang X, Wang L, Proud CG, Downes CP (2003) Muscarinic receptor-mediated activation of p70 S6 kinase 1 (S6K1) in 1321N1 astrocytoma cells: permissive role of phosphoinositide 3-kinase. *Biochem J* 374:137–143
  37. Alam H, Maizels ET, Park Y, Ghaey S, Feiger ZJ, Chandel NS, Hunzicker-Dunn M (2004) FSH activation of HIF-1 by the PI3-kinase/AKT/Rheb/mTOR pathway is necessary for induction of select protein markers of follicular differentiation. *J Biol Chem* 279:19431–19440
  38. Shaw RJ, Cantley LC (2006) Ras, PI(3)K and mTOR signalling controls tumour cell growth. *Nature* 441:424–430
  39. Suire S, Fontaine I, Guillou F (1995) Follicle stimulating hormone (FSH) stimulates transferrin gene transcription in rat Sertoli cells: cis and trans-acting elements involved in FSH action via cyclic adenosine 3', 5'-monophosphate on the transferrin gene. *Mol Endocrinol* 9:756–766
  40. Sun L, Peng Y, Sharrow AC, Iqbal J, Zhang Z, Papachristou DJ, Zaidi S, Zhu LL, Yaroslavskiy BB, Zhou H, Zallone A, Sairam MR, Kumar TR, Bo W, Braun J, Cardoso-Landa L, Schaffler MB, Moonga BS, Blair HC, Zaidi M (2006) FSH directly regulates bone mass. *Cell* 125:247–260
  41. Wang L, Proud CG (2002) Ras/Erk signaling is essential for activation of protein synthesis by Gq protein-coupled receptor agonists in adult cardiomyocytes. *Circ Res* 91:821–829
  42. Kam Y, Exton JH (2004) Role of phospholipase D1 in the regulation of mTOR activity by lysophosphatidic acid. *FASEB J* 18:311–319
  43. Polakiewicz RD, Schieferl SM, Gingras AC, Sonenberg N, Comb MJ (1998) mu-Opioid receptor activates signaling pathways implicated in cell survival and translational control. *J Biol Chem* 273:23534–23541
  44. Nechamen CA, Thomas RM, Cohen BD, Acevedo G, Poulikakos PI, Testa JR, Dias JA (2004) Human follicle-stimulating hormone (FSH) receptor interacts with the adaptor protein APPL1 in HEK 293 cells: potential involvement of the PI3K pathway in FSH signaling. *Biol Reprod* 71:629–636
  45. Nechamen CA, Thomas RM, Dias JA (2007) APPL1, APPL2, Akt2 and FOXO1a interact with FSHR in a potential signaling complex. *Mol Cell Endocrinol* 260–262:93–99
  46. Ferrari S, Bannwarth W, Morley SJ, Totty NF, Thomas G (1992) Activation of p70S6K is associated with phosphorylation of four clustered sites displaying Ser/Thr-Pro motifs. *Proc Natl Acad Sci USA* 89:7282–7286
  47. Duncan R, McConkey EH (1982) Rapid alterations in initiation rate and recruitment of inactive RNA are temporally correlated with S6 phosphorylation. *Eur J Biochem* 123:539–544
  48. DeWire SM, Kim J, Whalen EJ, Ahn S, Chen M, Lefkowitz RJ (2008) Beta-arrestin-mediated signaling regulates protein synthesis. *J Biol Chem* 283:10611–10620

### 3.1.3 Données complémentaires

Une originalité et un atout de ce travail de modélisation a été l'homogénéité des données qui a permis d'exprimer l'ensemble des quantités en  $nmol.\mu g^{-1}$  de protéines avec beaucoup de rigueur. J'ai choisi cette unité car toutes les expériences pouvaient s'y référer. En effet, nous avons trouvé une estimation de l'activité PI3K dans cette unité, ce qui nous a permis de calculer le 100% insuline correspondant. A ce stade nous avons fait l'hypothèse qu'il représentait les mêmes concentrations pour les deux forme de p70 et la rpS6. Enfin, connaissant les volumes et concentrations des autres mesures (FSH, récepteur et AMPc), leur conversion a pu être réalisée. Le détail des calculs pour l'AMPc qui était exprimé en  $nmol.L^{-1}$  au départ et a été converti en  $nmol.\mu g^{-1}$  est donné ci-dessous.

Tout d'abord, j'ai converti les  $nmol.L^{-1}$  en  $nmol.\mu g^{-1}$  de protéines. Pour les cellules en prolifération (5dpp), dans  $100\mu L$  on a  $53.65\mu g$  de protéines (pour les cellules différenciées(19dpp), on a  $57.58\mu g$  de protéines)

Soit

$$\begin{aligned} k_{5dpp} &= 1nmol.\mu g \\ &= (100 \cdot 10^{-6}L)/(53.65\mu g) \\ &= 1.86 \cdot 10^{-6}nmol.L^{-1} \end{aligned}$$

L'AMPc étant exprimée en  $mol$  par cellule, j'ai effectué une conversion du nombre de cellules en volume ( $L$ ).

$$100\%cAMP_{5dpp} = 4.5nmol.10^6cellules$$

Sachant que dans  $100 \cdot 10^{-6}L$  il y a  $410^6$  cellules, dans  $1L$  il y a  $(4 \cdot 10^4) \cdot 10^6$  cellules

$$\begin{aligned} 100\%cAMP_{5dpp} &= (4.5 \cdot nmol \cdot 10^6cellules) \cdot 410^4 \\ &= (18 \cdot 10^4)nmol.L^{-1} \\ &= (18 \cdot 10^4)nmol.L^{-1} \cdot kJ5 \\ &= 33.6 \cdot 10^{-2}nmol.\mu g \text{ de protéines} \end{aligned}$$

D'où

$$1\%cAMP_{5dpp} = 0.00336nmol.\mu g \text{ de protéines}$$

En procédant de la même façon pour les cellules différenciées, on obtient

$$1\%cAMP_{19dpp} = 0.00263nmol.\mu g \text{ de protéines}$$

### 3.1.4 Conclusion et perspectives

L'approche de modélisation a permis de formaliser, en un seul modèle, la dynamique de régulation des phosphorylations de p70S6K, et ce pour les deux stades de développement des cellules de Sertoli (prolifération et différenciation). Les paramètres ont été ajustés sur

les données contrôles mais le modèle permet aussi de simuler fidèlement les données perturbées (inhibitions de PI3K et inhibition de PKA). Enfin, ce travail a permis de fournir des constantes cinétiques ainsi que des simulations prédictives sur les concentrations des différentes formes de p70S6K impossibles à mesurer expérimentalement. Or, malgré ce niveau de détail, et les branchements catalytiques fins sur chaque régulation de phosphorylation, en réalisant l'ajustement paramétrique pour les deux stades de développement, un degré de liberté supplémentaire a du être ajouté : le paramètre  $k_7$ . Ce biais témoigne vraisemblablement de la présence d'un effecteur non identifié qui catalyserait la conversion de p70S6K en Pp70S6K{T389} dans les cellules en prolifération, ou plus probablement, l'existence d'autres sites phosphorylés que ceux déjà étudiés, et qui donneraient à la p70S6K ce comportement différent. La p70S6K comporte bien d'autres sites de phosphorylation, qui n'ont pas été détaillés par manque de données. De plus, le détail de l'ensemble des sites de phosphorylations aurait engendré une augmentation importante de la taille du modèle. En effet, pour trois sites de phosphorylation, on compte huit intermédiaires, et pour  $n$  sites,  $2^n$ . On peut se demander si cette complexité aurait été possible à prendre en compte dans notre formalisme simple de modélisation et si l'apport d'information, une observable par site phosphorylé, aurait suffi devant l'ajout de variables et de paramètres inconnus. La modélisation est bien souvent un équilibre entre détails et capacités. Nous avons déjà essayé d'élargir le modèle en détaillant les deux voies de signalisation mais l'environnement Scilab et les méthodes utilisées au moment de ce travail ne permettaient pas d'aller plus loin. Le besoin d'employer des méthodes plus performantes pour intégrer le système et ajuster les paramètres afin d'élargir le modèle nous a amené à envisager d'autres environnements.

### 3.2 Modélisation dynamique du réseau de signalisation du RAT<sub>1A</sub>

Les méthodes d'ajustement paramétrique utilisées dans la partie précédente ont montré leurs limites. En effet, comme présenté en introduction, la méthode d'ajustement paramétrique du gradient conjugué ne suffit pas pour des modèles plus complexes. D'autre part, l'environnement Biocham, dédié aux réseaux biochimiques, proposait un formalisme de propriétés temporelles pour décrire la *fonction objectif* qui permet d'intégrer des contraintes qualitatives à la recherche des paramètres, ce qui n'est pas possible avec un formalisme classique d'expression de la fonction objectif. Lorsque Biocham s'est doté d'une méthode d'ajustement paramétrique performante, CMAES, nous avons misé sur ces nouveaux formalismes pour réaliser un second modèle afin de répondre aux questions portant sur l'activation des ERK par le RAT<sub>1A</sub>.

Tout d'abord, comme le permet le formalisme QFLTL, nous avons utilisé les points de données temporelles munis de leurs écarts types ainsi que des contraintes originales, simplifiant parfois la recherche. Pour les jeux de données FRET d'observation en condition contrôle du DAG et de la PKC, nous avons intégré à la fonction objectif la cinétique rapide de départ entre 0 et 5mns tandis que pour les temps long, une contrainte d'encadrement a été ajoutée, comme illustré (Figure 3.2). Cela permet de ne pas augmenter le nombre de points de la fonction objectif tout en décrivant le comportement du DAG et de la PKC.

Des contraintes supplémentaires nous ont permis de canaliser les recherches en spécifiant les intervalles de recherche des paramètres et en obligeant les quantités inhibées à être inférieures aux quantités initiales. Ce formalisme nous a aussi permis de spécifier des contraintes originales et qualitatives telles que : le recrutement de  $\beta$ -arrestines doit s'effectuer préférentiellement au récepteur phosphorylé ; les récepteurs phosphorylés par GRKs 2 et 3 privilégient le recrutement des  $\beta$ -arrestines 1 de préférence aux  $\beta$ -arrestines 2.

Enfin, contrairement au modèle p70S6K dont les paramètres ont été ajustés en conditions contrôle, nous avons voulu aller plus loin en réalisant un ajustement sous plusieurs conditions simultanément. Ce qui nous a poussé à suivre cette direction était le jeu de données principales. Ces données, très homogènes, observent les ERK phosphorylés par l'angiotensine en condition contrôle et en plusieurs conditions perturbées : inhibition des  $\beta$ -arrestines, de la PKC, des GRKs 2, 3, 5 et 6. Pour cela nous avons veillé à garder commun entre deux conditions l'ensemble des lois et des paramètres du modèle et de ne modifier que les conditions initiales ou quantités totales modifiées par l'inhibition. Ainsi, au sein de la fonction objectif, nous intégrons cinq fois l'ODE pour les cinq conditions et nous recherchons les paramètres du modèle ainsi que les quantités initiales ( $\beta$ -arrestines 2 inhibées, PKC active et inactive initiales et inhibées, GRKs 2 et 3 inhibées et GRKs 5 et 6 inhibées) qui minimisent au mieux l'erreur sur les cinq modèles.

Avec ce modèle, simulant les cinétiques ERK du réseaux de signalisation de l'angiotensine II, nous avons répondu à plusieurs questions biologiques :

1] Comment les GRKs qui phosphorylent le récepteur contrôlent-elles l'équilibre entre les deux voies de signalisations ERK (dépendante des protéines G et dépendantes des  $\beta$ -arrestines) ?



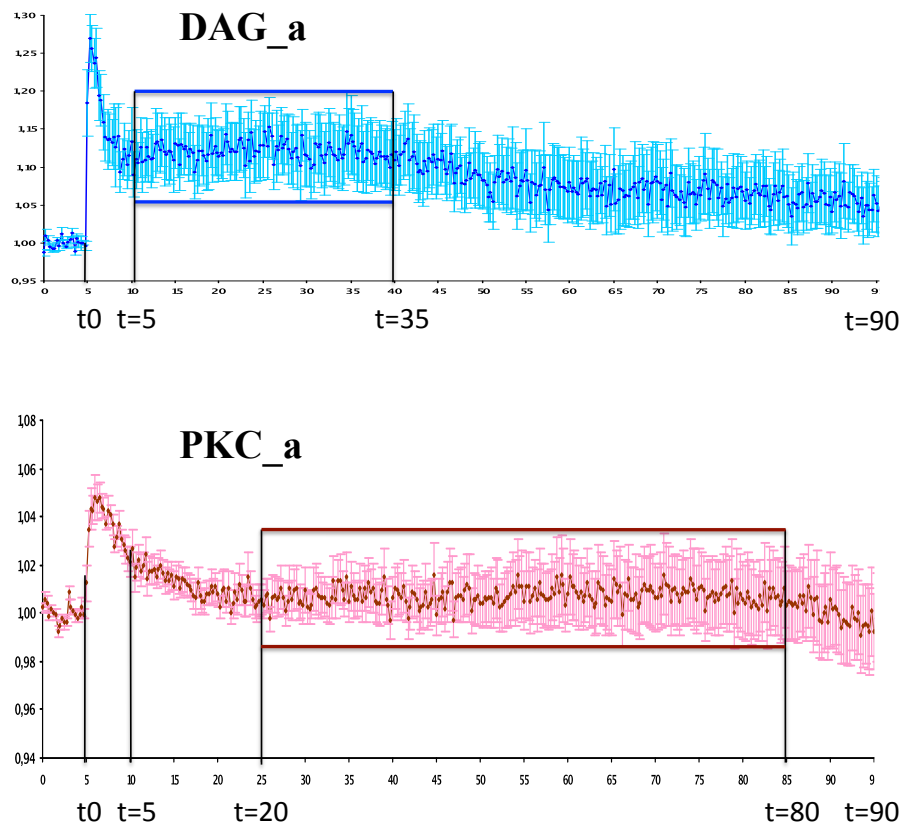


FIGURE 3.2 – Comme le propose le formalisme QFLTL de Biocham. Nous avons utilisé au sein de la fonction objectif, les points entre 0 et 5 mns ainsi que des majorants et des minorants entre 5mns et 35mns pour DAG et entre 20 et 80 mns pour la PKC.

- 2] Quelles sont les influences de chaque voie de signalisation sur les autres ?
- 3] Comment les  $\beta$ -arrestines jouent-elles leurs double rôle au sein de la désensibilisation et de la signalisation ?

Le modèle dynamique créé répond à ces questions, et remet en cause le paradigme d'activation des ERK par les  $\beta$ -arrestines. Habituellement présenté sous forme d'un complexe aux proportions 1-1 (1 récepteur, 1  $\beta$ -arrestine, 1 Raf, 1 MEK et 1 ERK), le modèle montre que cette stoechiométrie est impossible, il existe une amplification entre les différentes espèces. Cela a par la suite été vérifié expérimentalement. Enfin, l'ajout des taux d'inhibitions des GRKs aux paramètres recherchés pour les conditions perturbées a mis en lumière des spécificités du processus de phosphorylation des GRKs 2 et 3 opposées aux GRKs 5 et 6. Les GRK 2 et 3 agiraient indépendamment et pourraient se substituer l'une à l'autre tandis que les GRK 5 et 6 agiraient en série, la suppression d'une des deux formes suffit à

inhiber totalement la réponse ERK qui en découle.

### **3.2.1 Propriétés émergentes dans la transduction du signal par un R7TM révélées en combinant modélisation et expériences**

Dans l'article suivant, nous avons détaillé le modèle d'activation des ERK par l'angiotensine II qui comprend 18 EDOs, 7 lois de conservations et 56 paramètres dont 40 inconnus, recherchés grâce au couplage des méthodes QFLTL-CMAES. Ce modèle a été ajusté sur des données expérimentales de ERK phosphorylées en cinq conditions ainsi que sur des données DAG et PKC en conditions contrôle.

## Article à soumettre

**G protein-coupled receptor kinases: molecular switches controlling the balance between**

**G protein and  $\beta$ -arrestin-mediated signaling by G protein-coupled receptors.**

Domitille Heitzler<sup>1\*</sup>, Guillaume Durand<sup>1\*</sup>, Aurélien Rizk<sup>3</sup>, Seungkirl Ahn<sup>2</sup>, Jihee Kim<sup>2</sup>, Jonathan D. Violin<sup>2</sup>, Laurence Dupuy<sup>1</sup>, Christophe Gauthier<sup>1</sup>, Vincent Piketty<sup>1</sup>, Pascale Crépieux<sup>1</sup>, Anne Poupon<sup>1</sup>, Frédérique Clément<sup>4</sup>, François Fages<sup>3</sup>, Robert J. Lefkowitz<sup>2</sup>, Eric Reiter<sup>1#</sup>.

<sup>1</sup>BIOS group, INRA, UMR85, Unité Physiologie de la Reproduction et des Comportements, F-37380 Nouzilly, France ; CNRS, UMR6175, F-37380 Nouzilly, France ; Université François Rabelais, F-37041 Tours, France ; Haras Nationaux, F-37380 Nouzilly, France ;

<sup>2</sup>Howard Hughes Medical Institute, Departments of Medicine and Biochemistry, Duke University Medical Center, Durham, North Carolina 27710, USA ; <sup>3</sup>Contraintes team,

<sup>4</sup>Sisyphe team, INRIA Paris-Rocquencourt, BP 105, 78153 Le Chesnay, France.

**Key Words:** computational modeling, parameter optimization, signaling networks, seven-membrane-spanning receptor,  $\beta$ -arrestin, G protein-coupled receptor kinases.

**Running Title:** A computational model for ERK activation by AT<sub>1A</sub>R.

#To whom correspondence should be addressed

Eric Reiter, UMR 6175, 37380, Nouzilly, France.

Telephone: +33 2 47 42 77 83, Fax: +33 2 47 42 77 43

E-mail address: [Eric.Reiter@tours.inra.fr](mailto:Eric.Reiter@tours.inra.fr)

\* DH and GD contributed equally to this work

## SUMMARY

Seven-transmembrane receptors (7TMRs) are involved in nearly all aspects of chemical communications and represent major drug targets. Classically, 7TMRs transmit their signals *via* heterotrimeric G proteins. More recently, 7TMRs have also been shown to signal independently of G proteins through a family of adaptor proteins named  $\beta$ -arrestins, whose recruitment to the activated receptor is regulated by G protein-coupled receptor kinases (GRKs). Interestingly, the extracellular-signal-regulated kinases (ERK) are activated by both mechanisms but display distinct temporal (early *vs* late) and spatial (nuclear *vs* cytosolic) characteristics. To decipher the molecular mechanisms responsible for these differences as well as the hidden dynamics governing ERK activation by the angiotensin II 1A receptor (AT<sub>1A</sub>R), we combined experimental approaches with computational modeling. We built an abstracted, ODE-based model that captured the available knowledge and experimental data. Taking advantage of a newly established quantifier free linear time logic (QFLTL) approach, we inferred the unknown parameters by fitting data (*i.e.*: ERK phosphorylation, DAG production and PKC activity kinetics) obtained in HEK 293 cells transfected with AT<sub>1A</sub>R. In order to increase the constraints on the parameter learning process, both control and perturbed (*i.e.*: selective depletion of the different  $\beta$ -arrestins or GRKs using siRNA-mediated knock-down and PKC blockade using a chemical inhibitor) data were simultaneously fitted. The model predicts and experiments confirm that, for the AT<sub>1A</sub>R expressed in HEK293 cells: i) GRKs control the balance between G protein- and  $\beta$ -arrestin-dependent ERK activation by phosphorylating the C-terminal region of the activated receptor; ii) GRK2 not only mediates desensitization of G protein activation but also exerts a strong restraining influence on  $\beta$ -arrestin signaling; iii) GRK 5 or 6 exert little effect on G protein-stimulated ERK but are required for  $\beta$ -arrestin-mediated ERK activation; iv) the  $\beta$ -

arrestin-dependent ERK pathway undergoes both activation and deactivation through amplified enzymatic processes. Hence we demonstrate that, by virtue of their opposing effects on the activation of  $\beta$ -arrestin-dependent signaling, GRK2/3 and 5/6 regulate switching between the G protein and  $\beta$ -arrestin pathways as well as their distinct dynamics. Our results underline the value of using computational modeling to decipher the complex signaling mechanisms elicited by 7TMRs.

Abbreviations: 7TM, seven-transmembrane-spanning; siRNA, small interfering RNA; AT<sub>1A</sub>R, angiotensin II type 1A receptor; AngII, angiotensin II; ERK, Extracellular signal-regulated kinase; MAPK, mitogen-activated protein kinase; PKC, protein kinase C; GRK, G protein-coupled receptor kinase.

## INTRODUCTION

Seven transmembrane-spanning receptors (7TMRs) regulate nearly every known physiologic processes in mammals. They represent the largest class of cell surface receptors and are the target of up to 40% of current pharmaceuticals (1). Signaling by 7TMRs is classically mediated by receptor coupling to heterotrimeric G proteins which activate a variety of effectors, including second messengers and the mitogen-activated protein kinase (MAPK) cascades (2). G protein coupling is rapidly hampered by a two step process which begins with phosphorylation of the agonist-occupied receptor by G protein-coupled receptor kinases (GRK) (3). Cytoplasmic  $\beta$ -arrestins are subsequently recruited to the GRK-phosphorylated receptor and play key roles in both receptor desensitization and internalization (3-7).

Beside this classical paradigm, a growing body of evidence has revealed that  $\beta$ -arrestins also serve as signal transducers and adaptors (8, 9). So far, the best characterized signaling mechanism to be stimulated by  $\beta$ -arrestins is the extracellular signal-regulated kinase (ERK) signaling pathway.  $\beta$ -arrestins have been shown to scaffold components of this MAPK cascade (*i.e.*: Raf-1, MEK1 and ERK) in complexes with receptors, leading to activation of ERK (10-13). Inhibition of  $\beta$ -arrestin 2, using specific small interfering RNA (siRNA), impairs the angiotensin II type 1A receptor (AT<sub>1A</sub>R)-stimulated ERK signaling, while  $\beta$ -arrestin 1 depletion enhances AT<sub>1A</sub>R-mediated ERK activation (14, 15). Spatial distribution of ERK activated by the G protein and the  $\beta$ -arrestin-dependent mechanisms is distinct: the G protein-dependent pathway triggers the nuclear translocation of pERK while the  $\beta$ -arrestin 2-activated ERK is confined to the cytoplasm (11, 12, 16-18). The two pathways also have distinct activation kinetics: the rapid and transient G protein-dependent activation switches to the slower but persistent  $\beta$ -arrestin 2-mediated activation (17).

Interestingly, transient and sustained ERK activation have been shown to regulate cell fates such as growth and differentiation (19). In addition, the  $\beta$ -arrestin 2-dependent ERK pathway can be activated independently of G proteins as shown using the (DRY/AAY) mutant of AT<sub>1A</sub>R or a modified angiotensin II peptide (SII) (20).

Noteworthy, different GRK subtypes have been shown to play specialized regulatory functions for the two signaling mechanisms: activation of the  $\beta$ -arrestin 2-dependent ERK pathway by the AT<sub>1A</sub>R (21), V2 vasopressin (22),  $\beta_2$  adrenergic (23) and follicle-stimulating hormone (24) receptors specifically requires GRK5 and GRK6 action while second messenger generation by both V2 vasopressin (22) and H1 histamine (23) receptors is negatively regulated by GRK2 but unaffected by GRK5 or 6.

A number of studies have provided mathematical models of ERK activation by EGF (25-28), including the dynamics of transient and sustained ERK activation (19, 29). ERK response to a G protein-coupled 7TMR in yeast and mammals have also been modeled (30, 31). However, no mathematical model has been developed to address GRKs-mediated regulation of ERK activation by a 7TMR.

In this study, we used an integrated approach combining computational modeling, parameter optimization by fitting the models to dynamic data sets, *in silico* kinetic simulation and experimental validation. In order to limit data heterogeneity, we only considered data collected in a unique cellular model (i.e.: HEK293 cells transfected with the AT<sub>1A</sub>R). Indeed, signaling dynamics are well known to differ between cell lines as the expression levels of the signaling machinery's components are not comparable. On the computational side, we established and iteratively refined a simplified network's topology capturing the available knowledge and experimental data. An ODE-based dynamical model was achieved by representing interactions and reactions as either mass action laws or Michaelis-Menten kinetics. Some parameters were taken from the literature or were measured experimentally;



most of them however were determined by using a newly established parameter optimization approach combining quantifier free linear time logic (QFLTL) with covariance matrix adaptation evolution strategy (CMAES) (32). The model structure was reassessed back and forth until a satisfactory solution was reached. We validated the optimized model using data from the literature which had not been integrated to the parameter learning process. We also challenged the model predictions with original experimental data.

## RESULTS

### Construction of a minimal model for ERK activation by the AT<sub>1A</sub>R

Although some aspects of ERK activation by the AT<sub>1A</sub>R are understood, the mechanisms controlling the balance and distinct kinetic properties of the G protein- and  $\beta$ -arrestin-dependent pathways activating ERK remain largely unknown. To address this question, we have developed a deterministic dynamical model made of a system of ordinary differential equations (ODE).

First, we integrated the disparate knowledge available on the molecular mechanisms linking AT<sub>1A</sub>R activation to ERK phosphorylation into a formal network. However, attempts to consider all distinct biochemical species and interactions had previously been demonstrated to lead to combinatorial complexity associated with a dramatic increase in the number of unknown parameters and therefore to be impractical for dynamical simulation (33-35). Therefore, we chose to construct a basic minimal model encompassing the G protein- and  $\beta$ -arrestin-dependent transduction mechanisms while keeping key molecules and experimentally measurable variables. We generated an initial version of the minimal basic model amenable to numerical simulations and optimization of the unknown parameters by training with experimental data. Importantly, we found out that with the initial network structure, suitable fitting of the model to the data through the parameter optimization process was impossible. Moreover, as is generally the case in biological systems, the knowledge available on the signaling network was incomplete and a number of competing hypotheses existed. We took advantage of this situation by exploring a variety of model structure in order to minimize discrepancies between the fitted model and the experimental data. Importantly, the main structural predictions that arose during this process were experimentally validated

(see below). The final model resulting from this iterative process is presented in Systems Biology Graphical Notation (SBGN, CellDesigner) (Figure 1) (36, 37). The model consists of 28 kinetic reactions and 18 variables (Supplementary figure 1).

For simplification purpose and because the envisioned simulation horizon was rather short (*ie.*: 0 to 90 min) compared to these processes *in vivo* half-lives, neither degradation nor synthesis was taken into account in the model. This assumption led to 7 conservation laws (Supplementary figure 2) which significantly simplified the model. We used mass action laws throughout the model with the exception of GRK activities which were controlled by Michaelis-Menten kinetics.

*Classical G-protein coupling.* Signal transduction is known to be initiated by hormone (*ie.*: angiotensin) binding to the receptor. In the model, we aggregated this process in a single variable called **HR**. The hormone-receptor complex then couples to heterotrimeric G protein, alpha q subunit is activated (**G\_a**) by exchanging GDP with GTP and is subsequently released (3, 38-40). Since this process is rapid, transient and dependent upon receptor activation, in the model we expressed the whole heterotrimeric G protein activation/deactivation process as **G** being reversibly activated in **G\_a** under the catalytic control of **HR**. Since GRK-induced phosphorylation of the receptor only partially quenches G protein activation, **HRP1** also catalyses **G\_a** formation in the model (9). It is also well documented that 7TMRs induce G protein activation constitutively in the absence of agonist (39). We took this process into account in the model by adding the activation of **G** into **G\_a** in a parallel reaction independent of either HR or HRP1 catalysis.

*Desensitization and recycling of the receptor.* Complete desensitization is achieved upon  $\beta$ -arrestin binding to the GRK-phosphorylated receptor (9). In the model, **HRP1** interacts and forms a complex with either  $\beta$ -arrestin 1 (**barr1**) or  $\beta$ -arrestin 2 (**barr2**). Those two complexes are internalized; the receptor is then either degraded or dephosphorylated and

recycled back at the plasma membrane (41). As already explained we did not consider synthesis/degradation in the model. The multiple dephosphorylation/recycling steps were shortened in one single irreversible reaction departing from each complex and providing **barr1**, **barr2** and **HR** to the system.

*G protein-dependent signaling to ERK.* Once activated, G alpha q protein induces a second messenger signaling cascade sequentially involving PLC activation, IP3 and DAG accumulation from PIP2, calcium release and PKC activation (40). Since we were able to measure DAG accumulation and PKC activity, we included these variables in the model: **PIP2** was transformed into **DAG**, this reaction being indirectly catalyzed by **G\_a**. In turn, **DAG** catalyzed **PKC** activation in **PKC\_a**. Both DAG accumulation and PKC activation have been shown to be deactivated through complex mechanisms (42, 43), as a consequence both reactions were made reversible in the model. Activated PKC triggers the activation of many downstream targets including Ras and the ERK mitogen-activated protein kinase signaling module: Raf, Mek, ERK (20, 44, 45). ERK phosphorylated through G protein-dependent mechanisms was noted **GpERK**. Therefore, in the model, **PKC\_a** catalyzes **GpERK** formation from **ERK**. This reaction is reversible as a result of ERK dual phosphatase (46).

*$\beta$ -arrestin-dependent signaling to ERK.* Independently of G proteins,  $\beta$ -arrestin 2 contributes to ERK phosphorylation through the formation of a multiprotein complex with the activated GRK-phosphorylated receptor, Raf and Mek (9, 47). Consistently, in the model, a distinct GRK-phosphorylated form of the receptor **HRP2** associates with  $\beta$ -arrestin 2 (**barr2**) to form the complex **HRP2barr2**. **HRP2barr2** then promotes the reversible formation of  $\beta$ -arrestin-dependent phosphorylated ERK (**bpERK**) from the cellular pool of **ERK**. **GpERK** and **bpERK** were distinguished in the model since it is well documented that they exhibit distinct kinetics, subcellular localization and downstream targets (47, 48).

During the optimization process, we found that this  $\beta$ -arrestin-dependent pathway for ERK activation needed to be negatively regulated in order to fit the biological data. As a consequence, we made all the reactions leading to **bpERK** formation reversible. In addition, based on the actual quantities of receptors ( $0.099 \mu\text{mol.L}^{-1}$  (17, 21)) of  $\beta$ -arrestin 2 ( $0.513 \mu\text{mol.L}^{-1}$  (14)) and of phosphorylated ERK ( $1.265 \mu\text{mol.L}^{-1}$  (49)) measured in HEK293 cells, ERK phosphorylation within the  $\beta$ -arrestin scaffold could not work with a 1:1 stoichiometry. Therefore, we hypothesized the existence of an enzymatic amplification within the  $\beta$ -arrestin scaffold. Using a recently validated reverse phase protein array method (49), we were able to experimentally test this hypothesis. We measured the molar quantities of phosphorylated Mek and of phosphorylated ERK in the whole cell as well as within the  $\beta$ -arrestin 2 scaffold (*i.e.*: Ro-31-8425-treated cells in order to disrupt **GpERK**)(Figure 2A). In all cases we found substantially more phosphorylated ERK than phosphorylated Mek (5 to 36 fold, depending on the conditions, Figure 2B). With these data in hand, we switched from the equimolar scaffold representation (Figure 2C) to enzymatic catalysis of **bpERK** formation (Figure 1 and 2D).

*GRK-independent  $\beta$ -arrestin signaling.* Although GRK phosphorylation is well known to significantly enhance  $\beta$ -arrestins' affinity for the receptor (3), the main driving force for  $\beta$ -arrestin recruitment and activation is the agonist-induced transconformation of the receptor itself (9, 50). With that in mind and in order to provide more flexibility to the **bpERK** pathway, we made the assumption that  $\beta$ -arrestin 2 could be active even in the absence of GRK phosphorylation. We tested this hypothesis by comparing wild-type  $\text{AT}_{1\text{A}}\text{R}$  (WT) with a mutant (13A) in which all the 13 serines and threonines present in the C-tail of the receptor were replaced by alanines (Figure 3A). When expressed in HEK293 cells at similar levels, the two receptors behaved differently. In control conditions, the 13A triggered significantly more ERK phosphorylation than the WT, which possibly reflect the lack of

GRK-mediated desensitization (Figure 3B, Supplementary figure 4). As expected, siRNA-mediated  $\beta$ -arrestin 2 knock-down led to a dramatic decrease in ERK phosphorylation by the WT, especially at late time points (Figure 3C, Supplementary figures 3A and 4B).

Importantly, using the 13A,  $\beta$ -arrestin 2 depletion also led to a very significant inhibition of the late ERK response, strongly supporting our initial assumption that the non-phosphorylated receptor could still activate **bpERK** (Figure 3D, Supplementary figure 4C). As a consequence, we added a reversible complexation reaction between **HR** and **barr2** to form **HRbarr2**, which catalyses **bpERK** formation (Figure 1).

*GRK isoforms act differentially at the receptor.* Upon receptor activation, GRK2 and GRK3 are translocated to the plasma membrane where they form a complex with the free  $\beta\gamma$  subunits of heterotrimeric G protein, making these kinases' activation dependent on G protein  $\alpha/\beta\gamma$  subunits dissociation. On the other hand, GRK5 and GRK6 are constitutively bound to the plasma membrane and can therefore interact with and phosphorylate activated receptors independently of heterotrimeric G protein (3). Noteworthy, GRK2/3 and GRK5/6 have been shown to play specialized regulatory functions on signaling by various 7TMRs including the AT<sub>1A</sub>R: activation of the  $\beta$ -arrestin 2-dependent ERK pathway specifically requires GRK5 and GRK6 action while second messenger generation is negatively regulated by GRK2, but is unaffected by GRK5 or 6 (51) whereas GRK2 and GRK3 negatively impact on ERK phosphorylation (21, 22, 24). These four GRK isoforms have also been shown to participate in receptor phosphorylation, with GRK2 contributing the most (21-24). However, it was not known from the literature whether GRKs regulated  $\beta$ -arrestin-dependent ERK activation by phosphorylating the activated receptor or whether they acted on an unknown intermediate. This question was central to the construction of our model, especially since our experimental data showed that  $\beta$ -arrestin signaling also occurs in the absence of receptor phosphorylation in the C-tail (Figure 2). In order to sort out the two possibilities, we compared WT and 13A

AT<sub>1A</sub>R for their sensitivity to siRNA-mediated GRK depletion (Supplementary figure 3B, C, D). We reasoned that if GRKs acted on an intermediate, their depletion would be detrimental to ERK activation in cells expressing either the 13A or the WT AT<sub>1A</sub>R. On the contrary, if they acted by phosphorylating the receptor, they would lose their effect only in cells expressing the 13A. As expected, GRK2 and GRK5 differentially regulated ERK response by the WT receptor (Figure 4A, B, Supplementary figure 5A, B). In cells expressing the 13A however, GRK2 depletion was unable to enhance ERK phosphorylation whereas GRK5 knock down did not modify the ERK response (Figure 4C, D, Supplementary figure 5C, D). This experiment allowed us to better delineate GRKs' action and further supported the previously proposed concept of GRK2/3 and GRK5/6 phosphorylating distinct residue combinations on activated 7TMRs (21, 22, 24, 51). In the model, **HRP1** formation is catalyzed by **GRK23** and **HRP2** formation by **GRK56** using Michaelis-Menten kinetics, with **HRP1** and **HRP2** referring to receptor phosphorylated on different serine and/or threonine residue combinations. Since the available data (21) did not allow us to discriminate GRK2 from GRK3 or GRK5 from GRK6, we aggregated these kinases into two variables: **GRK23** and **GRK56** (Figure 1).

### Parameter estimation

The model comprises a total of 56 unknown parameters whose values had to be determined. Some of the model's parameters were directly taken from actual biological measurements, assuming the cell volume as 1pL (29, 52). Concentrations of expressed AT<sub>1A</sub>R (17, 21),  $\beta$ -arrestin 1 and 2 (14) as well as of total and phosphorylated ERK (49) were calculated from published experimental measurements made in HEK293 cells. In addition, a number of initial quantities were determined by making assumptions based on the available

knowledge and data. For instance, the initial amount of GRK-phosphorylated receptor (**HRP1** and **HRP2**) has been set to 0 (21) (the 16 experimentally and empirically determined parameters are listed in red in Supplementary figure 6).

As 40 parameters were experimentally unreachable, we optimized the missing parameters' value by training the model with experimental data. We used kinetic data on three variables of the model: previously published angiotensin-induced ERK phosphorylation data (21), DAG accumulation and PKC activity measured in real-time in AT<sub>1A</sub>R expressing HEK293 cells by using relevant FRET sensors (53). DAG and PKC data were available in control conditions (*i.e.*: 90 min stimulation with 100 pM AngII) whereas ERK data were obtained in control as well as in perturbed conditions (*i.e.*: 90 min stimulation of cells starved and depleted in either  $\beta$ -arrestin 2, GRK2, GRK3, GRK5 or GRK6 using siRNA as well as cells starved and treated with the PKC inhibitor Ro-31-8425). The ERK data obtained in GRK2 and 3 or GRK5 and 6-depleted cells were averaged to match with the aggregate variables used in the model.

To infer the unknown (kinetic and quantitative) parameter values from the experimental data obtained under various conditions, we expressed the objective function to be minimized in temporal logic properties with quantifier free LTL(R) formulae (QFLTL(R)) (32). This formalism allowed us to define, in addition to curve fitting, a number of constraints which were integrated to the objective function: i) each parameter value had to be within 0 and 200; ii) **HRP2** recruited at least 5 times more **barr2** than **HR** (parameter 41 > 5 \* parameter 40); iii) **HRP1** recruited more **barr1** than **barr2** (parameter 26 > parameter 27) (14); iv) inhibited parameter < control parameter (siRNA-depletions of **GRK23**, **GRK56** and **barr2** and **PKC** inhibition); v) constant boundaries (as defined by standard deviations) imposed for DAG (between 5 and 35 min) and PKC (between 20 and 80 min) values instead



of all the discrete time points and vi) basal levels of phosphorylated ERK were low (if  $\mathbf{HR} = 0$  then  $1\% < \mathbf{GpERK} + \mathbf{bpERK} < 6\%$ ).

We combined these QFLTL(R) constraints with the Covariance Matrix Adaptation Evolution Strategy CMAES, a previously validated non-linear optimization technique (54). As we experimentally measured the kinetics of only 3 variables out of 18 present in the model, the calibrated parameters were likely to be non-identifiable (55, 56). In order to avoid this problem, we developed a global data fitting approach in which multiple data, resulting from either control or perturbed biological conditions were fitted simultaneously to the model (Figure 5). We indeed observed that, by including additional experiments with different read-outs and known biological perturbations, we greatly reduced the degrees of freedom allowed to the parameter optimization process. Consistent with that view, when starting from random parameter values, the optimization algorithm converged to good fitting parameter sets in only 5 % of the cases. The whole parameter optimization process was carried out in the BIOCHAM modeling environment (see Supplementary figure 7 for the model in BIOCHAM code) (57-59).

This optimization strategy led us to identify a set of parameter values (in black in Supplementary figure 6) which allowed model simulations to closely match the experimental data (Figure 6). Interestingly, the values found for the perturbed parameters were dramatically diminished (more than 99 % inhibition) in all cases except GRK23 for which the reduction was only about 50 % of the control value (Supplementary figure 6, compare parameter 19 and 52). This result suggests that, unlike GRK5 and GRK6, GRK2 and GRK3 exert additive non-redundant effects on  $\beta$ -arrestin-induced ERK *in vivo*. Although the model with this parameter set could reproduce the data pretty accurately, further validations were necessary.

### **Model validations against independent experimental data sets.**

To further validate our model, it was challenged against published data sets which had not been used in our parameter estimation. First, using HEK293 cells expressing the AT<sub>1A</sub>R, Kim et al. reported that GRK2 or 3 overexpression led to decreased ERK phosphorylation upon angiotensin stimulation (21). Remarkably, the observed pattern was reminiscent of the one obtained in GRK5 or 6-depleted cells. Oppositely, GRK5 or 6 overexpression caused an increase in phosphorylated ERK at all time points similar to what was observed in GRK2 or 3-depleted cells (21). When increases in GRK23 (Figure 7A) or GRK56 (Figure 7B) levels were simulated, the model predicted phosphorylated ERK patterns which matched fairly well those experimentally obtained.

Next, we took advantage of another interesting set of data, generated by Ahn et al. in the same experimental conditions (17). Using either a mutant version of AT<sub>1A</sub>R (DRY-AAAY) or a biased ligand (SII), these authors showed that, in the absence of G protein coupling, ERK was still triggered and displayed kinetics virtually identical to the ones obtained in cells treated with the PKC inhibitor Ro-31-8425. We therefore inhibited G proteins in the model (Figure 7C) and that resulted in simulated levels of phosphorylated ERK very similar to PKC inhibition shown Figure 6B and in excellent agreement with the data reported by Ahn et al. (17).

Finally, another paper reported that, in the case of AT<sub>1A</sub>R,  $\beta$ -arrestin 1 acted as a dominant negative of  $\beta$ -arrestin 2 for ERK phosphorylation upon angiotensin stimulation (14). By simulating the effects of either  $\beta$ -arrestin 1 or  $\beta$ -arrestin 2 depletion on phosphorylated ERK response as a function of angiotensin concentration (Figure 7D), we observed effects very close to those previously reported (14).

Interestingly, our simulations provide quantitative estimations of the degrees of the different inhibition, depletion or over-expression tested.

### **Sensitivity analysis.**

We carried out a sensitivity analysis in order to test the model robustness. We examined how systematic perturbations of each of the optimized parameters affected the response of 5 different species: DAG, PKC, GpERK, bpERK and total pERK. The sensitivities were defined as the modification of the quantity of a given species (integrated for 90 min and expressed as a percentage of the initial value) caused by a 1 % change of a given reaction rate (29). Parameters 1 to 48 were sequentially decreased by 1 % and the impacts on each of the 5 variables were calculated and plotted (Figure 8). Positive value indicates that the quantity increases with a reaction rate decrease, while negative value indicates that the quantity decreases. Large variations, whether positive or negative, indicate potentially significant control mechanisms. This sensitivity analysis suggests that the model is relatively robust since the percentage of change in the analyzed species was always below  $\pm 1$  %. As expected, DAG, PKC and GpERK were more sensitive to variations in the parameters controlling G protein signaling than in those controlling  $\beta$ -arrestin signaling. Symmetrically, bpERK appeared tightly controlled by  $\beta$ -arrestin signaling while variations in G protein signaling had virtually no effect, outside of the number of occupied receptors. Consistent with that, total pERK was controlled by both pathways. Interestingly, a 1% variation in GRK23  $V_m$  or  $K_m$  (parameters 24 and 25) led to more important perturbation on bpERK than on GpERK, revealing that GRK23 exerts a strong negative effect on bpERK. These observations suggest that GRK23, rather than being restricted to the desensitization of G protein signaling as classically thought, might exert an equally important dampening action

on  $\beta$ -arrestin signaling through its competition with GRK56 for receptor phosphorylation. Since this sensitivity analysis only considered small perturbations, the possibility existed that the predictions would not hold for larger perturbations. Therefore, we evaluated how depletions in GRK23 and GRK56 levels would affect the ERK response.

### **Pivotal role of GRK23 in the control of ERK activation by $\beta$ -arrestin 2.**

When GRK23 was depleted, our model predicted an unexpected behavior of the signaling network: second messenger (DAG) (Figure 9A) and total pERK (Figure 9C) both increased as compared to control conditions. The conventional wisdom would have explained the increased ERK response as the reflection of a lack of G protein desensitization in the absence of GRK23. However, total pERK was minimally affected by PKC blockade. The simulations clearly predict that the contribution of GpERK is decreased instead and that bpERK activation is massively amplified. Importantly, these predictions were corroborated by experimental data. To reflect second messenger response, inositol uptake was measured in the presence of increasing concentrations of angiotensin. As in the simulation, second messenger accumulation was higher in GRK2-depleted cells than in control (Figure 9B). Moreover, we observed that ERK phosphorylation was strongly increased and that the PKC inhibitor Ro-31-8425 had very limited effect in the absence of GRK2 (Figure 9D, Supplementary figure 8). Although we were not able to achieve the double GRK2 and GRK3 siRNA-mediated depletion experimentally, the experimental observations matched the model simulations with good accuracy.

To explore the possibility that GRK23 exerts a strong negative effect on the  $\beta$ -arrestin-dependent pathway through its competition with GRK56 for receptor phosphorylation, we simulated the combined effect of GRK56 depletion and PKC inhibition

on ERK phosphorylation. At short stimulation time, GRK56 depletion had very little effect on the amount of ERK phosphorylation as compared to control condition. However, PKC inhibition combined with GRK56 depletion led to complete abolition of ERK response whereas significant, albeit reduced, ERK response was predicted by the model in PKC inhibited cells (Figure 10A). Again, these model simulations were closely corroborated by the experiments. We observed that ERK phosphorylation upon exposure to increasing doses of angiotensin was similar in control and in GRK5 or GRK6-depleted cells. Moreover, as suggested by the simulations, treatment with the PKC inhibitor Ro-31-8425 completely blocked ERK phosphorylation in GRK5 or GRK6-depleted cells while leading to only partial inhibition in control cells (Figure 10B and C, Supplementary figure 9). These data clearly indicate that GRK5 or 6 depletion switches ERK signaling exclusively towards the G protein-dependent transduction mechanism. Interestingly, siRNA-mediated depletion of either GRK5 or 6 is sufficient to trigger complete switching of the ERK signal. This suggests that the two kinases may have to phosphorylate a receptor on distinct residues (or combination of residues) in order to impart  $\beta$ -arrestin-mediated signaling competence to this receptor.

## **DISCUSSION**

Seven transmembrane domain receptor research is challenged with the question of how G protein and  $\beta$ -arrestin-dependent transduction mechanisms coordinate their actions to activate common downstream signaling intermediates with distinct temporal and spatial patterns. In this study, we combined computational modeling and experiments to address this question using the well documented dual ERK phosphorylation by AT<sub>1A</sub>R as a paradigm. We brought together previously published experimental data with our own new data and we assembled an original model inference and parameter optimization strategy to develop a

dynamical model. We chose to build a core model rather than a detailed mechanistic model encompassing all the molecular details because it had already been established that such an approach led to combinatorial increase of complexity which cannot currently be computationally handled (33-35). This model construction and analysis led to new insights into the regulation of ERK by the AT<sub>1A</sub>R.

A common limitation when modeling intracellular signaling networks is that the number of measurable read-outs is very limited with respect to the number of molecular intermediates actually involved in the pathways to be modeled (56, 60). This raises problems of identifiability since many of the parameters have to be determined through an optimization process dependent on the available experimental data (56). Therefore, a limited number of read-outs imply a relative lack of constraint on the parameter optimization process which in turn precludes the convergence of parameter sets separately optimized. In the present study, since we were limited to three read-outs (*i.e.*: pERK, DAG and PKC), we had to solve the model identifiability problem in an innovative fashion. We achieved a global fitting strategy which encompassed both control conditions and a number of targeted inhibitions within the network. We found out that, using this approach, the constraints on the parameter search were significantly higher. In addition, we took advantage of the BIOCHAM capabilities in LTL to combine the curve fitting algorithm with a number of qualitative constraints further decreasing the degrees of freedom. Altogether, this strategy allowed us to parameterize a model of 18 variables by actually measuring only three of them.

As anticipated, having the ability to simulate a dynamical model proved to be very efficient to decipher our signaling system and its key regulatory features. However, we also realized that the model-building process itself was very insightful and led to important findings on the network structure. More specifically, we found out that the goodness of fit resulting from the parameter optimization process was a very efficient way to rank different

network structures and to ultimately infer a model from prior knowledge and data. This is important since model inference remains a major challenge in systems biology (61, 62).

Other than AT<sub>1A</sub>R, a growing number of 7TMRs have been reported to activate ERK phosphorylation through dual G protein/ $\beta$ -arrestins transduction mechanisms which share many attributes with the AT<sub>1A</sub>R-dependent signaling network modeled in this study (22, 24, 48, 63-65). One aspect which seems to vary from one receptor to another is the role of  $\beta$ -arrestin 1. Some receptors require both  $\beta$ -arrestin isoforms to activate ERK while for others, including AT<sub>1A</sub>R,  $\beta$ -arrestin 1 antagonizes the action of  $\beta$ -arrestin 2 (14). Despite this, we believe that the structure of our core model could be easily adapted to other 7TMRs and/or other cellular contexts, granted that the parameters would have to be optimized in each case using proper data sets.

The main feature revealed by our modeling approach is that ERK activation by AT<sub>1A</sub>R is tightly controlled by the antagonistic actions of the two GRK subfamilies, namely GRK2 and 3 on one hand and GRK5 and 6 on the other. Our simulations and experimental data suggest that these kinases operate the major control gate regulating ERK activation in time and space. In the absence of GRK5 or 6, there is virtually no  $\beta$ -arrestin-dependent ERK activation whereas, more surprisingly, in the absence of GRK2 or 3,  $\beta$ -arrestin-dependent ERK activation clearly overwhelms G protein-mediated ERK. In the absence of GRK action, both pathways co-exist but the potential for regulation is much reduced. This reveals a new and unexpected function of GRK2 and 3: beyond their well documented role in G protein desensitization, these kinases strongly dampen  $\beta$ -arrestin signaling. Eventually, it is the interplay between GRK2,3 and 5,6 which defines the duration and intensities of each transduction mechanism. We propose that GRKs endow ERK activation (and potentially other signaling pathways) by 7TMRs with adaptability. The balance in the action of both GRK subfamilies will vary with the receptor (*i.e.*: different phosphorylation patterns present

in the receptors' C-terminus), the cell type and the cellular context (*i.e.*: differences in stoichiometries and subcellular localizations).

The fact that GRK2 and 3 on one hand and GRK5 and 6 on the other are both required for the regulatory process to operate emphasizes the importance of tightly controlling the balance between the G protein and the  $\beta$ -arrestin-dependent transduction mechanisms. Importantly, this is consistent with the mounting evidence that they both lead to distinct biological outcomes (47, 48).



## MATERIALS AND METHODS

*Materials---* Radiolabeled [<sup>125</sup>I] Tyr<sup>4</sup>-angiotensin II (AngII) was obtained from Perkin Elmer Life Sciences (Boston, MA). Ro-31-8425 was purchased from Calbiochem (Darmstadt, Germany). GeneSilencer reagent was from Gene Therapy Systems (San Diego, CA). All other reagents were purchased from Sigma (St. Louis, MO). Expression plasmid encoding the hemagglutinin epitope-tagged (HA-) wild type rat AT<sub>1A</sub>R has been previously described (20). All the 13 Ser and Thr residues present in the C-tail of the rat AT<sub>1A</sub>R have been mutated to alanine (13A) using the QuickChange site-directed mutagenesis kit (Stratagene).

*Synthesis of small interfering RNAs (siRNAs)---* Chemically synthesized double-stranded siRNAs, with 19-nt duplex RNA and 2-nt 3' dTdT overhangs were purchased from Dharmacon (Lafayette, CO) or Xeragon (Germantown, MD) in deprotected and desalted form. The siRNA sequences used in the present study to target the different human  $\beta$ -arrestins and GRKs have been previously validated (13, 21, 22, 24):  $\beta$ -arrestin 1, 5'-AAAGCCUUCUGCGCGGAGAAU-3' (positions 439-459 relative to the start codon);  $\beta$ -arrestin 2, 5'-AAGGACCGCAAAGUGUUUGUG-3' (positions 148-168); GRK2, 5'-AAGAAGUACGAGAAGCUGGAG-3' (positions 268-288); GRK 5, 5'-AAGCCGUGCAAAGAACUCUUU-3' (positions 406-426); GRK6, 5'-AACAGUAGGUUUGUAGUGAGC-3' (positions 724-744). A non-silencing RNA duplex (5'-AAUUCUCCGAACGUGUCACGU-3') was used as a control.

*Cell Culture and RNA Transfection---* Human embryonic kidney (HEK)-293 cells were cultured in minimum Eagle's medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. HEK-293 cells were transiently transfected as described (21, 24).

Briefly, 30-40% confluent, slow growing early passage (< 10) cells in 100-mm dishes were transfected simultaneously with 20  $\mu\text{g}$  of siRNA and plasmid encoding the AT<sub>1A</sub>R (wild-type and PKC-: 0.5  $\mu\text{g}$ ; 13A: 2.0  $\mu\text{g}$ ) using the GeneSilencer Transfection reagent (Gene Therapy System, San Diego, CA). Forty-eight hours after transfection, cells were divided into poly-D-lysine-coated 12-well plates (Becton Dickinson Labware, Bedford, MA) for receptor binding, or 12-well plates to prepare cell extracts. AT<sub>1A</sub>R expression was determined by radioligand binding assays, as described previously (66), and was 300-600 fmol/mg of protein in all experiments. In the GRK siRNA experiments, the amount of plasmid transfected was adjusted to obtain equivalent receptor levels at the plasma membrane than in the controls.

*Preparation of Cell Extracts and Immunoblotting---* Three days after transfection, cells were starved for at least 6 h in serum-free medium prior to stimulation. After stimulation, cells were solubilized in 2X SDS-sample buffer (pH 6.8), followed by sonication or boiling. Equivalent amounts of proteins were separated by SDS/PAGE on Tris-glycine polyacrylamide gels (Invitrogen), transferred to nitrocellulose, and immunoblotted with rabbit polyclonal antibodies. The phospho (1:2,000 dilution) and total (1:6,000 dilution) ERK antibodies were from Cell Signaling Technology (Beverly, MA) and Upstate (Charlottesville, VA), respectively. Endogenous  $\beta$ -arrestins and GRKs were detected as described (13, 21). Chemiluminescent detection was performed using the SuperSignal West Pico or West Femto reagent (Pierce, Rockford, IL) and were quantified by densitometry with a Fluor-S MultiImager (BioRad, Hercules, CA) or with ImageMaster 1D Elite version 4 software (Amersham Biosciences, Pittsburgh, PA).

*FRET---* HEK-293 cells expressing the AT<sub>1A</sub>R (600 fmol/mg of protein) were transiently transfected with either DAG (DAGR) or PKC (CKAR) plasmid-encoded FRET sensors (53). Twenty-four hours after transfection, cells were split and seeded on imaging dishes (BD Biocoat, BD Biosciences, San Jose, CA, USA). Twenty-four hours later, cells

were serum-starved for 90 min and imaged in the dark on a 37 °C temperature-controlled stage. Cells were imaged on a Leica DM IRBE (Leica Microsystems, Wetzlar, Germany) microscope with a CoolSnap fx cooled charge-coupled device camera (Roper Scientific, Ottobrunn, Germany) controlled by METAFLUOR 7.5 (Universal Imaging Corporation, Downingtown, PA). Dual emission ratio imaging used a 436DF10 excitation filter, a 436-510 DBDR dichroic mirror, and 480AF30 and 550AF30 emission filters for CFP and YFP, respectively. The filters were alternated by a Lambda 10-2 filter changer (Sutter Instruments, Novato, CA). Exposure time was 100-500 ms, and images were taken every 15 s. Typically, 15 sensor-positive cells and 5 non-specific areas were chosen in the field of the microscope. Base-line signals at the wavelength corresponding to YFP and CFP were recorded for 5 min before the hormone was added to the dish. Signals were recorded for the next 90 min. Fluorescent image background corrections were carried out by subtracting the intensity of non-specific areas. The FRET *ratios* were calculated for each individual cell. The means +/- SEM corresponding to individual cells from 3 different plates were plotted.

*RPPA*--- Serial dilutions of phosphorylated human recombinant ERK2 (Proteinkinase, Biaffin GmbH & Co KG, Germany) and of phosphorylated human recombinant MEK1 (Millipore, Billerica, MA) were prepared in RPPA lysis buffer. For cell lysates (1.7 mg/ml concentration in average), three two-fold serial dilutions were done in RPPA lysis buffer. Samples were printed on nitrocellulose-coated slides (Fast Slides, Whatman, Maidstone, UK) using a 32-pin manual arrayer (Glass Slide Microarrayer, VP478, V&P Scientific, San Diego, CA), immunodetected and revealed in near infrared according to previously reported procedures (49). Anti-pERK (Cell Signaling Technology Inc, Beverly, MA, 1/400), anti-pMEK (Cell Signaling Technology Inc, 1/300) and anti-ERK2 (Santa Cruz Biotechnology Inc, Santa Cruz, CA, 1/1000) antibodies were used. Finally, slides were air-dried by centrifugation at 2,500 rpm for 25 min and scanned at 700 nm with an Odyssey IR imaging

system (LI-COR Biosciences, Lincoln, NE) at a 42  $\mu\text{m}$  resolution. Scanned images of arrays were analyzed with GenePix Pro 6.0 software (Molecular Devices, Sunnyvale, CA) as described (49).

*Inositol Phosphate Determination*--- HEK-293 cells expressing the AT<sub>1A</sub>R were transiently transfected in 100-mm dishes with siRNA targeted against the different GRK subtypes. Two days post-transfection, cells were plated onto poly-D-lysine-coated 12-well plates (BD Biosciences Labware, San Jose, CA). To assay for inositol phosphate production, cells were incubated overnight at 37 °C in labeling medium (2.5  $\mu\text{Ci}$  of myo-[<sup>3</sup>H]inositol in 0.5 ml of minimum Eagle's medium with 10% fetal bovine serum/well). Cells were washed with Hank's balanced salt solution containing 20 mM LiCl for 15 min at 37 °C and then treated with agonist in Hank's balanced salt solution for 10 min. The reactions were terminated by the addition of perchloric acid, and total inositol phosphates were isolated by anion-exchange chromatography on Dowex AG1-X8 columns as described previously (67).




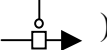

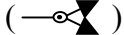


*Statistical analysis*--- Statistical analysis of the data was performed with two-way analysis of variance to compare entire curves (GraphPad Software Inc., San Diego, CA).

*Computational modeling analysis*---

## **ACKNOWLEDGEMENTS**

## FIGURE LEGENDS

**Figure 1:** Core model for ERK activation by the 7TMR AT<sub>1A</sub>R. The Cell Designer program has been used to represent the topology of the network in Systems Biology Graphical Notation (SBGN). The following semantic is used:

- proteins (  );
- active protein (  );
- receptor (  );
- catalysis (  );
- association (  );
- dissociation (  );
- phosphorylation (  );
- Boolean logic gate OR (  ).

Complexes are surrounded by a box. Colored species correspond to experimental read-outs.

**Figure 2:** Quantitative assessment of ERK1/2 and MEK1/2 phosphorylation in HEK293 cells expressing the AT<sub>1A</sub>R using a reverse phase protein array (RPPA) method. **(A)** pERK2 and pMEK1 detection by RPPA. Wide ranges of recombinant pERK2 and pMEK1 dilutions were arrayed in four replicates and revealed with specific primary antibody coupled to a previously published amplified near-infrared revelation system. HEK293 cells expressing the AT<sub>1A</sub>R, pre-treated or not with the Ro-31-8425 PKC inhibitor (1  $\mu$ M), were stimulated with AngII (100 nM) for 2 or 30 min as indicated. Cell lysates were arrayed in 3 dilutions (i.e.: 1/2, 1/4, 1/8). Four independent experiments spotted on the same slide are shown. **(B)** The molar quantities of pERK1/2 and of pMEK1/2 were calculated according to each standard curves and the pERK/pMEK ratios were determined for each conditions. All the

calculations were made in GraphPad Prism. (C) Representation of a AT<sub>1A</sub>R/ $\beta$ -arrestin 2/pERK1/2 equimolar complex in SBGN. (D) Representation of ERK phosphorylation being catalyzed by a AT<sub>1A</sub>R/ $\beta$ -arrestin 2 complex in SBGN.

**Figure 3:** Role of C-tail phosphorylation in ERK activation by the AT<sub>1A</sub>R. HEK-293 cells were transfected with either wild type AT<sub>1A</sub>R (WT) or a mutant AT<sub>1A</sub>R (13A) with the 13 Ser and Thr in the C-tail mutated in Ala. Control or  $\beta$ -arrestin 2-specific siRNAs were transfected simultaneously with the receptors. Three days after transfection, cells were serum-starved for 6 h and stimulated with AngII (100 pM) for increasing periods. (A) C-terminal sequence of the WT (Ser and Thr residues are shown in blue) and of the 13A (mutated residues are shown in red) AT<sub>1A</sub>R (B) Time course of AngII-mediated ERK phosphorylation in WT versus 13A-transfected cells. (C) Effect of siRNA-mediated  $\beta$ -arrestin 2 depletion on ERK phosphorylation in WT AT<sub>1A</sub>R-transfected cells. (D) Effect of siRNA-mediated  $\beta$ -arrestin 2 depletion on ERK phosphorylation in 13A AT<sub>1A</sub>R-transfected cells. Each data point represents the mean  $\pm$  S.E.M. from at least 4 independent experiments. Statistical comparison of the curves was performed by using a two-way ANOVA (\* p<0.05; \*\*\* p<0.001, as compared with controls).

**Figure 4:** Effects of siRNA-attenuated expression of GRK2 and GRK5 on the kinetic pattern of WT *versus* 13A AT<sub>1A</sub>R-mediated ERK activation. Cells were transiently transfected with WT or 13A expression plasmid and siRNA. Three days after transfection, cells were serum-starved and stimulated with 100 nM AngII for the indicated periods. Phosphorylation of ERK1/2 was analyzed and quantified by Western blotting as described. Values in graphs are expressed as percent of the phosphorylation of ERK1/2 obtained at 5 min of stimulation in control (CTL) siRNA-transfected cells. (A) Effect of GRK2 depletion

on AngII-induced ERK1/2 phosphorylation by WT AT<sub>1A</sub>R. **(B)** Effect of GRK5 depletion on AngII-induced ERK1/2 phosphorylation by WT AT<sub>1A</sub>R. **(C)** Effect of GRK2 depletion on AngII-induced ERK1/2 phosphorylation by 13A AT<sub>1A</sub>R. **(D)** Effect of GRK5 depletion on AngII-induced ERK1/2 phosphorylation by 13A AT<sub>1A</sub>R. Data represent the mean  $\pm$  S.E.M. from at least five independent experiments. Statistical comparison of the curves was performed by using a two-way ANOVA between GRK siRNA-transfected and control cells (\*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; ns, not significant).

**Figure 5:** Principle of the global parameter optimization strategy. One control and four perturbed data sets (in blue) were used to optimize the unknown parameters. Each of these data sets corresponds to a number of unknown parameters (in pink). Notably, only the parameter targeted by the biological perturbation was modified within the control parameter set, all the others being identical to control. A total of 45 parameters (40 for control conditions and 5 corresponding to the different perturbations) were optimized. The optimization algorithm minimizes the global error (*i.e.*: the distance of the simulated model with respect to the different experimental curves, either control or perturbed [in green]) through an iterative and auto-adaptative machine learning process.

**Figure 6:** Fitting of the model to the experimental data with the optimized parameter set. Experimental data are represented as point  $\pm$  S.E.M. All the curves were generated through model simulations. Data from **A**, **B**, **C** and **D** are from Kim et al. (21). Data from **E** and **F** were generated in living cells using FRET sensors for DAG and PKC activity respectively.



**Figure 7:** Model validation against published independent data sets generated in the same experimental conditions. Control conditions are presented as the blue curves. Conditions x, y and z represent gradual increase or decrease (as indicated) of the targeted parameter. **(A)** Simulation of GRK23 overexpression causes a decrease in phosphorylated ERK kinetics similar to previously reported experimental data (21). **(B)** Simulation of GRK56 overexpression causes an increase in phosphorylated ERK at all time points similar to previously reported experimental data (21). **(C)** Simulation of G protein inhibition results in decreased phosphorylated ERK levels, especially early, with a pattern in excellent agreement with the SII and DRY/AAY data reported by Ahn et al. (17). **(D)** Simulation of either  $\beta$ -arrestin 1 (red curve) or  $\beta$ -arrestin 2 (green curve) depletion on phosphorylated ERK response as a function of angiotensin concentration shows antagonistic actions of both  $\beta$ -arrestin isoforms consistent with previously published data (14).

**Figure 8:** Sensitivity analysis. Optimized parameters (1 to 48) were sequentially decreased by 1 % and the impacts on 5 variables (DAG, PKC, GpERK, bpERK and total pERK) are calculated and plotted. The modification of a given species integrated for 90 min and expressed as a percentage of the initial value is an indication of the parameter sensitivity. Positive value indicates that the quantity increases with a reaction rate decrease, while negative value indicates that the quantity decreases. Simulations are done with the model at 100 nM AngII concentration. Reactions are grouped (color code) according to the biological processes indicated above the top plot. Parameter numbers are indicated above the bottom plot.

**Figure 9:** Role of GRK2 in regulating second messenger generation and early ERK response by the  $AT_{1A}R$ . **(A)** DAG response was simulated as a function of AngII

concentration both in control (in blue) and GRK2/3-depleted (in purple) conditions using the core model and the optimized parameter set. **(B)** HEK293 cells were simultaneously transfected with AT<sub>1A</sub>R and either control (in blue) or GRK2 (in purple) siRNAs. Three days after transfection, cells were serum-starved for 6 h and stimulated for 10 min with increasing doses of AngII ( $10^{-10}$  to  $10^{-6}$  M). Inositol phosphate levels were measured as described in the materials and methods section **(C)** Total pERK response was simulated as a function of AngII concentration either in control (in blue), GRK2/3-depleted (in purple), or GRK2/3-depleted and PKC-inhibited (Ro, in black) conditions using the core model and the optimized parameter set. **(D)** HEK293 cells were simultaneously transfected with AT<sub>1A</sub>R and either control (in blue) or GRK2 (in purple and black) siRNAs. Three days after transfection, cells were serum-starved for 6 h, pre-incubated with DMSO (blue and purple) or the PKC inhibitor Ro-31-8425 (Ro, black) and stimulated for 2 min with increasing doses of AngII ( $10^{-10}$  to  $10^{-7}$  M). Phosphorylated ERK signals were normalized according to the amount of total ERK. For the experimental data, each point represents the mean  $\pm$  S.E.M. from at least 5 independent experiments.

**Figure 10:** Role of GRK 5 and 6 in early ERK response by the AT<sub>1A</sub>R. **(A)** Total pERK response was simulated as a function of AngII concentration either in control (in blue), GRK5/6-depleted (in purple), PKC-inhibited (in red) or GRK5/6-depleted and PKC-inhibited (in green) conditions using the core model and the optimized parameter set. **(B)** HEK293 cells were simultaneously transfected with AT<sub>1A</sub>R and either control (blue and red) or GRK5 (purple and green) siRNAs. Three days after transfection, cells were serum-starved for 6 h, pre-incubated with DMSO (blue and purple) or the PKC inhibitor Ro-31-8425 (black and green) and stimulated for 2 min with increasing doses of AngII ( $10^{-10}$  to  $10^{-7}$  M). **(C)** HEK293 cells were simultaneously transfected with AT<sub>1A</sub>R and either control (blue and red)

or GRK6 (purple and green) siRNAs. Three days after transfection, cells were serum-starved for 6 h, pre-incubated with DMSO (blue and purple) or the PKC inhibitor Ro-31-8425 (black and green) and stimulated for 2 min with increasing doses of AngII ( $10^{-10}$  to  $10^{-7}$  M). Phosphorylated ERK signals were normalized according to the amount of total ERK. For the experimental data, each point represents the mean  $\pm$  S.E.M. from at least 5 independent experiments.

## SUPPLEMENTARY FIGURE LEGENDS

**Supplementary figure 1:** Ordinary differential equation (ODE) set of the core model.

**Supplementary figure 2:** Conservation laws and observed variables of the core model.

**Supplementary figure 3:** Representative Western blots of HEK293 cells transiently transfected with the different siRNA used in this study. Endogenous proteins are shown. **(A)** Control *versus*  $\beta$ -arrestin 2-depleted cells. **(B)** Control *versus* GRK2-depleted cells. **(C)** Control *versus* GRK5-depleted cells. **(D)** Control *versus* GRK6-depleted cells.

**Supplementary figure 4:** Representative Western blots of HEK293 cells transiently co-transfected with wild-type (WT) or 13A AT<sub>1A</sub>R and either control (CTL) or  $\beta$ -arrestin 2 ( $\beta$ arr2) siRNAs. **(A)** Kinetics of ERK phosphorylation in WT *versus* 13A-transfected cells. **(B)** Kinetics of ERK phosphorylation in WT-transfected cells depleted or not in  $\beta$ -arrestin 2. **(C)** Kinetics of ERK phosphorylation in 13A-transfected cells depleted or not in  $\beta$ -arrestin 2.

**Supplementary figure 5:** Representative Western blots of HEK293 cells transiently co-transfected with wild-type (WT) or 13A AT<sub>1A</sub>R and either control (CTL), GRK2 or GRK5 siRNAs. **(A)** Effect of GRK2 depletion on the kinetics of ERK phosphorylation upon AngII stimulation in WT-transfected cells. **(B)** Effect of GRK5 depletion on the kinetics of ERK phosphorylation upon AngII stimulation in WT-transfected cells. **(C)** Effect of GRK2 depletion on the kinetics of ERK phosphorylation upon AngII stimulation in 13A-transfected

cells. **(D)** Effect of GRK5 depletion on the kinetics of ERK phosphorylation upon AngII stimulation in 13A-transfected cells.

**Supplementary figure 6:** Parameter number, description, name in the ODEs and values. Parameters in red were taken from the literature or experimentally measured. Parameters in black were optimized by fitting to the experimental data set. Parameters 49 to 53 represent the value of the indicated parameter after targeted perturbation. Parameters 54 to 56 correspond to the normalization factor applied to ERK, DAG and PKC data in order to transform relative values (arbitrary units) in absolute quantities (Molar).

**Supplementary figure 7:** The core model written with BIOCHAM semantics (<http://contraintes.inria.fr/BIOCHAM/>).

**Supplementary figure 8:** Representative Western blots of HEK293 cells transiently co-transfected with wild-type (WT) AT<sub>1A</sub>R and either control (CTL) or GRK2 siRNAs. **(A)** ERK phosphorylation as a function of AngII concentration ( $10^{-10}$  to  $10^{-7}$  M) in control conditions treated for 2 min. **(B)** ERK phosphorylation as a function of AngII concentration ( $10^{-10}$  to  $10^{-7}$  M) in GRK2-depleted cells treated for 2 min. **(C)** ERK phosphorylation as a function of AngII concentration ( $10^{-10}$  to  $10^{-7}$  M) in GRK2-depleted cells pre-incubated with Ro-31-8425 and treated for 2 min.

**Supplementary figure 9:** Representative Western blots of HEK293 cells transiently co-transfected with wild-type (WT) AT<sub>1A</sub>R and either control (CTL), GRK5 or GRK6 siRNAs. **(A)** ERK phosphorylation as a function of AngII concentration ( $10^{-10}$  to  $10^{-7}$  M) in control cells pre-treated with DMSO or Ro-31-8425 (as indicated) and stimulated for 2 min. **(B)** ERK

phosphorylation as a function of AngII concentration ( $10^{-10}$  to  $10^{-7}$  M) in GRK5-depleted cells pre-treated with DMSO or Ro-31-8425 (as indicated) and stimulated for 2 min. (C) ERK phosphorylation as a function of AngII concentration ( $10^{-10}$  to  $10^{-7}$  M) in GRK6-depleted cells pre-treated with DMSO or Ro-31-8425 (as indicated) and stimulated for 2 min.

## REFERENCES

1. P. Ma, R. Zimmel, *Nat Rev Drug Discov* 1, 571 (Aug, 2002).
2. R. A. Hall, R. T. Premont, R. J. Lefkowitz, *J Cell Biol* 145, 927 (May 31, 1999).
3. R. J. Lefkowitz, *J Biol Chem* 273, 18677 (Jul 24, 1998).
4. J. A. Pitcher, N. J. Freedman, R. J. Lefkowitz, *Annu Rev Biochem* 67, 653 (1998).
5. O. B. Goodman, Jr. *et al.*, *Nature* 383, 447 (Oct 3, 1996).
6. S. A. Laporte *et al.*, *Proc Natl Acad Sci U S A* 96, 3712 (Mar 30, 1999).
7. P. H. McDonald *et al.*, *J Biol Chem* 274, 10677 (Apr 16, 1999).
8. L. M. Luttrell, R. J. Lefkowitz, *J Cell Sci* 115, 455 (Feb 1, 2002).
9. R. J. Lefkowitz, S. K. Shenoy, *Science* 308, 512 (Apr 22, 2005).
10. P. H. McDonald *et al.*, *Science* 290, 1574 (Nov 24, 2000).
11. L. M. Luttrell *et al.*, *Proc Natl Acad Sci U S A* 98, 2449 (Feb 27, 2001).
12. K. A. DeFea *et al.*, *Proc Natl Acad Sci U S A* 97, 11086 (Sep 26, 2000).
13. S. Ahn, C. D. Nelson, T. R. Garrison, W. E. Miller, R. J. Lefkowitz, *Proc Natl Acad Sci U S A* 100, 1740 (Feb 18, 2003).
14. S. Ahn, H. Wei, T. R. Garrison, R. J. Lefkowitz, *J Biol Chem* 279, 7807 (Feb 27, 2004).
15. A. Tohgo, K. L. Pierce, E. W. Choy, R. J. Lefkowitz, L. M. Luttrell, *J Biol Chem* 277, 9429 (Mar 15, 2002).
16. K. A. DeFea *et al.*, *J Cell Biol* 148, 1267 (Mar 20, 2000).
17. S. Ahn, S. K. Shenoy, H. Wei, R. J. Lefkowitz, *J Biol Chem* 279, 35518 (Aug 20, 2004).
18. L. Ge, Y. Ly, M. Hollenberg, K. DeFea, *J Biol Chem* 278, 34418 (Sep 5, 2003).
19. S. Sasagawa, Y. Ozaki, K. Fujita, S. Kuroda, *Nat Cell Biol* 7, 365 (Apr, 2005).
20. H. Wei *et al.*, *Proc Natl Acad Sci U S A* 100, 10782 (Sep 16, 2003).
21. J. Kim *et al.*, *Proc Natl Acad Sci U S A* 102, 1442 (Feb 1, 2005).
22. X. R. Ren *et al.*, *Proc Natl Acad Sci U S A* 102, 1448 (Feb 1, 2005).
23. K. Iwata, J. Luo, R. B. Penn, J. L. Benovic, *J Biol Chem* 280, 2197 (Jan 21, 2005).
24. E. Kara *et al.*, *Mol Endocrinol* 20, 3014 (Nov, 2006).
25. N. Borisov *et al.*, *Mol Syst Biol* 5, 256 (2009).
26. B. N. Kholodenko, O. V. Demin, G. Moehren, J. B. Hoek, *J Biol Chem* 274, 30169 (Oct 15, 1999).
27. R. J. Orton *et al.*, *Biochem J* 392, 249 (Dec 1, 2005).
28. B. N. Kholodenko, J. F. Hancock, W. Kolch, *Nat Rev Mol Cell Biol* 11, 414 (Jun).
29. T. Nakakuki *et al.*, *Cell* 141, 884 (May 28).
30. N. Hao, M. Behar, T. C. Elston, H. G. Dohlman, *Oncogene* 26, 3254 (May 14, 2007).
31. D. Cserecsik, K. M. Hangos, G. M. Nagy, *J Theor Biol* 255, 119 (Nov 7, 2008).
32. A. Rizk, G. Batt, F. Fages, S. Soliman, paper presented at the Computational Methods in Systems Biology CMSB'08 2008.
33. N. M. Borisov, N. I. Markevich, J. B. Hoek, B. N. Kholodenko, *Biophys J* 89, 951 (Aug, 2005).
34. M. R. Birtwistle *et al.*, *Mol Syst Biol* 3, 144 (2007).
35. W. S. Hlavacek *et al.*, *Sci STKE* 2006, re6 (Jul 17, 2006).
36. H. Kitano, A. Funahashi, Y. Matsuoka, K. Oda, *Nat Biotechnol* 23, 961 (Aug, 2005).

37. A. Funahashi *et al.*, paper presented at the Proceedings of the IEEE 2008.
38. A. De Lean, J. M. Stadel, R. J. Lefkowitz, *J Biol Chem* 255, 7108 (Aug 10, 1980).
39. P. Samama, S. Cotecchia, T. Costa, R. J. Lefkowitz, *J Biol Chem* 268, 4625 (Mar 5, 1993).
40. G. Milligan, E. Kostenis, *Br J Pharmacol* 147 Suppl 1, S46 (Jan, 2006).
41. R. J. Lefkowitz, E. J. Whalen, *Curr Opin Cell Biol* 16, 162 (Apr, 2004).
42. M. Kato, S. Kawai, T. Takenawa, *J Biol Chem* 262, 5696 (Apr 25, 1987).
43. A. C. Newton, *J Biol Chem* 270, 28495 (Dec 1, 1995).
44. E. Rozengurt, *J Cell Physiol* 213, 589 (Dec, 2007).
45. L. M. Luttrell, *J Mol Neurosci* 26, 253 (2005).
46. R. J. Dickinson, S. M. Keyse, *J Cell Sci* 119, 4607 (Nov 15, 2006).
47. S. M. DeWire, S. Ahn, R. J. Lefkowitz, S. K. Shenoy, *Annu Rev Physiol* 69, 483 (2007).
48. L. M. Luttrell, D. Gesty-Palmer, *Pharmacol Rev* 62, 305 (Jun).
49. L. Dupuy *et al.*, *Proteomics* 9, 5446 (Dec, 2009).
50. V. V. Gurevich, J. L. Benovic, *J Biol Chem* 268, 11628 (Jun 5, 1993).
51. E. Reiter, R. J. Lefkowitz, *Trends Endocrinol Metab* 17, 159 (May-Jun, 2006).
52. B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, G. Muller, *Nat Biotechnol* 20, 370 (Apr, 2002).
53. J. D. Violin, J. Zhang, R. Y. Tsien, A. C. Newton, *J Cell Biol* 161, 899 (Jun 9, 2003).
54. N. Hansen, A. Ostermeier, *Evol Comput* 9, 159 (Summer, 2001).
55. T. Maiwald, J. Timmer, *Bioinformatics* 24, 2037 (Sep 15, 2008).
56. A. Raue *et al.*, *Bioinformatics* 25, 1923 (Aug 1, 2009).
57. L. Calzone, F. Fages, S. Soliman, *Bioinformatics* 22, 1805 (Jul 15, 2006).
58. F. Fages, S. Soliman, paper presented at the Formal Methods in Systems Biology 2008.
59. F. Fages, S. Soliman, paper presented at the International School on Formal Methods for the Design of Computer, Communication and Software Systems: Computational Systems Biology, Bertinoro, Italy 2008.
60. I. Swameye, T. G. Muller, J. Timmer, O. Sandra, U. Klingmuller, *Proc Natl Acad Sci U S A* 100, 1028 (Feb 4, 2003).
61. S. Nelander *et al.*, *Mol Syst Biol* 4, 216 (2008).
62. T. R. Xu *et al.*, *Sci Signal* 3, ra20.
63. S. K. Shenoy *et al.*, *J Biol Chem* 281, 1261 (Jan 13, 2006).
64. D. Gesty-Palmer *et al.*, *J Biol Chem* 281, 10856 (Apr 21, 2006).
65. T. A. Kohout *et al.*, *J Biol Chem* 279, 23214 (May 28, 2004).
66. S. A. Laporte *et al.*, *Mol Pharmacol* 49, 89 (Jan, 1996).
67. S. Cotecchia, J. Ostrowski, M. A. Kjelsberg, M. G. Caron, R. J. Lefkowitz, *J Biol Chem* 267, 1633 (Jan 25, 1992).



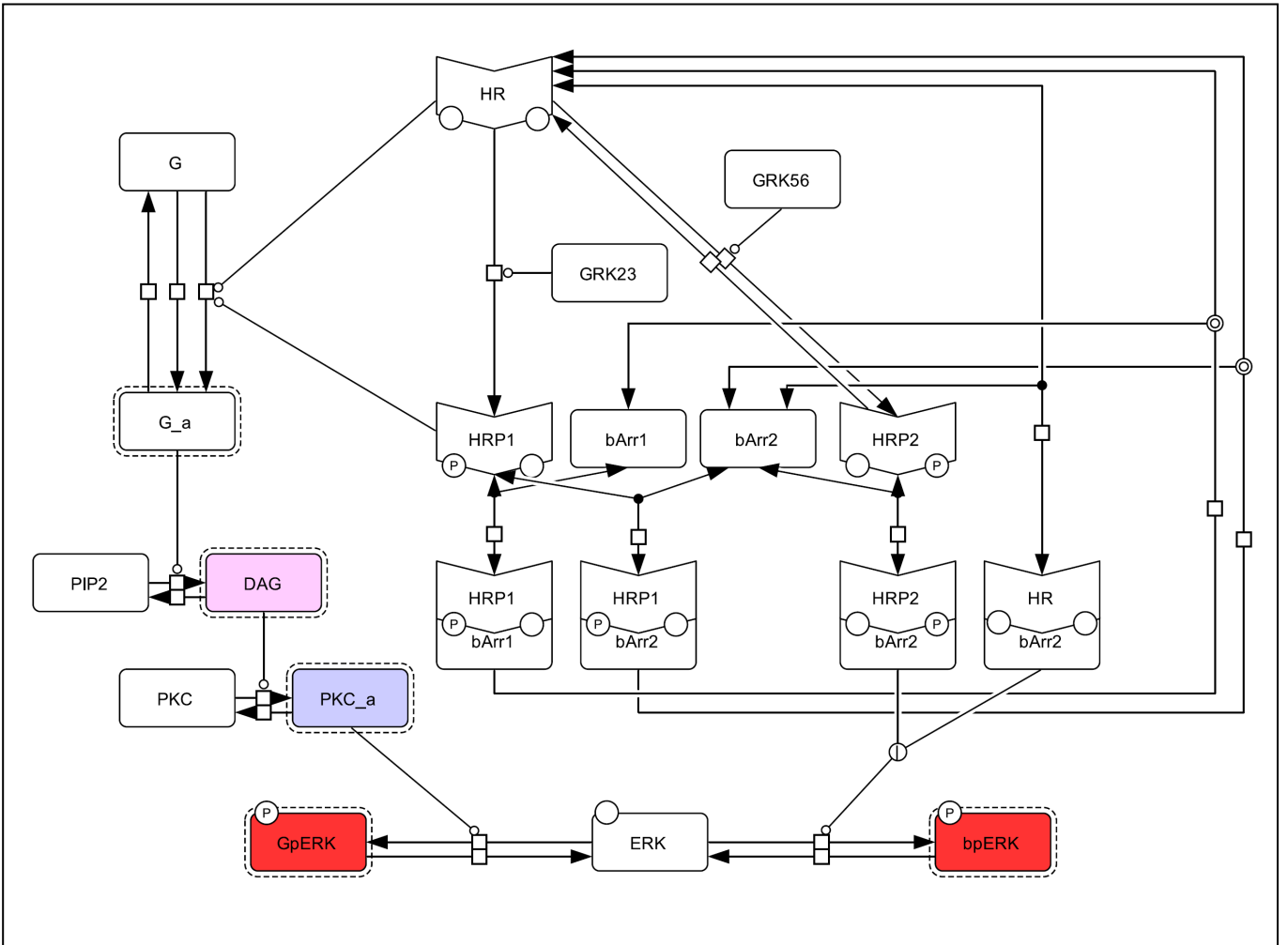


Figure 1

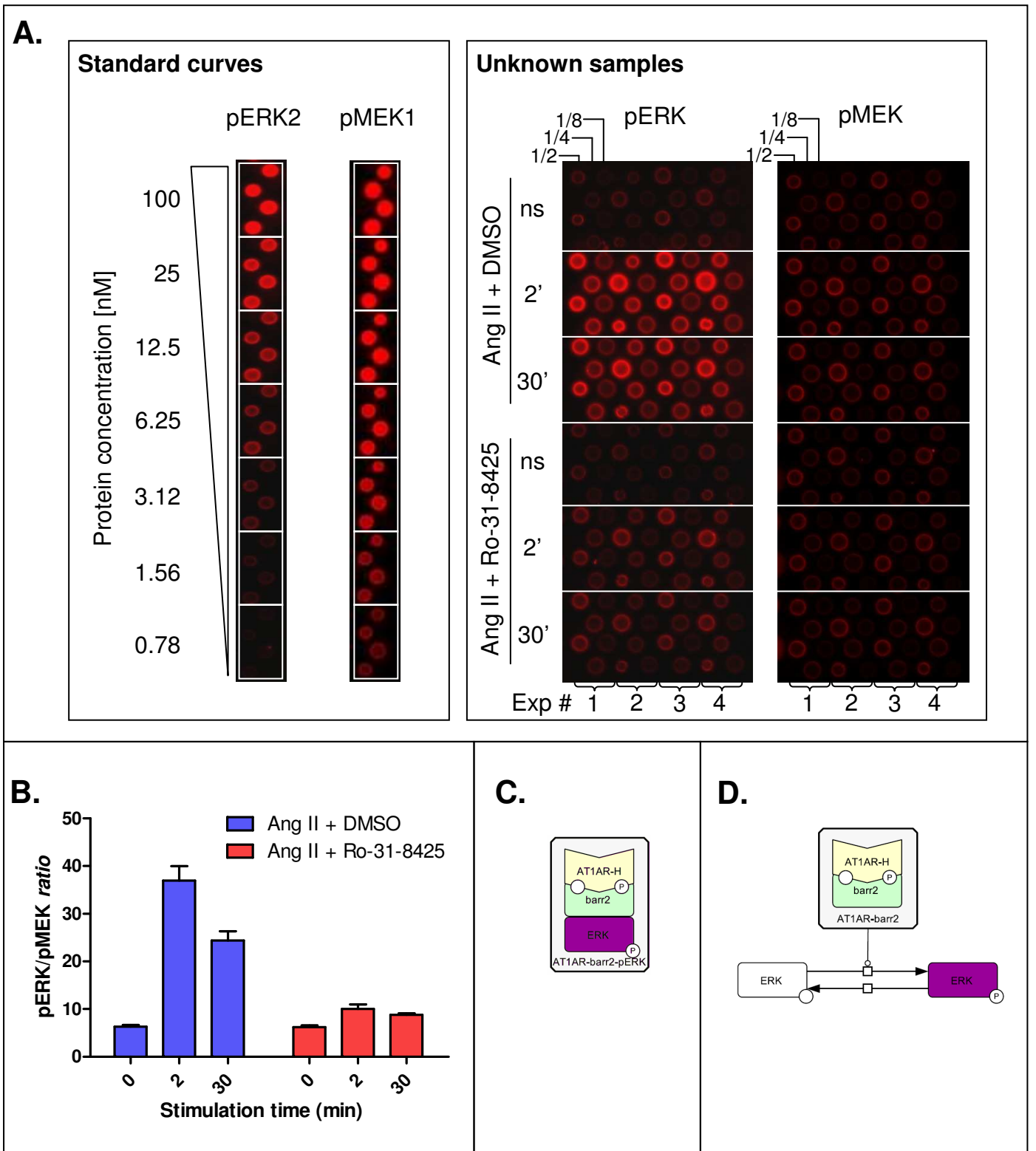
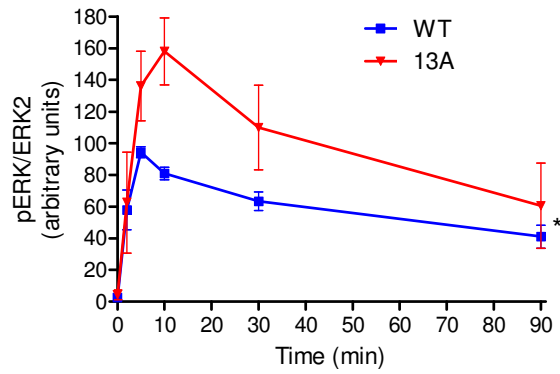


Figure 2

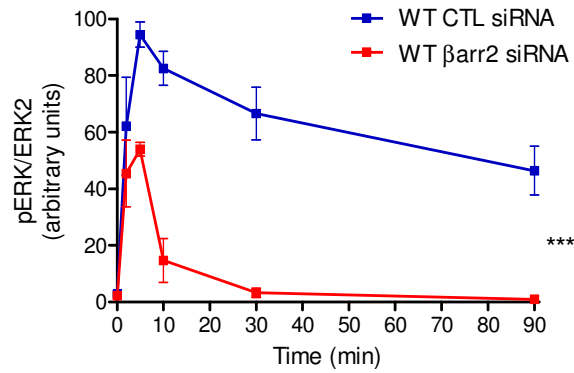
**A.**

**AT1AR** KSHSSLSTKMSTLSYRPSDNMSSSAKKPASC  
**13A** -A-AA-AA--AA-A--A---AAA-----A-

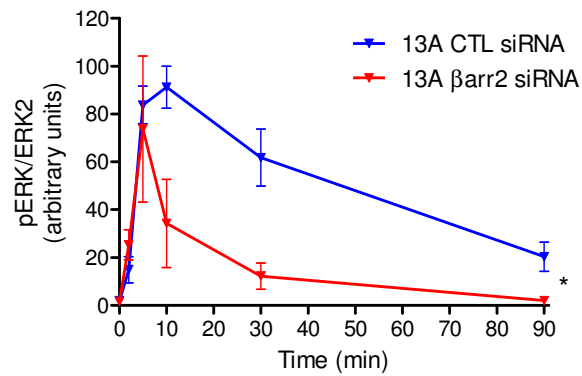
**B.**



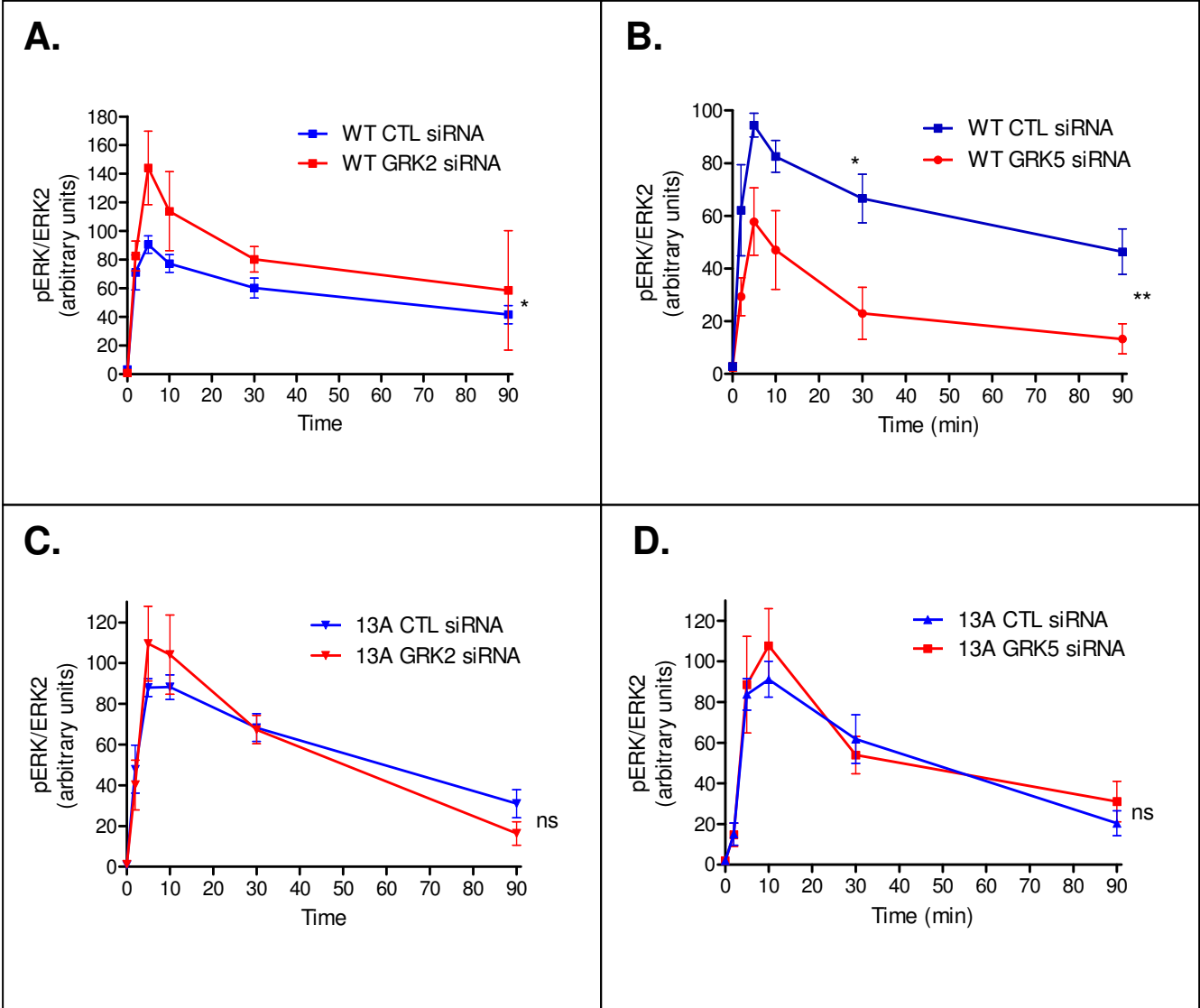
**C.**



**D.**



**Figure 3**



**Figure 4**

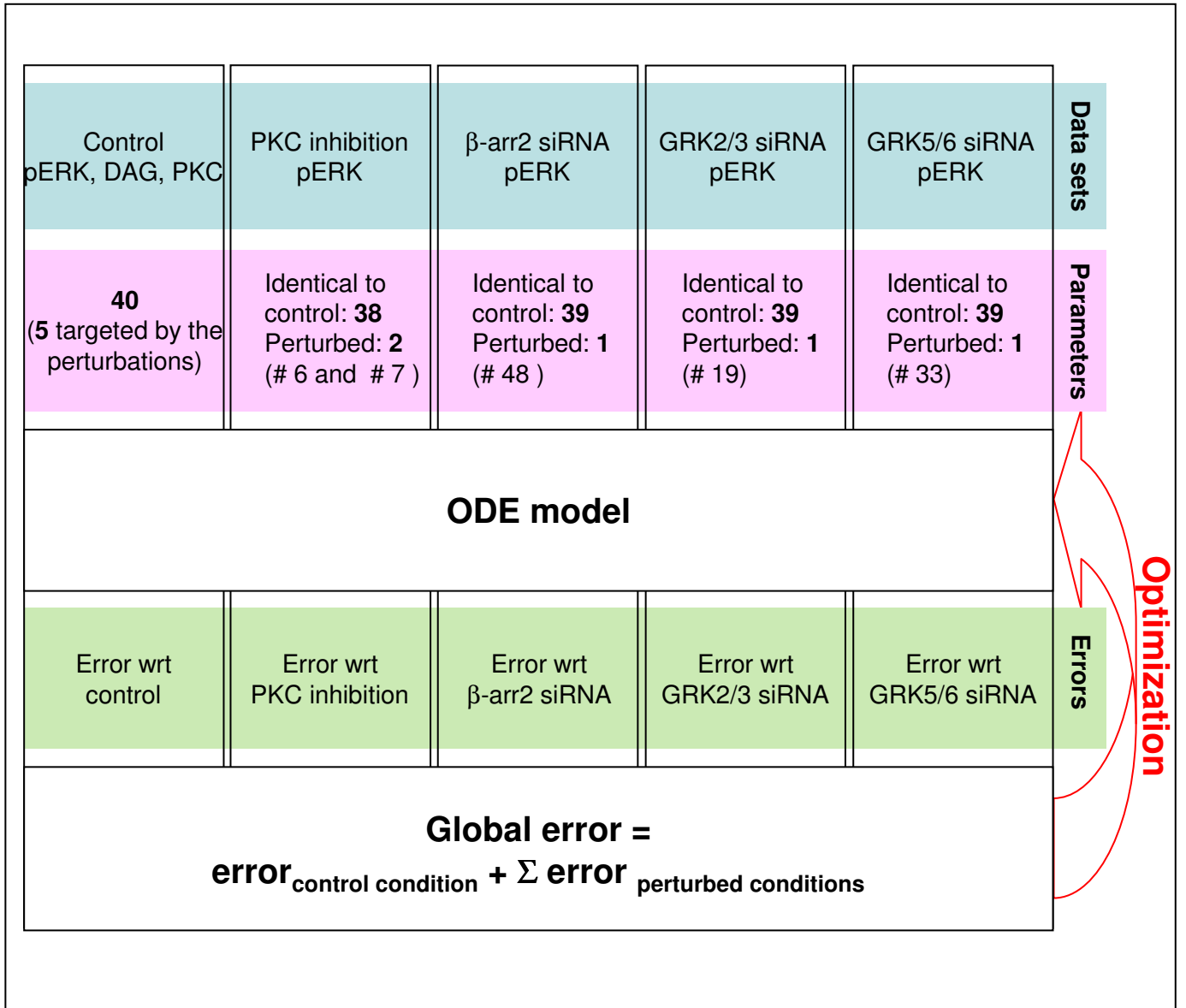


Figure 5

## Data fitting

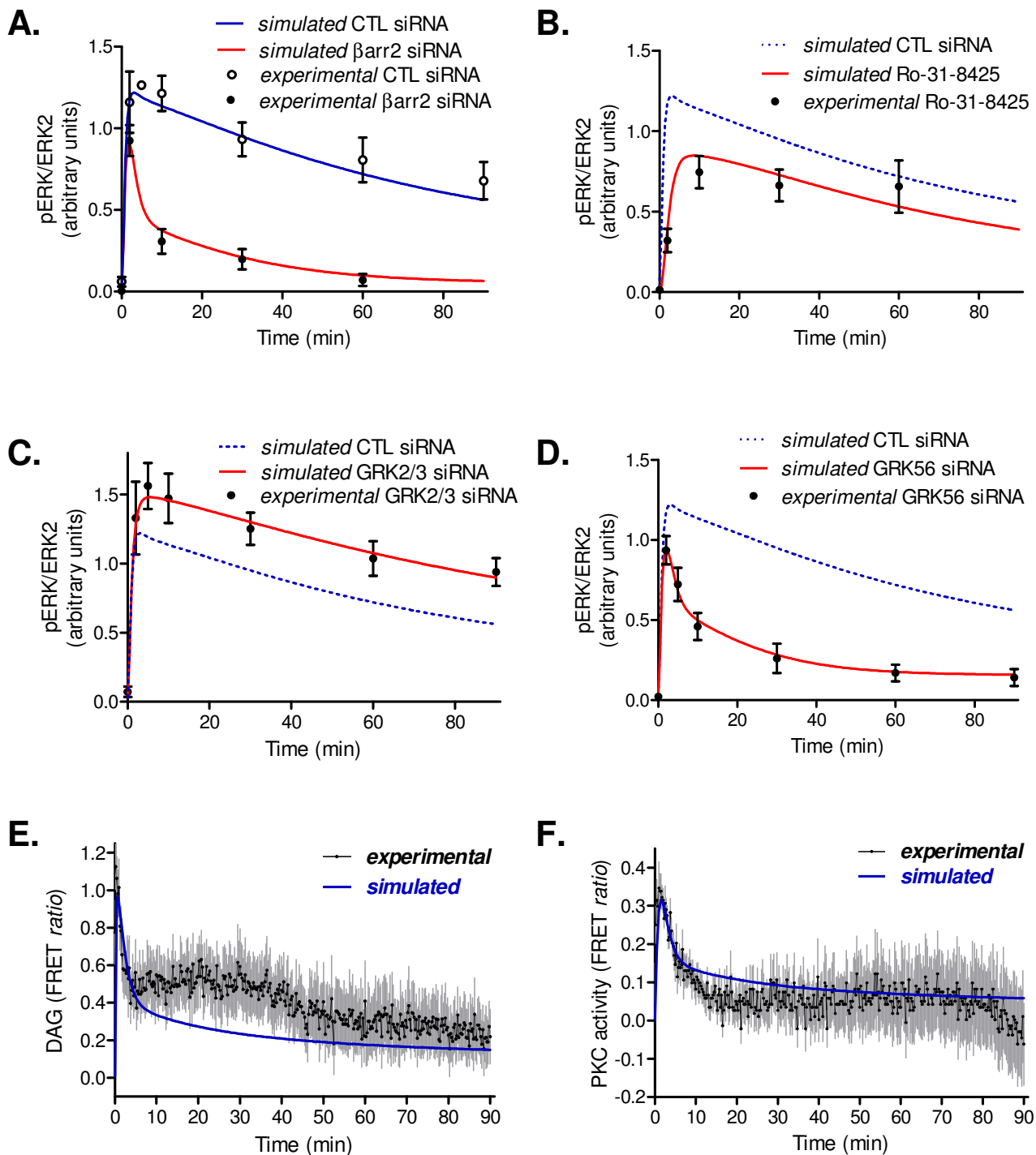
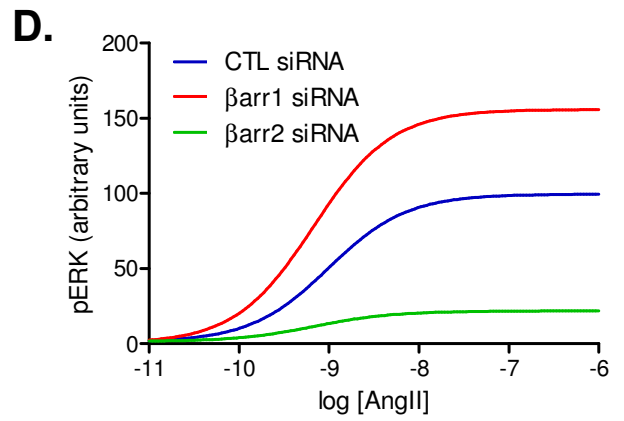
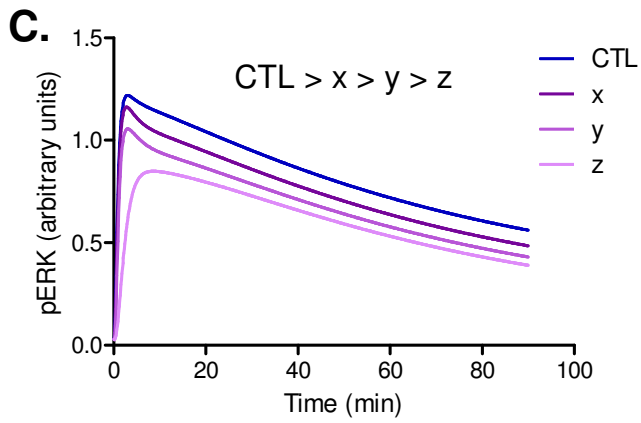
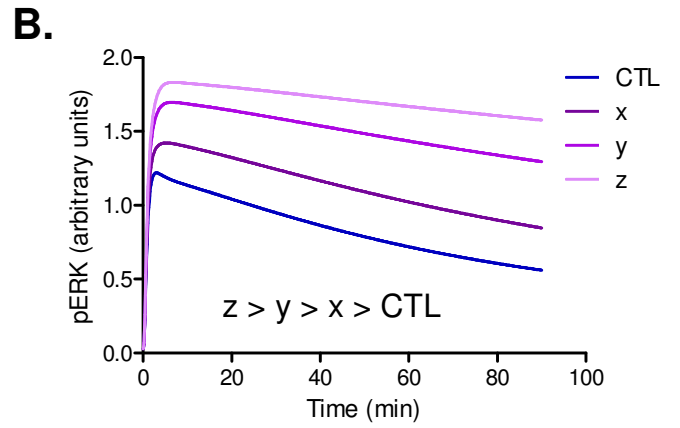
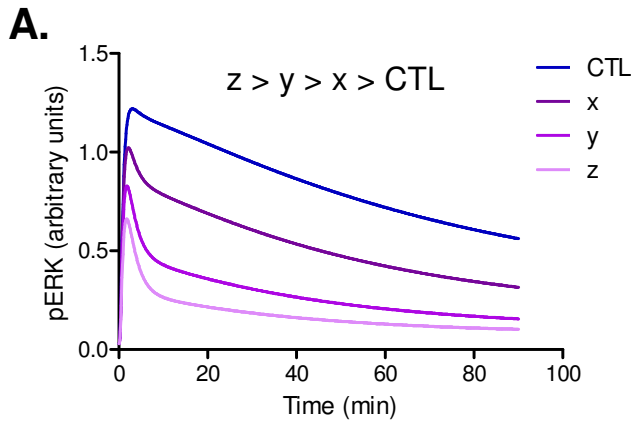


Figure 6

**Simulated**



**Figure 7**

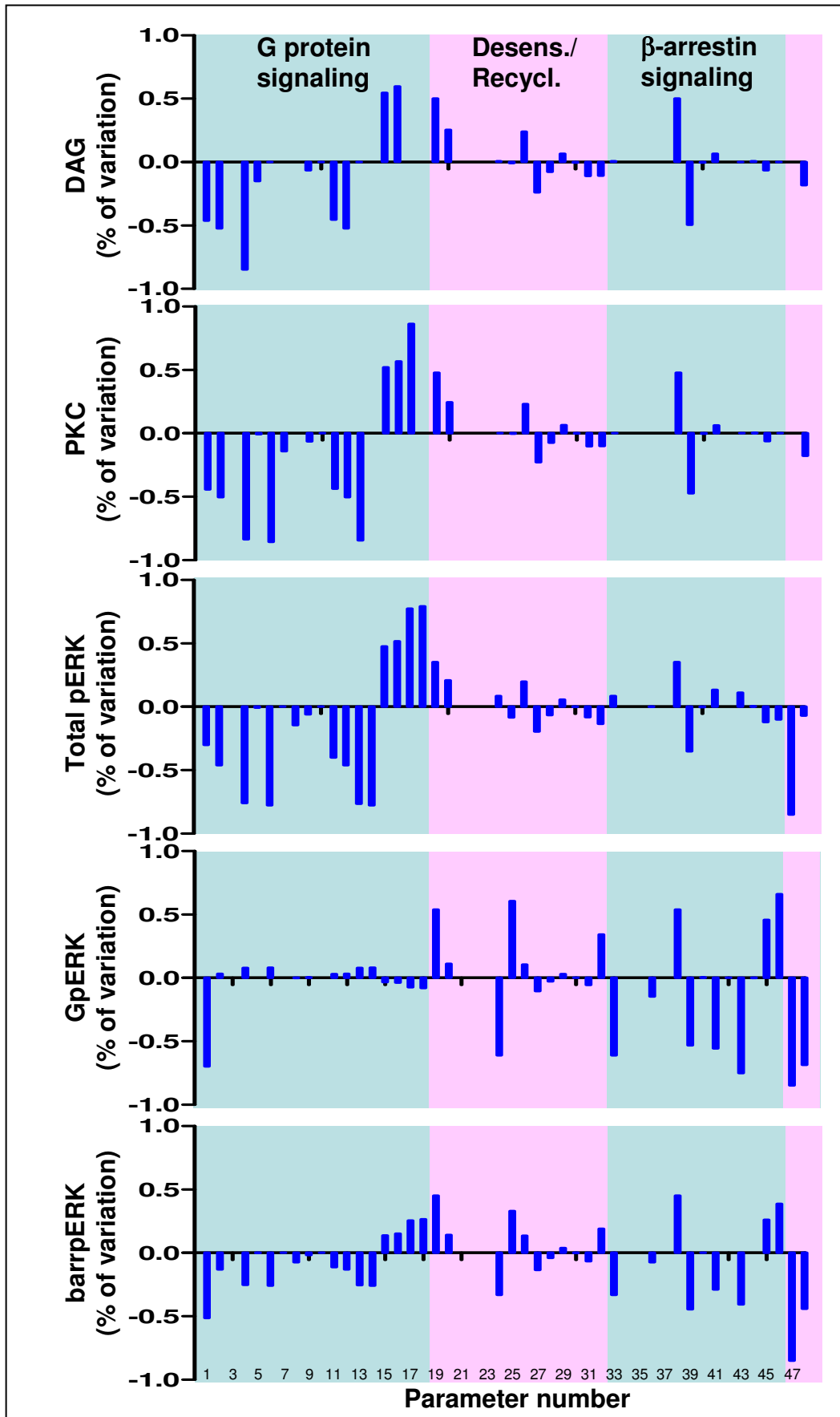


Figure 8



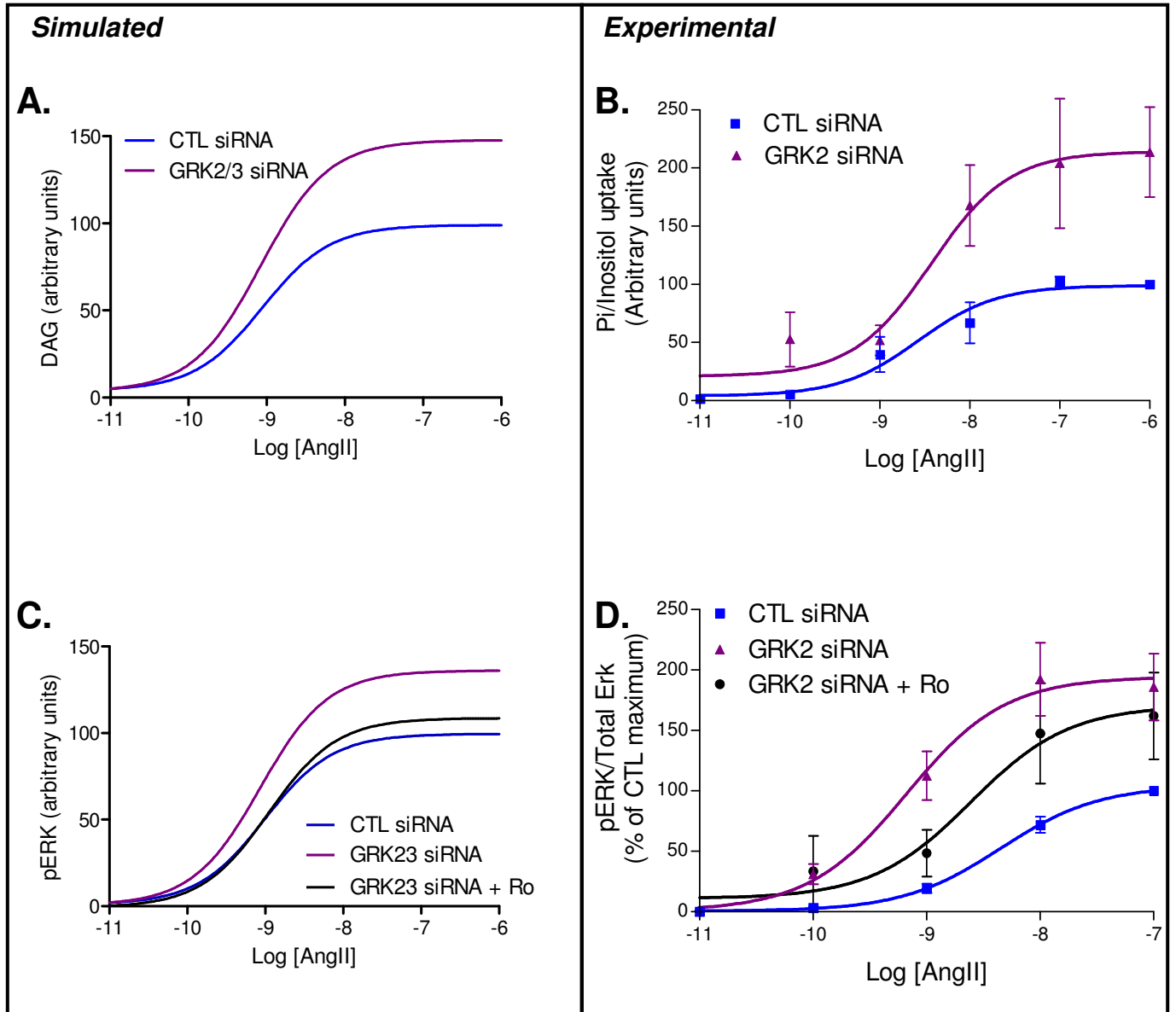
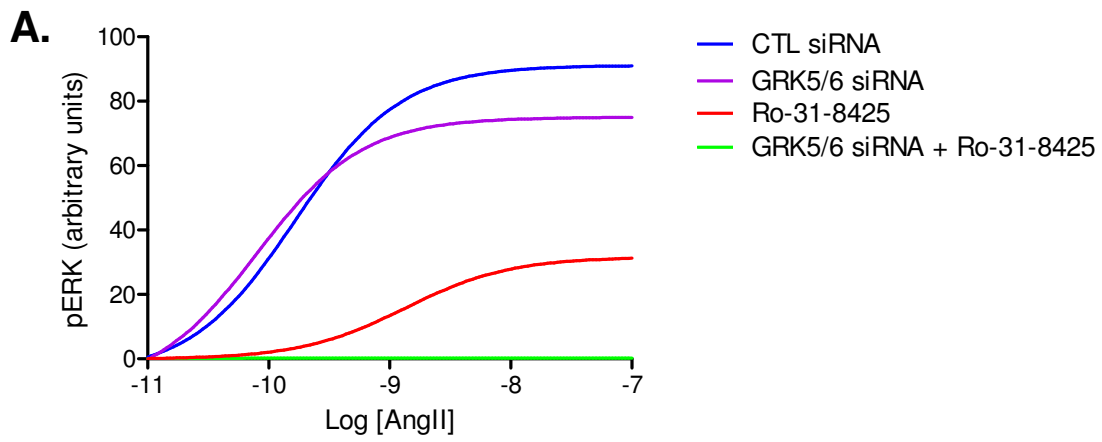


Figure 9

## Simulated



## Experimental

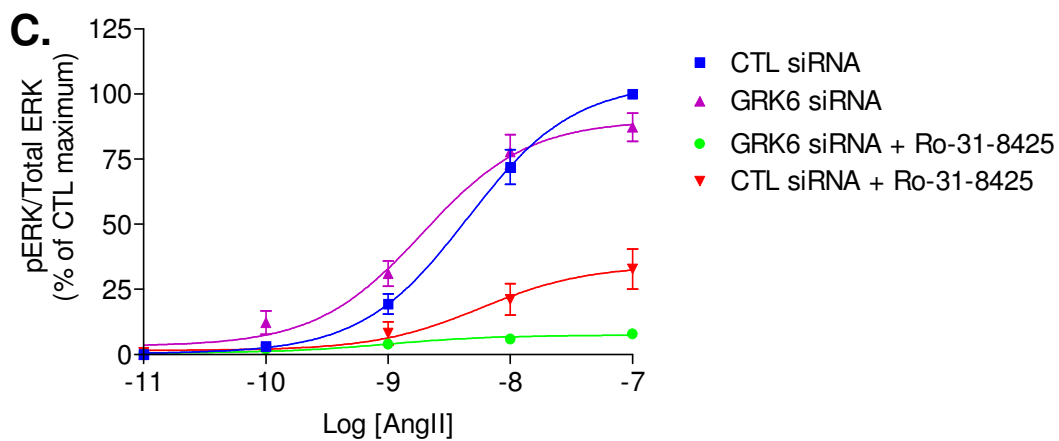
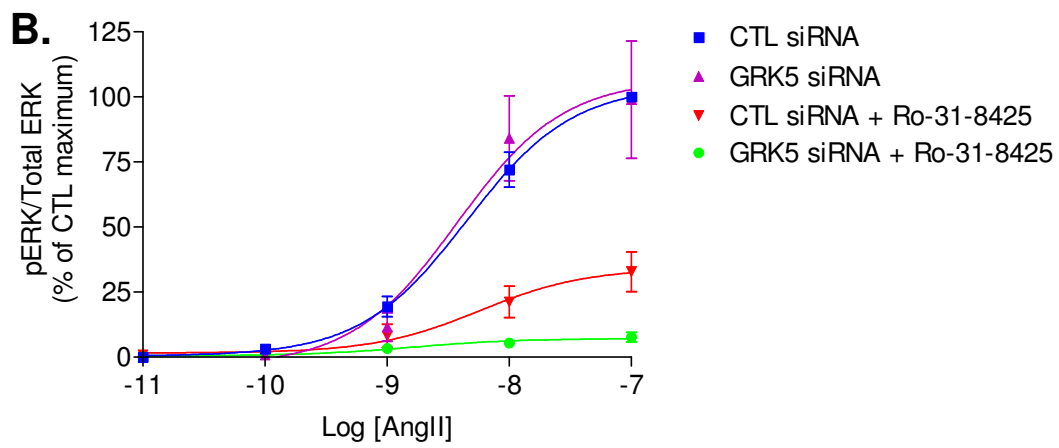


Figure 10

$$\begin{aligned}
\frac{d[HR]}{dt} &= k14 \cdot [HRP1barr1] + (k17 \cdot [HRP1barr2] + (k20 \cdot [HRbarr2] + k25 \cdot [HRP2] - k11 \cdot [barr2] \cdot [HR] - \frac{Vm18 \cdot GRK56 \cdot [HR]}{Km19 + [HR]})) - \frac{Vm9 \cdot GRK23 \cdot [HR]}{Km10 + [HR]} \\
\frac{d[HRP1]}{dt} &= \frac{Vm9 \cdot GRK23 \cdot [HR]}{Km10 + [HR]} + k13 \cdot [HRP1barr1] + k16 \cdot [HRP1barr2] - k15 \cdot [barr2] \cdot [HRP1] - k12 \cdot [barr1] \cdot [HRP1] \\
\frac{d[HRP2]}{dt} &= \frac{Vm18 \cdot GRK56 \cdot [HR]}{Km19 + [HR]} + k22 \cdot [HRP2barr2] - k25 \cdot [HRP2] - k21 \cdot [barr2] \cdot [HRP2] \\
\frac{d[G]}{dt} &= k0 \cdot [G_a] - k26 \cdot [G] - k2 \cdot [G] \cdot [HR] - k1 \cdot [G] \cdot [HRP1] \\
\frac{d[G_a]}{dt} &= k1 \cdot [G] \cdot [HRP1] + k2 \cdot [G] \cdot [HR] + k26 \cdot [G] - k0 \cdot [G_a] \\
\frac{d[PIP2]}{dt} &= k4 \cdot [DAG] - k3 \cdot [PIP2] \cdot [G_a] \\
\frac{d[DAG]}{dt} &= k3 \cdot [PIP2] \cdot [G_a] - k4 \cdot [DAG] \\
\frac{d[PKC]}{dt} &= k6 \cdot [PKC_a] - k5 \cdot [DAG] \cdot [PKC] \\
\frac{d[PKC_a]}{dt} &= k5 \cdot [DAG] \cdot [PKC] - k6 \cdot [PKC_a] \\
\frac{d[ERK]}{dt} &= k8 \cdot [GpERK] + k27 \cdot [bpERK] - k24 \cdot [ERK] \cdot [HRP2barr2] - k23 \cdot [ERK] \cdot [HRbarr2] - k7 \cdot [ERK] \cdot [PKC_a] \\
\frac{d[GpERK]}{dt} &= k7 \cdot [ERK] \cdot [PKC_a] - k8 \cdot [GpERK] \\
\frac{d[barr1]}{dt} &= (k13 + k14) \cdot [HRP1barr1] - k12 \cdot [barr1] \cdot [HRP1] \\
\frac{d[HRP1barr1]}{dt} &= k12 \cdot [barr1] \cdot [HRP1] - (k13 + k14) \cdot [HRP1barr1] \\
\frac{d[barr2]}{dt} &= (k16 + k17) \cdot [HRP1barr2] + (k20 \cdot [HRbarr2] + k22 \cdot [HRP2barr2] - k21 \cdot [barr2] \cdot [HRP2] - k11 \cdot [barr2] \cdot [HR] - k15 \cdot [barr2] \cdot [HRP1]) \\
\frac{d[HRP1barr2]}{dt} &= k15 \cdot [barr2] \cdot [HRP1] - (k16 + k17) \cdot [HRP1barr2] \\
\frac{d[HRbarr2]}{dt} &= k11 \cdot [barr2] \cdot [HR] - k20 \cdot [HRbarr2] \\
\frac{d[HRP2barr2]}{dt} &= k21 \cdot [barr2] \cdot [HRP2] - k22 \cdot [HRP2barr2] \\
\frac{d[bpERK]}{dt} &= k23 \cdot [ERK] \cdot [HRbarr2] + k24 \cdot [ERK] \cdot [HRP2barr2] - k27 \cdot [bpERK]
\end{aligned}$$

**Supplementary figure 1**

Conservation laws:

$$H_{total} = HR + HRP1 + HRP2 + HRbarr2 + HRP1barr1 + HRP1barr2 + HRP2barr2$$

$$G_{total} = G + G_a$$

$$DAG_{total} = PIP2 + DAG$$

$$PKC_{total} = PKC + PKC_a$$

$$barr1_{total} = barr1 + HRP1barr1$$

$$barr2_{total} = barr2 + HRbarr2 + HRP1barr2 + HRP2barr2$$

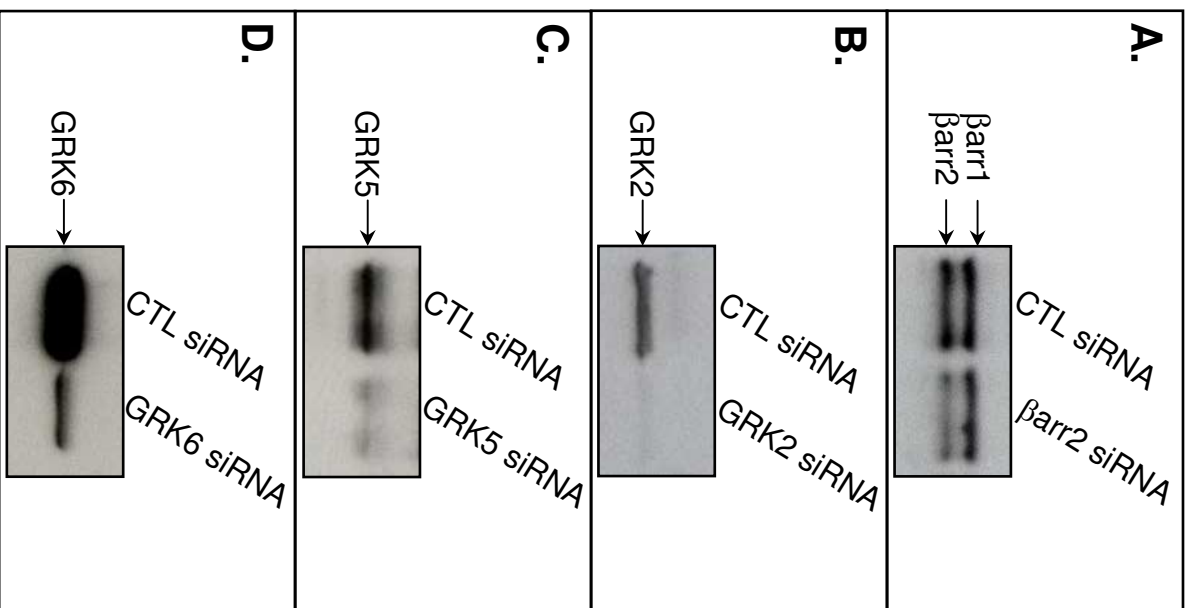
$$ERK_{total} = ERK + GpERK + bpERK$$

Observed variables:

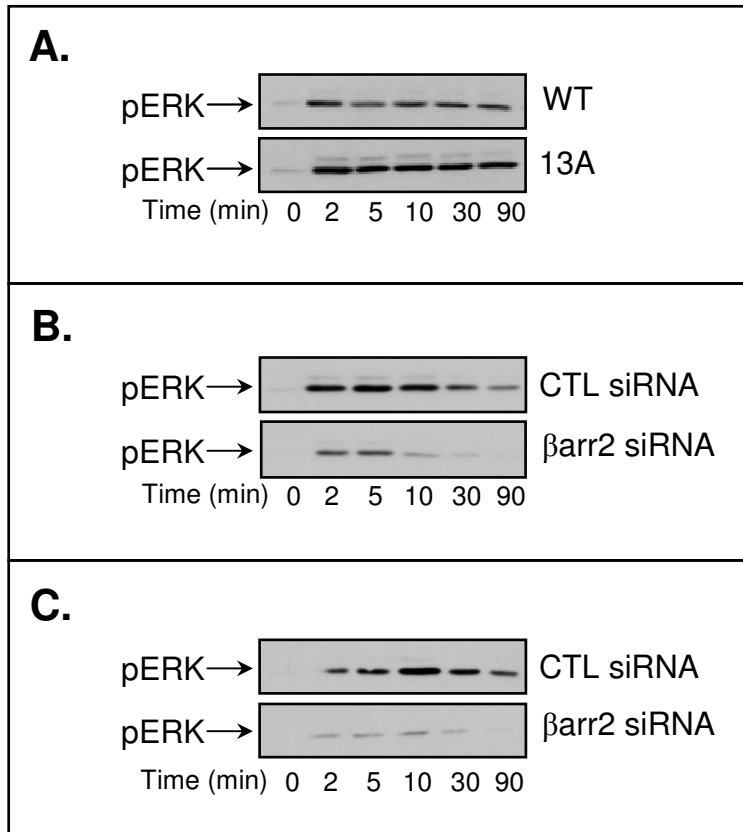
$$pERK_{total\_observed} = \frac{[GpERK] + [bpERK]}{erknorm}$$

$$DAG_{observed} = \frac{[DAG] - D0_a}{DAGnorm} + 1$$

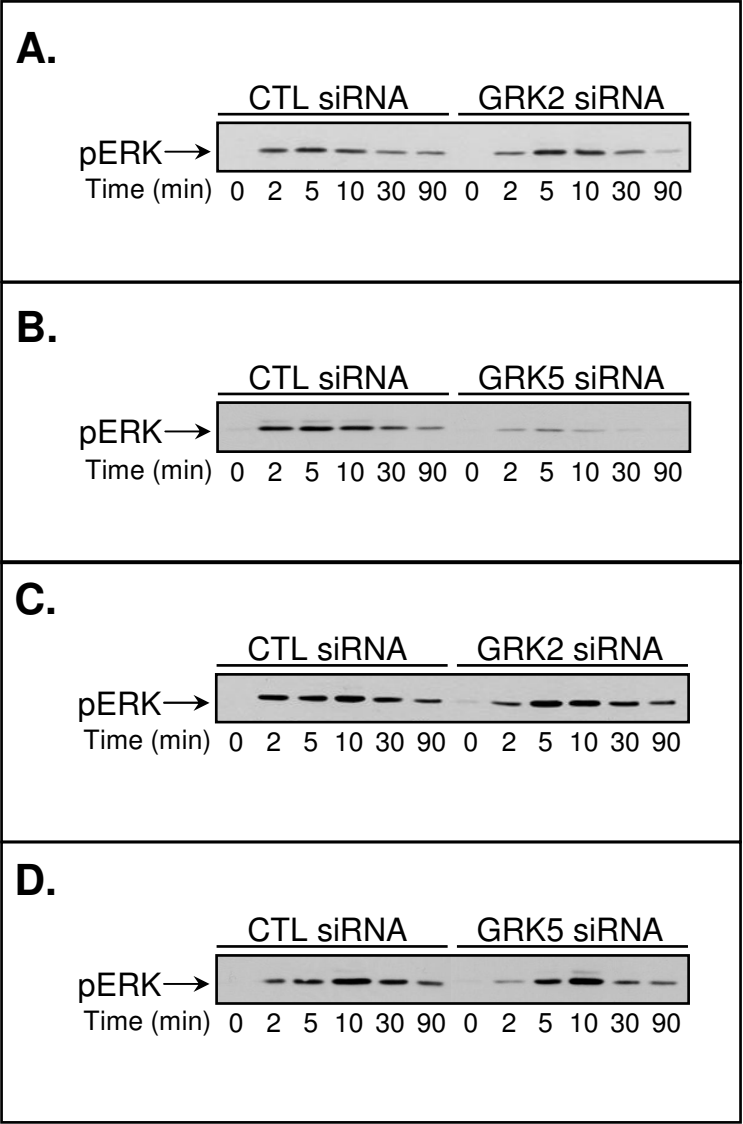
$$PKC_{observed} = \frac{[PKC_a] - P0_a}{PKCnorm} + 1$$



Supplementary figure 3



Supplementary figure 4



Supplementary figure 5

Parameter #	Description	Name in ODE	Value (unit)
1	Initial bound receptor quantity (HR)		0.08 ( $\mu\text{mol L}^{-1}$ )
2	Initial G protein quantity (G)		56.9917983 ( $\mu\text{mol L}^{-1}$ )
3	Initial activated G protein quantity (G_a)		0 ( $\mu\text{mol L}^{-1}$ )
4	Initial PIP2 quantity		0.996936631 ( $\mu\text{mol L}^{-1}$ )
5	Initial DAG quantity	D0_a	0.009 ( $\mu\text{mol L}^{-1}$ )
6	Initial PKC quantity		8.8420289 ( $\mu\text{mol L}^{-1}$ )
7	Initial activated PKC quantity (PKC_a)	P0_a	0.002 ( $\mu\text{mol L}^{-1}$ )
8	Initial GpERK quantity		0.015 ( $\mu\text{mol L}^{-1}$ )
9	G auto activation	k26	0.00031132215 (s-1)
10	Activation of G by HRP1	k1	0.0179507395 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
11	Activation of G by HR	k2	7.5982987 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
12	Activation of DAG by active G protein	k3	4.6298141 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
13	Activation of PKC by DAG	k5	0.078716779 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
14	Phosphorylation of ERK by PKC	k7	2.64869098 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
15	Deactivation of G	k0	5.09854239 (s-1)
16	Deactivation of DAG	k4	0.46054447 (s-1)
17	Deactivation of PKC	k6	1.7662961 (s-1)
18	Dephosphorylation of GpERK	k8	3.03854703 (s-1)
19	Initial GRK-2 and -3 quantity	GRK23	0.899447579 ( $\mu\text{mol L}^{-1}$ )
20	Initial $\beta$ -arrestin 1 quantity (barr1)		0.858 ( $\mu\text{mol L}^{-1}$ )
21	Initial HRP1barr1 quantity		0 ( $\mu\text{mol L}^{-1}$ )
22	Initial HRP1barr2 quantity		0 ( $\mu\text{mol L}^{-1}$ )
23	Initial HRP1 quantity		0 ( $\mu\text{mol L}^{-1}$ )
24	GRK-2 and -3 phosphorylation rate	Vm18	1.35620014 (s-1)
25	GRK-2 and -3 phosphorylation rate	Km19	2.29886703 ( $\mu\text{mol L}^{-1}$ )
26	Association rate of HRP1 with barr1	k12	2.60941119 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
27	Association rate of HRP1 with barr2	k15	2.59345866 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
28	Dissociation rate of HRP1 with barr1	k13	0.00618527927 (s-1)
29	Dissociation rate of HRP1 with barr2	k16	0.0311141026 (s-1)
30	HRP1barr1 recycling rate	k14	6.53966254e - 05 (s-1)
31	HRP1barr2 recycling rate	k17	0.0723469566 (s-1)
32	HRP2 dephosphorylation	k25	0.0664816309 (s-1)
33	Initial GRK-5 and -6 quantity	GRK56	1.5180818 ( $\mu\text{mol L}^{-1}$ )
34	Initial HRbarr2		0 ( $\mu\text{mol L}^{-1}$ )
35	Initial HRP2barr2		0 ( $\mu\text{mol L}^{-1}$ )
36	Initial bpERK quantity		0.015 ( $\mu\text{mol L}^{-1}$ )
37	Initial HRP2		0 ( $\mu\text{mol L}^{-1}$ )
38	GRK-5 and -6 phosphorylation rate	Vm9	6.18288645 (s-1)
39	GRK-5 and -6 phosphorylation rate	Km10	2.72407145 ( $\mu\text{mol L}^{-1}$ )
40	Association rate of HR with barr2	k11	0.2045906 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
41	Association rate of HRP2 with barr2	k21	1.04254722 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
42	Phosphorylation of ERK by HRbarr2	k23	0.000419664616 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
43	Phosphorylation of ERK by HRP2barr2	k24	14.4389992 ( $\mu\text{mol}^{-1} \text{L s}^{-1}$ )
44	Dissociation rate of HR with barr2	k20	1.04519973 (s-1)
45	Dissociation rate of HRP2 with barr2	k22	0.346933098 (s-1)
46	bpERK dephosphorylation	k27	0.762264567 (s-1)
47	Initial ERK quantity		4.243 ( $\mu\text{mol L}^{-1}$ )
48	Initial $\beta$ -arrestin 2 quantity (barr2)		0.483 ( $\mu\text{mol L}^{-1}$ )
49	Inhibited inactive PKC quantity		0.00173556731 ( $\mu\text{mol L}^{-1}$ )
50	Inhibited active PKC quantity		2.09075206e - 05 ( $\mu\text{mol L}^{-1}$ )
51	Inhibited initial $\beta$ -arrestin 2 quantity		0.000112306963 ( $\mu\text{mol L}^{-1}$ )
52	Inhibited GRK-2 and -3 quantity		0.478865092 ( $\mu\text{mol L}^{-1}$ )
53	Inhibited GRK-5 and -6 quantity		0.00103512491 ( $\mu\text{mol L}^{-1}$ )
54	ERK norm	erknorm	0.01265 ( $\mu\text{mol L}^{-1}$ )
55	DAG norm	DAGnorm	4.11974423 ( $\mu\text{mol L}^{-1}$ )
56	PKC norm	PKCnorm	7.20901786 ( $\mu\text{mol L}^{-1}$ )

Supplementary figure 6



$Vm9 * GRK23 * [R-H] / (Km10 + [R-H])$  for  $R-H \Rightarrow R\sim\{P1\}-H$ .  
 $Vm18 * GRK56 * [R-H] / (Km19 + [R-H])$  for  $R-H \Rightarrow R\sim\{P2\}-H$ .

MA(k1) for  $G = [R\sim\{P1\}-H] \Rightarrow G\_a$ .  
MA(k2) for  $G = [R-H] \Rightarrow G\_a$ .  
MA(k0) for  $G\_a \Rightarrow G$ .  
MA(k26) for  $G \Rightarrow G\_a$ .

MA(k3) for  $PIP2 = [G\_a] \Rightarrow DAG$ .  
MA(k4) for  $DAG \Rightarrow PIP2$ .

MA(k5) for  $PKC = [DAG] \Rightarrow PKC\_a$ .  
MA(k6) for  $PKC\_a \Rightarrow PKC$ .

MA(k7) for  $ERK = [PKC\_a] \Rightarrow GpERK$ .  
MA(k8) for  $GpERK \Rightarrow ERK$ .

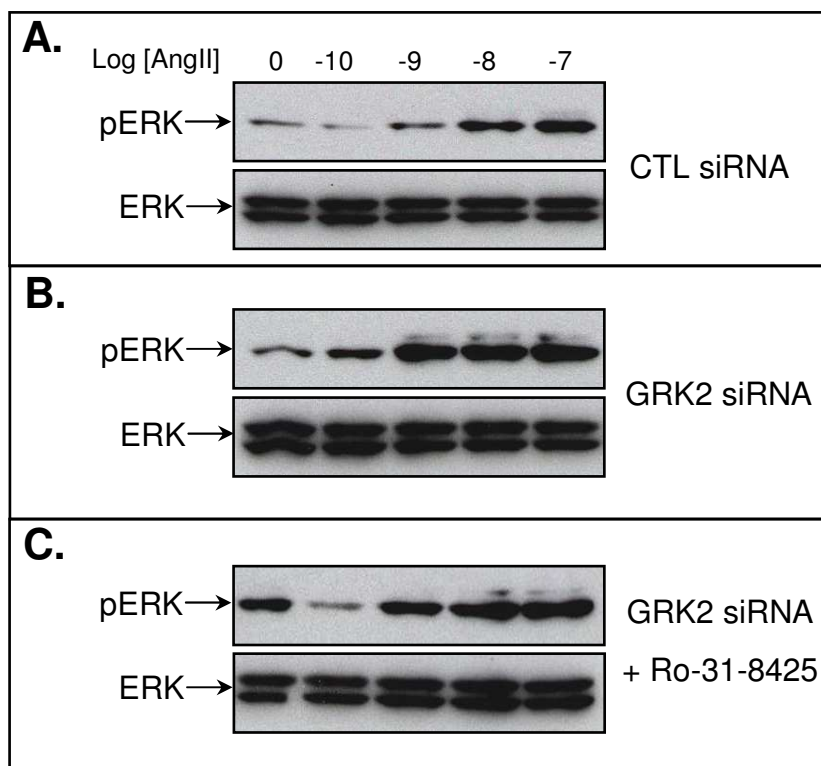
(MA(k12), MA(k13)) for  $R\sim\{P1\}-H + barr1 \Leftrightarrow R\sim\{P1\}-H-barr1$ .  
MA(k14) for  $R\sim\{P1\}-H-barr1 \Rightarrow R-H + barr1$ . % recyclage  
(MA(k15), MA(k16)) for  $R\sim\{P1\}-H + barr2 \Leftrightarrow R\sim\{P1\}-H-barr2$ .  
MA(k17) for  $R\sim\{P1\}-H-barr2 \Rightarrow R-H + barr2$ . % recyclage

(MA(k11), MA(k20)) for  $barr2 + R-H \Leftrightarrow barr2-R-H$ .  
(MA(k21), MA(k22)) for  $barr2 + R\sim\{P2\}-H \Leftrightarrow barr2-H-R\sim\{P2\}$ .  
MA(k23) for  $ERK = [barr2-R-H] \Rightarrow bpERK$ .  
MA(k24) for  $ERK = [barr2-H-R\sim\{P2\}] \Rightarrow bpERK$ .

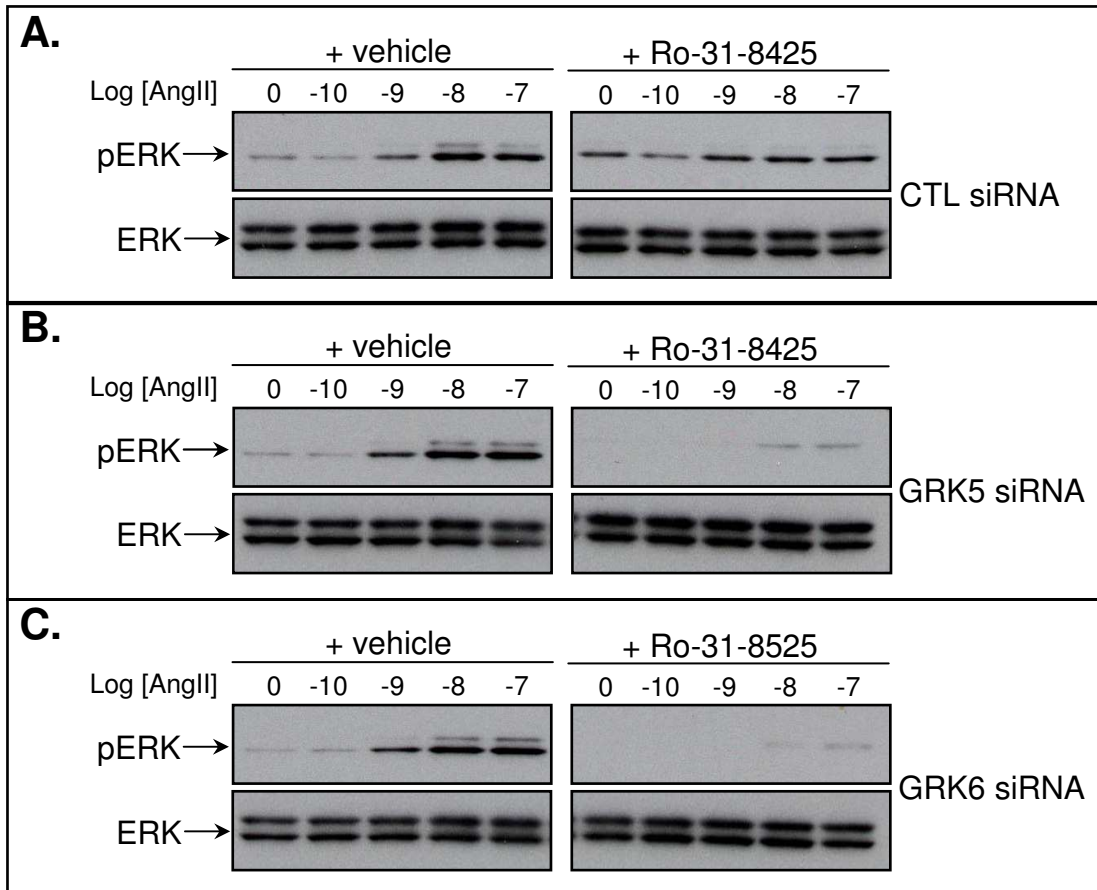
MA(k27) for  $bpERK \Rightarrow ERK$ .

MA(k25) for  $R\sim\{P2\}-H \Rightarrow R-H$ .

**Supplementary figure 7**



Supplementary figure 8



Supplementary figure 9

### 3.2.2 Conclusion et perspectives

L'environnement Biocham ainsi que son formalisme d'expression de la fonction objectif en QFLTL nous a permis d'ajuster un modèle qui a pu être validé par plusieurs données complémentaires. De plus, nous avons participé à l'amélioration d'un logiciel qui n'était pas prévu au départ pour l'ajustement paramétrique en plusieurs conditions. Cependant, au cours de ce travail, des interrogations sont apparues concernant les jeux de données nécessaires à la réalisation de tels modèles et la méthode d'optimisation utilisée. L'environnement Biocham n'étant pas conçu pour ce type de tests à grande échelle, nous avons développé un programme en langage C qui, certes, ne permet pas l'utilisation des propriétés temporelles, mais permet une grande flexibilité et une bonne rapidité de calcul.

### 3.3 Mise au point et évaluation d'une méthode hybride d'optimisation paramétrique

Cette dernière partie des résultats a pour objectif de déterminer les meilleures conditions biologique et la meilleure méthode d'optimisation paramétrique permettant la construction d'un modèle dynamique ODE. Lorsque l'on veut étudier un réseau de signalisation à l'aide d'une modélisation dynamique ODE, il faut connaître l'ensemble des valeurs des paramètres de l'ODE (valeurs initiales, constantes cinétique, normes, voir chapitre 1.5). Or ces paramètres sont généralement inconnus. Pour les déterminer plusieurs procédures d'ajustements paramétriques peuvent s'appliquer (chapitre 1.5). Cependant une multitude de questionnement subsiste. Est-il possible d'obtenir des résultats en initialisant la recherche par des valeurs aléatoires pour les paramètres inconnus? Quelle quantités de données fournir? de quelle nature? combien de point de temps cinétiques sont nécessaires? et quelle variabilité peut être acceptée? Ces *entrées biologiques* varient selon les réseaux observés et peuvent être fonction l'une de l'autre. Si peu d'espèces sont observées, alors beaucoup de points de temps peuvent compenser. D'autre part, au cours des recherches pour le modèle d'activation des ERKs, nous avons constaté que la méthode d'optimisation CMAES ne convergeait pas toujours. En revanche, l'algorithme génétique propose toujours une solution. Combien de fois faut-il lancer l'ajustement pour obtenir un nombre significatif de jeux de paramètres pouvant contenir la solution recherchée? Combien d'itération par recherche sont suffisante? Avec le modèle précédent, nous avons observé que lorsque CMAES ne convergeait pas, le nombre d'itération n'y changeait rien. Les solutions d'ajustement provenant de l'algorithme génétique sont-elle comparables aux solutions CMAES? Est-il intéressant de coupler les deux méthodes pour obtenir à coup sûr une solution et de bonne qualité? Avant de les coupler que retient-on de la première méthode? la moyenne des jeux? Le meilleur?

Pour répondre à ces questions, nous nous sommes basés sur cinq modèles pouvant correspondre à des réseaux biologiques et ayant une complexité croissante. Avec ces modèles tests, décrits et illustrés ci-dessous, nous avons comparer les ajustements paramétriques réalisés par l'algorithme génétique, l'algorithme CMAES puis le couplage algorithme génétique - CMAES dans différentes situations. Les paramètres des modèles tests sont entièrement connus, la méthode d'étude consiste à simuler ces modèles sous différentes conditions, à ajouter du bruit (5%, 10% ou 20%) à une partie de ces données puis à demander un ajustement paramétrique des modèles à partir de paramètres aléatoires sur ces données partielles « bruitées ».

La dernière méthode hybride s'avère la plus performante et parfaitement adaptée à la paramétrisation des réseaux de signalisation. Enfin, et pour mettre à l'épreuve ces résultats obtenus sur des fonctions tests, nous avons décidé de reprendre le modèle FSH focalisé sur les phosphorylations de la p70S6K. C'est le sujet de l'article en cours d'écriture qui suit.

Article à soumettre

**HYPE: HYbrid method for Parameter Estimation in molecular systems.**

Domitille Heitzler<sup>1</sup>, Jérôme Azé<sup>2</sup>, Pascale Crépieux<sup>1</sup> Eric Reiter<sup>1</sup> Anne Poupon<sup>1,\*</sup>

**1** BIOS group, INRA, UMR85, Unité Physiologie de la Reproduction et des Comportements, F-37380 Nouzilly, France; CNRS, UMR6175, F-37380 Nouzilly, France ; Université François Rabelais, 37041 Tours, France

**2** Bioinformatics group – AMIB INRIA – Laboratoire de Recherche en Informatique, Université Paris-Sud, 91405 Orsay, France

**Key Words:** computational modeling, ODE, parameter optimization, evolutionary algorithm, signaling networks.

\* To whom correspondence should be addressed

Anne Poupon, UMR 6175, 37380, Nouzilly, France.

Telephone: +33 2 47 42 75 05 , Fax: +33 2 47 42 77 43 E-mail address:

Anne.Poupon@tours.inra.fr

### 3.4 Conclusion

Au cours de ses trois projets d'envergures, nous avons pu apporter une vision rationnelle à deux réseaux de signalisation qui pourrait se généraliser à la signalisation d'autres R7TMs. De plus, nous avons réussi, avec des modèles de petites tailles à prédire des comportements et des aspects structurels originaux. En parallèle, nous avons soulevé les problèmes d'ajustements paramétriques et nous avons proposé une solution fiable pour les réseaux de signalisation. Dans la partie qui suit, nous allons discuter certains éléments, prendre un peu de recul sur le travail déjà effectué et proposer des perspectives à explorer.

Quatrième partie

Discussion et Perspectives





---

Chez l'homme, les R7TMs couvrent des domaines très variés et sont responsables du maintien des grandes fonctions physiologiques. Leur dysfonctionnement conduit à de nombreuses pathologies. D'autre part, les R7TMs sont ciblés par plus de 30% des médicaments actuels. Cependant, les concepts pharmacologiques et les méthodes de criblage classiques ont de plus en plus de mal à identifier de nouveaux médicaments. En effet, la complexité combinatoire des voies de signalisation induites par les R7TMs couplée à l'afflux de données biologique soulève un besoin d'approches mathématiques et informatiques pour rationaliser les études. En particulier, la recherche de ligands biaisés capable d'activer sélectivement certaines parties du réseau de signalisation. Ces études pourraient être réalisées méthodiquement et conduire à de nouveaux médicaments potentiellement plus sélectifs et porteur de moins d'effets secondaires.

Au cours de cette thèse, j'ai pu mesurer les nombreuses limitations qui rendent problématique la modélisation mathématique des voies de signalisation induites par les R7TMs. Dans certains cas, nous avons pu contourner les problèmes ou leur apporter des solutions originales. Nous avons également pu identifier des domaines sensibles où des progrès devront être réalisés en priorité.

Tout d'abord, les données biologiques disponibles dans la littérature sont souvent lacunaires et hétérogènes. Les études ne sont pas les mêmes pour l'ensemble des R7TMs et les différentes expériences portant sur un R7TM ne sont pas faites à l'identique pour les différents types cellulaires. On a donc pour un même R7TM, des résultats variables selon les auteurs, les approches expérimentales et les types cellulaires. De plus, l'hétérogénéité peut être présente au sein d'un jeu de données : il arrive que les contrôles des conditions perturbées ne soient pas comparables entre eux. En effet, pour l'observation des ERK phosphorylés du modèle angiotensine, il y a deux contrôles pour les deux conditions :  $\beta$  arrestines 2 siRNA et Ro-31-84-25. Pour la première condition, le contrôle consiste en des cellules transfectées avec un siRNA inactif tandis que le contrôle de la seconde condition consiste à traiter les cellules avec le véhicule seul (DMSO). Il aurait été plus rigoureux de procéder comme suit :

- contrôle = siRNA contrôle + solvant (DMSO)
- siRNA barr2 = siRNA barr2 + solvant (DMSO)
- Ro 31-8425 = siRNA contrôle + solvant contenant Ro.

Le modèle mathématique, lui, traite l'ensemble des quantités globalement et permet des inhibitions à la source sans perturbateurs extérieurs. Dans certains cas, le modèle peut s'adapter aux hétérogénéités des expériences et décrire une réalité biaisée. Dans d'autres cas, l'impossibilité de rendre compte des données avec le modèle révèle l'existence de problème et peut montrer où l'hétérogénéité des données impacte.

Au delà de l'hétérogénéité des données, leur interprétation et le passage en quantités molaires est un travail fastidieux. En effet, nous avons eu l'idée de traduire en quantité molaire les expériences FRET de manière proportionnelle. Le passage du quotient des intensités normées (données FRET) à une concentration s'effectue en soustrayant la quantité basale puis en normalisant (section 2.1). Cette interprétation est-elle correcte ? Rien ne prouve que le quotient des intensités lumineuses est proportionnelle à une concentration. Existe-t-il toujours une interprétation en terme de concentration molaire capable d'être formalisée en une observable dépendante des variables ?

---

L'interprétation des données FRET nécessite la connaissance des quantités basales ainsi que la norme à associer à la cinétique. Ces grandeurs ne sont pas mesurées et doivent être estimées. Les expériences de « Western blots » mesurent la quantité initiale mais nécessitent aussi la recherche d'une norme. En effet, les résultats sont fonction d'un 100% égal au maximum de la réponse. Ainsi, pour les quantités de PIP3, p70S6K et rpS6, qui ont servi au modèle FSH, nous avons trouvé une mesure de PIP3 au sein de notre système et nous en avons déduit les concentrations de PIP3 au cours du temps. Pour les deux autres observables : p70S6K et rpS6 nous avons supposé que les contrôles insuline étaient équivalents d'une espèce à l'autre. Or, ce n'est pas forcément le cas, on pourrait supposer qu'il existe des disparités de concentrations entre espèces. Ce problème de normes est un vaste problème en tant que tel. Ajouter une norme inconnue à la recherche des paramètres semble ouvrir un degré de liberté à l'ajustement. On remarque souvent que pour chaque valeur de norme, il existe un ajustement des autres paramètres de l'ODE. Au delà, existe-t-il tout de même une valeur de norme, proche de la réalité, qui donnerait de meilleurs ajustements? Ces différents ajustements ne sont pas linéairement dépendants car l'ODE est non-linéaire. Plusieurs paramètres de normes inconnues compliquent la recherche des autres paramètres et engendrent des difficultés de convergence. Il serait intéressant de tester le réel impact d'une ou plusieurs normes au sein d'un ajustement paramétrique. Ces normes encore très peu mesurées expérimentalement ont une importance indéniable au sein des systèmes dynamiques quantitatifs. À l'avenir, tendre vers des méthodes biologiques quantitatives permettrait de s'affranchir de ce type de problèmes.

Une observation de nombreuses molécules du réseau en parallèle et leur quantification représenteront une grande avancée pour le développement de modèles dynamiques. Cet objectif est maintenant atteignable. En effet, la mise au point de la méthode RPPA (voire Annexe A) permet d'observer une espèce dans un grand nombre de conditions expérimentales (jusqu'à 12 000 spots par puce). De plus, comme le dépôt des échantillons sur la puce est reproductible et ne nécessite que des quantités infimes d'extraits cellulaires, les mêmes jeux d'expériences peuvent être analysés avec de nombreux anticorps différents, pour observer différentes espèces. Ce type d'approche va fortement améliorer la vitesse d'obtention des données, l'homogénéité de celles-ci ainsi que leur fiabilité (répétitions à la demande et gamme de détection dynamique plus large).

D'autre part, cette méthode couplée d'une gamme étalon pour les espèces observées a permis la quantification des observations et ouvert de nouveaux horizons. Ainsi, comme suggéré *in silico*, la comparaison des différents éléments de la cascade MAPK ERK a révélé une amplification significative y compris au sein de la voie  $\beta$ -arrestine. Précédemment, sans connaissance au préalable des normes à relier aux ERKs phosphorylés, nous avons admis que le complexe *bpERK* était en conditions stoechiométrique 1 ( $bpERK = 1$  récepteur + 1  $\beta$ -arrestine + 1 Raf + 1 MEK + 1 ERK). Et, lorsque la quantification de *bpERK* est apparue, la construction en complexe s'est effondrée car il y avait 15 fois moins de récepteur que de *bpERK*. La modélisation du réseau d'interaction de l'angiotensine a ainsi mis en lumière une amplification au niveau de la voie  $\beta$ -arrestine. À l'aide des puces RPPA, l'amplification entre les MEK et les ERK de la voie  $\beta$ -arrestine a pu être quantifiée. D'une part, il devrait être possible à l'avenir, d'étendre ces quantifications sur puces à de nombreux nœuds du réseau de signalisation. D'autre part, lorsqu'il y a une amplification, plusieurs possibilités peuvent co-exister. En effet, on a vu en introduction que le récepteur angioten-

---

sine recrutait les  $\beta$ -arrestines une fois phosphorylé. Cependant, [DeWire et al., 2007] précise qu'un mutant du récepteur angiotensine auquel on a retiré les sites de phosphorylation est toujours capable de recruter les  $\beta$ -arrestines mais de manière différente : rapide et brève au lieu de soutenue. On pourrait donc imaginer que les différentes possibilités de recrutement existent au sein du réseau de signalisation. L'observation expérimentale relateraient en fait une somme d'amplification des différentes voies coexistantes. Comment les quantifier chacune indépendamment ? N'est-ce pas encore un défi pour la modélisation ?

Outre les amplifications, bien d'autres éléments de la structure du modèle sont délicats à spécifier et une multitude de possibilités existe. En quête d'une solution, [Nelander et al., 2008] ont inséré une contrainte pour déterminer la structure minimale (contenant un minimum de réaction) au sein de leur ajustement paramétrique sur les données expérimentale. Cependant, il s'avère souvent que les réactions d'un réseau de signalisation sont un enchevêtrement de régulations pas forcément minimales mais nécessaires à garantir un équilibre en toute circonstance. Parfois, les données expérimentales nous donnent des indications ou nous aide à éliminer les structures impossibles mais bien souvent, l'interprétation et le choix arbitraire prennent le dessus. Dans d'autre cas, l'ajustement paramétrique identifie clairement une structure qui offre de meilleures solutions au vu des données. Ce sont ces allers/retours entre la structure du modèle et l'ajustement paramétrique qui nous ont permis d'inférer les modèles FSH et angiotensine (sections 3.1 et 3.2). Cependant, dans certains cas, des expériences dédiées doivent être réalisées pour discriminer ou valider une nouvelle structure du réseau suggérée par l'approche empirique décrite ci-dessus. À l'avenir, nous aimerions construire le réseau par inférence automatique afin de se soustraire à une quelconque interprétation. C'est le sujet d'une thèse en cours au sein de l'équipe BIOS à l'INRA. Plus précisément, les différentes structures d'un réseau de signalisation seront proposées à partir de l'ensemble des connaissances ordonnées au sein d'une base de données. Un programme d'inférence déduira des expériences, les règles possibles entre les différentes espèces du réseau sans *a priori* du biologiste ni du modélisateur.

Cependant, si une espèce n'est pas explicitée au sein de la base de données, la méthode d'inférence ne pourra pas la placer au sein du réseau de signalisation. Nombre d'intermédiaires et de catalyses restent inconnus mais parfois cela peut dépendre de l'exploitation des données qui reste un travail critique en amont de la base de données. Par exemple, pour l'établissement du modèle FSH - p70S6K, nous étions persuadé d'observer deux espèces de p70S6K celle qui était phosphorylée sur le site {T389} et celle phosphorylée sur le site {S421/424}. Ces p70S6K passaient d'un état à l'autre comme si ces états étaient disjoints. Or, en disséquant la p70S6K, nous avons trouvé quatre formes : p70S6K, Pp70S6K~{T389}, Pp70S6K~{T421/S424} et Pp70S6K~{T389,T421/S424} et nous avons découvert que nos observables partageaient la forme Pp70S6K~{T389,T421/S424}. C'est ainsi que l'apparente stabilité de l'observation des formes phosphorylé sur le site {T389} serait due à la somme d'une augmentation de Pp70S6K~{T389} compensée par une diminution de P70S6K~{T389,T421/S424} [Musnier et al., 2009].

Malgré ce niveau de détail et les catalyses précises pour chaque phosphorylation, en réalisant l'ajustement paramétrique pour les deux stades de développement, un degré de liberté supplémentaire a du être ajouté : le paramètre  $k7$  qui peut varier d'un stade à l'autre. Ce biais, suggère la présence d'un effecteur non identifiés qui catalyserait la conversion de p70S6K en Pp70S6K{T389} dans les cellules en prolifération ou, plus proba-

---

blement, l'existence d'autres sites phosphorylés que les deux déjà étudiés qui donnerait à la p70S6K ce comportement différent en fonction du stade de développement. La p70S6K comporte d'autres sites de phosphorylation, qui n'ont pas été détaillé par manque de données ou d'outils expérimentaux. Cependant, le détail de l'ensemble des sites de phosphorylations aurait engendré une augmentation importante de la taille du modèle. En effet pour trois sites de phosphorylation, on compte huit intermédiaires, et pour  $n$  sites,  $2^n$ . On peut se demander si cette complexité aurait été compatible avec notre formalisme simple de modélisation et si l'apport d'information, une observable par site phosphorylé, aurait suffi devant l'ajout de variables et de paramètres inconnus. La modélisation nécessite bien souvent de trouver un équilibre entre détails et capacités.

En effet, pour le modèle angiotensine, les combinaisons précises de sites phosphorylés n'ont donc pas eu besoin d'être détaillées pour permettre au modèle de rendre compte de l'action différentielle des GRKs. Nous avons supposés que deux couples de GRKs phosphorylaient le récepteur sur deux combinaisons de sites : {P1} pour GRKs 2 et 3, {P2} pour GRKs 5 et 6. Selon sa combinaison de phosphorylation, le récepteur angiotensine est orienté vers la désensibilisation ou la signalisation dépendante des  $\beta$ -arrestines. De plus et malgré l'agrégation des GRKs 2 et 3 en une seule espèce et l'agrégation des GRKs 5 et 6 en une deuxième espèce, nous avons obtenu une précision inattendue sur les modes d'action des GRKs : les ajustements paramétriques indiquent que lorsque les GRKs 2 ou 3 sont inhibées *in vivo*, leurs quantités correspondante *in silico*, *GRK23*, diminuent de moitié tandis que pour GRKs 5 ou 6, *GRK56* est quasi-nulle. Cette propriété a été retrouvée pour les 5 jeux de paramètres trouvés par Biocham-CMAES en partant d'un jeu aléatoire. Ces résultats suggèrent que les GRKs 2 et 3 agissent « en parallèle » pour phosphoryler le récepteur tandis que les GRKs 5 et 6 « en série ». Ou, une autre possibilité est d'imaginer que GRK 2 phosphorylerait les mêmes sites sur le récepteur que GRK 3 tandis que les GRKs 5 et 6 phosphoryleraient des sites différents. En outre, sans la combinaisons des sites phosphorylés par GRKs 5 et 6 le récepteur ne pourrait pas poursuivre sa voie de signalisation par le recrutement des  $\beta$ -arrestines puis la formation de ERK. Ces différentes possibilités restent à vérifier expérimentalement. Ce mécanisme agrégé semblent pouvoir se généraliser et s'adapter à d'autres R7TMs comme la combinaison de deux signalisations : une rapide et brève et une autre plus lente mais soutenue. Ces cinétiques se retrouvent pour les voies G et  $\beta$ -arrestines pour d'autres R7TMs [Luttrell and Gesty-Palmer, 2010]. En adaptant les espèces, les observables et les jeux de données, notre structure minimale de modèle pourraient servir de bases à d'autres signalisations intracellulaires. Enfin, sachant que les différentes voies de signalisation engendrent des réponses physiologiques distinctes, les modèles dynamiques intégrant les différentes voies pourraient-ils faire le lien entre les mécanismes de transduction du signal et leurs conséquences physiologiques ?

Pour conclure sur le modèle angiotensine, on notera trois points.

Premièrement, l'ajustement paramétrique actuel ne permet pas de relier certains effets doses portant sur la comparaison des deux voies de transduction (dépendantes des protéines G et des  $\beta$ -arrestines) en fonction du temps (données non publiées). Dans un premier temps, nous pourrions penser que nous n'avons pas la bonne solution au niveau des paramètres et élargir les jeux solutions en effectuant d'autres ajustements paramétriques. Ceci a déjà été réalisé, en particulier en partant de valeurs aléatoires (voir paragraphe sur l'identifiabilité) et n'a donné que des résultats corrélés à notre jeux précédent ne permettant

---

pas non plus l'explication des effets doses supplémentaires. Comme nous le pensions, notre modèle est donc bien posé pour la fonction objectif et les données complémentaires utilisées au sein des résultats. Cependant, pour les effets doses supplémentaires, que faut-il modifier ? Un élément de la structure ? Pourrait-on le déterminer en ajoutant, comme [Nelander et al., 2008], une recherche sur la structure à l'ajustement paramétrique ? Pourquoi la structure serait-elle en cause ? A-t-on choisit un mauvais modèle réduit ?

Deuxièmement, nous avons un modèle détaillé du réseau de régulation des ERKs par l'angiotensine (Annexe 3.2). Serait-il possible de passer de ce modèle détaillé au modèle réduit en appliquant la méthode décrite récemment par [Gay et al., 2010] ? Si non, quels doit être les caractéristiques des modèles détaillés pour qu'ils admettent comme réduction le modèle réduit ? Inversement, serait-il possible de trouver une réduction rendrait compte des données utilisées au sein des résultats ainsi que des effets doses supplémentaires non publiés ?

Troisièmement, bien que qu'il existe différents compartiments cellulaires impliqués dans la signalisation du récepteur angiotensine, nous n'en avons pas tenu compte. Au niveau de l'impact des différents volumes des compartiments sur les concentrations, on peut dire qu'il a été agrégé au sein des constantes cinétiques. Quant aux réactions de transfert entre les différents compartiments, elles sont absentent du modèle. Dans un tel modèle, nous devrions répartir le récepteur transmembranaire dans deux compartiments distinct puisqu'il est à la fois intégré à la membrane plasmique et en partie dans le cytoplasme. Une solution pourrait être de définir un compartiment de travail comprenant la membrane plasmique et son voisinage cytoplasmique proche comme un seul compartiment. Avec ces compartiments et les réactions de transfert supplémentaires, comment les dynamiques du modèles vont-elle être affectées ? La dimension du modèle restera-t-elle raisonnable pour un ajustement paramétrique ?

Un constat important réalisé au cours de cette thèse est que certaines méthodes d'optimisation atteignent leur limite lorsque le modèle est trop large. Par la suite, nous avons mis en évidence d'autres facteurs de difficultés pour l'ajustement paramétrique comme les différents ordres de grandeur entre les paramètres, les paramètres de normes, les quantités totales et initiales. En particulier, au cœur de l'optimisation, il y a le calcul de la distance simulation - données au sein de la fonction objectif. Pourrait-on comparer les performances d'un ajustement paramétrique selon la distance utilisée ? En particulier, pour les fonctions objectifs qui peuvent s'exprimer en propriétés QFLTL et à l'aide d'une fonction usuelle, quelle distance permet le meilleur et le plus rapide ajustement paramétrique ? D'autre part, nous avons étudié les performances de deux méthodes d'optimisation : l'algorithme génétique et CMAES seuls puis couplés. Le couplage apparaît comme la meilleure procédure (article section 3.3). Pourquoi, CMAES seul, un algorithme très performant, se perd-il dans sa recherche ? Existe-t-il un paramétrage de CMAES plus adéquat à nos modèles ? Pour aller plus vite que le couplage de 30 algorithmes génétique avec 30 CMAES, peut-on créer un algorithme génétique auto-adaptatif qui utiliserait CMAES comme mutation pour les jeux de paramètres parents dont l'erreur est faible ? Toutefois, nous pensons que certaines difficultés pour ajuster un modèle sont en lien avec l'identifiabilité du modèle.

**Définition 3.4.0.1.** Un modèle est dit identifiable s'il existe un unique jeu de paramètre solution pour les données expérimentales associées.

---

Tout comme l'ajustement paramétrique, l'identifiabilité relève d'un domaine mathématique en soi. Il existe des résultats concernant les ODEs, cependant, pour le cas non-linéaire qui nous concerne, le problème de l'identifiabilité n'est pas totalement résolu. C'est pourquoi, nous avons au cours de cette thèse simplement effleuré le sujet. Cependant, il serait intéressant, de connaître les propriétés d'identifiabilité des modèles « biochimiques ». Pour les modèles non-linéaires constitués uniquement de termes constants, linéaires et bilinéaires, des conditions nécessaires et suffisantes existent. Pourrait-on les tester facilement et automatiquement ? Quels sont les principes à appliquer en construisant nos modèles pour minimiser la non-identifiabilité ? Quels sont les termes et les conditions qui posent le plus problème ? En conséquence de tout cela quelles sont les possibilités expérimentales existantes pour résoudre certains manques au niveau des paramètres ?

Par ailleurs et de manière empirique, nous avons remarqué qu'en réalisant, à partir de paramètres aléatoires, 100 ajustements paramétriques Biocham-CMAES, seules cinq solutions acceptables ont été trouvées. De plus, il est apparu que ces cinq jeux de solutions étaient fortement corrélées. Peut-on dire que ce modèle muni de sa fonction objectif est empiriquement identifiable ? Les modèles réellement identifiables ont-ils aussi de telles propriétés ? Cela dépend-il de l'algorithme d'optimisation des paramètres ?

Ces procédures d'ajustement paramétrique répétées nécessitent beaucoup de temps de calculs et cela est en partie dû à l'intégration numérique réalisée pour simuler le modèle ODE. Existerait-il des procédures simplifiées plus rapides pour certains types de modèles ?

Dans un autre esprit, pourrait-on remplacer les ODEs par un formalisme plus léger tout en gardant les propriétés temporelles et quantitatives très importantes en signalisation cellulaire ? Le formalisme de simulations stochastiques est-il une solution ? Les simulations stochastiques ne nécessitent pas d'intégration numérique. Certains formalismes stochastiques sont coûteux en temps de calculs mais en existe-t-il d'autres plus rapides que l'intégration numérique des ODEs ? Plus précisément, un modèle stochastique associe à chaque concentration une quantité dénombrable de molécules. À chaque pas de temps, les différentes possibilités de transitions sont calculées et mise en parallèle pour déterminer l'état du modèle au pas de temps suivant. Lorsque les molécules de chaque espèce sont en grand nombre, les simulations stochastiques ressemblent aux simulations ODE, ce qui n'est pas le cas des petites quantités. Ce formalisme serait-il plus réaliste au sein de la cellule où les récepteurs sont en petit nombre ? Et, permettrait-il de surveiller les molécules et préciser ou découvrir comment certaines espèces interagissent. Par exemple, nous savons qu'au sein des cellules, il y a une amplification entre le récepteur et la protéine G. Comment le modèle l'explique-t-il ? Est-ce que peu de protéines G interagissent avec l'ensemble des récepteurs par un mécanisme de recyclage des protéines G rapide et efficace ou beaucoup de protéines G se relayent pour interagir avec les récepteurs car leur temps de recyclage est non-négligeable ?

# Annexes





**Annexe A**

**Méthode RPPA**

## RESEARCH ARTICLE

# A highly sensitive near-infrared fluorescent detection method to analyze signalling pathways by reverse-phase protein array

Laurence Dupuy<sup>1,2,3,4</sup>, Christophe Gauthier<sup>1,2,3,4</sup>, Guillaume Durand<sup>1,2,3,4</sup>, Astrid Musnier<sup>1,2,3,4</sup>, Domitille Heitzler<sup>1,2,3,4</sup>, Adrien Herledan<sup>5</sup>, Vehary Sakanyan<sup>5,6</sup>, Pascale Crépieux<sup>1,2,3,4</sup> and Eric Reiter<sup>1,2,3,4</sup>

<sup>1</sup> BIOS group, INRA, Unité Physiologie de la Reproduction et des Comportements, Nouzilly, France

<sup>2</sup> CNRS, Nouzilly, France

<sup>3</sup> Université François Rabelais, Tours, France

<sup>4</sup> Haras Nationaux, Nouzilly, France

<sup>5</sup> Laboratoire de Biotechnologie, Université de Nantes, Nantes, France

<sup>6</sup> ProtNeteomix, Nantes, France

The comprehensive and quantitative analysis of the protein phosphorylation patterns in different cellular context is of considerable and general interest. The ability to quantify phosphorylation of discrete signalling proteins in large collections of biological samples would greatly favour the development of systems biology in the field of cell signalling. Reverse-phase protein array (RPPA) potentially represents a very attractive approach to map signal transduction networks with high throughput. In the present report, we describe an improved detection method for RPPA combining near-infrared with one or two rounds of tyramide-based signal amplification. The LOQ was lowered from 6.84 attomoles with a direct detection protocol to 0.21 attomole with two amplification steps. We validated this method in the context of intracellular signal transduction triggered by follicle-stimulating hormone in HEK293 cells. We consistently detected phosphorylated proteins in the sub-attomole range from less than 1 ng of total cell extracts. Importantly, the method correlated with Western blot analysis of the same samples while displaying excellent intra- and inter-slide reproducibility. We conclude that RPPA combined with amplified near-infrared detection can be used to capture the subtle regulations intrinsic to signalling network dynamics at an unprecedented throughput, from minute amounts of biological samples.

Received: March 27, 2009  
Revised: July 9, 2009  
Accepted: September 7, 2009

**Keywords:**

Cell signalling / FSH / Near-infrared fluorescence / Reverse-phase protein arrays

## 1 Introduction

Amongst post-translational modifications, reversible protein phosphorylation is centrally involved in signal transmission within cells [1]. The comprehensive and quantitative analy-

sis of the protein phosphorylation patterns in different cellular background is therefore of paramount and general interest. Nevertheless, current techniques able to capture the subtle and highly dynamic nature of phosphorylation cascades have limited throughput.

**Correspondence:** Dr. Eric Reiter, INRA, UMR85, Unité Physiologie de la Reproduction et des Comportements, F-37380 Nouzilly, France

**E-mail:** eric.reiter@tours.inra.fr

**Fax:** +33-2-47-42-77-43

**Abbreviations:** Ab, antibody; AF680-streptavidin, AlexaFluor-conjugated streptavidin; 95% CI, 95% confidence interval;

FCF, fast green; FSH, follicle-stimulating hormone; FSH-R, FSH receptor; HEK, human embryonic kidney; HEK293 FSH-R, HEK293 cells stably expressing FSH-R; HRP-streptavidin, HRP-conjugated streptavidin; MEM, Minimum Essential Medium Eagle; NIR, near-infrared; pERK, phosphorylated extracellular signal-regulated kinase; RPPA, reverse-phase protein array; RT, room temperature; WB, Western blot

Over the last decade, protein microarray-based methods, such as reverse-phase protein arrays (RPPA), have emerged for analyzing the proteome [2, 3]. RPPA requires small quantities of biological material while allowing to analyze a given protein in large collection of samples using phosphospecific antibodies [2]. Moreover, the simultaneous and automated processing of numerous samples under the same conditions, which is possible with RPPA, drastically reduces data heterogeneity and variability. To date, RPPA has been successfully applied to biomarker profiling [2, 4–6]. Attempts have also been made to analyze signalling network dynamics but applications to this field have been limited so far [7–11]. What makes the application of RPPA to signalling systems particularly challenging is that: (i) phosphorylation often accounts for a small proportion of the studied protein; and (ii) even subtle changes in phosphorylation levels can be meaningful and have to be captured. Those two points emphasize the need to develop very sensitive and quantitative detection of low abundance target proteins contained in the samples to be analyzed.

In RPPA, the protein of interest is usually detected by chromogen-based colorimetry [2, 3, 12], chemiluminescence [12, 13] or fluorescence [3, 8, 10, 12]. In addition, at least two commercially available bioanalytical systems, based on electrochemiluminescence detection [14], and on evanescent field fluorescent detection [15], have displayed excellent sensitivities (attomole and zeptomole range, respectively) and a wide dynamic range up to 5–6 logs. However, these technologies can only be utilized with proprietary devices and consumables, which most academic laboratories cannot afford to get. The colorimetry and chemiluminescence detection methods both exhibit good sensitivity, but poor dynamic range, which necessitates the samples to be spotted in serial dilutions for analysis, hence significantly reducing the throughput of the method. Fluorescent detection provides better dynamic range but is significantly less sensitive even though the use of near-infrared (NIR) dyes proved to dramatically increase the *S/N* [16]. Indeed, the NIR region of the spectrum provides a low-noise response from membrane supports and spotted biological molecules possessing visible auto-fluorescence [17]. We previously showed that a greater *S/N* permits to detect less than 50 attomoles of target proteins in lysates arrayed onto nitrocellulose membrane by using NIR dye-labelled DNA or protein probes in binding assays [16, 18].

Here, we describe an improved detection method for RPPA combining NIR with one or two rounds of signal amplification. We validated this method in the context of intracellular signal transduction triggered by follicle-stimulating hormone (FSH) in HEK293 cells. We repeatedly detected phosphorylated proteins in the sub-attomole range in less than 1 ng of total cell extracts. Importantly, the method proved to correlate perfectly with Western blot (WB) analysis of the same samples. Our data highlight that RPPA combined with amplified NIR detection can be used to improve the throughput of signalling pathways assessment.

## 2 Materials and methods

### 2.1 Cell culture and stimulation

HEK293 cells stably expressing FSH-R (HEK293 FSH-R) were grown in Minimum Essential Medium eagle (MEM) (Sigma-Aldrich, St Louis, MO, USA), supplemented with 10% FBS (Sigma-Aldrich), antibiotics (100 U/mL penicillin, 100 µg/mL streptomycin) (Sigma-Aldrich) and G418 (400 µg/mL) (Invivogen, San Diego, CA, USA) at 37°C with 5% CO<sub>2</sub> in a humidified atmosphere. Cells were serum-starved for 4 h with MEM containing 0.1% BSA, 10 mM HEPES and antibiotics before stimulation with 3 nM FSH. FSH was purified by Dr. J. Closset (Université de Liège, Liège, Belgium). The medium was then replaced by 400 µL of RPPA lysis buffer (50 mM Tris pH 6.8, 2% SDS, 5% glycerol, 1% β-mercaptoethanol, 5 mM EDTA, 1% phosphatase inhibitor cocktail 1 and 2 (Sigma-Aldrich)). Plates were frozen at –20°C until use. Once thawed at room temperature (RT), cells were harvested and boiled for 30 min. Cell lysates were centrifuged at 13 000 rpm for 10 min and the supernatants were stored at –20°C until use.

### 2.2 RPPA

Serial two-fold dilutions of active (phosphorylated) human recombinant pERK<sub>2</sub> (Protein kinase, BiAffin GmbH, KG, Germany) were prepared in RPPA lysis buffer. For cell lysates (1.7 mg/mL concentration in average), three two-fold serial dilutions were done in RPPA lysis buffer. Diluted samples were then aliquoted into 384-well plates and stored at –20°C. Plates were thawed/frozen three times at maximum to avoid sample degradation. Desiccated nitrocellulose-coated slides (Fast Slides, Whatman, Maidstone, UK) were printed with samples, using a 32-pin manual arrayer (Glass Slide Microarrayer, VP478, V&P Scientific, San Diego, CA, USA) and desiccated again overnight. Then, slides were either immunodetected or stored at –20°C. Our immunodetection procedure was adapted from Chan *et al.* [8]. All antibodies were pre-cleared in FBS for 1 h at 37°C prior to use. After rehydration with PBS-Tween 20 0.1% (PBST), slides were blocked for 3 h at RT with 3% casein in PBST. The optimal dilution of anti-phosphorylated ERK<sub>1/2</sub> (pERK<sub>1/2</sub>) antibody (Cell Signaling Technology, Beverly, MA, USA) was determined. The dilution of 1:400 was found to give the best *S/N* on slide. Slides were probed overnight at 4°C with anti-pERK<sub>1/2</sub> antibody (1:400) or anti-ERK<sub>2</sub> (C-14) (1:1000) (Santa Cruz Biotechnology, Santa Cruz, CA, USA) in PBST with 20% FBS (PBST-FBS). Slides then were washed three times for 5 min with three rinses with PBST between washes and then probed for 45 min at RT with either AlexaFluor® 680 goat anti-rabbit IgG (AlexaFluor680-labelled secondary antibody, 1:400, Invitrogen Life Technologies, Carlsbad, CA, USA) in PBST-FBS for a direct detection or with Peroxidase-conjugated F(ab')<sub>2</sub> Fragment

Donkey anti-rabbit IgG (1:2000) (Jackson ImmunoResearch Laboratories) diluted in PBST-FBS for amplified detection. For signal amplification, slides were subsequently incubated for 10 min at RT with BioRad Amplification Reagent (Amplified Opti-4CN Substrate Kit, BioRad, Hercules, CA, USA). Then, slides were rinsed six times, washed once for 5 min in PBST with 20% DMSO (PBST-DMSO) and washed twice for 5 min and rinsed in PBST as above. Slides were then probed for 1 h at RT with streptavidin AlexaFluor<sup>®</sup> 680 conjugate (0.2 µg/mL) (AF680-streptavidin) (Invitrogen Life Technologies) in PBST-FBS. Double signal amplifications were achieved by probing slides for 30 min at RT with streptavidin-conjugated HRP (HRP-streptavidin) (1:1000) in PBST-FBS following the first round of amplification. After three 5 min washes and rinses with PBST as described above, a second round of amplification was performed. After rinsing and washing, slides were probed for 1 h at RT with AF680-streptavidin (0.2 µg/mL) in PBST-FBS. Finally, slides were air-dried by centrifugation at 2500 rpm for 25 min and scanned at 700 nm with the Odyssey IR imaging system (LI-COR Biosciences, Lincoln, NE, USA) at a 42 µm resolution.

### 2.3 On-slide protein quantification

Fast Green (FCF, Sigma-Aldrich) was used for on-slide protein staining using a procedure adapted from Loebke *et al.* [10]. A standard curve of BSA ranging from 0.2 to 2 ng/nL was printed on slides in parallel with cell lysates. Slides were rehydrated in PBS and stained with FCF (0.001% w/w in water) for 3 h. Slides were washed three times for 5 min with PBS and then air-dried by centrifugation at 2500 rpm for 25 min and scanned as described above.

### 2.4 Western blotting

An equal volume of 2 × Laemmli buffer (32 mM Tris pH 6.8, 8% SDS, 5% β-mercaptoethanol, 10% glycerol, 0.02% bromophenol blue) was added to cell lysates or to pERK<sub>2</sub>, as prepared for RPPA. Next, samples were boiled for 10 min, resolved by 10% SDS-PAGE and transferred to nitrocellulose membrane (Protran<sup>®</sup>, Whatman, Maidstone, UK). Membranes were saturated with TBST-BSA (Tris 10 mM, pH 8, 150 mM NaCl with 0.1% Tween 20 and 4% BSA) and probed with primary antibodies: anti-ERK<sub>2</sub> (1:10 000) or anti-pERK<sub>1/2</sub> (1:3000) in TBST-BSA. WB was revealed with Supersignal West Pico Chemiluminescent Substrate (Pierce Biotechnology, Rockford, IL, USA). Films (Amersham hyperfilm<sup>™</sup> ECL) were scanned and the optical density of the signals was measured with ImageMaster 1D Elite version 4 software (Amersham Biosciences). Signal intensities were determined for each point with subtraction of minimum profile background. The whole sample collection was revealed on the same film. Different exposure times

were compared for each WB experiment in order to avoid signal saturation.

## 2.5 Data analysis

Scanned images of arrays were analyzed with GenePix Pro 6.0 software (Molecular Devices, Sunnyvale, CA, USA). The median value of intensity was determined for each spot with subtraction of the local background intensity. The signal intensity was considered significant when the *S/N* was higher than 1.5-fold. Signals were normalized according to the protein amount *per* spot as determined on an identical slide. The results were expressed as median intensities mean ± SEM of duplicate experiments for cell lysates and of six replicates for recombinant pERK<sub>2</sub>. The inter- and intra-slide reproducibility and correlation between RPPA and WB signals were analyzed by Pearson correlation analysis and statistical analysis was performed using one-way ANOVA with Bonferroni's post-test (GraphPad Prism 5 Software, San Diego, CA, USA). Values with *p* < 0.05 were considered as significantly different. Mean square deviations were calculated to compare intra-assay reproducibility of WB and RPPA. The theoretical LOD of the three methods reported here was determined as the mean signal of the blanks (obtained at zero analyte concentration) incremented by the two-fold standard deviations of the blanks [15]. We considered the lower LOQ as being the lowest pERK<sub>2</sub> concentration within the linear range of detection and with a *S/N* > 1.5. The upper LOQ was defined by the saturation threshold of the Odyssey scanner for fluorescence detection at 700 nm.

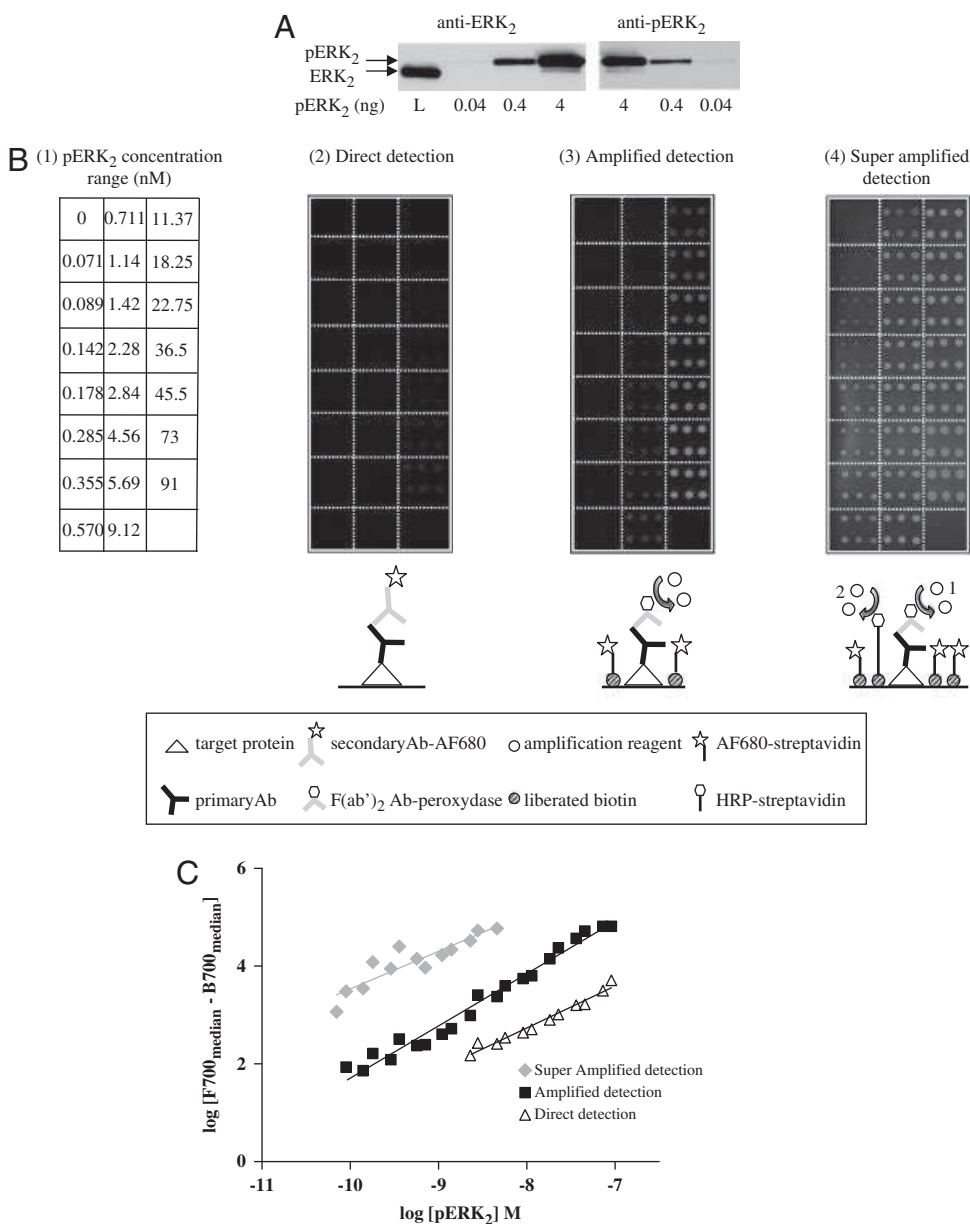
## 3 Results and discussion

### 3.1 Improvement of fluorescence-based detection method for RPPA

Fluorescence-based detection methods provide unmatched dynamic range but suffer from a relative lack of sensitivity when compared with colorimetric and chemiluminescent approaches. Our primary objective was to improve the sensitivity of fluorescence-based detection procedure for RPPA. First, as previously reported [10, 19], we found that detection in NIR dramatically improved *S/N*, when compared with visible wavelengths. The background measured with a visible fluorescent dye (AlexaFluor647) was 16 000 RFUs (relative fluorescent units) *versus* 100–500 RFUs for NIR fluorescent dye (AlexaFluor680) in identical experiments (data not shown). Second, we added one or two rounds of tyramide amplifications to the previously described direct detection protocol in NIR [10]. We then compared the three detection modes side-by-side, systematically monitoring the binding specificity and accuracy, the LOD and the LOQ for the three procedures. We tested

whether secondary antibody and AF680-streptavidin unspecifically bound to the purified recombinant pERK<sub>2</sub> or total proteins in cell lysates. In the absence of primary antibody, non-specific signal was very low (data not shown). We prepared standard curves using pERK<sub>2</sub> as a means to compare the different detection methods on a quantitative basis. The purity of pERK<sub>2</sub> was confirmed to be >95% by WB with anti-pERK<sub>1/2</sub> antibody raised against the phosphorylated sites, Thr 185 and Tyr 187 (Fig. 1A). No dephosphorylated form of pERK<sub>2</sub> was detected whereas a minor band, likely corresponding to single-phosphorylated form of pERK<sub>2</sub>, was observed. A wide range (71 pM to 91 nM) of pERK<sub>2</sub> serial dilutions was printed on nitrocellulose-coated slides (Fig. 1B, (1)). Then, pERK<sub>2</sub> was detected on three

identical slides, using the same anti-pERK<sub>1/2</sub> primary antibody (1:400 dilution) but with three different immunodetection procedures: an AlexaFluor680-labelled secondary antibody for a direct detection (Fig. 1B (2), direct detection); following secondary antibody incubation, one round of tyramide-based amplification was performed, liberating biotin, which deposited on the spot and was detected with AF680-streptavidin (Fig. 1B (3), amplified detection); following the first round of tyramide-based amplification, biotin was detected with HRP-streptavidin and a second round of amplification was performed (Fig. 1B (4), super-amplified detection). The results showed a linear relationship between the concentration of pERK<sub>2</sub> and the intensity of the signal, whatever the detection method used



**Figure 1.** New detection methods for RPPA combining NIR fluorescence and tyramide amplification. (A) Determination of pERK<sub>2</sub> integrity by WB. Increasing concentrations of pERK<sub>2</sub> were immunostained with either anti-ERK<sub>2</sub> or with anti-pERK<sub>1/2</sub> antibody. HEK293 cell lysate (L) served as control. (B) pERK<sub>2</sub> detection by RPPA. A wide range of recombinant pERK<sub>2</sub> dilutions was printed in six replicates. The spotting pattern and concentrations of pERK<sub>2</sub> are shown in (1). Three detection methods were compared side-by-side: direct detection (2); amplified detection (3); and super-amplified detection (4). A schematic representation of each method is presented. (C) Images of three slides were simultaneously acquired using the same scanner settings so that signal intensities can be compared. We plotted the log [spot fluorescence intensity (F) at 700 nm – local background intensity (B) at 700 nm] against the log [pERK<sub>2</sub> concentration (M)]. Linear regressions were calculated for each detection method using Graphpad PRISM.

( $r^2 = 0.77$ – $0.97$ , Table 1, Fig. 1C and Supporting Information Table S1 for the raw data). For the three detection methods, the theoretical LOD (see Section 2) reached a sub-attomole range (Table 1). However, striking differences were observed with respect to the LOQ. The amplified and super-amplified detections were clearly more sensitive with a lower LOQ (0.27 and 0.21 attomole, respectively) than the direct detection displaying a lower LOQ at 6.84 attomoles (Table 1). The upper LOQ followed an opposite trend with the direct and amplified detections being linear up to 273 attomoles against 13.8 attomoles for the super-amplified detection (Table 1). The signal intensity obtained with the amplified detection system was proportional to the concentration of pERK<sub>2</sub>, whatever the concentration used (slope of  $1.069 \pm 0.016$ ) with a linear range of detection encompassing three logs (Table 1, Fig. 1C). On the contrary, the

**Table 1.** LOD and LOQ of the three detection methods

	Direct	Amplified	Super-amplified
$r^2$	0.9640	0.9717	0.7792
LOD (amol)	0.1	0.15	0.013
Lower LOQ (amol)	6.84	0.27	0.21
Upper LOQ (amol)	273	273	13.8
Linear range of detection (log)	1.6	3	1.7

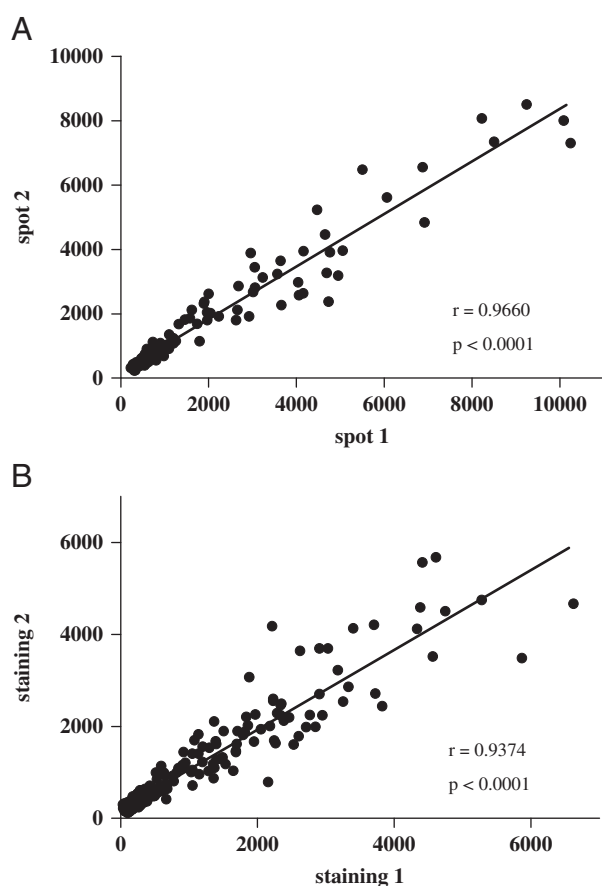
super-amplified detection displayed a lower slope ( $0.770 \pm 0.047$ ) associated with a weaker dynamic range of about 1.7 logs. The direct detection had an intermediary slope ( $0.858 \pm 0.019$ ), and displayed the weakest (1.6 logs) dynamic range.

This level of sensitivity makes NIR detection comparable with the widely used colorimetric-based detection [2, 5, 19–21]. Consistent with previous reports, we found that fluorescence detection had a dynamic range encompassing 3–4 logs [8, 13], while colorimetric-based detection was limited to 1–2 logs [2, 22]. The extended dynamic range of fluorescence-based detection potentially allows samples to be assayed in one single dilution instead of several two-fold serial dilutions, typically required to achieve reliable quantifications out of colorimetric-based detection. This feature makes fluorescence less tedious, hence suited to high-throughput analysis.

We describe here two amplification procedures, which lower the LOQ up to 32-fold when compared with the previously described direct detection method (Table 1) [10]. Importantly, this improvement in detection sensitivity appears to be essential to reliably measure variations in signalling proteins' phosphorylation levels. Indeed, most of the recorded signals were below the threshold of quantification for direct detection. In the course of this study, the amplified protocol happened to be sensitive enough to fully analyze the regulation of the selected signalling intermediates. Furthermore, the super-amplified protocol provides an easy mean to study less abundant proteins.

### 3.2 Reproducibility of RPPA

Next, we checked the intra- and inter-slide reproducibility of RPPA with the above described amplified-detection procedure. We used time-courses of FSH-induced ERK<sub>1/2</sub> phosphorylation using HEK293 FSH-R as a model. Three dilutions of samples corresponding to six independent FSH kinetics experiments were printed in duplicate. Slides were stained with anti-pERK<sub>1/2</sub> antibody combined with amplified NIR-detection. To assess intra-slide reproducibility, we analyzed 102 sample duplicates printed on the same slide. A Pearson's correlation coefficient of 0.97 (95% CI 0.95–0.98,  $p < 0.0001$ ) was obtained (Fig. 2A). Inter-slide variability was measured by comparing the pERK<sub>1/2</sub> signals



**Figure 2.** Intra- and inter-slide reproducibility of RPPA. Lysates from cells stimulated with FSH (3 nM) for increasing times were used. Three two-serial dilutions of six experiments with six stimulation points were printed in duplicate on slides and pERK<sub>1/2</sub> was detected using the amplified procedure. (A) Intra-slide reproducibility. Each replicate was compared with its counterpart so that the analysis was performed with 102 sample pairs. (B) Inter-slide reproducibility. Two slides were printed with the same samples and stained simultaneously. The analysis was performed with 204 sample pairs.

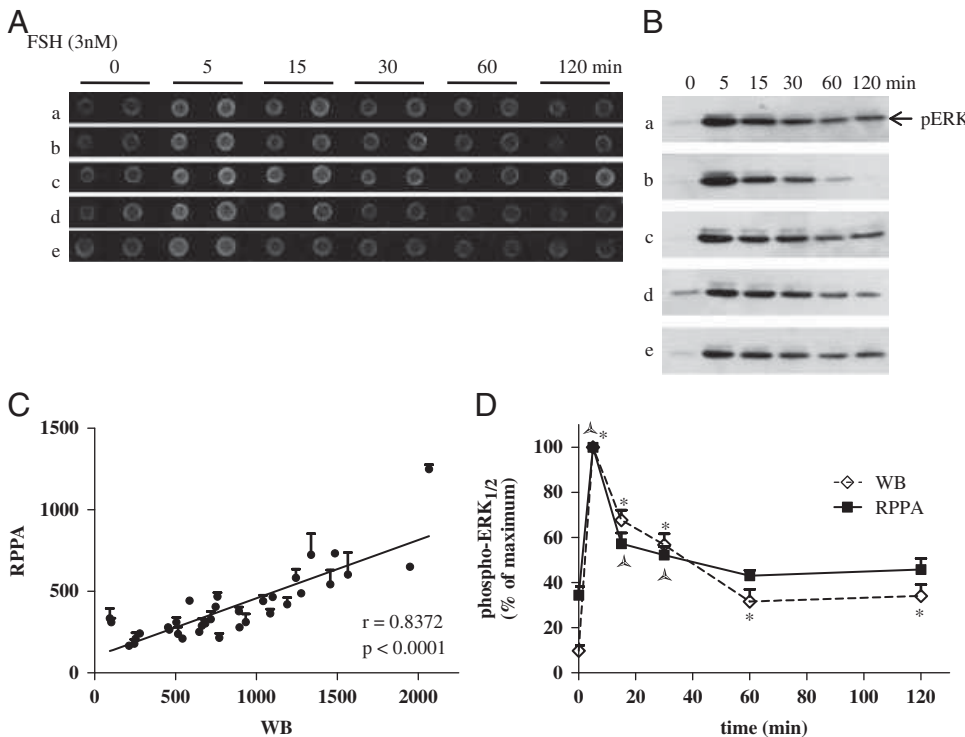
from two slides printed with the same sample collection and simultaneously stained. The analysis was performed with 204 sample pairs and gave a Pearson's correlation coefficient of 0.94 (95% CI 0.92–0.95,  $p < 0.0001$ ) (Fig. 2B).

The stability of the spotted slides was also assessed. We compared the pERK<sub>1/2</sub> signal of three identical slides printed the same day and stained immediately or after 1 and 6 months at  $-20^{\circ}\text{C}$ . Our results showed that printed slides could be frozen and stored at  $-20^{\circ}\text{C}$  for extended periods without significantly affecting measurements' accuracy (see Supporting Information Fig. S1).

### 3.3 Correlation between RPPA and WB data

A key prerequisite to the broad implementation of RPPA is to clearly establish its assay performance, robustness and reliability when compared to classical methods. For this reason, we carried out a side-by-side comparison of RPPA and WB using FSH-induced signalling pathways in HEK293 FSH-R as a model. Identical cell lysates were analyzed in parallel with both techniques (Fig. 3A and B). We found that signal intensities measured by the two methods correlated nicely (Pearson's correlation coefficient of 0.83; 95% CI 0.67–0.91,  $p < 0.0001$ ) (Fig. 3C). Furthermore, intra- and inter-assay reproducibility of WB was appreciated. Inter-assay variability was measured by comparing in each sample the pERK<sub>1/2</sub> signal from two independent WB. This analysis was performed with 34 sample pairs and gave a Pearson's correlation coefficient of 0.93 (95% CI 0.87–0.96,  $p < 0.0001$ ).

To assess the respective intra-assay variabilities of the two methods, we calculated the mean square deviations. RPPA was found more reproducible than WB with mean square deviation of 12.36 for RPPA against 18.83 for WB. Interestingly, the kinetic profile of pERK<sub>1/2</sub> measured upon FSH stimulation was similar with both techniques, the maximal effect being observed at 5 min of stimulation (Fig. 3D) and corresponding to the kinetic profile described in our previous work [23]. The only notable discrepancy was that the basal pERK<sub>1/2</sub> was significantly lower in WB analysis than in RPPA analysis (9.7% versus 34.45% for WB and RPPA, respectively), likely resulting from signal approaching the limit of WB detection in basal conditions. However, Calvo *et al.* had previously described that, in RPPA, the difference of background depends on the antibody, on the target protein and on the abundance of target proteins [24]. The absence of protein separation for RPPA also likely plays a role since non-specific signals accumulate and cannot be discriminated from specific detection. This observation, reminiscent of previous reports, seems to be frequently encountered when using RPPA [24, 25]. Based on the standard curve of pERK<sub>2</sub> spotted on the same slide, we calculated the proportion of total pERK<sub>2</sub> in serum-starved cells. Two slides spotted with the same sample collection and with pERK<sub>2</sub> concentration range were immunostained with anti-pERK<sub>1/2</sub> antibody for the first one and, with total anti-ERK<sub>2</sub> antibody for the second one. Our results showed that 7.6% of the total ERK<sub>2</sub> pool (0.96 attomole *per* equivalent cell) were phosphorylated in serum-starved HEK293 cells. In addition, 29.4% of the total ERK<sub>2</sub> pool (2.53 atto-



**Figure 3.** Correlation between pERK<sub>1/2</sub> signals measured by RPPA and WB. Cells were stimulated with FSH (3nM) for increasing times. All samples were analyzed by the amplified procedure of RPPA, and by WB. Representative RPPA (A) and WB (B) immunostained with anti-pERK<sub>1/2</sub> antibody. (C) Data obtained by the two methods were plotted and Pearson's correlation was calculated. (D) Kinetic profiles of pERK<sub>1/2</sub> upon FSH stimulation. Signals were expressed as percentage of maximal phosphorylation obtained at 5 min. Each data point represents mean  $\pm$  SEM of five independent experiments. \*(WB),  $\blacktriangle$  (RPPA):  $p < 0.05$ , as compared with non-stimulated controls.



moles *per* equivalent cell) were phosphorylated after 5 min of FSH stimulation.

### 3.4 On-slide protein quantification

At that stage, we hypothesized that part of the variability might be due to the preparation of samples and to spotting heterogeneity. To further improve the reliability of our procedure, we developed an on-slide quantification method for total proteins. Previous work described that total protein amount on membrane can be determined by direct biotinylation of protein and detected by streptavidin-coupled fluorophore [26]. However, this method presents at least two limitations: (i) it requires two different fluorophores emitting at two different wavelengths; and (ii) the biotinylation might affect the affinity between the protein of interest and its antibody. Instead, in the present study, we have stained proteins with FCF, which quantitatively interacts with proteins, has a maximum absorption at 625 nm and is

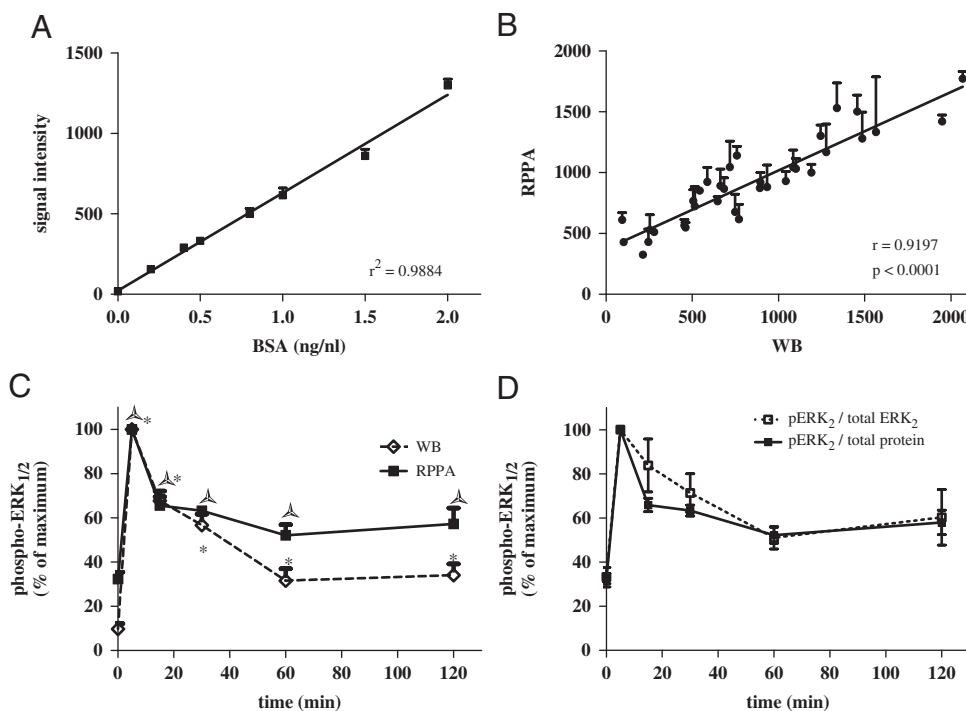
detectable in NIR (700 nm) [10, 27, 28]. To precisely quantify the total protein amount, a standard curve of BSA was printed on slides. A linear relationship between BSA concentrations and signal intensities was observed ( $r^2 = 0.98$ , Fig. 4A). On-slide FCF-staining allowed us to detect less than 1 ng of total protein from cell extract, corresponding to two equivalent cells printed *per* spot (Table 2). We then calculated the protein concentrations in each spot and used them to normalize pERK<sub>1/2</sub> signal intensities as measured by RPPA. By doing this, we improved the Pearson's correlation coefficient from 0.83 (Fig. 3C) to 0.92 (95% CI 0.844–0.9595,  $p < 0.0001$ ) (Fig. 4B). After the normalization, we re-examined the kinetic profile of pERK<sub>1/2</sub> and noticed that the sustained response, which is not visible on Fig. 3D, now appears as significant over basal level (Fig. 4C) [23].

To further validate this FCF-based on-slide protein quantification method, we normalized the amounts of pERK<sub>2</sub> with those of total ERK<sub>2</sub>. When we compared the kinetic profile of pERK<sub>2</sub> obtained with both normalization

**Table 2.** Estimate of protein amounts and equivalent cell *per* array pin touch

	Dilution 1:2	Dilution 1:4	Dilution 1:8
Protein (ng/nL)	0.8612 ± 0.022	0.44 ± 0.014	0.2716 ± 0.009
Protein (ng)/array pin touch (3–13 nL)	2.58–11.19	1.32–5.72	0.81–3.51
Equivalent cells/array pin touch (3–13 nL)	4–16	2–8	1–4

According to the manufacturer's indications, 3–13 nL of sample were spotted on Fast-slide *per* array pin touch. Protein amounts were expressed as the mean ± SEM of six experiments in duplicate.



**Figure 4.** Normalization of signal intensities according to on-slide protein amount. (A) BSA was printed in parallel with the samples on the same slide. Total protein quantification of each spot was performed with FCF staining. (B) Data from Fig. 3C were normalized with respect to normal protein amounts. (C) Kinetic profile of pERK<sub>1/2</sub> upon FSH stimulation after normalization with the FCF-based method. (D) Kinetic profile of pERK<sub>1/2</sub> upon FSH stimulation after normalization with total ERK<sub>2</sub> versus FCF-based method. The results were expressed as percentage of maximal stimulation obtained at 5 min. Each data point represents mean ± SEM of five independent experiments. \*(WB),  $\blacktriangle$  (RPPA):  $p < 0.05$ , as compared with non-stimulated controls.

procedures, no significant difference was observed (Fig. 4D; two-way ANOVA with Bonferroni's post-test,  $p > 0.05$  for each point of stimulation).

#### 4 Concluding remarks

The proof of principle that protein array technology could lead to new applications for immunoassays, potentially expanding the scope of protein expression and protein interaction analysis, has been described more than 15 years ago [29]. However, RPPA is still in its early life since many questions regarding validity, sensitivity and standardization remain to be addressed [2]. In this study, we showed that RPPA combined with tyramide-amplified NIR detection was adapted to capture subtle variations of signalling proteins' phosphorylation in cell lysates. Importantly, the fluorescent detection method reported here warrants an exquisite dynamic range. Therefore, much less sample dilutions need to be processed and analyzed, without requiring any heavy equipment or costly reagent. Moreover, the range of sensitivity reported here is comparable with colorimetric detection methods. These procedures are therefore widely applicable.

Together, these features are very attractive in the perspective of high-throughput analyses of signalling pathways and could provide a solution to one important bottleneck encountered by systems biology of cellular signalling.

*The authors thank members of the BIOS team for their advices and support throughout this project. L. D. was an INRA-MRT post-doctoral fellow. D. H. was funded by a fellowship from the INRA (ASC). A. M. was funded by a fellowship from the Region Centre and INRA and G. D. by a fellowship from the French MRT. This work was supported by the INRA AIP AgroBI, AE INRIA-INRA REGATE and ANR-07-RIB-012-01 "PhosphoProteomeArray".*

*The authors have declared no conflict of interest.*

#### 5 References

- [1] Mann, M., Jensen, O. N., Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 2003, 21, 255–261.
- [2] Paweletz, C. P., Charboneau, L., Bichsel, V. E., Simone, N. L. *et al.*, Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* 2001, 20, 1981–1989.
- [3] Spurrier, B., Ramalingam, S., Nishizuka, S., Reverse-phase protein lysate microarrays for cell signaling analysis. *Nat. Protoc.* 2008, 3, 1796–1808.
- [4] Wulfkühle, J. D., Aquino, J. A., Calvert, V. S., Fishman, D. A. *et al.*, Signal pathway profiling of ovarian cancer from human tissue specimens using reverse-phase protein microarrays. *Proteomics* 2003, 3, 2085–2090.
- [5] Tibes, R., Qiu, Y., Lu, Y., Hennessy, B. *et al.*, Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. *Mol. Cancer Ther.* 2006, 5, 2512–2521.
- [6] Jiang, R., Mircean, C., Shmulevich, I., Cogdell, D. *et al.*, Pathway alterations during glioma progression revealed by reverse phase protein lysate arrays. *Proteomics* 2006, 6, 2964–2971.
- [7] Nishizuka, S., Charboneau, L., Young, L., Major, S. *et al.*, Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays. *Proc. Natl. Acad. Sci. USA* 2003, 100, 14229–14234.
- [8] Chan, S. M., Ermann, J., Su, L., Fathman, C. G., Utz, P. J., Protein microarrays for multiplex analysis of signal transduction pathways. *Nat. Med.* 2004, 10, 1390–1396.
- [9] Winters, M. E., Mehta, A. I., Petricoin, E. F., 3rd, Kohn, E. C., Liotta, L. A., Supra-additive growth inhibition by a celecoxib analogue and carboxyamido-triazole is primarily mediated through apoptosis. *Cancer Res.* 2005, 65, 3853–3860.
- [10] Loebke, C., Sueltmann, H., Schmidt, C., Henjes, F. *et al.*, Infrared-based protein detection arrays for quantitative proteomics. *Proteomics* 2007, 7, 558–564.
- [11] Majidi, M., Al-Wadei, H. A., Takahashi, T., Schuller, H. M., Nongenomic beta estrogen receptors enhance beta1 adrenergic signaling induced by the nicotine-derived carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in human small airway epithelial cells. *Cancer Res.* 2007, 67, 6863–6871.
- [12] Espina, V., Woodhouse, E. C., Wulfkühle, J., Asmussen, H. D. *et al.*, Protein microarray detection strategies: focus on direct detection technologies. *J. Immunol. Methods* 2004, 290, 121–133.
- [13] Schutz-Geschwender, A., Zhang, Y., Holt, T., McDermitt, D., Olive, M., Quantitative, Two-Color Western blot detection with infrared fluorescence. Application Note. *LI-COR Biosciences.* 2004.
- [14] Chung, J. Y., Lee, S. J., Kris, Y., Braunschweig, T. *et al.*, A well-based reverse-phase protein array applicable to extracts from formalin-fixed paraffin-embedded tissue. *Proteomics Clin. Appl.* 2008, 2, 1539–1547.
- [15] Pawlak, M., Schick, E., Bopp, M. A., Schneider, M. J. *et al.*, Zeptosens' protein microarrays: a novel high performance microarray platform for low abundance protein analysis. *Proteomics* 2002, 2, 383–393.
- [16] Snapyan, M., Lecocq, M., Guevel, L., Arnaud, M. C. *et al.*, Dissecting DNA-protein and protein-protein interactions involved in bacterial transcriptional regulation by a sensitive protein array method combining a near-infrared fluorescence detection. *Proteomics* 2003, 3, 647–657.
- [17] Sakanyan, V., Yeretssian, G., Near infrared fluorescence detection of antigen-antibody interactions on microarrays. in Sem, D. (Ed.), *Spectral Techniques in Proteomics.* CRC Press, Boca Raton, London, New York 2007, pp. 185–205.

- [18] Ghochikyan, A., Karaivanova, I. M., Lecocq, M., Vusio, P. *et al.*, Arginine operator binding by heterologous and chimeric ArgR repressors from *Escherichia coli* and *Bacillus stearothermophilus*. *J. Bacteriol.* 2002, *184*, 6602–6614.
- [19] Calvert, V. S., Tang, Y., Boveia, V., Wulfschlegel, J. *et al.*, Development of multiplexed protein profiling and detection using near infrared detection of reverse-phase protein microarrays. *Clin. Proteom. J.* 2004, *1*, 81–90.
- [20] Wulfschlegel, J. D., Speer, R., Pierobon, M., Laird, J. *et al.*, Multiplexed cell signaling analysis of human breast cancer applications for personalized therapy. *J. Proteome. Res.* 2008, *7*, 1508–1517.
- [21] Sheehan, K. M., Calvert, V. S., Kay, E. W., Lu, Y. *et al.*, Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma. *Mol. Cell. Proteomics* 2005, *4*, 346–355.
- [22] Rimm, D. L., What brown cannot do for you. *Nat. Biotechnol.* 2006, *24*, 914–916.
- [23] Kara, E., Crepieux, P., Gauthier, C., Martinat, N. *et al.*, A phosphorylation cluster of five serine and threonine residues in the C-terminus of the follicle-stimulating hormone receptor is important for desensitization but not for beta-arrestin-mediated ERK activation. *Mol. Endocrinol.* 2006, *20*, 3014–3026.
- [24] Calvo, K. R., Dabir, B., Kovach, A., Devor, C. *et al.*, IL-4 protein expression and basal activation of Erk in vivo in follicular lymphoma. *Blood* 2008, *112*, 3818–3826.
- [25] Liotta, L. A., Espina, V., Mehta, A. I., Calvert, V. *et al.*, Protein microarrays: meeting analytical challenges for clinical applications. *Cancer Cell* 2003, *3*, 317–325.
- [26] Chung, J. Y., Braunschweig, T., Baibakov, G., Galperin, M. *et al.*, Transfer and multiplex immunoblotting of a paraffin embedded tissue. *Proteomics* 2006, *6*, 767–774.
- [27] Korf, U., Löbke, C., Haller, F., Sülthmann, H., Poutska, A., Infrared-based protein detection arrays for quantitative proteomics. *Expert Opin. Drug Discov.* 2008, *3*, 273–283.
- [28] Luo, S., Wehr, N. B., Levine, R. L., Quantitation of protein on gels and blots by infrared fluorescence of Coomassie blue and Fast Green. *Anal. Biochem.* 2006, *350*, 233–238.
- [29] Ekins, R. P., Chu, F. W., Multianalyte microspot immunoassay-microanalytical “compact disk” of the future. *Clin. Chem.* 1991, *37*, 1955–1967.

## Annexe B

# Modèles avec Biocham

### B.1 Petit mémo de la syntaxe BIOCHAM

Récepteur inactif R

Récepteur complexé à l'hormone R-H

Complexe phosphorylé sur le site P1 :  $R^{\sim}\{P1\}\text{-H}$

Récepteur à la membrane R::mb

Complexation

$R::mb + H::mb \Rightarrow R\text{-H}::mb.$

Décomplexation

$R\text{-H}::mb \Rightarrow R::mb + H::mb.$

Phosphorylation catalysée

$R\text{-H-Ga-Gb-Gg-GDP}::mb = [\text{GRK2/3}::mb] \Rightarrow R^{\sim}\{P1\}\text{-H}::mb + \text{Ga-Gb-Gg-GDP}::mb.$

Synthèse

$\_ \Rightarrow A.$

Dégradation

$R^{\sim}\{P1\}\text{-H-bArr}_2::mb \Rightarrow \_.$

Transport

$\text{Ca}::ret \Rightarrow \text{Ca}::cyt.$

## B.2 Modèle détaillé de réseau de signalisation Angiotensine

```

%-----%
% Activation récepteur - protéine G                                     %
%-----%

R::mb + H::mb <=> R-H::mb.
R-H::mb + Ga-Gb-Gg-GDP::mb <=> R-H-Ga-Gb-Gg-GDP::mb.
R-H-Ga-Gb-Gg-GDP::mb <=> R-H::mb + Ga-GTP::mb + Gb-Gg::mb.

R~{P1}-H::mb + Ga-Gb-Gg-GDP::mb <=> R~{P1}-H-Ga-Gb-Gg-GDP::mb.
R~{P1}-H-Ga-Gb-Gg-GDP::mb <=> R~{P1}-H::mb + Ga-GTP::mb + Gb-Gg::mb.

Ga-Gb-Gg-GDP::mb <=> Ga-GTP::mb + Gb-Gg::mb. %faiblement

Ga-GTP::mb => Ga-GTP::cyt.
Ga-GTP::cyt => Ga-GDP::cyt.
Ga-GDP::cyt => Ga-GDP::mb.

Ga-GDP::mb + Gb-Gg::mb => Ga-Gb-Gg-GDP::mb.

%-----%
% VOIE G : PLC IP3 DAG PIP2 Ca PKC => ERK::ker                       %
%-----%
PLCb_inactive::cyt =[Ga-GTP::cyt]=> PLCb_active::cyt.
PLCb_active::cyt => PLCb_inactive::cyt.
PIP2::cyt =[PLCb_active::cyt]=> DAGa::cyt + IP3::cyt.
DAGa::cyt + IP3::cyt => PIP2::cyt.

% Cascade MAPK %
Ca::ret =[IP3::cyt]=> Ca::cyt.
PKC::cyt + DAGa::cyt =[Ca::cyt]=> PKCa::cyt + DAGa::cyt.
Ca::cyt => Ca::ret.
PKCa::cyt => PKC::cyt.

Raf::cyt =[PKCa::cyt] => Raf~{P6}::cyt.
Mek::cyt =[Raf~{P6}::cyt]=> Mek~{P7}::cyt.
ERK::cyt =[Mek~{P7}::cyt]=> ERK~{P}::cyt.
% Cascade MAPK % FIN

% GpERK %
ERK~{P}::cyt => ERK~{P}::ker.
ERK~{P}::ker => ERK::ker.
ERK::ker => ERK::cyt.
% GpERK % FIN

%-----%
% Recrutement des bArr1/2 à la membrane                               %
%-----%
bArr1~{P3}::cyt =[R-H::mb]=> bArr1::mb.
bArr2~{P4}::cyt =[R-H::mb]=> bArr2::mb.

```

## B.2. MODÈLE DÉTAILLÉ DE RÉSEAU DE SIGNALISATION ANGIOTENSINE

```
%-----%
% VOIE GRK2/3 %
%-----%
R-H::mb =[GRK23::mb]=> R~{P1}-H::mb.
R~{P1}-H::mb + bArr1::mb <=> R~{P1}-H-bArr1::mb.
R~{P1}-H::mb + bArr2::mb <=> R~{P1}-H-bArr2::mb.

R~{P1}-H-bArr1::mb <=> R~{P1}-H-bArr1~{P3}::cyt.
R~{P1}-H-bArr2::mb <=> R~{P1}-H-bArr2~{P4}::cyt.

% Recyclage des récepteurs RP1
R~{P1}-H-bArr1~{P3}::cyt => R-H::cyt + bArr1~{P3}::cyt.
R~{P1}-H-bArr2~{P4}::cyt => R-H::cyt + bArr2~{P4}::cyt.

R-H::cyt => R-H::mb.

% ou dégradation des récepteurs RP1
R~{P1}-H-bArr1~{P3}::cyt => _ .
R~{P1}-H-bArr2~{P4}::cyt => _ .

%-----%
% VOIE GRK5/6 %
%-----%
R-H::mb =[GRK56::mb]=> R~{P2}-H::mb.

% RP2 et barr2 signalent %
R~{P2}-H::mb + bArr2::mb <=> R~{P2}-H-bArr2::mb.
R~{P2}-H-bArr2::mb <=> R~{P2}-H-bArr2~{P4}::cyt.
src::cyt + Raf::cyt + Mek::cyt =[R~{P2}-H-bArr2~{P4}::cyt]=> src-Raf-
Mek::cyt.
src-Raf-Mek::cyt => src::cyt + Raf::cyt + Mek::cyt

% R-H et barr2 signalent faiblement %
R-H::mb + bArr2::mb <=> R-H-bArr2::mb.
R-H-bArr2::mb <=> R-H-bArr2~{P4}::cyt.
src::cyt + Raf::cyt + Mek::cyt =[R-H-bArr2~{P4}::cyt]=> src-Raf-
Mek::cyt.
% R-H et barr2 signalent faiblement % FIN

%R-Raf-Mek-Erk %
%Phosphorylation en chaine des voisins (de proche en proche) %
src-Raf-Mek::cyt <=> src~{P5}-Raf-Mek::cyt.
src~{P5}-Raf-Mek::cyt <=> src~{P5}-Raf~{P6}-Mek::cyt.
src~{P5}-Raf~{P6}-Mek::cyt <=> src~{P5}-Raf~{P6}-Mek~{P7}::cyt.
%Phosphorylation en chaine des voisins (de proche en proche) % FIN
%R-Raf-Mek-Erk % FIN

ERK::cyt =[src~{P5}-Raf~{P6}-Mek~{P7}::cyt]=> ERK~{P}::cyt.
% barr 2 signal % FIN
```



# Bibliographie

- Ahn, S., Shenoy, S. K., Wei, H., and Lefkowitz, R. J. (2004a). Differential kinetic and spatial patterns of beta-arrestin and g protein-mediated erk activation by the angiotensin ii receptor. *J Biol Chem*, 279(34) :35518–35525.
- Ahn, S., Wei, H., Garrison, T. R., and Lefkowitz, R. J. (2004b). Reciprocal regulation of angiotensin receptor-activated extracellular signal-regulated kinases by beta-arrestins 1 and 2. *J Biol Chem*, 279(9) :7807–7811.
- Aittomäki, K., Lucena, J. L., Pakarinen, P., Sistonen, P., Tapanainen, J., Gromoll, J., Kaskikari, R., Sankila, E. M., Lehväslaiho, H., Engel, A. R., Nieschlag, E., Huh-taniemi, I., and de la Chapelle, A. (1995). Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. *Cell*, 82(6) :959–968.
- Ajay, S. M. and Bhalla, U. S. (2004). A role for erkii in synaptic pattern selectivity on the time-scale of minutes. *Eur J Neurosci*, 20(10) :2671–2680.
- Alam, H., Maizels, E. T., Park, Y., Ghaey, S., Feiger, Z. J., Chandel, N. S., and Hunzicker-Dunn, M. (2004). Follicle-stimulating hormone activation of hypoxia-inducible factor-1 by the phosphatidylinositol 3-kinase/akt/ras homolog enriched in brain (rheb)/mammalian target of rapamycin (mTOR) pathway is necessary for induction of select protein markers of follicular differentiation. *J Biol Chem*, 279(19) :19431–19440.
- Arey, B. J., Yanofsky, S. D., Pérez, M. C., Holmes, C. P., Wrobel, J., Gopalsamy, A., Stevis, P. E., López, F. J., and Winneker, R. C. (2008). Differing pharmacological activities of thiazolidinone analogs at the fish receptor. *Biochem Biophys Res Commun*, 368(3) :723–728.
- Aubin, C. and Menninger, D. J. (2009). Evaluation des politiques de prévention des grossesses non désirées et de prise en charge des interruptions volontaires de grossesse suite à la loi du 4 juillet 2001. Technical report, Inspection générale des affaires sociales.
- Audet, N., Paquin-Gobeil, M., Landry-Paquet, O., Schiller, P. W., and Piñeyro, G. (2005). Internalization and src activity regulate the time course of erk activation by delta opioid receptor ligands. *J Biol Chem*, 280(9) :7808–7816.
- Auger, A. (2004). *Contribution théoriques et numériques à l'optimisation continue par algorithmes évolutionnaires*. PhD thesis, Université de Paris 6.
- Azzi, M., Charest, P. G., Angers, S., Rousseau, G., Kohout, T., Bouvier, M., and Piñeyro, G. (2003). Beta-arrestin-mediated activation of mapk by inverse agonists reveals



- distinct active conformations for g protein-coupled receptors. *Proc Natl Acad Sci U S A*, 100(20) :11406–11411.
- Baker, J. G., Hall, I. P., and Hill, S. J. (2003). Agonist and inverse agonist actions of beta-blockers at the human beta 2-adrenoceptor provide evidence for agonist-directed signaling. *Mol Pharmacol*, 64(6) :1357–1369.
- Barthet, G. (2007). *Régulation des voies de signalisation dépendante et indépendante des protéines G activées par le récepteur 5-HT<sub>4</sub>*. PhD thesis, Université Montpellier I.
- Bhalla, U. S. and Iyengar, R. (1999). Emergent properties of networks of biological signaling pathways. *Science*, 283(5400) :381–387.
- Birtwistle, M. R., Hatakeyama, M., Yumoto, N., Ogunnaike, B. A., Hoek, J. B., and Kholodenko, B. N. (2007). Ligand-dependent responses of the erbb signaling network : experimental and modeling analyses. *Mol Syst Biol*, 3 :144.
- Bockaert, J., Perroy, J., Bécamel, C., Marin, P., and Fagni, L. (2010). Gpcr interacting proteins (gips) in the nervous system : Roles in physiology and pathologies. *Annu Rev Pharmacol Toxicol*, 50 :89–109.
- Bockaert, J. and Pin, J. P. (1999). Molecular tinkering of g protein-coupled receptors : an evolutionary success. *EMBO J*, 18(7) :1723–1729.
- Bockaert, J., Roussignol, G., Bécamel, C., Gavarini, S., Joubert, L., Dumuis, A., Fagni, L., and Marin, P. (2004). Gpcr-interacting proteins (gips) : nature and functions. *Biochem Soc Trans*, 32(Pt 5) :851–855.
- Bogerd, J., Granneman, J. C. M., Schulz, R. W., and Vischer, H. F. (2005). Fish fish receptors bind lh : how to make the human fish receptor to be more fishy ? *Gen Comp Endocrinol*, 142(1-2) :34–43.
- Bohn, L. M., Lefkowitz, R. J., Gainetdinov, R. R., Peppel, K., Caron, M. G., and Lin, F. T. (1999). Enhanced morphine analgesia in mice lacking beta-arrestin 2. *Science*, 286(5449) :2495–2498.
- Borisov, N. M., Chistopolsky, A. S., Faeder, J. R., and Kholodenko, B. N. (2008). Domain-oriented reduction of rule-based network models. *IET Syst Biol*, 2(5) :342–351.
- Bridge, L. J., King, J. R., Hill, S. J., and Owen, M. R. (2010). Mathematical modelling of signalling in a two-ligand g-protein coupled receptor system : agonist-antagonist competition. *Math Biosci*, 223(2) :115–132.
- Brown, D. I. and Griendling, K. K. (2009). Nox proteins in signal transduction. *Free Radic Biol Med*, 47(9) :1239–1253.
- Broyden, C. G. (1970). The convergence of a class of double-rank minimization algorithms. *Journal of the Institute of Mathematics and Its Applications*, 6 :7–90.
- Calzone, L., Fages, F., and Soliman, S. (2006). Biocham : an environment for modeling biological systems and formalizing experimental knowledge. *Bioinformatics*, 22(14) :1805–1807.
- Cappé, O., Guillin, A., Marin, J.-M., Robert, C. P., and Roberty, C. P. (2004). Population monte carlo. *Journal of Computational and Graphical Statistics*, 13 :907–929.
- Carlone, D. L. and Richards, J. S. (1997). Evidence that functional interactions of creb and sf-1 mediate hormone regulated expression of the aromatase gene in granulosa cells and constitutive expression in r2c cells. *J Steroid Biochem Mol Biol*, 61(3-6) :223–231.

- Chabrier-Rivier, N., Chiaverini, M., Danos, V., Fages, F., and Schächter, V. (2003). Modeling and querying biomolecular interaction networks. *Theoretical Computer Science*, 325 :25–44.
- Chen, K. C., Calzone, L., Csikasz-Nagy, A., Cross, F. R., Novak, B., and Tyson, J. J. (2004). Integrative analysis of cell cycle control in budding yeast. *Mol Biol Cell*, 15(8) :3841–3862.
- Cheng, X., Ji, Z., Tsalkova, T., and Mei, F. (2008). Epac and pka : a tale of two intracellular camp receptors. *Acta Biochim Biophys Sin (Shanghai)*, 40(7) :651–662.
- Choi, J.-H., Chen, C.-L., Poon, S. L., Wang, H.-S., and Leung, P. C. K. (2009). Gonadotropin-stimulated epidermal growth factor receptor expression in human ovarian surface epithelial cells : involvement of cyclic amp-dependent exchange protein activated by camp pathway. *Endocr Relat Cancer*, 16(1) :179–188.
- Clerc, M. and Kennedy, J. (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*, 6 :58–73.
- Clément, F., Monniaux, D., Stark, J., Hardy, K., Thalabard, J. C., Franks, S., and Claude, D. (2001). Mathematical model of fsh-induced camp production in ovarian follicles. *Am J Physiol Endocrinol Metab*, 281(1) :E35–E53.
- Conner, D. A., Mathier, M. A., Mortensen, R. M., Christe, M., Vatner, S. F., Seidman, C. E., and Seidman, J. G. (1997). beta-arrestin1 knockout mice appear normal but demonstrate altered cardiac responses to beta-adrenergic stimulation. *Circ Res*, 81(6) :1021–1026.
- Conti, M., Toscano, M. V., Petrelli, L., Geremia, R., and Stefanini, M. (1982). Regulation of follicle-stimulating hormone and dibutyryl adenosine 3',5'-monophosphate of a phosphodiesterase isoenzyme of the sertoli cell. *Endocrinology*, 110(4) :1189–1196.
- Cottom, J., Salvador, L. M., Maizels, E. T., Reierstad, S., Park, Y., Carr, D. W., Davare, M. A., Hell, J. W., Palmer, S. S., Dent, P., Kawakatsu, H., Ogata, M., and Hunzicker-Dunn, M. (2003). Follicle-stimulating hormone activates extracellular signal-regulated kinase but not extracellular signal-regulated kinase through a 100-kda phosphotyrosine phosphatase. *J Biol Chem*, 278(9) :7167–7179.
- Crépieux, P., Marion, S., Martinat, N., Fafeur, V., Vern, Y. L., Kerboeuf, D., Guillou, F., and Reiter, E. (2001). The erk-dependent signalling is stage-specifically modulated by fsh, during primary sertoli cell maturation. *Oncogene*, 20(34) :4696–4709.
- Crépieux, P., Martinat, N., Marion, S., Guillou, F., and Reiter, E. (2002). Cellular adhesion of primary sertoli cells affects responsiveness of the extracellular signal-regulated kinases 1 and 2 to follicle-stimulating hormone but not to epidermal growth factor. *Arch Biochem Biophys*, 399(2) :245–250.
- Csercsik, D., Farkas, I., Szederkényi, G., Hrabovszky, E., Liposits, Z., and Hangos, K. M. (2010). Hodgkin-huxley type modelling and parameter estimation of gnrh neurons. *Biosystems*, 100(3) :198–207.
- Csercsik, D., Hangos, K. M., and Nagy, G. M. (2008). A simple reaction kinetic model of rapid (g protein dependent) and slow (beta-arrestin dependent) transmission. *J Theor Biol*, 255(1) :119–128.

- Dattatreyaumurty, B., Figgs, L. W., and Reichert, L. E. (1987). Physical and functional association of follitropin receptors with cholera toxin-sensitive guanine nucleotide-binding protein. *J Biol Chem*, 262(24) :11737–11745.
- De Lean, A., Stadel, J. M., and Lefkowitz, R. J. (1980). A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled beta-adrenergic receptor. *J Biol Chem*, 255(15) :7108–7117.
- DeWire, S. M., Ahn, S., Lefkowitz, R. J., and Shenoy, S. K. (2007). Beta-arrestins and cell signaling. *Annu Rev Physiol*, 69 :483–510.
- Douc, R., Guillin, A., m. Marin, J., and Robert, C. P. (2005). Minimum variance importance sampling via population monte carlo. Technical report, Ecole Polytechnique - Université Paris Dauphine.
- Draws, J. (2000). Drug discovery : a historical perspective. *Science*, 287(5460) :1960–1964.
- Dupont, G., Combettes, L., and Leybaert, L. (2007). Calcium dynamics : spatio-temporal organization from the subcellular to the organ level. *Int Rev Cytol*, 261 :193–245.
- El-Benna, J., Dang, P. M.-C., Gougerot-Pocidalo, M. A., Marie, J. C., and Braut-Boucher, F. (2009). p47phox, the phagocyte nadph oxidase/nox2 organizer : structure, phosphorylation and implication in diseases. *Exp Mol Med*, 41(4) :217–225.
- Escamilla-Hernandez, R., Little-Ihrig, L., Orwig, K. E., Yue, J., Chandran, U., and Zeleznik, A. J. (2008). Constitutively active protein kinase a qualitatively mimics the effects of follicle-stimulating hormone on granulosa cell differentiation. *Mol Endocrinol*, 22(8) :1842–1852.
- Eskola, V., Ryhänen, P., Savisalo, M., Rannikko, A., Kananen, K., Sprengel, R., and Huhtaniemi, I. (1998). Stable transfection of the rat follicle-stimulating hormone receptor complementary dna into an immortalized murine sertoli cell line. *Mol Cell Endocrinol*, 139(1-2) :143–152.
- Fages, F. and Soliman, S. (2008a). Formal cell biology in biocham. In *SFM'08 : Proceedings of the Formal methods for the design of computer, communication, and software systems 8th international conference on Formal methods for computational systems biology*, pages 54–80, Berlin, Heidelberg. Springer-Verlag.
- Fages, F. and Soliman, S. (2008b). From reaction models to influence graphs and back : A theorem. In *FMSB*, pages 90–102.
- Fan, Q. R. and Hendrickson, W. A. (2005). Structure of human follicle-stimulating hormone in complex with its receptor. *Nature*, 433(7023) :269–277.
- Fisher, A., Heldman, E., Gurwitz, D., Haring, R., Barak, D., Meshulam, H., Marciano, D., Brandeis, R., Pittel, Z., and Segal, M. (1993). Selective signaling via unique m1 muscarinic agonists. *Ann N Y Acad Sci*, 695 :300–303.
- Fletcher, R. (1970). A new approach to variable metric algorithms. *Computer journal*, 13 :327–322.
- Frank, G. D., Saito, S., Motley, E. D., Sasaki, T., Ohba, M., Kuroki, T., Inagami, T., and Eguchi, S. (2002). Requirement of ca(2+) and pkcdelta for janus kinase 2 activation by angiotensin ii : involvement of pyk2. *Mol Endocrinol*, 16(2) :367–377.

- Fredriksson, R., Lagerström, M. C., Lundin, L.-G., and Schiöth, H. B. (2003). The g-protein-coupled receptors in the human genome form five main families. phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol*, 63(6) :1256–1272.
- Galandrin, S. and Bouvier, M. (2006). Distinct signaling profiles of beta1 and beta2 adrenergic receptor ligands toward adenylyl cyclase and mitogen-activated protein kinase reveals the pluridimensionality of efficacy. *Mol Pharmacol*, 70(5) :1575–1584.
- Galandrin, S., Oligny-Longpré, G., and Bouvier, M. (2007). The evasive nature of drug efficacy : implications for drug discovery. *Trends Pharmacol Sci*, 28(8) :423–430.
- Galés, C., Durm, J. J. V., Schaak, S., Pontier, S., Percherancier, Y., Audet, M., Paris, H., and Bouvier, M. (2006). Probing the activation-promoted structural rearrangements in preassembled receptor-g protein complexes. *Nat Struct Mol Biol*, 13(9) :778–786.
- Gay, S., Soliman, S., and Fages, F. (2010). A graphical method for reducing and relating models in systems biology. *Bioinformatics*, 26(18) :i575–i581.
- Gbahou, F., Rouleau, A., Morisset, S., Parmentier, R., Crochet, S., Lin, J.-S., Ligneau, X., Tardivel-Lacombe, J., Stark, H., Schunack, W., Ganellin, C. R., Schwartz, J.-C., and Arrang, J.-M. (2003). Protean agonism at histamine h3 receptors in vitro and in vivo. *Proc Natl Acad Sci U S A*, 100(19) :11086–11091.
- Glendinning, P. (1994). *Stability, instability and chaos*. Cambridge University Press.
- Goldberg, D. E. (1989). *Genetic algorithms in search optimization and machine learning*. Addison Wesley.
- Goldfarb, D. (1970). A family of variable metric updates derived by variational means. *Mathematics of Computation*, 24 :23–26.
- Gomez, C., Delebecque, F., Bunks, C., Chancelier, J.-P., Steer, S., and Nikoukhah, R., editors (1998). *Engineering and Scientific Computing with Scilab with Cdrom*. Birkhauser Boston.
- Gonzalez-Robayna, I. J., Falender, A. E., Ochsner, S., Firestone, G. L., and Richards, J. S. (2000). Follicle-stimulating hormone (fsh) stimulates phosphorylation and activation of protein kinase b (pkb/akt) and serum and glucocorticoid-induced kinase (sgk) : evidence for a kinase-independent signaling by fsh in granulosa cells. *Mol Endocrinol*, 14(8) :1283–1300.
- Gorban, A., Radulescu, O., and Zinovyev, A. (2010). Asymptotology of chemical reaction networks. *Chemical Engineering Science*, 65(7) :2310 – 2324. International Symposium on Mathematics in Chemical Kinetics and Engineering.
- Grady, E. F. (2007). Cell signaling. beta-arrestin, a two-fisted terminator. *Science*, 315(5812) :605–606.
- Griendling, K. K., Sorescu, D., and Ushio-Fukai, M. (2000). Nad(p)h oxidase : role in cardiovascular biology and disease. *Circ Res*, 86(5) :494–501.
- Han, M., Gurevich, V. V., Vishnivetskiy, S. A., Sigler, P. B., and Schubert, C. (2001). Crystal structure of beta-arrestin at 1.9 a : possible mechanism of receptor binding and membrane translocation. *Structure*, 9(9) :869–880.
- Hansen N, Müller SD, K. P. (2003). Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (cma-es). *Evolutionary Computation*, 11(1) :1–18.

- He, J. C., Neves, S. R., Jordan, J. D., and Iyengar, R. (2006). Role of the go/i signaling network in the regulation of neurite outgrowth. *Can J Physiol Pharmacol*, 84(7) :687–694.
- Heffner, L. J. (2003). *Reproduction Humaine*. Sciences Médicales - De Boeck Université.
- Hirsch, J. A., Schubert, C., Gurevich, V. V., and Sigler, P. B. (1999). The 2.8 Å crystal structure of visual arrestin : a model for arrestin's regulation. *Cell*, 97(2) :257–269.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, Ann Arbor, Michigan.
- Hooke, R. and Jeeves, T. A. (1961). "direct search" solution of numerical and statistical problems. *J. ACM*, 8(2) :212–22.
- Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P., and Kummer, U. (2006). Copasi—a complex pathway simulator. *Bioinformatics*, 22(24) :3067–3074.
- Horowitz, A., Menice, C. B., Laporte, R., and Morgan, K. G. (1996). Mechanisms of smooth muscle contraction. *Physiol Rev*, 76(4) :967–1003.
- Ingber, A. L. and Ingber, L. (1989). Very fast simulated re-annealing. *Mathl. Comput. Modelling*, 12(8) :967–973.
- Ingber, L. (1993). Simulated annealing : Practice versus theory. *Mathl. Comput. Modelling*, 18 :29–57.
- Jain, P. and Bhalla, U. S. (2009). Signaling logic of activity-triggered dendritic protein synthesis : an mtor gate but not a feedback switch. *PLoS Comput Biol*, 5(2) :e1000287.
- Jiang, Y., Ma, W., Wan, Y., Kozasa, T., Hattori, S., and Huang, X. Y. (1998). The g protein g alpha12 stimulates bruton's tyrosine kinase and a rasgap through a conserved ph/bm domain. *Nature*, 395(6704) :808–813.
- Jin, J., Smith, F. D., Stark, C., Wells, C. D., Fawcett, J. P., Kulkarni, S., Metalnikov, P., O'Donnell, P., Taylor, P., Taylor, L., Zougman, A., Woodgett, J. R., Langeberg, L. K., Scott, J. D., and Pawson, T. (2004). Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization. *Curr Biol*, 14(16) :1436–1450.
- Jin, L., Ying, Z., Hilgers, R. H. P., Yin, J., Zhao, X., Imig, J. D., and Webb, R. C. (2006). Increased rhoa/rho-kinase signaling mediates spontaneous tone in aorta from angiotensin ii-induced hypertensive rats. *J Pharmacol Exp Ther*, 318(1) :288–295.
- Kaper, H. G. and Kaper, T. J. (2006). Asymptotic analysis of two reduction methods for systems of chemical reactions.
- Kara, E., Crépieux, P., Gauthier, C., Martinat, N., Piketty, V., Guillou, F., and Reiter, E. (2006). A phosphorylation cluster of five serine and threonine residues in the c-terminus of the follicle-stimulating hormone receptor is important for desensitization but not for beta-arrestin-mediated erk activation. *Mol Endocrinol*, 20(11) :3014–3026.
- Kayampilly, P. P. and Menon, K. M. J. (2004). Inhibition of extracellular signal-regulated protein kinase-2 phosphorylation by dihydrotestosterone reduces follicle-stimulating hormone-mediated cyclin d2 messenger ribonucleic acid expression in rat granulosa cells. *Endocrinology*, 145(4) :1786–1793.

- Kenakin, T. (2002). Drug efficacy at g protein-coupled receptors. *Annu Rev Pharmacol Toxicol*, 42 :349–379.
- Kholodenko, B. N., Demin, O. V., Moehren, G., and Hoek, J. B. (1999). Quantification of short term signaling by the epidermal growth factor receptor. *J Biol Chem*, 274(42) :30169–30181.
- Kim, J., Ahn, S., Ren, X.-R., Whalen, E. J., Reiter, E., Wei, H., and Lefkowitz, R. J. (2005). Functional antagonism of different g protein-coupled receptor kinases for beta-arrestin-mediated angiotensin ii receptor signaling. *Proc Natl Acad Sci U S A*, 102(5) :1442–1447.
- Kimura, S. and Konagaya, A. (2003). High dimensional function optimization using a new genetic localsearch suitable for parallel computers. In *Proc. IEEE Int Systems, Man and Cybernetics Conf*, volume 1, pages 335–342.
- Kitano, H., Funahashi, A., Matsuoka, Y., and Oda, K. (2005). Using process diagrams for the graphical representation of biological networks. *Nat Biotechnol*, 23(8) :961–966.
- Kovacs, J. J., Hara, M. R., Davenport, C. L., Kim, J., and Lefkowitz, R. J. (2009). Arrestin development : emerging roles for beta-arrestins in developmental signaling pathways. *Dev Cell*, 17(4) :443–458.
- Kroeze, W. K., Sheffler, D. J., and Roth, B. L. (2003). G-protein-coupled receptors at a glance. *J Cell Sci*, 116(Pt 24) :4867–4869.
- Krupnick, J. G., Gurevich, V. V., Schepers, T., Hamm, H. E., and Benovic, J. L. (1994). Arrestin-rhodopsin interaction. multi-site binding delineated by peptide inhibition. *J Biol Chem*, 269(5) :3226–3232.
- Lange, A., Paull, G. C., Coe, T. S., Katsu, Y., Urushitani, H., Iguchi, T., and Tyler, C. R. (2009). Sexual reprogramming and estrogenic sensitization in wild fish exposed to ethinylestradiol. *Environ. Sci. Technol.*, 43(4) :1219–1225.
- Laporte, S. A., Oakley, R. H., Holt, J. A., Barak, L. S., and Caron, M. G. (2000). The interaction of beta-arrestin with the ap-2 adaptor is required for the clustering of beta 2-adrenergic receptor into clathrin-coated pits. *J Biol Chem*, 275(30) :23120–23126.
- Lassègue, B. and Griendling, K. K. (2010). NADPH oxidases : functions and pathologies in the vasculature. *Arterioscler Thromb Vasc Biol*, 30(4) :653–661.
- Layman, L. C., Porto, A. L. A., Xie, J., da Motta, L. A. C. R., da Motta, L. D. C., Weiser, W., and Sluss, P. M. (2002). Fsh beta gene mutations in a female with partial breast development and a male sibling with normal puberty and azoospermia. *J Clin Endocrinol Metab*, 87(8) :3702–3707.
- Lefkowitz, R. J. (1998). G protein-coupled receptors. iii. new roles for receptor kinases and beta-arrestins in receptor signaling and desensitization. *J Biol Chem*, 273(30) :18677–18680.
- Lefkowitz, R. J. and Shenoy, S. K. (2005). Transduction of receptor signals by beta-arrestins. *Science*, 308(5721) :512–517.
- Levallet, G., Levallet, J., Bouraïma-Lelong, H., and Bonnamy, P.-J. (2007). Expression of the camp-phosphodiesterase pde4d isoforms and age-related changes in follicle-stimulating hormone-stimulated pde4 activities in immature rat sertoli cells. *Biol Reprod*, 76(5) :794–803.

- Limbird, L. E. (2004). The receptor concept : a continuing evolution. *Mol Interv*, 4(6) :326–336.
- Liu, X., DePasquale, J. A., Griswold, M. D., and Dias, J. A. (1994). Accessibility of rat and human follitropin receptor primary sequence (r265-s296) in situ. *Endocrinology*, 135(2) :682–691.
- Loganzo, F. and Fletcher, P. W. (1992). Follicle-stimulating hormone increases guanine nucleotide-binding regulatory protein subunit alpha i-3 mrna but decreases alpha i-1 and alpha i-2 mrna in sertoli cells. *Mol Endocrinol*, 6(8) :1259–1267.
- Lohse, M. J., Benovic, J. L., Codina, J., Caron, M. G., and Lefkowitz, R. J. (1990). beta-arrestin : a protein that regulates beta-adrenergic receptor function. *Science*, 248(4962) :1547–1550.
- Luttrell, L. M. and Gesty-Palmer, D. (2010). Beyond desensitization : physiological relevance of arrestin-dependent signaling. *Pharmacol Rev*, 62(2) :305–330.
- Luttrell, L. M. and Lefkowitz, R. J. (2002). The role of beta-arrestins in the termination and transduction of g-protein-coupled receptor signals. *J Cell Sci*, 115(Pt 3) :455–465.
- Luttrell, L. M., Roudabush, F. L., Choy, E. W., Miller, W. E., Field, M. E., Pierce, K. L., and Lefkowitz, R. J. (2001). Activation and targeting of extracellular signal-regulated kinases by beta-arrestin scaffolds. *Proc Natl Acad Sci U S A*, 98(5) :2449–2454.
- Lécureuil, C., Kara, E., Guillou, F., Monniaux, D., and Crépieux, P. (2007). [does fsh signaling have a gender?]. *Med Sci (Paris)*, 23(1) :75–80.
- Lécureuil, C., Tesseraud, S., Kara, E., Martinat, N., Sow, A., Fontaine, I., Gauthier, C., Reiter, E., Guillou, F., and Crépieux, P. (2005). Follicle-stimulating hormone activates p70 ribosomal protein s6 kinase by protein kinase a-mediated dephosphorylation of thr 421/ser 424 in primary sertoli cells. *Mol Endocrinol*, 19(7) :1812–1820.
- Maizels, E. T., Cottom, J., Jones, J. C., and Hunzicker-Dunn, M. (1998). Follicle stimulating hormone (fsh) activates the p38 mitogen-activated protein kinase pathway, inducing small heat shock protein phosphorylation and cell rounding in immature rat ovarian granulosa cells. *Endocrinology*, 139(7) :3353–3356.
- Maltsev, V. A. and Lakatta, E. G. (2010). A novel quantitative explanation for the autonomic modulation of cardiac pacemaker cell automaticity via a dynamic system of sarcolemmal and intracellular proteins. *Am J Physiol Heart Circ Physiol*, 298(6) :H2010–H2023.
- Marion, S., Fralish, G. B., Laporte, S., Caron, M. G., and Barak, L. S. (2007). N-terminal tyrosine modulation of the endocytic adaptor function of the beta-arrestins. *J Biol Chem*, 282(26) :18937–18944.
- Marion, S., Kara, E., Crépieux, P., Piketty, V., Martinat, N., Guillou, F., and Reiter, E. (2006). G protein-coupled receptor kinase 2 and beta-arrestins are recruited to fsh receptor in stimulated rat primary sertoli cells. *J Endocrinol*, 190(2) :341–350.
- Marion, S., Robert, F., Crepieux, P., Martinat, N., Troispoux, C., Guillou, F., and Reiter, E. (2002). G protein-coupled receptor kinases and beta arrestins are relocalized and attenuate cyclic 3',5'-adenosine monophosphate response to follicle-stimulating hormone in rat primary sertoli cells. *Biol Reprod*, 66(1) :70–76.

- Mehta, P. K. and Griendling, K. K. (2007). Angiotensin ii cell signaling : physiological and pathological effects in the cardiovascular system. *Am J Physiol Cell Physiol*, 292(1) :C82–C97.
- Metropolis, N. and Ulam, S. (1949). The monte carlo method. *Journal of the American Statistical Association*, 44 :335–341.
- Milligan, G. and Kostenis, E. (2006). Heterotrimeric g-proteins : a short history. *Br J Pharmacol*, 147 Suppl 1 :S46–S55.
- Mishra, J. and Bhalla, U. S. (2002). Simulations of inositol phosphate metabolism and its interaction with insp(3)-mediated calcium release. *Biophys J*, 83(3) :1298–1316.
- Moore, C. A. C., Milano, S. K., and Benovic, J. L. (2007). Regulation of receptor trafficking by grks and arrestins. *Annu Rev Physiol*, 69 :451–482.
- Musnier, A., Heitzler, D., Boulo, T., Tesseraud, S., Durand, G., Lécureuil, C., Guillou, H. G., Poupon, A., Reiter, E., and Crépieux, P. (2009). Developmental regulation of p70 s6 kinase by a g protein-coupled receptor dynamically modeled in primary cells. *Cell Mol Life Sci*, 66(21) :3487–3503.
- Nakakuki, T., Birtwistle, M. R., Saeki, Y., Yumoto, N., Ide, K., Nagashima, T., Brusch, L., Ogunnaike, B. A., Okada-Hatakeyama, M., and Kholodenko, B. N. (2010). Ligand-specific c-fos expression emerges from the spatiotemporal control of erbb network dynamics. *Cell*, 141(5) :884–896.
- Nelander, S., Wang, W., Nilsson, B., She, Q.-B., Pratilas, C., Rosen, N., Gennemark, P., and Sander, C. (2008). Models from experiments : combinatorial drug perturbations of cancer cells. *Mol Syst Biol*, 4 :216.
- Nelson, C. D., Perry, S. J., Regier, D. S., Prescott, S. M., Topham, M. K., and Lefkowitz, R. J. (2007). Targeting of diacylglycerol degradation to m1 muscarinic receptors by beta-arrestins. *Science*, 315(5812) :663–666.
- Neves, S. R., Ram, P. T., and Iyengar, R. (2002). G protein pathways. *Science*, 296(5573) :1636–1639.
- Novère, N. L., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A., Demir, E., Wegner, K., Aladjem, M. I., Wimalaratne, S. M., Bergman, F. T., Gauges, R., Ghazal, P., Kawaji, H., Li, L., Matsuoka, Y., Villéger, A., Boyd, S. E., Calzone, L., Courtot, M., Dogrusoz, U., Freeman, T. C., Funahashi, A., Ghosh, S., Jouraku, A., Kim, S., Kolpakov, F., Luna, A., Sahle, S., Schmidt, E., Watterson, S., Wu, G., Goryanin, I., Kell, D. B., Sander, C., Sauro, H., Snoep, J. L., Kohn, K., and Kitano, H. (2009). The systems biology graphical notation. *Nat Biotechnol*, 27(8) :735–741.
- Oakley, R. H., Laporte, S. A., Holt, J. A., Caron, M. G., and Barak, L. S. (2000). Differential affinities of visual arrestin, beta arrestin1, and beta arrestin2 for g protein-coupled receptors delineate two major classes of receptors. *J Biol Chem*, 275(22) :17201–17210.
- Orton, R. J., Sturm, O. E., Vyshemirsky, V., Calder, M., Gilbert, D. R., and Kolch, W. (2005). Computational modelling of the receptor-tyrosine-kinase-activated mapk pathway. *Biochem J*, 392(Pt 2) :249–261.
- Ostermeier, A., Gawelczyk, A., and Hansen, N. (1994). Step-size adaptation based on non-local use of selection information. In *Parallel Problem Solving from Nature - PPSN III*, pages 189–198. Springer.



- Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006). How many drug targets are there? *Nat Rev Drug Discov*, 5(12) :993–996.
- Penela, P., Ribas, C., and Mayor, F. (2003). Mechanisms of regulation of the expression and function of g protein-coupled receptor kinases. *Cell Signal*, 15(11) :973–981.
- Perry, S. J., Baillie, G. S., Kohout, T. A., McPhee, I., Magiera, M. M., Ang, K. L., Miller, W. E., McLean, A. J., Conti, M., Houslay, M. D., and Lefkowitz, R. J. (2002). Targeting of cyclic amp degradation to beta 2-adrenergic receptors by beta-arrestins. *Science*, 298(5594) :834–836.
- Ponsioen, B., Zhao, J., Riedl, J., Zwartkruis, F., van der Krogt, G., Zaccolo, M., Moolenaar, W. H., Bos, J. L., and Jalink, K. (2004). Detecting camp-induced epac activation by fluorescence resonance energy transfer : Epac as a novel camp indicator. *EMBO Rep*, 5(12) :1176–1180.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C : The Art of Scientific Computing. Second Edition*. Cambridge University Press.
- Radu, A., Pichon, C., Camparo, P., Antoine, M., Allory, Y., Couvelard, A., Fromont, G., Hai, M. T. V., and Ghinea, N. (2010). Expression of follicle-stimulating hormone receptor in tumor blood vessels. *N Engl J Med*, 363(17) :1621–1630.
- Reiter, E. and Lefkowitz, R. J. (2006). Grks and beta-arrestins : roles in receptor silencing, trafficking and signaling. *Trends Endocrinol Metab*, 17(4) :159–165.
- Ren, X.-R., Reiter, E., Ahn, S., Kim, J., Chen, W., and Lefkowitz, R. J. (2005). Different g protein-coupled receptor kinases govern g protein and beta-arrestin-mediated signaling of v2 vasopressin receptor. *Proc Natl Acad Sci U S A*, 102(5) :1448–1453.
- Rizk, A., Batt, G., Fages, F., and Soliman, S. (2008). On a continuous degree of satisfaction of temporal logic formulae with applications to systems biology. In *In M.Heiner and A.M.Uhrmacher (eds.), CMSB 08 : Proc. of the sixth international conference on Computational Methods in Systems Biology*, pages 251–268. Springer.
- Rizk, A., Batt, G., Fages, F., and Soliman, S. (2009). A general computational method for robustness analysis with applications to synthetic gene networks. *Bioinformatics*, 25(12) :i169–i178.
- Rizk, A., Batt, G., Fages, F., and Soliman, S. (2010). Continuous valuations of temporal logic specifications with applications to parameter optimization and robustness measures. *Theoretical Computer Science*, To appear .
- Rosenbrock, H. H. (1963). Some general implicit processes for the numerical solution of differential equations. *The Computer Journal*, 5(4) :329–330.
- Salvador, L. M., Park, Y., Cottom, J., Maizels, E. T., Jones, J. C., Schillace, R. V., Carr, D. W., Cheung, P., Allis, C. D., Jameson, J. L., and Hunzicker-Dunn, M. (2001). Follicle-stimulating hormone stimulates protein kinase a-mediated histone h3 phosphorylation and acetylation leading to select gene activation in ovarian granulosa cells. *J Biol Chem*, 276(43) :40146–40155.
- Samama, P., Cotecchia, S., Costa, T., and Lefkowitz, R. J. (1993). A mutation-induced activated state of the beta 2-adrenergic receptor. extending the ternary complex model. *J Biol Chem*, 268(7) :4625–4636.

- Sasagawa, S., Ichi Ozaki, Y., Fujita, K., and Kuroda, S. (2005). Prediction and validation of the distinct dynamics of transient and sustained erk activation. *Nat Cell Biol*, 7(4) :365–373.
- Scobey, M., Bertera, S., Somers, J., Watkins, S., Zeleznik, A., and Walker, W. (2001). Delivery of a cyclic adenosine 3',5'-monophosphate response element-binding protein (creb) mutant to seminiferous tubules results in impaired spermatogenesis. *Endocrinology*, 142(2) :948–954.
- Seeger, R., Hanoach, T., Rosenberg, R., Dantes, A., Merz, W. E., Strauss, J. F., and Amsterdam, A. (2001). The erk signaling cascade inhibits gonadotropin-stimulated steroidogenesis. *J Biol Chem*, 276(17) :13957–13964.
- Shanno, D. F. (1970). Conditioning of quasi-newton methods for function minimization. *Mathematics of Computation*, 24 :647–656.
- Shirai, H., Takahashi, K., Katada, T., and Inagami, T. (1995). Mapping of g protein coupling sites of the angiotensin ii type 1 receptor. *Hypertension*, 25(4 Pt 2) :726–730.
- Solis, F. J. and Wets, R. J.-B. (1981). Minimization by Random Search Techniques. *MATHEMATICS OF OPERATIONS RESEARCH*, 6(1) :19–30.
- Stokoe, D., Stephens, L. R., Copeland, T., Gaffney, P. R., Reese, C. B., Painter, G. F., Holmes, A. B., McCormick, F., and Hawkins, P. T. (1997). Dual role of phosphatidylinositol-3,4,5-trisphosphate in the activation of protein kinase b. *Science*, 277(5325) :567–570.
- Tapanainen, J. S., Aittomäki, K., Min, J., Vaskivuo, T., and Huhtaniemi, I. T. (1997). Men homozygous for an inactivating mutation of the follicle-stimulating hormone (fsh) receptor gene present variable suppression of spermatogenesis and fertility. *Nat Genet*, 15(2) :205–206.
- Timossi, C., Maldonado, D., Vizcaíno, A., Lindau-Shepard, B., Conn, P. M., and Ulloa-Aguirre, A. (2002). Structural determinants in the second intracellular loop of the human follicle-stimulating hormone receptor are involved in g(s) protein activation. *Mol Cell Endocrinol*, 189(1-2) :157–168.
- Touyz, R. M. and Schiffrin, E. L. (2000). Signal transduction mechanisms mediating the physiological and pathophysiological actions of angiotensin ii in vascular smooth muscle cells. *Pharmacol Rev*, 52(4) :639–672.
- Traynor, J. (2010). Regulator of g protein-signaling proteins and addictive drugs. *Ann N Y Acad Sci*, 1187 :341–352.
- Troispoux, C., Guillou, F., Elalouf, J. M., Firsov, D., Iacovelli, L., Blasi, A. D., Combarrous, Y., and Reiter, E. (1999). Involvement of g protein-coupled receptor kinases and arrestins in desensitization to follicle-stimulating hormone action. *Mol Endocrinol*, 13(9) :1599–1614.
- Uehata, M., Ishizaki, T., Satoh, H., Ono, T., Kawahara, T., Morishita, T., Tamakawa, H., Yamagami, K., Inui, J., Maekawa, M., and Narumiya, S. (1997). Calcium sensitization of smooth muscle mediated by a rho-associated protein kinase in hypertension. *Nature*, 389(6654) :990–994.

- Urban, J. D., Clarke, W. P., von Zastrow, M., Nichols, D. E., Kobilka, B., Weinstein, H., Javitch, J. A., Roth, B. L., Christopoulos, A., Sexton, P. M., Miller, K. J., Spedding, M., and Mailman, R. B. (2007). Functional selectivity and classical concepts of quantitative pharmacology. *J Pharmacol Exp Ther*, 320(1) :1–13.
- Ushio-Fukai, M., Alexander, R. W., Akers, M., Lyons, P. R., Lassègue, B., and Griendling, K. K. (1999). Angiotensin ii receptor coupling to phospholipase d is mediated by the betagamma subunits of heterotrimeric g proteins in vascular smooth muscle cells. *Mol Pharmacol*, 55(1) :142–149.
- Ushio-Fukai, M., Griendling, K. K., Akers, M., Lyons, P. R., and Alexander, R. W. (1998). Temporal dispersion of activation of phospholipase c-beta1 and -gamma isoforms by angiotensin ii in vascular smooth muscle cells. role of alphaq/11, alpha12, and beta gamma g protein subunits. *J Biol Chem*, 273(31) :19772–19777.
- Varty, G. B., Cohen-Williams, M. E., and Hunter, J. C. (2003). The antidepressant-like effects of neurokinin nk1 receptor antagonists in a gerbil tail suspension test. *Behav Pharmacol*, 14(1) :87–95.
- Vidal, A., Médigue, C., Malpoux, B., and Clément, F. (2009). Endogenous circannual rhythm in luteinizing hormone secretion : insight from signal analysis coupled with mathematical modelling. *Philos Transact A Math Phys Eng Sci*, 367(1908) :4759–4777.
- Violin, J. D., DiPilato, L. M., Yildirim, N., Elston, T. C., Zhang, J., and Lefkowitz, R. J. (2008). beta2-adrenergic receptor signaling and desensitization elucidated by quantitative modeling of real time camp dynamics. *J Biol Chem*, 283(5) :2949–2961.
- Violin, J. D. and Lefkowitz, R. J. (2007). Beta-arrestin-biased ligands at seven-transmembrane receptors. *Trends Pharmacol Sci*, 28(8) :416–422.
- Wayne, C. M., Fan, H.-Y., Cheng, X., and Richards, J. S. (2007). Follicle-stimulating hormone induces multiple signaling cascades : evidence that activation of rous sarcoma oncogene, ras, and the epidermal growth factor receptor are critical for granulosa cell differentiation. *Mol Endocrinol*, 21(8) :1940–1957.
- Wehbi, V., Decourtye, J., Piketty, V., Durand, G., Reiter, E., and Maurel, M.-C. (2010a). Selective modulation of follicle-stimulating hormone signaling pathways with enhancing equine chorionic gonadotropin/antibody immune complexes. *Endocrinology*, 151(6) :2788–2799.
- Wehbi, V., Tranchant, T., Durand, G., Musnier, A., Decourtye, J., Piketty, V., Butnev, V. Y., Bousfield, G. R., Crépieux, P., Maurel, M.-C., and Reiter, E. (2010b). Partially deglycosylated equine lh preferentially activates beta-arrestin-dependent signaling at the follicle-stimulating hormone receptor. *Mol Endocrinol*, 24(3) :561–573.
- Wei, H., Ahn, S., Shenoy, S. K., Karnik, S. S., Hunyady, L., Luttrell, L. M., and Lefkowitz, R. J. (2003). Independent beta-arrestin 2 and g protein-mediated pathways for angiotensin ii activation of extracellular signal-regulated kinases 1 and 2. *Proc Natl Acad Sci U S A*, 100(19) :10782–10787.
- Weiss, J., Morgan, P., Lutz, M., and Kenakin, T. (1996). The cubic ternary complex receptor-occupancy model .1. model description. *JOURNAL OF THEORETICAL BIOLOGY*, 178(2) :151–167.

- Wettschureck, N. and Offermanns, S. (2005). Mammalian g proteins and their cell type specific functions. *Physiol Rev*, 85(4) :1159–1204.
- Woodroffe, P. J., Bridge, L. J., King, J. R., Chen, C. Y., and Hill, S. J. (2010). Modelling of the activation of g-protein coupled receptors : drug free constitutive receptor activity. *J Math Biol*, 60(3) :313–346.
- Xiao, K., McClatchy, D. B., Shukla, A. K., Zhao, Y., Chen, M., Shenoy, S. K., Yates, J. R., and Lefkowitz, R. J. (2007). Functional specialization of beta-arrestin interactions revealed by proteomic analysis. *Proc Natl Acad Sci U S A*, 104(29) :12011–12016.
- Xiao, K., Sun, J., Kim, J., Rajagopal, S., Zhai, B., Villén, J., Haas, W., Kovacs, J. J., Shukla, A. K., Hara, M. R., Hernandez, M., Lachmann, A., Zhao, S., Lin, Y., Cheng, Y., Mizuno, K., Ma'ayan, A., Gygi, S. P., and Lefkowitz, R. J. (2010). Global phosphorylation analysis of beta-arrestin-mediated signaling downstream of a seven transmembrane receptor (7tmr). *Proc Natl Acad Sci U S A*, 107(34) :15299–15304.
- Yin, G., Yan, C., and Berk, B. C. (2003). Angiotensin ii signaling pathways mediated by tyrosine kinases. *Int J Biochem Cell Biol*, 35(6) :780–783.
- Zeleznik, A. J., Saxena, D., and Little-Ihrig, L. (2003). Protein kinase b is obligatory for follicle-stimulating hormone-induced granulosa cell differentiation. *Endocrinology*, 144(9) :3985–3994.
- Zou, Y., Komuro, I., Yamazaki, T., Kudoh, S., Aikawa, R., Zhu, W., Shiojima, I., Hiroi, Y., Tobe, K., Kadowaki, T., and Yazaki, Y. (1998). Cell type-specific angiotensin ii-evoked signal transduction pathways : critical roles of gbetagamma subunit, src family, and ras in cardiac fibroblasts. *Circ Res*, 82(3) :337–345.





## Domitille HEITZLER

### Modélisation dynamique des mécanismes de signalisation cellulaire induits par l'hormone folliculo-stimulante et l'angiotensine

#### Résumé :

La signalisation cellulaire induite par les récepteurs à sept domaines transmembranaires (R7TM) contrôle les principales fonctions physiologiques humaines. Ces R7TMs sont cibles de médicaments et initient de larges réseaux d'interactions. Nous avons modélisé dynamiquement les réseaux de signalisation du récepteur à l'hormone folliculo-stimulante (FSH) régulant la fonction de reproduction et du récepteur angiotensine, un R7TM modèle régulant la tension pour comprendre le fonctionnement de ces réseaux et prédire des données inaccessibles expérimentalement. Notre modélisation a utilisé des équations différentielles ordinaires, en assimilant une variable par espèce et un paramètre par constante cinétique. Les paramètres manquant ont été déterminés par optimisation paramétrique. Puis, nous avons développé un environnement afin de comparer plusieurs algorithmes d'optimisation et créer une nouvelle méthode hybride plus performante et adaptée à la paramétrisation des réseaux de signalisation.

**Mots clés :** Signalisation intracellulaire, R7TM, FSH, angiotensine, modèles mathématiques, ODEs, optimisation continue.

#### Abstract :

Seven transmembrane receptor (7TMR) signaling controls the main human physiological functions. R7TMs are targeted by drugs and initiate large and complex networks responsible for physiological effects. In this thesis, we dynamically modelled the FSHR-induced network that have critical role in reproduction and the angiotensin receptor-induced network which regulates blood pressure and is considered as a model 7TMR with the objective to understand their mechanisms of functioning and to predict experimentally unreachable data. Our modeling, based on ODE formalism, assumes each species as a variable and each kinetic rate as a parameter. Some parameters were unknown and required an adjustment. This led us to develop an environment allowing the comparison of existing adjustment methods and to create a novel and efficient hybrid method well-adapted to parametrization of signaling networks.

**Keywords :** Intracellular signaling, FSH, angiotensin, mathematic modeling, ODEs, continuous optimization.