Comparison of rule- and ordinary differential equation-based dynamic model of DARPP-32 signalling network

Emilia M. Wysocka[1], **Matthew Page**[2], **James Snowden**[2], **and T. Ian Simpson**[1]

[1]**School of Informatics, University of Edinburgh, Edinburgh**
[2]**UCB Celltech, 208 Bath Road, Slough, Berkshire SL1 3WE**

Corresponding author:

Emilia M. Wysocka

Email address: emilia.m.wysocka@gmail.com

ABSTRACT

Dynamic modelling has considerably improved our understanding of complex molecular mechanisms. Ordinary differential equations (ODEs) are the most detailed and popular approach to modelling the dynamics of molecular systems. However, their application in signalling networks, characterised by multi-state molecular complexes, can be prohibitive. Contemporary modelling methods, such as rule-based (RB) modelling, have addressed these issues. The advantages of RB modelling over ODEs have been presented and discussed in numerous reviews. In this study, we conduct a direct comparison of the time courses of a molecular system founded on the same reaction network but encoded in the two frameworks. To make such a comparison, a set of reactions that underlie an ODE model was manually encoded in the Kappa language, one of the RB implementations. A comparison of the models was performed at the level of model specification and dynamics, acquired through model simulations. In line with previous reports, we confirm that the Kappa model recapitulates the general dynamics of its ODE counterpart with minor differences. These occur when molecules have multiple sites binding the same interactor. Furthermore, activation of these molecules in the RB model is slower than in the ODE one. As reported for other molecular systems, we find that, also for the DARPP-32 reaction network, the RB representation offers a more expressive and flexible syntax that facilitates access to fine details of the model, easing model reuse. In parallel with these analyses, we report a refactored model of the DARPP-32 interaction network that can serve as a canvas for the development of more complex dynamic models to study this important molecular system.

INTRODUCTION

Computational dynamic modelling probe mechanistic and quantitative aspects of molecular interactions, which can grant the development of mechanism-based therapies with more predictive power on outcomes of therapeutic interventions (Xie et al., 2014). Such interventions often target molecular signalling (Li and Mansmann, 2014; Jia et al., 2019; Volkow and Boyle, 2018), characterised as a complex system of coupled interacting components, often proteins, forming networks that activity leads to non-additive effects (Kitano, 2002). Defining molecular reactions as a set of coupled ODEs has traditionally enabled dynamic modelling of molecular pathways (Sible and Tyson, 2007). ODE-based modelling is a powerful and acclaimed formalism with a long tradition. It has established standards, and various software tools that support and facilitate the formulation and analysis of ODE models (Dräger et al., 2008; Hoops et al., 2006; Sible and Tyson, 2007). However, the explicit enumeration of molecular species as variables required by the equation-based formulation excludes the representation of molecules that assemble into multivalent protein complexes; typically these have multiple functionally divergent states, a common characteristic of molecules involved in cell signalling (Hlavacek et al., 2003). Computational modelling methods effectively addressed these challenges of expressivity, increasing complexity of the modelled systems. An example of such a method is RB modelling, designed to model interacting proteins. Compared to the

46 ODE-based paradigm, which represents the molecular system as concentrations of molecular species and
47 focuses on their reaction kinetics, RB modelling is an agent-centred method in which the distribution
48 of molecular compositions can be studied along with their abundance (Faeder et al., 2003; Blinov et al.,
49 2004; Danos and Laneve, 2004).

50 **Previous comparisons**   The potential of the RB paradigm has been extensively discussed (Danos,
51 2007; Chylek et al., 2013, 2014, 2015), exemplified (Danos, 2007) and used to answer novel biological
52 questions across signalling, regulatory and metabolic networks (Wilson-Kanamori et al. 2015; Antunes
53 et al. 2016; Di Camillo et al. 2016; Santibáñez et al. 2020; Chattaraj et al. 2021; Nosbisch et al. 2022;
54 see reviews for earlier models published before 2006 - Hlavacek et al. 2006; and from 2007 to 2013 -
55 Chylek et al. 2014). These models are often based on chemical reaction networks previously developed
56 for ODEs. Though the RB paradigm was implemented to match the concept of chemical reactions, rules
57 are generators of reaction networks that may lead to diverging results. As the modelling paradigm can
58 affect the underlying model specification, it would be informative to compare simulations of RB and
59 ODE models defined by the same reaction network, originally built to be solved with ODEs. RB and
60 ODE models have been compared on various aspects before. For instance, Blinov et al. (2006) compared
61 network-like model of epidermal growth factor receptor (EGFR) based on reactions first defined in the
62 ODE model of Kholodenko et al. (1999), with simplified pathway-like structure. However, several
63 assumptions underlying the original model were purposefully modified what changed the underlying
64 interaction network, such as the dissociation of EGFR dimers when phosphorylated or bound to other
65 molecules. In addition to contrasting rules to reactions, another intensively studied aspect, finalised with
66 positive conclusions, was whether stochastic simulation can reveal new properties of a system previously
67 modelled as deterministic one (Vlysidis and Kaznessis, 2018; Hahl and Kremling, 2016; Bustos et al.,
68 2018). As simplistic as it may sound, none of the comparisons mentioned were intended to painstakingly
69 disassemble reaction using the rules of a relatively medium-sized model and compare simulation results.

70 **Objectives**   This study compares the existing ODE model with the new RB model, where both are based
71 on the same definition of chemical reactions. We specifically ask: (1) how closely the dynamics of an
72 ODE-based model can be replicated with an RB one? (2) if differences between the two are observed,
73 what is the underlying cause? (3) if the dynamics are indeed replicated, what advantages are there to
74 using an RB model? We present the results of the models' comparison at the level of notation and model
75 dynamics generated under different conditions. The advantages and disadvantages of the two model
76 representations are discussed, alongside suggestions for future research. We hope that this small-scale
77 attempt may nevertheless bring value to the modelling community to better choose between the two
78 different approaches.

BACKGROUND

80 Dynamic computational modelling frameworks consist of model specification and simulation methods.
81 The model specification is a set of equations or instructions written in a machine-interpretable language
82 that define relationships between variables. Models are run as numerical simulations using algorithms
83 that calculate changes in the variable based on the model's formulation. By adopting a suitable level of
84 abstraction and with the use of a sufficiently expressive language, systems can be modelled such that
85 experimentally-derived evidence can be incorporated to improve the model's quality. A formal approach
86 to the generation of models is desirable to (i) encode facts in an unambiguous and explicit manner, (ii)
87 facilitate the understanding of models, (iii) allow easy modification of models to accommodate more than
88 one hypothesis, (iv) aid interpretation of the underlying biological phenomena, and (v) provide a standard
89 approach to the integration of novel data sources (Kitano, 2002). These features are especially important
90 because model generation often requires knowledge spanning multiple disciplines; the existence of formal
91 modelling frameworks enforces a common understanding of the explicit meaning of model components.
92   In the first step of building a kinetic model, transitions from reactants into products are defined as
93 *chemical reactions* between molecules. Quantitative evaluation of model behaviour over time has been
94 commonly achieved by converting coupled chemical reactions to a set of ODEs that are solved with
95 numerical procedures (Wilkinson, 2006). Each rate equation expresses the change of concentration of a
96 single molecular species over time, formulated with reaction rates that directly take part in the creation and
97 elimination of this species (Sible and Tyson, 2007; Klipp et al., 2005; Hlavacek et al., 2006). Each reaction
98 rate is weighted by a reaction-specific *rate constant*. Time courses obtained by solving ODE models are

continuous and deterministic, characterised by smooth and gradual change of species concentrations over time (Wilkinson, 2009). Although, this setup does not reflect the actual characteristics of subcellular events driven by random collisions between discrete molecules (Gillespie, 1976), this approach is correct as long as abundances of reactants are large enough to render random fluctuations as negligible (Chen et al., 2010). If this condition is violated, the current development of machine coding formats for biological models (e.g. SBML) allows deterministic solvers and stochastic simulators to be applied to the same model specification (Gillespie, 1977; Hoops et al., 2006). A more critical shortcoming of ODE-based models lies in the requirement for explicit enumeration of all molecular species in signalling networks (Hlavacek et al., 2003; Danos, 2007). This drawback precludes mechanistic modelling of systems with multi-state promiscuous molecules that can adopt combinatorially complex states (Seshacharyulu et al., 2012; Chen et al., 2016; Mayer et al., 2009; Suderman and Deeds, 2013), where only a small fraction can be represented with ODEs (Chylek et al., 2014). However, the development of formal methods in computer science has expanded the number of observed properties of biological systems that can be dynamically and quantitatively modelled (Bartocci and Lió, 2016) (reviewed in Machado et al. 2011; Tenazinha and Vinga 2011; Bartocci and Lió 2016; Ji et al. 2017; Le Novère 2015). As this study examines an alternative to ODEs, the focus lies on non-spatial single-scale mechanistic methods. Of those that fit this characterisation (Baeten, 2005; Ciocchetta and Hillston, 2008; Regev et al., 2001; Guerriero et al., 2007; Dematté et al., 2010) we chose rule-based modelling (RBM) as a suitable framework for representing systems with multistate combinatorial interactions. Among two major RBM implementations (Stefan et al., 2014), the very first being BioNetGen (BNG) (Faeder et al., 2003; Blinov et al., 2004; Faeder et al., 2009), the Kappa framework (Danos and Laneve, 2004) was chosen for this comparison. Although there are some notational differences between the two frameworks, the reactions are coded as rules in virtually the same way (Suderman and Hlavacek, 2017). An earlier study by Suderman and Deeds (2013) also showed equality of simulation results between both frameworks.

RBM is based on a formal concept of *graph rewriting*, where molecules are abstracted as structured graph objects referred to as *agents* with *sites* that express *internal* and *binding states*, together constituting *interface* of an agent (Feret et al., 2009a). Reactions are defined as graph transformations encoded as *rules*, that are instructions for local and conditioned transformations. The reaction conditions are encoded in the agent interfaces. Depending on the completeness of these conditions, a rule can represent either a set of reactions that match the *pattern* expressed in a rule, or an exact reaction instance with complete information about the agents' states (Feret et al., 2009b). Thus, a rule can express an infinite number of reactions with a small and finite number of generalised rules (Chylek et al., 2013). During the system simulation agents form *mixture* of *molecular species* that contain a complete description of their states and site occupancy, captured by *snapshots* that provide information on molecular species and their quantities at specified time-points. The trajectories resulting from the model simulation are obtained by declaring variables, called *observables*. If declared with an incomplete pattern, an observable is the sum of the trajectories of the many molecular species that fit the pattern. The Kappa language also allows to induce perturbations during the simulation, for example by updating rate constants or adding molecules. The Kappa simulator, KaSim (Danos, 2007; Krivine et al., 2009), is based on Gillespie's Stochastic Simulation Algorithm (SSA) (Gillespie, 1976, 1977) that produces individual stochastic trajectories of molecules (Gillespie, 1977; Wilkinson, 2009), and whose performance is independent of the size of the reaction network (Yang and Hlavacek, 2011). This RB formulation gives a sufficiently expressive system to capture the principal mechanisms of signalling processes (e.g. biding, dissociation, synthesis, degradation, and state change; Liu and Thiagarajan 2012), providing site-specific details of molecular interactions (e.g. affinities, dynamics of post-translational modifications, domain availability, competitive binding, causality) and the structure of interaction networks.

METHODS

To directly compare models' simulations, we selected an ODE model available in a machine-readable format that can be numerically simulated in non-obsolete software, thus satisfying the reproducibility criterion. A model of the immediate interactors of dopamine- and cAMP-regulated neuronal phosphoprotein with molecular weight 32 kDa (DARPP-32) network by Fernandez et al. (2006) satisfies this requirement. The model was also considered as a solid core for the construction of larger and more complex models in the community interested in the modelling of dopamine (DA)-dependent synaptic plasticity (Manninen et al., 2011). Finally, it is a study widely cited not only by modellers (Nair et al., 2016; Nakano et al.,
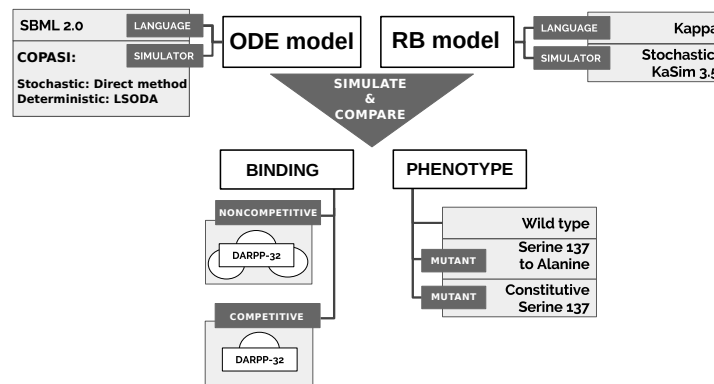
Figure 1. Approach to comparison of ODE and RB modelling frameworks.
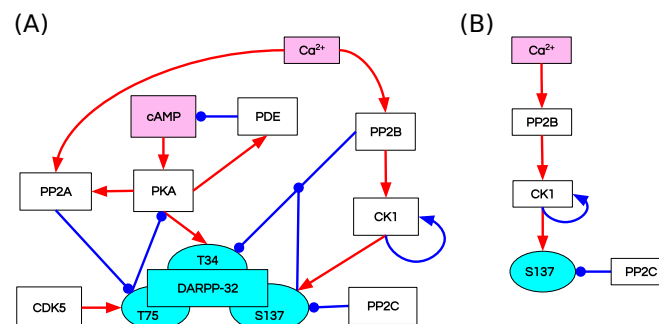
Figure 2. Reaction diagrams showing (A) the DARPP-32 network included in the ODE model by Fernandez et al. (2006) Nodes: DARPP-32 *(cyan)*, second messengers *(magenta)*, kinases/phosphatases *(white)*. Edges: inhibition reactions *(blue)*, activation reactions *(red)*.; (B) the observables with the greatest divergence between trajectories of the ODE and RB models. These observables are connected in a chain of mutually dependent activation reactions triggered by the influx of calcium ions (Ca^{2+}).

2010; Mattioni and Le Novère, 2013) but also by experimentalists (Bales et al., 2011; Kim et al., 2015; Buesa et al., 2016). It should be noted that the stochastic simulation for this particular system should not show any significant differences from the ODE model, as the number of particle copies is sufficiently large. The reactions underlying the ODE model of Fernandez et al. (2006) ("model B") were encoded into the Kappa language, version 3.5 (Feret and Krivine, 2012). Then, models were simulated in different variants to obtain time courses of equivalent observables to compare (Figure 1). First, however, we briefly introduce the DARPP-32 interaction network.

DARPP-32 and its interaction network

DARPP-32, Protein Phosphatase 1 Regulatory Inhibitor subunit 1B isoform 1 (PPP1R1B) (NCBI Resource Coordinators, 2016), is an important multistate and intrinsically disordered protein (Marsh et al., 2010) regulating DA-dependent synaptic plasticity in medium spiny projection neurons (MSPN), the predominant striatal cell type, where signalling cascades are activated simultaneously by glutamatergic and dopaminergic stimuli (Beninger and Gerdjikov, 2005). The malfunction of DARPP-32 relates to a number of neurological disorders, e.g. Alzheimer's disease (Cho et al., 2015), addiction (Philibin et al., 2011), affective disorders (Kunii et al., 2014), and schizophrenia (Kunii et al., 2014; Wang et al., 2015). Glutamatergic and dopaminergic signals are integrated by DARPP-32 which is involved in a complex network of interactions regulated by its multiple phosphorylation sites, of which 4 are known to have an impact on DARPP-32 itself (Yger and Girault, 2011): Threonine 34 (Thr34), Threonine 75 (Thr75), Serine 137 (Ser137), and Serine 102 (Ser102). The Threonine sites (Thr34, Thr75) have major regulatory roles in signal processing. Whereas the Serine sites (Ser137, Ser102) reinforce Thr34 signal.

Fernandez et al. (2006) studied the integrative effect and sensitivities of DA- and glutamate (Glu)-mediated signals on the DARPP-32 network (Figure 2A). Their model examined the particular effect of

175 cyclic adenosine monophosphate (cAMP)-pulse followed by Ca^{2+} spike trains showing that DARPP-32
176 is a robust integrator, far more complex than a bistable switch between DA and Glu signals. The authors
177 included two main pathways that mediate these signals, cAMP–PKA–DARPP-32 phosphorylated at
178 Threonine 34 (D34) and Ca^{2+}–PP2B–DARPP-32 phosphorylated at Threonine 75 (D75). Contrary to the
179 majority of previous models of the Glu and DA signal integration (Lindskog et al., 2006; Gutierrez-Arenas
180 et al., 2014; Nair et al., 2014, 2016), DARPP-32 included three phosphorylation sites: Thr34, Thr75 and
181 Ser137. The authors performed two *in silico* mutagenesis experiments modifying the role of Ser137. The
182 first mutation inhibits site phosphorylation by changing Serine to Alanine at the 137 position (Ser137Ala).
183 The second mutation leads to permanent phosphorylation of the Ser137 site (constSer137).

Model translation

185 As a single reaction can be written as a single rule, the translation could have been accomplished in a
186 one-to-one manner. However, to fully benefit from rule patterns, multiple reactions can be condensed
187 into fewer rules by removing irrelevant context, i.e. *decontextualised*. The context of a reaction in the RB
188 model is defined as the information about the agent's binding sites, partners and internal states. Based
189 on this definition of reaction context, the following criteria guided decisions about condensing reactions
190 into rules. Given a set of reactions of the same type (forward, backward, or catalytic) between the same
191 reactants (agents), if the difference between reactions lies in agent states (internal or binding) that do not
192 change after the transition from reactants to products, and reaction constants (rates) in all these reactions
193 have the same values, then information about agent states does not define reaction conditions; hence, it
194 can be removed from the reaction notation, i.e. a set of reactions become a rule pattern.
195 Among the least intuitive cases in encoding reactions into rules are complex substrate activations
196 of PKA and PP2B. PKA is activated by the binding of 4 cAMP molecules, whereas PP2B activation
197 requires 4 Ca^{2+} ions. In other words, multiple molecules of the same type bind substrates on different sites
198 that have to be uniquely named, which requires explicit encoding of all possible binding combinations
199 on 4 different sides (Danos et al. 2011), called hereafter as *combinatorial binding*. Translation from
200 deterministic to stochastic rate constants and molecular concentrations to copy-numbers were performed
201 in a standard way (Feret and Krivine, 2012; Sekar and Faeder, 2012). Lastly, the cAMP pulse and the Ca^{2+}
202 spiking are reproduced by the addition of molecular copy numbers and modification of rate constants
203 during the simulation, respectively.

Approach to comparison of models

205 The qualitative comparison of the results of two dynamic models required the simulation of both models
206 in a stochastic scheme and the alignment of the trajectory of the corresponding observables under varying
207 conditions to comprehensively compare the frameworks (Figure 1).

208 **Selecting and pairing observables**   The plots in the original publication show aggregated variables
209 that are summed trajectories of multiple molecular species. For instance, "D34" denotes DARPP-32
210 phosphorylated at Thr34, regardless of its state of binding or other phosphorylation sites. The concept
211 of aggregated variables corresponds to *observables* in RB modelling, and therefore, we use the term
212 *observable* hereafter to denote aggregated variables. Observables of the ODE model were aggregated
213 here based on their names matching partial strings representing the observables of interest. To verify this
214 approach, obtained observables of the ODE model with this method were qualitatively compared with the
215 6 observables plotted in the original publication. The choice of other observables follows these principles:
216 (1) if an agent has internal states, the activated state is set as its observable form, e.g. "CK1u"; (2) if
217 an agent is created and degraded over the simulation, its observable is set to its least specific form, e.g.
218 "PKA"; (3) if an agent is not created and degraded during the simulation, i.e. its level remains constant
219 throughout the simulation and has no internal states, its observable is set to its bound form, e.g. "_CDK5"
220 (see Table S1 for the complete list of RB and ODE observables with definitions).

221 **Model simulation**   We simulate both models in the stochastic scheme but within different simulation
222 environments. The RB model was simulated with KaSim. The SBML format of the ODE model was
223 run with COPASI (version 4.20), a common simulation environment for Systems Biology Markup
224 Language (SBML)-formated models (Hoops et al., 2006), using the deterministic solver (LSODA) and
225 stochastic simulator (*direct method* Gillespie 1977). To match units of molecular abundances across
226 all simulation setups, the results of the SBML model were set to "particle numbers", equivalent to copy
227 numbers in the RBM naming convention.

ODE model			RB model	
Model Component	Total counts	Total counts	Model Component
Reaction instances	152	132	Reaction rules
Concentration-based rate constants	152	62	Stochastic rate constants
Initial concentrations	75	8	Initial copy numbers
Molecular species	75	91/137	Molecular species
Stimuli events	21	21	Stimuli events

Table 1. The specifications of the ODE and RB models can be broken down into elements, the number of which can be compared.

Model perturbations   To induce the first type of site-directed mutations, Ser137Ala and constSer137, we modified rate constants during the simulation. In both cases, the alteration of the model involved the inactivation of 4 reactions by zeroing their rate constants. In the RB model, these reactions are represented by 1 rule, thus a change of a single constant induced each mutagenesis.

We additionally tested the RB model with two different binding schemes, applicable only to the RB model, and further called *noncompetitive* and *competitive binding* (Figure 1). In the noncompetitive binding, all interactors of DARPP-32 can bind simultaneously to three different sites. The competitive binding assumes one interaction with DARPP-32 at a time, which reflects the ODE model assumption.

RESULTS

Comparison of models was performed on two levels, model notation and simulation results. The model notation was analysed by dividing the model into components and comparing their sizes. We expect the set of reactions underlying the ODE model to be represented with fewer rules since a single one can constitute a pattern representing several reactions. The comparison of simulation results involved the alignment of equivalent time courses obtained by model simulations. We performed the comparison of time courses between three variants of each model: (1) base-line condition (wild-type) and two site-directed mutations: (2) Ser137Ala and (3) constSer137. Finally, we compared two RB model variants, representing: (1) DARPP-32 with a single binding site; and (2) DARPP-32 with three independent binding sites.

Rule patterns reduce reaction number in a certain type of model components

Table 1 juxtaposes the total counts of model components in each model. The RB model represents 152 reactions with 132 rules, each parameterised by one of 62 unique rate constants. This number is lower than the total number of rate constants used to parameterise the ODE model (152). The final rule set is more than twice as large as the unique number of rate constants, meaning that more than one rule is parameterised by the same rate constant. The number of molecular species in the RB model, obtained with snapshots capturing the state of the molecular mixture over simulation time (every 10000th event), is 91 for the competitive RB model, and 137 for the non-competitive one. In both cases, the sum of molecular species is higher than in the ODE model (75). As expected, the number of rules corresponding to reactions is lower, and the number of molecular species is much higher, confirming that expression patterns reduce the number of rules needed to represent a reaction system. Nonetheless, the number of rules is only slightly lower than the number of reactions (152 to 131). If we closely compare models by parts representing more general molecular mechanisms, rule representation reduces the reaction number in some components but extends it in others (Table 2). The reduction occurred only in "DARPP-32 phosphorylation" and "PP2A activation by Ca^{2+}" components, where combinations of states of DARPP-32 phosphorylation sites do not have to be explicitly named. In contrast, the increase in the number of reactions compared to the reactions occurred in the components "PKA activation" and "PP2B activation". They both have 4 sites that bind the same molecules, Ca^{2+} and cAMP, respectively.

RB model recapitulates dynamics of ODE model with minor discrepancies

The RB model recapitulates the principal dynamics of the ODE model, albeit there are some observable differences (compare Figures 3B and C). For instance, during the relaxation phase (after the 600th time point), "D34" in the ODE model needs 100 more time steps to reach the second peak, and it is weaker than its RB counterpart ("D34*"). Worth noting is that the standard deviation in the stochastically simulated ODE model reveals a distinctive variation in abundance of the "D34" observable during the relaxation

	Model component	Reactions	Rules	Unique rate constants
1.	DARPP-32 phosphorylation	84	27	27
2.	CK1 phosphorylation	4	4	4
3.	PDE phosphorylation	4	4	4
4.	PP2A phosphorylation	4	4	4
5.	PP2B activation	4	24	4
6.	PKA activation	12	56	7
7.	cAMP & Ca^{2+} degradation	8	8	8
8.	PP2A activation by Ca^{2+}	32	4	4

Table 2. The list of reactions in the Fernandez et al. (2006) publication was divided into components based on more general molecular processes represented by subsets of reactions, such as phosphorylation or activation. We can closely examine the relationship between reaction and rule by comparing models by component. The table shows the number of reaction rules versus reaction instances and a unique number of rates per model component. It is noticeable that the reduction in the number of reaction instances due to the translation of reactions into Kappa language occurred in only two model components (1. & 8.), while in two others, it resulted in an expansion of the rule number (5. & 6.).
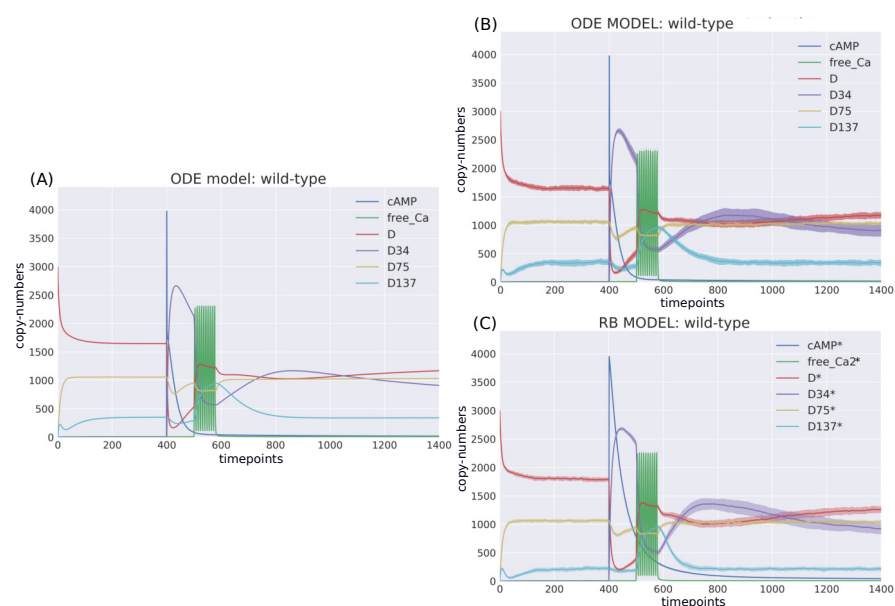
Figure 3. Time-courses of the ODE model for DARPP-32 isoforms triggered by a pulse of cAMP followed by a train of Ca2+ spikes obtained with (A) a deterministic solver, and (B) a stochastic simulation. Trajectories of the stochastic simulation were obtained from calculating mean value (*line*) and standard deviation (*shade*) based on 40 simulations. (C) RB model (stochastic simulation). Variable isoforms of DARPP-32: "D" - unphosphorylated; "D137" - Ser137 phosphorylated; "D75" - Thr75 phosphorylated; "D34" - Thr34 phosphorylated.
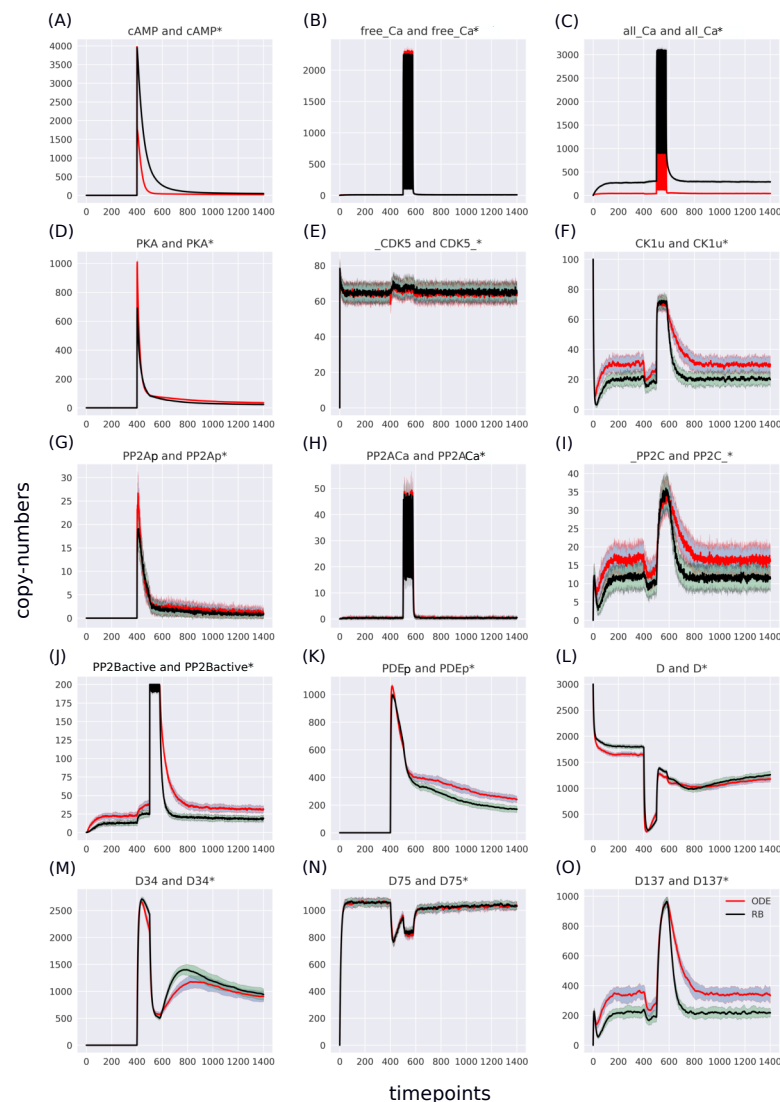
Figure 4. Overlaid time courses of the ODE stochastic and the RB stochastic in the baseline condition. Note that the scales on the y-axis are different to closely compare the traces of the observables. Trace colour: ODE (*red*), RB (*black*).

phase. We further use the stochastic trajectories of the ODE model for comparison with the RB model. For a closer examination, traces of 15 observables (defined in Table S1) obtained from ODE and RB simulations were paired and superimposed (Figure 4). Next to the clear matches (e.g. Figure 4: B, E, H, N), there are discrepancies between paired curves. Five of these 15 observables (Figure 4: C, F, I, J, O) are examples of the largest divergence between models following a similar pattern of behaviour. They are directly connected in a chain of activation reactions that begins with Ca^{2+} (Figure 2B). Higher abundance of all Ca^{2+} ions present in the system of the ODE model (Figure 4C) could explain differences between the remaining 4 observables. However, the trajectory of "all_Ca" remains at the 0 level during steady states rising only in the spiking interval, which resembles the abundance of free Ca^{2+} (Figure 4B). Ca^{2+} activates PP2B represented by the trajectory "PP2Bactive". The higher level of "PP2Bactive" is consistent with the other three observables, suggesting that this is a factor generating the divergences between the models. Based on the curves of the ODE model, we can reason that a stronger activation of PP2B results in proportionally more copies of the unphosphorylated CK1 and phosphorylated D137. This, in turn, increases substrate availability for PP2C; therefore, more copy-numbers of its bound form. This effect is inverted in the trajectories of the RB model.

RB language allows for detailed dissection of observed molecular species

As the "all_Ca" observable trajectory produced by the ODE model is much lower than in the RB model at the steady state, and "PP2Bactive" appears to dictate the higher effect on the other three observables ("D137", "CK1u", "_PP2C"), "all_Ca" and "PP2Bactive" are closer analysed in further steps.

According to the reaction system underlying both models, the activated PP2B is a complex of 4 Ca^{2+} ions and PP2B. This detail is not explicitly stated in the variable name of the ODE model. Therefore, to obtain the trajectory of all Ca^{2+} ions, the sum of the copy numbers of the molecular species whose variable names contain "Ca^{2+}", would have to be replaced by a more thorough analysis of the relevant reaction context of the ODE model. This is not the case in the RB model, where an observable of interest is obtained with an automated procedure that sums the trajectories of molecular species containing the specified expression pattern. Since the trajectory of all Ca^{2+} in the RB model includes the ions bound to PP2B, the comparison of "all_Ca" to "all_Ca*" is inaccurate due to a difference in the molecular species included in these observables. A similar inaccuracy, related to the naming of the observables, explains the discrepancy between the time courses of the total number of cAMP observables (Figure 3). In contrast to the RB model, the multiple copies of cAMP bound to R2C2 are not included in this ODE model time course (Figure 4A).

To determine the identity of molecular species whose trajectories were summed up in "all_Ca*", all species containing Ca^{2+} in the RB model simulation were isolated from snapshot data ammounting to 24 compared to only 13 in the ODE model ("all_Ca"). These 13 species correspond in molecular composition to 18 of the 24 RB species sampled in total. The six absent species in the "all_Ca" observable are composed of an active form of PP2B containing 4 Ca^{2+} ions, either free or bound to phosphorylated CK1, or DARPP-32 in 4 different combinations of phosphorylation states. The number of species comprising "all_Ca" observable in the RB model is higher by five because the half-active form of PP2B (bound to two Ca^{2+} ions) in the RB model exists in 6 variants. Whilst in the ODE model, it is represented as a single species, named "PP2BinactiveCa2". By rerunning the RB model with a new set of observables matching those in the ODE model and superimposing only one of 6 trajectories of "PP2BinactiveCa2", we can obtain a close match between the "all_Ca" observables (Figure S1C). This demonstrates that the differences between the "all_Ca" observables of the two models can be explained by the difference in the number of representations of the molecular species. As the six trajectories have the same dynamics and average levels of abundances, choosing one of them is arbitrary (Figure S2). Moreover, the distinction between locations of two Ca^{2+} ions on the numbered sites of PP2B is irrelevant since all 4 sites are functionally indistinguishable.

Rate constants of reactions formulated with "combinatorial-binding" notation should be increased to match ODE trajectories

The largest discrepancy between trajectories can be observed in "PP2BinactiveCa2" (Figure 5M) that for the RB time course was obtained by summing 6 entities representing a half-active PP2B into one. If divided by 6, representing a single variant of half-active PP2B, the trajectory of the RB model becomes lower than the one of the ODE model (Figure S3M). These six forms of the half-active PP2B suggest that a better fit between the two models can be achieved by decreasing the constant rate of rules that represent the binding of Ca^2+ to free PP2B. However, the decrease of this rate constant of PP2B intermediate form would lead to a further decrease of copy numbers of other coupled observables that trajectories are lower than in the ODE model in the current parameter setting, e.g. the fully active form of PP2B (Figure 4J). To further examine this observation, we can return to the comparison of model specifications (Table 2). The 4 reactions of PP2B activation are represented by 24 rules, explicit in site-specific detail that includes all combinations of Ca^{2+} positioning on 4 sites of PP2B. Moreover, instead of 3 distinct PP2B species in the reaction representation (inactive, semi-active, activated), in the RB model there are 8 different molecular species. Thus, the fine-grained representation of species in the molecular mixture slows down the transition from inactive to active PP2B, despite equal presence of the two-step transition encoded in both formalisms. The same can be observed in the second example that required a much larger number of rules, i.e. an activation of PKA. Figure 4D shows that the RB trajectory of the "PKA*" observable also reaches a much lower peak than its ODE counterpart. Accordingly, values of rate constants of rules that the number increased due to the "combinatorial binding" notation in the RB model, i.e. are represented by more variants of species, should be increased to closely match the ones in the ODE model.
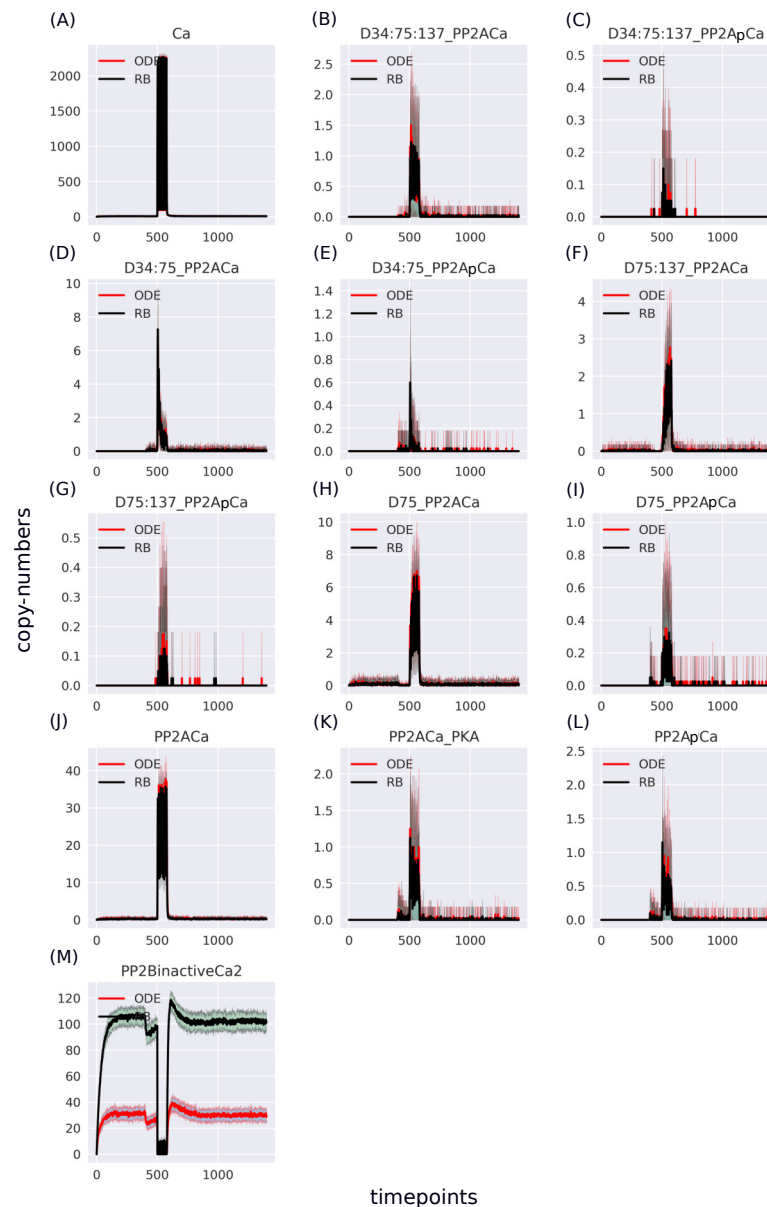
Figure 5. Traces of 13 pairs of molecular species containing Ca^{2+}, selected to match the ODE model. The largest disparity lies in the "PP2BinactiveCa2" variable - summation result of 6 entities representing an inactive form of PP2B in the RB model.
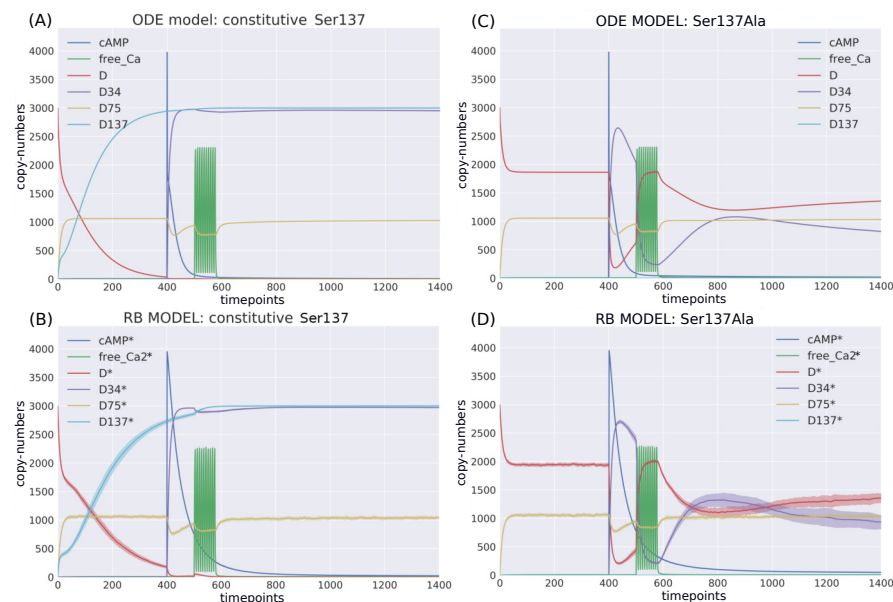
Figure 6. Comparison of the constitutive Ser137 mutation induced in (A) ODE model in deterministic setting; (B) RB model in stochastic setting; and the Ser137Ala mutation in (C) ODE model in deterministic setting; (D) RB model in stochastic setting; The same interference performed on rate constants of the two models caused similar dynamics.

RB language facilitates modifications of dynamic models

Site-directed mutations   The ODE model was analysed under two site-directed mutagenesis conditions affecting the Ser137 site. When these perturbations were applied in the RB model, we could observe a close fit in the initial conditions and a general pattern of dynamics based on six key observables (Figure 6). Thus, RBM allows to emulate experimentally observed perturbations similarly to ODEs.

Competitive and non-competitive site binding   We also compared two variants of the RB model with different binding site specifications to establish whether the dynamics of the model are affected when DARPP-32s binds multiple partners at once. The first specification is a competitive variant of the model with one binding site (oBS). In the second, a non-competitive variant, the partners bind simultaneously (three-binding-sites, tBS). This type of comparison between ODEs and RBs was previously presented by Blinov et al. (2006). In contrast to their results, the superimposed trajectories of two models (Figure S4) demonstrate no effect on model dynamics. As the direct consequence of this modification is an increase in the size of the complexes to more than two proteins, it seems that larger complexes are rarely formed during the simulation. This interpretation is confirmed in a direct examination of species counts (Figure S5).

DISCUSSION

ODE-based modelling is a classical and commonly used method for creating detailed dynamic models of biological systems (Lotka, 1920; Hodgkin and Huxley, 1952; Lisman, 1989). It is frequently a point of reference and comparison to newly proposed modelling methods (Morris et al., 2010; Ciocchetta and Hillston, 2008; Chaouiya, 2007; Chylek et al., 2015; Danos and Laneve, 2004). Nevertheless, modelling of signalling systems with ODE poses difficulties due to complexities underlying molecular interactions (Kholodenko, 2006; Stefan et al., 2014). RBM was proposed as a solution to this problem.

Numerous reviews (Danos, 2007; Chylek et al., 2014; Hlavacek et al., 2006) and studies discussed the advantages of RB modelling over ODE. This paper presents results of encoding reactions underlying an ODE model to the RB language and a comparison of their specification and simulation results. The manual translation of this ODE model into any RB language was necessary despite the existence of a method (Tapia and Faeder, 2013) for automated translation of the SBML-format encoding ODE-based models to an RB model format. The translation failed to generate executable model with correctly identified agents (Supplement 1).

Effects of the RB framework on the model notation   Encoding reactions into rules slightly reduced the size of model specification and increased counts of molecular species, which confirmed the well-known advantage of rule representation (Hlavacek et al., 2006; Chylek et al., 2014). Closer analysis of reaction subsets representing more general molecular mechanisms showed that reduction in the reaction number is only true for reactions occurring between the same reactants, describe the same transformation, and are parameterised with the same values of rate constants, but differ only concerning binding or internal state of reactants. In this type of reaction, the number of unique reaction rates was equal to the number of rules. An increase in the number of rules representing reactions occurred where the same partner binds an agent at multiple sites. In such reactions, all possible positions and stages of the binding process had to be explicitly encoded. The "combinatorial binding" notation is not a general property of the RB language but characteristic of the Kappa syntax. In the alternative to Kappa RB framework, BNG, a rule is definable with identical sites' names. This implies that the rule pattern defined for one applies to the others, effectively shortening the rule description (Sekar and Faeder, 2012).

Discrepancies in dynamics between models   As in other similar study (Blinov et al., 2006), the comparison of trajectories showed agreement between model dynamics, with some discrepancies. Their source is not due to the issue of ambiguous molecularity that can arise in Kappa, which requires control on the part of the modeller, as none of the rules in the DARPP-32 model allow the formation of closed rings between agents. The discrepancies mainly appeared due to a lack of precision in the variable names of the ODE model caused by differences in molecular species comprising tracked observables. When the simulation of the RB model was performed with observables exactly matching the ones in the ODE model, almost all paired trajectories fitted perfectly. The only problematic observable involved the "combinatorial binding" notation. Further analysis suggested that activation of proteins encoded with fine detail was slower than in the ODE model. More specifically, although half-active PP2B was more abundant in the RB model, its active form had lower levels compared to the ODE model. This observation may apply to RBM in general because the molecular mixture when simulated with NFsim (Sneddon et al., 2011), a variant of the network-free simulator for BNGL, will implicitly contain the same number of molecular species with different binding variants as it is in KaSim. Further, based on the example of antibodies with identical antigen-binding sites, the BNGL developers demonstrated that the equivalent set of rules in both RBM frameworks are parameterised by the same rate constant (Suderman and Hlavacek, 2017). In the BNG framework, however, this rate is automatically scaled by a factor equivalent to the multiplicity of indistinguishable ways to obtain the reaction product (Faeder et al., 2005, 2009). Moreover, comparison of the time traces of Kappa- and BNGL-defined rules shows equivalent results under stochastic realisations (Suderman and Hlavacek, 2017). Thus, we argue that "combinatorial binding" expressed as a single rule in BNGL will not change the dynamics observed with Kappa. These specific discrepancies should be taken into account when the reactions and rate constants of RB models are derived from ODE models, and preferably determined experimentally for these reaction types.

A shift in modelling focus with the use of the RB framework   The process of encoding reactions into rules turns attention to questions such as how many binding partners can simultaneously bind a protein. The translation process has shown that information about interfaces of interacting proteins and their alternative states would considerably ease the process of model development by guiding decisions on agents' signatures. Therefore, data resources that could support RB modelling, such as a source of protein interaction interfaces, post-translational modifications (PTMs), and protein domains. For instance, proteins containing phosphatase catalytic domains are enzymes of dephosphorylation reactions (Sacco et al., 2012). However, such detailed information is not accessible for most molecular agents.

Facilitation of modification and reuse of dynamic models   Molecular species defined in the ODE framework are fixed and definite, whereas, in the RB model, they are a subject of investigation. The determination of created species and their abundances is enabled with snapshots, state vectors of molecular mixture sampled during the simulation. Snapshots have been used as a tool for model exploration in other studies (Sorokina et al., 2011; Suderman and Deeds, 2013), and their visualisation is included as a standard tool in the Kappa framework (Boutillier et al., 2018). Although, unlike Kappa, BNG allows counting the number of species through network generation, this option is limited by the size of the model. Then, if not sampled, the number of specie types is calculated analytically (Suderman and Deeds, 2013). While this is an approximate approach, by using snapshots, we were able to study the emerging molecular species in detail and precisely track the observables of interest. Furthermore, the automatic identification

and merging of time courses into observables makes the RB framework particularly advantageous over the ODE framework, as it avoids errors in the identification of molecules in the modelled system. This is important when the system consists of many molecules and their states are to be analysed in detail. In ODE models, on the other hand, obtaining the details of the molecule composition, hidden in the arbitrary names of the individual molecular species, would require further deconstruction of the reaction system. Though, the precise identification of molecules within species could be performed by parsing the SBML-model encoding, the Fernandez et al. (2006) model web page in the BioModels website[1], shows incomplete annotations of molecular species behind variable names, both concerning the actual counts of interactions (e.g. Ca^{2+}) and their components.

The Fernandez et al. (2006) study does not discuss if DARPP-32 partners bind to the same or different active sites, though DARPP-32 is an intrinsically disordered protein with an unknown binding interface (Dancheck et al., 2008; Mollica et al., 2016; Engmann et al., 2015; Choy et al., 2012). The ODE model specification demonstrates that DARPP-32 forms at most heterodimers. This type of modification in the ODE model would require the arduous extension to much more complex model by enumeration of additional molecular species, the addition of new equations and updating the existing ones. Contrary to this, a definition of such a binding scenario in the RB notation requires the same number of rules, provided that concurrently bound interactors do not influence each other. The RB model was tested with two types of site-directed modifications, demonstrating the framework's flexibility to reproduce experimentally conducted perturbations. Though the binding site modification effectively changed the model reaction network, it did not affect the model response. Nevertheless, this intervention demonstrated the ease of performing such alterations within the RB framework. Additionally, the pattern notation improves model clarity and provides an intuitive representation of a model akin to a set of chemical reactions rather than equations, potentially improving the learning curve for a modeller-to-be.

Note on simulation time   The ODE simulation remains incomparably faster than the RB simulation. In both deterministic and stochastic settings, COPASI returns ODE results in seconds, while KaSim 3.5 requires ~40min[2] to return the RB model results (single core on ThinkPad Core i7-4700MQ; RAM: 16GB). With more recent versions of KaSim (≥ 4), the execution time could be much shorter thanks to a new solution for the update of observable instances (Boutillier et al., 2017). Nevertheless, with a sufficiently simple model, network-based simulation can be performed far more efficiently using BNGL, bearing in mind that more complex models will have a higher runtime (Sneddon et al., 2011). When generating large-scale models, the type of RB model simulation can be scheduled on a case-by-case basis (Santibáñez et al., 2020).

Further explorations of the DARPP-32 RB model   There are two main routes for further explore the RB model. The first one is a modification of parameters defining different phases of combinatorially bound Ca^{2+} ions to PP2B and cAMP to R2C2. A particular task would be to identify factors by which the binding constants could be modified to counteract the many intermediate variants of these complexes and the lower copies of their activated final forms. Next, it would be interesting to identify conditions under which we could observe a difference in model dynamics after the addition of binding sites to DARPP-32. In the current setup, the lack of difference might be caused by the similarity in occupancy between a single site and all three sites together, as they do not counter each other's binding properties. The probability of a site being connected depends on copy numbers of reactants and the strength of binding affinities. Reactions in the model are classified as weak, with dissociation rates in the range of μM. Low-affinity bindings generally lead to lower levels of site occupancy. Moreover, the amount of DARPP-32 molecules exceeds the total counts of all its interactors. Thus, with the current proportions of reactants, all three DARPP-32 sites cannot be saturated. To expose the potential differences in dynamics between the two binding scenarios, the site occupancy could be modified by increasing the size of the reactant pools or reducing the level of DARPP-32. As mentioned by Fernandez et al. (2006), levels of DARPP-32 vary considerably, between μM to tens of μMs, in the striatum. With the greater availability of single-cell techniques for protein quantification (Lo et al., 2015) it would be worth establishing more precisely the range of DARPP-32 even at the resolution of a dendritic spine (Otmakhov and Lisman, 2012). Estimating variability between cells could also be used to compare the varying levels of phosphorylated DARPP-32 at Thr34 that were observed in the stochastic simulations.

---

[1] http://www.ebi.ac.uk/biomodels-main/BIOMD0000000153
[2] The total CPU time measured with the "time" command on Linux OS

Extending the model to match evolving knowledge on the DARPP-32 interaction network The main advantage of RBM over ODE is the flexibility to extend and manipulate the model, rather than to rewrite it into a new one. Through a single representation, the community interested in DARPP-32-related molecular systems could pool their knowledge. We believe that the refactored model of the DARPP-32 in a rule-based framework will facilitate the creation of such an updated model in future. It should be noted that the "DARPP-32 events" pathway (R-HSA-180024.3) specified in the REACTOME database is in fact the model of Fernandez et al. (2006) used in this study. So far, only early signalling events of DARPP-32 have been modelled, localised mainly in the cytosol. The Ser102 site regulates nuclear transportation of DARPP-32 representing late signalling events Stipanovich et al. (2008). The site was omitted in the Fernandez et al. (2006) model due to the lack of evidence that it may be affected by DA or Glu signalling (Girault et al., 1989; Svenningsson et al., 2004). However, a recent study by Nishi et al. (2017) suggests that Glu can decrease the effect of DA signalling by dephosphorylating DARPP-32 at Ser102 causing the accumulation of DARPP-32 in the nucleus, the effect known to be promoted by drugs of abuse (Stipanovich et al., 2008). Moreover, the role of the DARPP-32 interaction network in the nucleus appears to be an important future goal of the recently published ODE model incorporating DARPP-32 (Yapo et al., 2018), indicating the need to explicitly demonstrate the contribution of DARPP-32 to the observed switch-like behaviour of PKA in the nucleus. The inclusion of the Ser102 site seems inevitable to achieve this goal, but extending the current model with 2 phosphorylation sites may prove to be quite complex. Encouragingly, to calibrate this extended model, measurements experimentally obtained by Nishi et al. (2017) for all 4 phosphorylation sites could be fed into Pleione (Santibáñez et al., 2019), a recently developed tool for this task in RBM. Lastly, it would be worth to establish if stochastic simulation can demonstrate new behaviour of this particular molecular system that deterministic simulation has so far not revealed.

Need for formal prioritisation methods of emerging molecular species RB modelling offers tools for the dissection of emerging molecular species during simulations. As the number of such molecular species increases it becomes difficult to identify which model components are of particular importance to the system. There is great potential to exploit commonly used methods for model exploration, such as sensitivity analysis, to identify critical model features including parameters and model output variables. It would be advantageous to support the modeller's assumptions with automated methods to prioritise model outputs for downstream analysis and greater insight into the underlying biological systems.

CONCLUSIONS

Dynamic molecular modelling has become increasingly important in uncovering and integrating our dispersed knowledge of molecular mechanisms. Choosing the best formalism meet this challenge is a difficult task. In this work, we have presented a detailed and systematic comparison of two major formal approaches to quantitative modelling. We demonstrated the advantages and disadvantages of RB modelling to the preeminent ODE approach. We confirm, after other similar studies, that RB modelling is more detailed and flexible way to represent biological molecular systems, enabling exploration of individual molecular entities, model extension, and future reuse. These conclusions confirm the potential of the RB formalism and hopefully will embolden future exploration and research in this topic.

CODE ACCESSIBILITY

The code reproducing the figures and the encoding of the Kappa models can be found on github : `https://github.com/ewysocka/rb_vs_ode_model_of_darpp-32`.

LIST OF ACRONYMS

BNG BioNetGen

Ca^{2+} calcium ions

cAMP cyclic adenosine monophosphate

DA dopamine

DARPP-32 dopamine- and cAMP-regulated neuronal phosphoprotein with molecular weight 32 kDa

EGFR epidermal growth factor receptor

Glu glutamate

MSPN medium spiny projection neurons

ODE ordinary differential equation

PTM post-translational modification

RB rule-based

RBM rule-based modelling

SBML Systems Biology Markup Language

Ser102 Serine 102

Ser137 Serine 137

SSA Stochastic Simulation Algorithm

Thr34 Threonine 34

Thr75 Threonine 75

REFERENCES

Antunes, G., Roque, A. C., and Simoes-de Souza, F. M. (2016). Stochastic Induction of Long-Term Potentiation and Long-Term Depression. *Nature Scientific Reports*, 6(1):30899.

Baeten, J. (2005). A brief history of process algebra. *Theoretical Computer Science*, 335(2-3):131–146.

Bales, J. W., Yan, H. Q., Ma, X., Li, Y., Samarasinghe, R., and Dixon, C. E. (2011). The dopamine and cAMP regulated phosphoprotein, 32kDa (DARPP-32) signaling pathway: A novel therapeutic target in traumatic brain injury. *Experimental Neurology*, 229(2):300–307.

Bartocci, E. and Lió, P. (2016). Computational Modeling, Formal Analysis, and Tools for Systems Biology. *PLoS Computational Biology*, 12(1):e1004591.

Beninger, R. J. and Gerdjikov, T. V. (2005). Dopamine-Glutamate Interactions in Reward-Related Incentive Learning. In *Dopamine and Glutamate in Psychiatric Disorders*, pages 319–354. Humana Press, Totowa, NJ.

Blinov, M. L., Faeder, J. R., Goldstein, B., and Hlavacek, W. S. (2004). BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics*, 20(17):3289–3291.

Blinov, M. L., Faeder, J. R., Goldstein, B., and Hlavacek, W. S. (2006). A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial complexity. *Bio Systems*, 83(2-3).

Boutillier, P., Ehrhard, T., and Krivine, J. (2017). Incremental Update for Graph Rewriting. In Yang, H., editor, *Programming Languages and Systems*, Lecture Notes in Computer Science, pages 201–228, Berlin, Heidelberg. Springer.

Boutillier, P., Maasha, M., Li, X., Medina-Abarca, H. F., Krivine, J., Feret, J., Cristescu, I., Forbes, A. G., and Fontana, W. (2018). The kappa platform for rule-based modeling. *Bioinformatics*, 34:i583–i592.

Buesa, I., Aira, Z., Azkue, J. J., Li, J., Zhang, X., and Chen, X. (2016). Regulation of Nociceptive Plasticity Threshold and DARPP-32 Phosphorylation in Spinal Dorsal Horn Neurons by Convergent Dopamine and Glutamate Inputs. *PLOS ONE*, 11(9):e0162416.

Bustos, Á., Fuenzalida, I., Santibáñez, R., Pérez-Acle, T., and Martin, A. J. M. (2018). Rule-Based Models and Applications in Biology. In von Stechow, L. and Santos Delgado, A., editors, *Computational Cell Biology*, volume 1819, pages 3–32. Springer New York, New York, NY. Series Title: Methods in Molecular Biology.

Chaouiya, C. (2007). Petri net modelling of biological networks. *Briefings in bioinformatics*, 8(4):210–9.

Chattaraj, A., Blinov, M. L., and Loew, L. M. (2021). The solubility product extends the buffering concept to heterotypic biomolecular condensates. *eLife*, 10:e67176. Publisher: eLife Sciences Publications, Ltd.

Chen, J., Zeng, F., Forrester, S. J., Eguchi, S., Zhang, M.-Z., and Harris, R. C. (2016). Expression and Function of the Epidermal Growth Factor Receptor in Physiology and Disease. *Physiological Reviews*, 96(3):1025–1069.

Chen, Y., Cunningham, F., Rios, D., McLaren, W. M., Smith, J., Pritchard, B., Spudich, G. M., Brent, S., Kulesha, E., Marin-Garcia, P., Smedley, D., Birney, E., and Flicek, P. (2010). Ensembl variation resources. *BMC Genomics*, 11(1):293.

Cho, K., Cho, M. H., Seo, J. H., Peak, J., Kong, K. H., Yoon, S. Y., and Kim, D. H. (2015). Calpain-mediated cleavage of DARPP-32 in Alzheimer's disease. *Aging Cell*, 14(5):878–886.

Choy, M. S., Page, R., and Peti, W. (2012). Regulation of protein phosphatase 1 by intrinsically disordered proteins. *Biochemical Society transactions*, 40(5):969–74.

Chylek, L. A., Harris, L. A., Faeder, J. R., and Hlavacek, W. S. (2015). Modeling for (physical) biologists: an introduction to the rule-based approach. *Physical biology*, 12(4):045007.

Chylek, L. A., Harris, L. A., Tung, C.-S., Faeder, J. R., Lopez, C. F., and Hlavacek, W. S. (2014). Rule-based modeling: a computational approach for studying biomolecular site dynamics in cell signaling systems. *Wiley interdisciplinary reviews. Systems biology and medicine*, 6(1):13–36.

Chylek, L. A., Stites, E. C., Posner, R. G., and Hlavacek, W. S. (2013). Innovations of the Rule-Based Modeling Approach. In Prokop, A. and Csukás, B., editors, *Systems Biology: Integrative Biology and Simulation Tools*, pages 273–300. Springer Netherlands, Dordrecht.

Ciocchetta, F. and Hillston, J. (2008). Process algebras in systems biology. In Bernardo, M., Degano, P., and Zavattaro, G., editors, *Formal Methods for Computational Systems Biology*, pages 265–312, Berlin, Heidelberg. Springer Berlin Heidelberg.

Dancheck, B., Nairn, A. C., and Peti, W. (2008). Detailed structural characterization of unbound protein phosphatase 1 inhibitors. *Biochemistry*, 47(47):12346–56.

Danos, V. (2007). *Rule-Based Modelling of Cellular Signalling*, volume 4703. Springer Berlin Heidelberg, Berlin, Heidelberg.

Danos, V., Koeppl, H., and Wilson-Kanamori, J. (2011). Cooperative assembly systems. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6937 LNCS:1–20.

Danos, V. and Laneve, C. (2004). Formal Molecular Biology. *Theoretical Computer Science*, 325:69–110.

Dematté, L., Larcher, R., Palmisano, A., Priami, C., and Romanel, A. (2010). Programming Biology in BlenX. In Choi, S., editor, *Systems Biology for Signaling Networks*, chapter 31, pages 777–820. Springer-Verlag New York, 1 edition.

Di Camillo, B., Carlon, A., Eduati, F., and Toffolo, G. M. (2016). A rule-based model of insulin signalling pathway. *BMC systems biology*, 10(1):38.

Dräger, A., Hassis, N., Supper, J., Schröder, A., and Zell, A. (2008). SBMLsqueezer: a CellDesigner plug-in to generate kinetic rate equations for biochemical networks. *BMC systems biology*, 2(1):39.

Engmann, O., Giralt, A., Gervasi, N., Marion-Poll, L., Gasmi, L., Filhol, O., Picciotto, M. R., Gilligan, D., Greengard, P., Nairn, A. C., Hervé, D., and Girault, J.-A. (2015). DARPP-32 interaction with adducin may mediate rapid environmental effects on striatal neurons. *Nature communications*, 6:10099.

Faeder, J. R., Blinov, M. L., Goldstein, B., and Hlavacek, W. S. (2005). Rule-based modeling of biochemical networks. *Complexity*, 10(4):22–41.

Faeder, J. R., Blinov, M. L., and Hlavacek, W. S. (2009). Rule-Based Modeling of Biochemical Systems with BioNetGen. In Maly, I. V., editor, *Systems Biology*, Methods in Molecular Biology, pages 113–167. Humana Press, Totowa, NJ.

Faeder, J. R., Hlavacek, W. S., Reischl, I., Blinov, M. L., Metzger, H., Redondo, A., Wofsy, C., and Goldstein, B. (2003). Investigation of early events in fcɛri-mediated signaling using a detailed mathematical model. *The Journal of Immunology*, 170.

Feret, J., Danos, V., Krivine, J., Harmer, R., and Fontana, W. (2009a). Internal coarse-graining of molecular systems. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16):6453–8.

Feret, J., Danos, V., Krivine, J., Harmer, R., and Fontana, W. (2009b). Internal coarse-graining of molecular systems (Supporting Information). *Proceedings of the National Academy of Sciences of the United States of America*, 106(16):6453–8.

Feret, J. and Krivine, J. (2012). KaSim3 reference manual.

Fernandez, É., Schiappa, R., Girault, J. A., and Le Novère, N. (2006). DARPP-32 is a robust integrator of dopamine and glutamate signals. *PLoS Computational Biology*, 2(12):1619–1633.

Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434.

Gillespie, D. T. (1977). Exact Stochastic Simulation of Coupled Chemical Reactions. *Journal of Physical Chemistry*, 81(1):2340–2361.

Girault, J. A., Hemmings, H. C. J., Williams, K. R., Nairn, A. C., and Greengard, P. (1989). Phosphorylation of DARPP-32, a Dopamine- and cAMP-regulated Phosphoprotein, by Casein Kinase II. *Journal of Biological Chemistry*, 264(36):21748–21759.

Guerriero, M. L., Priami, C., and Romanel, A. (2007). Modeling static biological compartments with beta-binders. In Anai, H., Horimoto, K., and Kutsia, T., editors, *Algebraic Biology*, pages 247–261, Berlin, Heidelberg. Springer Berlin Heidelberg.

Gutierrez-Arenas, O., Eriksson, O., and Kotaleski, J. H. (2014). Segregation and crosstalk of D1 receptor-mediated activation of ERK in striatal medium spiny neurons upon acute administration of psychostimulants. *PLoS computational biology*, 10(1):e1003445.

Hahl, S. K. and Kremling, A. (2016). A Comparison of Deterministic and Stochastic Modeling Approaches for Biochemical Reaction Systems: On Fixed Points, Means, and Modes. *Frontiers in Genetics*, 7.

Hlavacek, W. S., Faeder, J. R., Blinov, M. L., Perelson, A. S., and Goldstein, B. (2003). The complexity of complexes in signal transduction. *Biotechnology and bioengineering*, 84(7):783–94.

Hlavacek, W. S., Faeder, J. R., Blinov, M. L., Posner, R. G., Hucka, M., and Fontana, W. (2006). Rules for modeling signal-transduction systems. *Science's STKE : signal transduction knowledge environment*, 2006(344):re6.

Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500–544.

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.

Ji, Z., Yan, K., Li, W., Hu, H., and Zhu, X. (2017). Mathematical and Computational Modeling in Complex Biological Systems. *BioMed Research International*, 2017:1–16.

Jia, L., Piña-Crespo, J., and Li, Y. (2019). Restoring wnt/β-catenin signaling is a promising therapeutic strategy for alzheimer's disease. *Molecular Brain*, 12(1).

Kholodenko, B. N. (2006). Cell-signalling dynamics in time and space. *Nature reviews. Molecular cell biology*, 7(3):165–76.

Kholodenko, B. N., Demin, O. V., Moehren, G., and Hoek, J. B. (1999). Quantification of Short Term Signaling by the Epidermal Growth Factor Receptor*. *Journal of Biological Chemistry*, 274(42):30169–30181.

Kim, J., Ryu, I. S., Seo, S. Y., and Choe, E. S. (2015). Activation of Protein Kinases and Phosphatases Coupled to Glutamate Receptors Regulates the Phosphorylation State of DARPP32 at Threonine 75 After Repeated Exposure to Cocaine in the Rat Dorsal Striatum in a Ca2+-Dependent Manner. *The international journal of neuropsychopharmacology / official scientific journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)*, 18(12):992–999.

Kitano, H. (2002). Systems Biology: A Brief Overview. *Science*, 295(5560).

Klipp, E., Herwig, R., Kowald, A., Wierling, C., and Lehrach, H. (2005). *Systems Biology in Practice: Concepts, Implementation and Application*. Wiley-VCH.

Krivine, J., Danos, V., and Benecke, A. (2009). Modelling Epigenetic Information Maintenance: A Kappa Tutorial. In Bouajjani, A. and Maler, O., editors, *Computer Aided Verification*, pages 17–32, Berlin, Heidelberg. Springer-Verlag.

Kunii, Y., Hyde, T. M., Ye, T., Li, C., Kolachana, B., Dickinson, D., Weinberger, D. R., Kleinman, J. E., and Lipska, B. K. (2014). Revisiting DARPP-32 in postmortem human brain: changes in schizophrenia and bipolar disorder and genetic associations with t-DARPP-32 expression. *Molecular Psychiatry*, 19:192–199.

Le Novère, N. (2015). Quantitative and logic modelling of molecular and gene networks. *Nature Reviews Genetics*, 16(3):146–158.

Li, J. and Mansmann, U. R. (2014). A molecular signaling map and its application. *Cellular Signalling*, 26(12):2834–2842.

Lindskog, M., Kim, M., Wikström, M. A., Blackwell, K. T., and Kotaleski, J. H. (2006). Transient calcium and dopamine increase PKA activity and DARPP-32 phosphorylation. *PLoS computational biology*, 2(9):e119.

Lisman, J. (1989). A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory. *Proceedings of the National Academy of Sciences of the United States of America*, 86(23):9574–8.

Liu, B. and Thiagarajan, P. S. (2012). Modeling and analysis of biopathways dynamics. *Journal of bioinformatics and computational biology*, 10(4).

Lo, C. A., Kays, I., Emran, F., Lin, T. J., Cvetkovska, V., and Chen, B. E. (2015). Quantification of Protein Levels in Single Living Cells. *Cell Reports*, 13(11):2634–2644.

Lotka, A. J. (1920). Analytical note on certain rhythmic relations in organic systems. *Proceedings of the National Academy of Sciences*, 6(7):410–415.

Machado, D., Costa, R. S., Rocha, M., Ferreira, E. C., Tidor, B., and Rocha, I. (2011). Modeling formalisms in Systems Biology. *AMB Express, Springer Open Journal*, 1(45).

Manninen, T., Hituri, K., Toivari, E., and Linne, M.-L. (2011). Modeling signal transduction leading to synaptic plasticity: evaluation and comparison of five models. *EURASIP journal on bioinformatics & systems biology*, 2011:797250.

Marsh, J. A., Dancheck, B., Ragusa, M. J., Allaire, M., Forman-Kay, J. D., and Peti, W. (2010). Structural diversity in free and bound states of intrinsically disordered protein phosphatase 1 regulators. *Structure (London, England : 1993)*, 18(9):1094–103.

Mattioni, M. and Le Novère, N. (2013). Integration of Biochemical and Electrical Signaling-Multiscale Model of the Medium Spiny Neuron of the Striatum. *PLoS ONE*, 8(7):e66811.

Mayer, B. J., Blinov, M. L., and Loew, L. M. (2009). Molecular machines or pleiomorphic ensembles:

715   signaling complexes revisited. *Journal of biology*, 8(9):81.

716   Mollica, L., Bessa, L. M., Hanoulle, X., Jensen, M. R., Blackledge, M., and Schneider, R. (2016). Binding
717   Mechanisms of Intrinsically Disordered Proteins: Theory, Simulation, and Experiment. *Frontiers in*
718   *molecular biosciences*, 3:52.

719   Morris, M. K., Saez-Rodriguez, J., Sorger, P. K., and Lauffenburger, D. A. (2010). Logic-based models
720   for the analysis of cell signaling networks. *Biochemistry*, 49(15):3216–24.

721   Nair, A. G., Bhalla, U. S., Kotaleski, J. H., and Saucerman, J. J. (2016). Role of DARPP32 and ARPP21
722   in the Emergence of Temporal Constraints on Striatal Calcium and Dopamine Integration. *PLOS*
723   *Computational Biology PLoS Comput Biol*, 12(9).

724   Nair, A. G., Gutierrez-Arenas, O., Eriksson, O., Jauhiainen, A., Blackwell, K. T., and Kotaleski, J. H.
725   (2014). Modeling intracellular signaling underlying striatal function in health and disease. *Progress in*
726   *Molecular Biology and Translational Science*, 123:277–304.

727   Nakano, T., Doi, T., Yoshimoto, J., and Doya, K. (2010). A kinetic model of dopamine- and calcium-
728   dependent striatal synaptic plasticity. *PLoS Computational Biology*, 6(2):e1000670.

729   NCBI Resource Coordinators (2016). Database resources of the national center for biotechnology
730   information. *Nucleic Acids Research*, 44(D1):D7.

731   Nishi, A., Matamales, M., Musante, V., Valjent, E., Kuroiwa, M., Kitahara, Y., Rebholz, H., Greengard,
732   P., Girault, J.-A., and Nairn, A. C. (2017). Glutamate Counteracts Dopamine/PKA Signaling via
733   Dephosphorylation of DARPP-32 Ser-97 and Alteration of Its Cytonuclear Distribution. *The Journal*
734   *of biological chemistry*, 292(4):1462–1476.

735   Nosbisch, J. L., Bear, J. E., and Haugh, J. M. (2022). A kinetic model of phospholipase C-γ1 linking
736   structure-based insights to dynamics of enzyme autoinhibition and activation. *Journal of Biological*
737   *Chemistry*, 298(5):101886.

738   Otmakhov, N. and Lisman, J. (2012). Measuring CaMKII concentration in dendritic spines. *Journal of*
739   *neuroscience methods*, 203(1):106–14.

740   Philibin, S. D., Hernandez, A., Self, D. W., and Bibb, J. A. (2011). Striatal signal transduction and drug
741   addiction. *Frontiers in Neuroanatomy*, 5(60).

742   Regev, A., Silverman, W., and Shapiro, E. (2001). Representation and simulation of biochemical processes
743   using the pi-calculus process algebra. *Pacific Symposium on Biocomputing. Pacific Symposium on*
744   *Biocomputing*, pages 459–70.

745   Sacco, F., Perfetto, L., Castagnoli, L., and Cesareni, G. (2012). The human phosphatase interactome: An
746   intricate family portrait. *FEBS letters*, 586(17):2732–9.

747   Santibáñez, R., Garrido, D., and Martin, A. J. M. (2020). Atlas: automatic modeling of regulation of
748   bacterial gene expression and metabolism using rule-based languages. *Bioinformatics*, 36(22-23):5473–
749   5480.

750   Santibáñez, R., Garrido, D., and Martin, A. J. M. (2019). Pleione: A tool for statistical and multi-objective
751   calibration of Rule-based models. *Scientific Reports*, 9(1):15104.

752   Sekar, J. A. P. and Faeder, J. R. (2012). Rule-based modeling of signal transduction: a primer. *Methods in*
753   *molecular biology (Clifton, N.J.)*, 880:139–218.

754   Seshacharyulu, P., Ponnusamy, M. P., Haridas, D., Jain, M., Ganti, A. K., and Batra, S. K. (2012).
755   Targeting the EGFR signaling pathway in cancer therapy. *Expert opinion on therapeutic targets*,
756   16(1):15–31.

757   Sible, J. C. and Tyson, J. J. (2007). Mathematical modeling as a tool for investigating cell cycle control
758   networks. *Methods (San Diego, Calif.)*, 41(2):238–47.

759   Sneddon, M. W., Faeder, J. R., and Emonet, T. (2011). Efficient modeling, simulation and coarse-graining
760   of biological complexity with nfsim. *Nature methods*, 8:177–183.

761   Sorokina, O., Sorokin, A., and Armstrong, J. D. (2011). Towards a quantitative model of the post-synaptic
762   proteome. *Molecular bioSystems*, 7(10):2813–2823.

763   Stefan, M. I., Bartol, T. M., Sejnowski, T. J., and Kennedy, M. B. (2014). Multi-state Modeling of
764   Biomolecules. *PLoS Computational Biology*, 10(9):e1003844.

765   Stipanovich, A., Valjent, E., Matamales, M., Nishi, A., Ahn, J.-H., Maroteaux, M., Bertran-Gonzalez, J.,
766   Brami-Cherrier, K., Enslen, H., Corbillé, A.-G., Filhol, O., Nairn, A. C., Greengard, P., Hervé, D., and
767   Girault, J.-A. (2008). A phosphatase cascade by which rewarding stimuli control nucleosomal response.
768   *Nature*, 453(7197):879–884.

769   Suderman, R. and Deeds, E. J. (2013). Machines vs. Ensembles: Effective MAPK Signaling through

770   Heterogeneous Sets of Protein Complexes. *PLoS Computational Biology*, 9(10):e1003278.

771 Suderman, R. and Hlavacek, W. S. (2017). Truml: A translator for rule-based modeling languages.

772 Svenningsson, P., Nishi, A., Fisone, G., Girault, J.-A., Nairn, A. C., and Greengard, P. (2004). DARPP-32:
773   an integrator of neurotransmission. *Annual review of pharmacology and toxicology*, 44:269–96.

774 Tapia, J.-J. and Faeder, J. R. (2013). The Atomizer: Extracting Implicit Molecular Structure from Reaction
775   Network Models. *Proceedings of the International Conference on Bioinformatics, Computational
776   Biology and Biomedical Informatics (BCB'13)*, pages 726–727.

777 Tenazinha, N. and Vinga, S. (2011). A Survey on Methods for Modeling and Analyzing Integrated
778   Biological Networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 8(4):943–
779   958.

780 Vlysidis, M. and Kaznessis, Y. N. (2018). On Differences between Deterministic and Stochastic Models
781   of Chemical Reactions: Schlögl Solved with ZI-Closure. *Entropy*, 20(9).

782 Volkow, N. D. and Boyle, M. (2018). Neuroscience of addiction: Relevance to prevention and treatment.
783   *The American journal of psychiatry*.

784 Wang, H., Farhan, M., Xu, J., Lazarovici, P., Zheng, W., Wang, H., Farhan, M., Xu, J., Lazarovici, P.,
785   Zheng, W., Wang, H., Farhan, M., Xu, J., Lazarovici, P., and Zheng, W. (2015). The involvement of
786   DARPP-32 in the pathophysiology of schizophrenia. *Oncotarget*, 5(0).

787 Wilkinson, D. (2006). *Stochastic Modelling for Systems Biology*. Chapman & Hall/CRC Mathematical &
788   Computational Biology. Taylor & Francis.

789 Wilkinson, D. J. (2009). Stochastic modelling for quantitative description of heterogeneous biological
790   systems. *Nature reviews. Genetics*, 10(2):122–33.

791 Wilson-Kanamori, J., Danos, V., Thomson, T., and Honorato-Zimmer, R. (2015). Kappa rule-based
792   modeling in synthetic biology. *Methods in Molecular Biology (Clifton, N.J.)*, 1244:105–135.

793 Xie, L., Ge, X., Tan, H., Xie, L., Zhang, Y., Hart, T., Yang, X., and Bourne, P. E. (2014). Towards
794   Structural Systems Pharmacology to Study Complex Diseases and Personalized Medicine. *PLoS
795   Computational Biology*, 10(5):e1003554.

796 Yang, J. and Hlavacek, W. S. (2011). The efficiency of reactant site sampling in network-free simulation
797   of rule-based models for biochemical systems. *Physical biology*, 8(5):055009.

798 Yapo, C., Nair, A. G., Hellgren Kotaleski, J., Vincent, P., and Castro, L. R. V. (2018). Switch-like PKA
799   responses in the nucleus of striatal neurons. *Journal of Cell Science*, 131(14):jcs216556.

800 Yger, M. and Girault, J.-A. (2011). DARPP-32, Jack of All Trades... Master of Which? *Frontiers in
801   Behavioral Neuroscience*, 5(56).