Complex Functional Rates in the Modeling of Nano Devices (Extended Abstract)

Cristian Versari[1]

Lille University, Laboratoire d'Informatique Fondamentale de Lille (LIFL, CNRS UMR 8022), Cité Scientifique, Batiment M3, 59655 Villeneuve d'Ascq CEDEX, France

Gianluigi Zavattaro[2]

Dipartimento di Informatica, Università di Bologna, Mura Anteo Zamboni 7, I-40127 Bologna, Italy

Abstract

We give an overview of recent work on the rule-based modeling of nano devices. In particular, our experience in the modeling of a nanoscale elevator suggested us to enhance rule-based modeling with *complex functional rates* that can be used to express rates that depend on the current state of the entire complexes in which the reacting molecules reside.

Keywords: Rule-based modeling, nano devices, functional rates

1 The CompReNDe project

In this overview we briefly describe *CompReNDe* (Compositional and executable Representations of Nano Devices), an interdisciplinary project of the Chemistry and Computer Science departments of the University of Bologna aimed at combining the expertises of two groups, one specialized in the design and construction of devices and machines of molecular size [3,2] and the other one qualified in formal models, based on the theory of process calculi, for describing and analyzing molecular systems [14]. Such expertises have been joined together in order to deliver a programming model for describing molecular machines that is also amenable to automated simulations and verifications by means of existing algorithms.

---

[1] Email: cristian.versari@lifl.fr

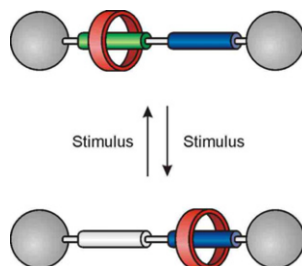[2] Email: gianluigi.zavattaro@unibo.it

Fig. 1.   Schematic representation of a two-station rotaxane and its operation as a controllable molecular shuttle.

The CompReNDe research activity started with the initial goal of formalizing a *[2]rotaxane* [29] with the Kappa-calculus [14] in order to simulate its behavior. [2]rotaxanes [29] (simply rotaxanes in the following) are systems composed of a molecular axle surrounded by a ring-type (macrocyclic) molecule. Bulky chemical moieties ("stoppers") are placed at the extremities of the axle to prevent the disassembly of the system. In rotaxanes containing two different recognition sites on the axle ("stations"), it is possible to switch the position of the macrocyclic ring between the two stations by an external energy input as illustrated in Figure 1. Several rotaxanes of this kind, known as *molecular shuttles*, have been developed (see [7] and the references therein) and used for building more complex systems [21,2].

The Kappa-calculus is a formal language idealizing molecular interactions as a particular kind of graph-rewriting. Molecules are nodes with fixed numbers of sites and molecular bonds are arcs connecting sites. Complexes are connected graphs built over such nodes and bonds. Reactions are modeled by rewriting rules that can modify the internal state of nodes, create bonds to represent *complexations*, and destroy bonds to represent *decomplexations*.

One of the distinct features of Kappa is the *"don't care, don't write"* approach: in a reaction, the reactants are not mandatorily fully described, but they can be identified by a pattern, i.e. an abstract description that can be matched by several different concrete molecules. In this way, a rule contains the description of only those parts of the complexes that are actually involved in a reaction.

The Kappa-calculus was selected as the best candidate for the modeling of the rotaxane for two main reasons: the graph-based modeling approach allowed for a natural representation of the rotaxane components and their bindings, and the rewriting rules could be used to easily represent both the chemical reaction representing the stimulus and the subsequent mechanical movement. In particular, reactions could be expressed according to the *"don't care, don't write"* approach thus focusing only on those sub-elements of the rotaxane that are actually involved.

In [9,10] we have reported about our experience in the modeling of the rotaxane by using the Kappa-calculus. One of the most interesting observations was that Kappa was not expressive enough to model complexes in which the modification of the internal state of one molecule influenced the behavior of other molecules in the same complex. To overcome this limitation, we introduced nano-K, an extension of Kappa with *instantaneous reactions* used to implement instantaneous protocols
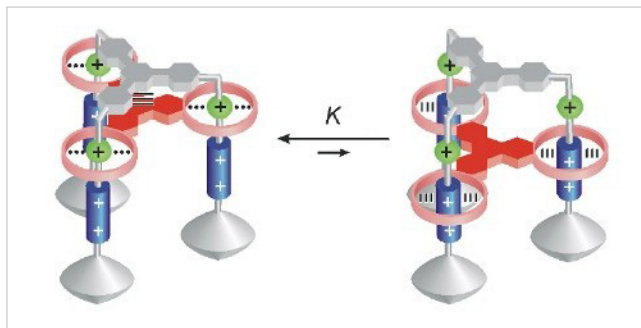
Fig. 2. Schematic representation of the nanoscale elevator in [2]. The platform (red structure) is connected to the rings of three 2-stations rotaxanes (vertical legs) which are fused together on their top parts (grey structure). The platform is moved by the stimuli that shift the rotaxane rings from one station to the other one.

used to notify to molecules the occurrence of a reaction involving other molecules in the same complex. In this way, after a reaction taking place, all the molecules of the involved complexes could possibly update their internal state accordingly. In [9,10] we also discussed how to perform simulations of the behavior of systems modeled with nano-Kappa.

After the experience with the modeling and simulation of the rotaxane, we moved to a more sophisticated nano device: the nanoscale elevator presented in [2]. This molecular machine (schematically depicted in Figure 2) is obtained by integration of three rotaxanes: their top stopper are fused together and a "platform" is connected to the three rings. In this way, the movement of the rings between the two stations has the effect of moving the platform. The modeling of such system immediately revealed challenging because laboratory experiments demonstrated that the current internal state of one rotaxane had an influence on the rates of the reactions of the other rotaxanes. The modeling of this phenomenon in nano-K required molecules with a rather rich internal state and a significant amount of instantaneous reactions for keeping such internal states updated. The resulting model was not satisfactory mainly because the description of the interdependencies among the various components of the same nano elevator was fragmented in the internal states of the molecules and the instantaneous reactions. This revealed a problem because some rates, as well as some potential interdependencies, were neither known nor observable in laboratory. For this reason we had to frequently modify the model in order to validate, by means of simulations, different assumptions about the possible interdependencies among the nano elevator components.

The problem derived from the fact that the *"don't care, don't write"* approach considered in Kappa allowed us to express only "local" properties that characterize the dynamics of the system: indeed, the kinetics of reactions must depend only on the part of reactants matched by the corresponding patterns. In this way, it is not possible to take into account properties that still influence the kinetics of the system but regard molecular complexes in their entirety. One common example is the influence of the mass on reactivity: the higher the mass of a complex, the lower the reactivity of its components, because of the slower velocity at which the

complex drifts by brownian motion.

To take into account this kind of non-local effects, in [31] we proposed *Kappa with complex functional rates* (κ_F for short), an extension of the Kappa-calculus in which the rate is a function of the complexes in which the reacting molecules actually reside. The *"don't care, don't write"* approach is still used to abstractly specify the reactants via patterns, but the rate can now depend on any property that emerges from whole complexes and the localization of the actual reacting molecules inside such complexes.

In order to be able to ease the representation of any kind of non-local effect at the level of molecular complexes, κ_F reactions are enriched with *colors*: we use a color to identify each of the reacting molecules inside their molecular complexes. Reaction rates are expressed as functions of colored complexes. More precisely, once the reacting molecules are detected in the solution, they are colored and the relevant complexes are obtained by transitive closure following their bindings. The functional rate is then applied to such colored complexes to compute the actual rate of the reaction.

From a syntactic point of view, κ_F is a slight modification of Kappa: the constant reaction rate is replaced by a functional rate. From a semantic point of view, on the contrary, the modifications are relevant (see [31] for the details). Due to this significant difference, it is not trivial to modify the Kappa simulation and analysis tools to use them on κ_F. For this reason, and in order to obtain results that could experimentally justify the introduction of the new calculus, we have studied in [31] a translation from κ_F to standard chemical reaction networks. This can be done by associating to each κ_F complex a chemical species, and then by considering for each κ_F reaction rule all its possible instantiations on those species. For all the cases where the so-called "combinatorial explosion" (arising from the many internal states and the many ways in which molecules could bind to each other) is not prohibitive, the translation to chemical reaction networks is a reasonable and easy to use way to simulate and analyze biochemical systems, as it allows the modeler to exploit all the tools already available for traditional chemistry.

In the remainder of this paper we recall the syntax of κ_F (the reader interested in the semantics and its translation to chemical reaction networks can refer to [31]) and we show how the rotaxane and the nano elevator systems can be easily modeled in κ_F. We finally conclude with some discussions about the related literature and possible lines for future work.

2   The κ_F calculus

As in the Kappa-calculus, the basic entities of κ_F models are molecules, each belonging to some particular *chemical species*. We consider a countable set of species names, ranged over by A. Species are sorted according to the number of *fields* and *sites* they possess: fields are used to denote the internal states of the molecules, while sites keep track of the formation of chemical bonds between molecules. To denote fields and sites, we introduce two functions $\mathfrak{s}_f(\cdot)$ and $\mathfrak{s}_s(\cdot)$ from A to natural

numbers; the integers 1, 2, ..., $\mathfrak{s}_f(A)$ and 1, 2, ..., $\mathfrak{s}_s(A)$ are respectively the fields and the sites of A (in particular, $\mathfrak{s}_f(A) = 0$ means there is no field, $\mathfrak{s}_s(A) = 0$ means there is no site).

The chemical bonds formed by molecules are denoted by a countable set of *bond* identifiers ranged over by x, y, ... Sites may be either *bound* to other sites or *unbound*, i.e. not connected to other sites. The state of sites are defined by injective maps, called *interfaces* and ranged over by σ, ρ, ... Given a species A, its interfaces are partial functions from $\{1, \ldots, \mathfrak{s}_s(A)\}$ to the set of bond names or a special empty value ε. A site a is bound with bond x in σ if $\sigma(a) = x$; it is unbound if $\sigma(a) = \varepsilon$. For instance, if A is a species with three sites, $(2 \mapsto x, 3 \mapsto \varepsilon)$ is one of its interfaces. In order to ease the reading, we write this map as $2^x + 3$ (the empty value is always omitted). This interface σ does not define the state of the site 1, which may be bound or not. In the following, when we write $\sigma + \sigma'$ we assume that the domains of σ and σ' are disjoint. We require interfaces to be injective in order to ensure that two sites belonging to the same molecule cannot be bound: this reflects the impossibility for single molecules to form self-complexes. In other words, we impose that the endpoints of a bond cannot belong to the same molecule.

The values of fields are defined by maps, called *evaluations*, and ranged over by u, v, ... For instance, if A is a species with three fields, $[1 \mapsto 5, 2 \mapsto 0, 3 \mapsto 4]$ is an evaluation of its fields. We write this map as $1^5 + 2^0 + 3^4$. We assume there are finitely many internal states, that is every field h is mapped to values in $\{0, \ldots, n_h\}$. In the following, we use partial evaluations and, when we write the union of evaluations $u + v$, we implicitly assume that the domains of u and v are disjoint.

A κ_F system is composed by sets of molecules interacting and changing internal state: such sets are called *solutions*. The syntax of κ_F solutions is defined by the following grammar:

$$
\begin{array}{rcl}
S & ::= & \text{Mol} \quad | \quad S, S \\
\text{Mol} & ::= & A[u](\sigma)
\end{array}
$$

with "," associative. According to the previously introduced notation, $A[u](\sigma)$ denotes a molecule of species A, with evaluation u and interface σ.

The dynamic evolution of a κ_F system over time is governed by a set of chemical rules which define the interaction and state change capabilities of the molecules. Each rule is associated with a rate which roughly denotes its "speed" or, if we follow the stochastic interpretation of κ_F models, the probability of its application to the current chemical solution. Rates are not necessarily fixed, but can depend on the internal state and chemical bonds of both the interacting molecules and the chemical complexes where these molecules are placed.

The expression of functional rule rates in κ_F relies on the use of *colors*, which allow the modeler to specify which, between those in the chemical solution, are the reacting molecules. Colors are formalized as vectors of identifiers associated to solutions: in order to denote them, we introduce a denumerable set of color identifiers $\mathcal{C}$, with $\epsilon \in \mathcal{C}$ denoting the empty color. A color is a tuple $\tilde{c} = (c_1, \ldots, c_n)$

of color identifiers $c_i \in \mathcal{C}$, such that an identifier different from ϵ can appear only once, namely, if $c_i \neq \epsilon$ then $c_i \neq c_j$ $\forall i \neq j$, with $1 \leq i, j \leq n$. A *colored solution* $S^{\tilde{c}}$ is a pair $(S, \tilde{c})$ where S is a solution $S = \text{MOL}_1, \dots, \text{MOL}_n$ and $\tilde{c} = (c_1, \dots, c_n)$ is a color.

The chemical reactions that specify the evolution over time of a κ_F systems are formalized as triples $(M_1^{\tilde{c}_1}, f_\rho, M_2^{\tilde{c}_2})$ usually written as $\rho = M_1^{\tilde{c}_1} \xrightarrow{f_\rho} M_2^{\tilde{c}_2}$ where $M_1^{\tilde{c}_1}$ and $M_2^{\tilde{c}_2}$ are patterns specifying the possible *reactants* and the corresponding *products*, respectively, and f_ρ is a function from colored solutions to non-negative numbers (representing actual rates).

Notice that the definition of reaction is essentially the same as in Kappa, with the unique difference that reactants and products are colored and a functional rate is considered. The functional rate f_ρ is responsible for checking the complexes in which the reacting molecules reside, and according to their structure and the localization of the reactants (identified by the colors), a corresponding rate is computed.

3   Modeling the rotaxane and the nano elevator in κ_F

The rotaxanes used in [2] to build the molecular elevator have two stations, an ammonium/amine molecule (*Nh* in the following) green colored in Figure 1 and 2, and a bipyridinium molecule (*Bipy* in the following) colored in blue. The *Nh* molecule can be protonated and deprotonated by adding acid or base to the solution: when it is protonated the stable position for the ring is on the *Nh* station (as depicted in Figure 2), while it is on the *Bipy* station when it is deprotonated.

The behavior of such rotaxane has been modeled in [10] by using a Kappa-like language extended with instantaneous reactions. The latter were used to immediately communicate to all the molecules belonging to the same rotaxane the occurred (de)protonation of the *Nh*. This is no longer needed in κ_F as functional rates allow the modeler to express the influence of the internal state of the *Nh* molecule on the behavior of the entire rotaxane.

The rotaxane can be modeled by considering three distinct molecules for representing the *Nh*, the *Bipy* and the ring, respectively. The two stations are connected by a permanent bond, while the ring has a switchable binding to one of them (such bond indicates the current location of the ring). The chemical solution corresponding to such a rotaxane with the ring on the *Nh* station can be written as follows:
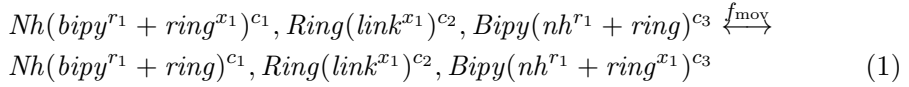
$$Nh[h^1](bipy^{r_1} + ring^{x_1}), Ring(link^{x_1}), Bipy(nh^{r_1} + ring)$$

where we use mnemonic names to represent sites and fields. The *Nh* molecules have one field h: the field holds 0 when the *Nh* is deprotonated, it holds 1 otherwise. The switch of the ring from the *Nh* to the *Bipy* station leads instead to the following configuration:

$$Nh[h^1](bipy^{r_1} + ring), Ring(link^{x_1}), Bipy(nh^{r_1} + ring^{x_1})$$

Such a switch between these two stable configurations can be easily expressed with

the following reversible reaction, formally corresponding to two κ_F reactions:

$$Nh(bipy^{r_1} + ring^{x_1})^{c_1}, Ring(link^{x_1})^{c_2}, Bipy(nh^{r_1} + ring)^{c_3} \xleftrightarrow{f_{\text{mov}}}$$
$$Nh(bipy^{r_1} + ring)^{c_1}, Ring(link^{x_1})^{c_2}, Bipy(nh^{r_1} + ring^{x_1})^{c_3} \qquad (1)$$

with

$$f_{\text{mov}}(S^{\tilde{c}}) \;=\; \text{let } Nh[h^x](ring^y + \sigma)^{c_1} \in S^{\tilde{c}} \qquad \text{in}$$
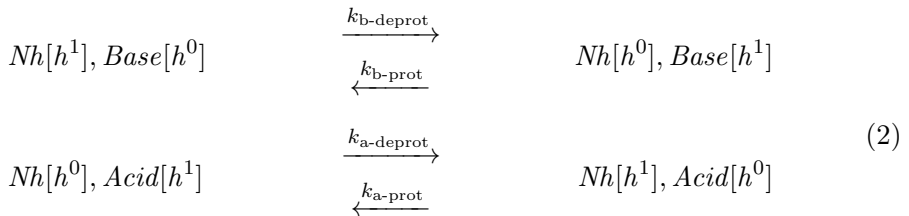
$$\text{if } y = \varepsilon \text{ then}$$

$$\text{if } x = 0 \text{ then } \lambda^1_{\text{mov}} \text{ else } \lambda^2_{\text{mov}}$$

$$\text{else}$$

$$\text{if } x = 0 \text{ then } \lambda^3_{\text{mov}} \text{ else } \lambda^4_{\text{mov}}$$

The rate of the ring movement from one station to the other one depends on the protonated/deprotonated state of the *Nh*. We model this dependency by using a functional rate, returning one of four possible rates λ^i_{mov}, depending on the combination of two distinct factors: whether the *Nh* is protonated or not, and whether the *Ring* is moving from the *Nh* to the *Bipy* or vice versa. Thanks to κ_F functional rate, we obtain for the rotaxane a very succint model with respect to other calculi such as Kappa.

The modeling of protonation/deprotonation process in the rotaxane does not need instead functional rates:

$$Nh[h^1], Base[h^0] \qquad \xrightleftharpoons[k_{\text{b-prot}}]{k_{\text{b-deprot}}} \qquad Nh[h^0], Base[h^1]$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2)$$

$$Nh[h^0], Acid[h^1] \qquad \xrightleftharpoons[k_{\text{a-prot}}]{k_{\text{a-deprot}}} \qquad Nh[h^1], Acid[h^0]$$

for some kinetic constants $k_{\text{b-deprot}}, k_{\text{b-prot}}, k_{\text{a-deprot}}, k_{\text{a-prot}}$, We use *Base* and *Acid* molecules with a field h which holds 0 or 1 to denote whether the molecule is ready to receive or donate a proton, respectively.

The model of the nanoscale elevator is obtained by joining together the models of its three rotaxanes by means of an additional complex, composed of three *Top* molecules. According to the structure depicted in Figure 2 we add to each of the three rotaxanes a *Top* molecule connected to the *Nh* station, and we bind together the three *Top* molecules. Moreover, the three rings are connected together to represent the platform. The complete representation of the elevator is then as
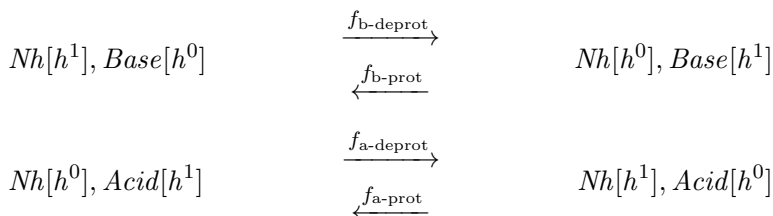
follows:

$$Top(l^{t_1} + nh^{s_1} + r^{t_2}), Top(l^{t_2} + nh^{s_2} + r^{t_3}), Top(l^{t_3} + nh^{s_3} + r^{t_1}),$$
$$Nh[h^1](top^{s_1} + bipy^{r_1} + ring^{x_1}), Ring(l^{p_1} + link^{x_1} + r^{p_2}), Bipy(nh^{r_1} + ring),$$
$$Nh[h^1](top^{s_2} + bipy^{r_2} + ring^{x_2}), Ring(l^{p_2} + link^{x_2} + r^{p_3}), Bipy(nh^{r_2} + ring),$$
$$Nh[h^1](top^{s_3} + bipy^{r_3} + ring^{x_3}), Ring(l^{p_3} + link^{x_3} + r^{p_1}), Bipy(nh^{r_3} + ring)$$

In the first line we present the three *Top* molecules each one connected to a left and a right *Top* molecule. Moreover, each *Top* is connected to the *Nh* molecule of one rotaxane. The three rotaxanes are represented in the subsequent three lines. Notice that we assume that the *Ring* molecules are connected to the *Nh* station, and that each *Ring* is connected to a left and a right *Ring* molecule. We assume the *Nh* molecules initially protonated.

We now move to the representation of the dynamics of the system. As for the rotaxane, two kinds of reactions are used: those for protonation/deprotonation between the *Nh* and an acid-base molecule, and those for switching the bond between the ring and the two stations.

Thanks to the *"don't care, don't write"* approach inherited from Kappa, the modularity of the structure of the nano elevator allows us to reuse the chemical reactions of Eq.(1) without any change for the movement of the platform. Despite also the second group of reactions, i.e. those concerned with the proton exchange between the *Nh* and the acid-base molecules, could be reused as well, we considered a modified version. The rate of these reactions is indeed influenced by an interesting phenomenon observed on the behavior of the nanoscale elevator. The (de)protonation of the three *Nh* molecules of an elevator follows three distinct processes. Upon addition of acid-base to the solution, the (de)protonation effect does not distribute homogeneously among the *Nh* molecules, but among the elevators. Namely, the *"first equivalent of base does not lead to a statistical mixture of differently protonated species but rather causes the first deprotonation process to occur"*. One likely cause of this phenomenon is that the (de)protonation rate of the *Nh* is influenced by the current (de)protonated state of the other two *Nh* molecules in the same elevator. According to this interpretation, the *dependence* of the protonation/deprotonation process on the total number of protonated *Nh* molecules of the elevator requires the following modification to the reactions of Eq.(2) which were used for the single rotaxane:

$$Nh[h^1], Base[h^0] \quad \underset{f_{\text{b-prot}}}{\overset{f_{\text{b-deprot}}}{\rightleftharpoons}} \quad Nh[h^0], Base[h^1]$$

$$Nh[h^0], Acid[h^1] \quad \underset{f_{\text{a-prot}}}{\overset{f_{\text{a-deprot}}}{\rightleftharpoons}} \quad Nh[h^1], Acid[h^0]$$
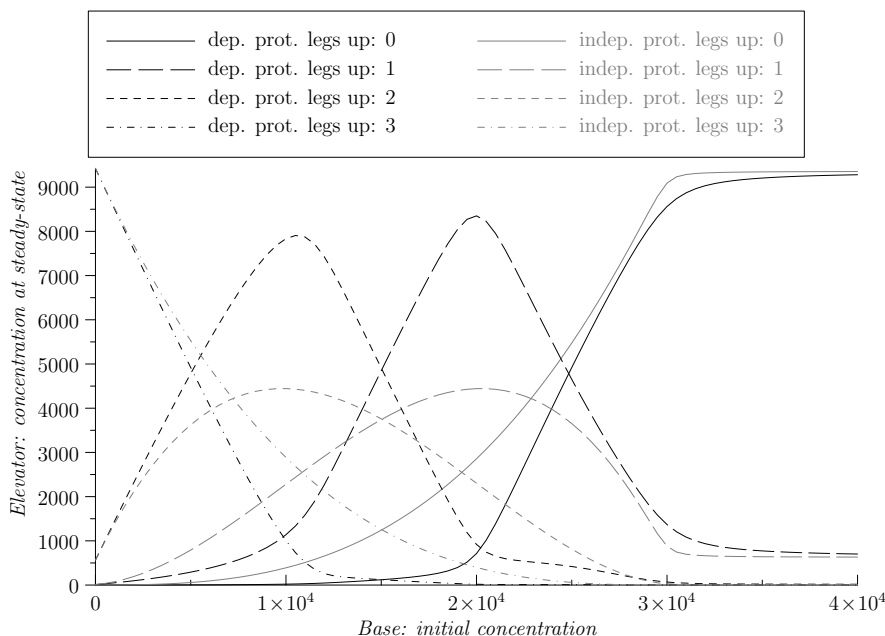
Fig. 3. Comparison of possible behaviors of the elevator model at steady state in relation to the assumption of dependence or independence of the protonation/deprotonation process on the total number of already protonated *Nh* molecules. On the horizontal axis the initial number of $Base[h^0]$ molecules in the solution: each point of the graph represents the result of a run with different initial number of base molecules. On the vertical axis the corresponding number of molecules at steady state for different configurations and functional rates of the elevator, starting with an initial concentration of 10^4 elevators each with all the 3 *Nh* molecules protonated. The four species listed on the left of the legend represent the number of elevators with respectively 0 to 3 "legs up" (i.e. number of rotaxanes whose ring is bound to the *Nh* molecule) in the case of movement dependent on the number of protonated *Nh* molecules as described in Eq.(3), with $\lambda_{\text{b-deprot}} = s_{\text{b-deprot}} = 1, \lambda_{\text{b-prot}} = 10^{-2}, s_{\text{b-prot}} = -1$. The four species on the right represent the same elevator states but in the case of movement independent of the total number of protonated *Nh* molecules, that is with $s_{\text{b-deprot}} = s_{\text{b-prot}} = 0$ and $\lambda_{\text{b-deprot}} = 10^2, \lambda_{\text{b-prot}} = 10^{-2}$. For both the dependent and the independent cases, we have $\lambda_{\text{mov}}^1 = \lambda_{\text{mov}}^4 = 20, \lambda_{\text{mov}}^2 = \lambda_{\text{mov}}^3 = 10^3$.

with

$$f_{\text{k}}(S^{\tilde{c}}) = \text{let } P = \sum\nolimits_{Nh[h^x](\sigma) \in S^{\tilde{c}}} x \quad \text{in} \quad (\lambda_{\text{k}}) \cdot 10^{s_{\text{k}} \cdot P} \qquad (3)$$

for $\text{k} \in \{\text{b-deprot}, \text{b-prot}, \text{a-deprot}, \text{a-prot}\}$. In this case, the functional rate modifies a base rate λ_{k} according to the number P of *Nh* protonated in the same elevator.

Thanks to the translation of κ_F to standard chemical reaction networks given in [31] we were able to simulate numerically the above formalization of the molecular elevator. Simulation results are shown in Figure 3 at steady state under the hypotheses of dependence and independence of the protonation/deprotonation process on the total number of protonated *Nh* molecules.

4 Related and future work

The *"don't care, don't write"* approach adopted in the Kappa-calculus, as well as in other rule-based languages like BioNetGen [4], opened the way for introducing compositional modeling in rule-based process calculi, and provides very compact and readable descriptions of biochemical systems in the presence of sophisticated molecule bindings. While compositional modeling represents in general a desirable

advantage in the hands of the modeler, it becomes a limit when important properties of the system cannot be described in a compositional calculus because of their intrinsic non-compositionality.

Based on our experience in the modeling of nano devices within the project CompReNDe, we have introduced the notion of *complex functional rates* and applied it to the Kappa-calculus, thus obtaining an extended calculus that we called κ_F. It is worth noting that complex functional rates and the corresponding techniques that we have developed for κ_F are of general applicability, and can be adapted to extend also other process calculi for biochemical modeling in order to take into account non-compositional properties (physical, chemical, etc.) without losing the advantage of a compositional description. Their applicability could be also extended to other process calculi with binding capabilities like, e.g., [18,26,27,28].

The simulator for the Kappa language KaSim [1] already includes the possibility to associate to a reaction a pair of rates, the first one to be used when the reactants are freely floating and the second one for the case in which they are part of the same complex. This latter mechanism is useful to resolve the ambiguity of Kappa rules among two reactants A and B that could be applied in a context where A and B are sometimes already connected and sometimes disconnected. Indeed, this would lead to an inconsistency in the definition of the kinetic rate which should have a volume dependency in the former case and no volume dependency in the latter. Nevertheless, the introduction of physical or chemical properties influencing reaction rates as functions of the whole involved complexes makes unfeasible any attempt of modeling in Kappa.

It is also worth to mention the variant of Kappa presented in [11], where the so-called *rule refinement* approach is presented. The idea of rule refinement is to replace a rule with a set of rules, each one strengthening the conditions under which the initial rule can be applied. Our approach is different because we do not add rules to a Kappa model, but we simply allow for the definition of the reaction rate as a function of the complexes in which the reactants actually reside. We consider our approach more appropriate for the modeling of systems in which some specific physico-chemical properties of the complexes hosting the reactants have an impact on the system kinetics. On the contrary, the rule refinement approach revealed appropriate to study the distribution of already known rates from a Kappa rule to its refinements. In fact, one of the main contribution of [11] is the definition of a mechanism for inferring the rates of the refined rules in such a way that the kinetics of the initial non refined system is preserved.

The stochastic simulator NFsim [30], based on an extension of the BioNetGen language, allows the expression of rate functions which can depend on properties either global (at the level of the system) or "local" (at the level of the molecular complexes involved in the reaction). While the first kind of properties is not directly included in κ_F and should be encoded manually by the modeler, the latter kind makes NFsim capabilities closer to κ_F. However, the adoption of *colors* in κ_F semantics allows the modeler to take into account more sophisticated properties which depend not only on the number of molecules of any kind that form each

molecular complex, but also on the way they are arranged to form the complex. In other words, only in κ_F rate functions can exploit the information pertaining the graph-like structure of each complex involved in the reaction and the position of reacting molecules inside them.

Despite the expressiveness of κ_F, in [31] we provided its formal translation in traditional chemistry and proved the correctness. Thanks to this translation it is possible to apply to (some classes of) κ_F models the efficient verification techniques (such as simulation by ordinary differential equations, as well as by efficient stochastic algorithms [19,5]) and reuse, at least in principle, the existing software tools developed for traditional chemistry (e.g [8], but in general any tool supporting languages comparable to traditional chemistry, like SBML [20]).

However, the superior expressiveness of process calculi like Kappa with respect to traditional chemistry can lead to translations with an infinite number of chemical species and rules. In general, as the Kappa-calculus is Turing complete, the problem of checking whether a Kappa model can generate only finitely many complexes is undecidable. Nevertheless, there are fragments of Kappa for which this problem turns out to be decidable [17], and also techniques —based on abstract interpretation— which are capable of computing an over-approximation of the set of reachable complexes [13] that can be used to prove, in some cases, that this set is finite.

Future work directions point at several aims. First, it must be investigated to what extent the introduction of functional rates in bio-oriented process calculi semantics can be pushed, in particular in those calculi equipped with high-level structural rearrangement primitives (for example, calculi with nested compartments [6,28]). Moreover, it is still unclear how this approach can be adapted to cope with more sophisticated calculi like [23,22], where functional rates have been already introduced but with a different technique, based on communication constraints. A special case is the React(C) language [24], that can be regarded as an extension of Kappa as well as of κ_F, but differently from κ_F the functional rate takes under consideration the entire system (i.e. the entire solution). We found the κ_F approach more appropriate when one wants to specify models which are modular at least at the level of complexes, even if not modular at the level of the single molecules. It must be investigated how and to what extent the technique used for the translation of κ_F to traditional chemistry can be generalized in order to be applied to React(C) as well, thus giving the possibility to introduce also in React(C) a notion of complex. More sophisticated mappings to traditional chemistry (e.g. along the line of [25]) may be helpful in this regard.

Last, more efficient translation techniques may allow us to widen the class of models manageable in practice. For example, abstract interpretation has been already applied to Kappa [12] to reduce, under some circumstances, the number of chemical species and reactions resulting from a translation from Kappa to standard chemistry similar to the one we have presented in this paper. In order to apply such techniques in κ_F, it is necessary to check their applicability in the presence of the functional rate.

Acknowledgement

We are grateful to all our collaborators in the CompReNDe project: Alberto Credi, Marco Garavelli, Cosimo Laneve, Sylvain Pradalier and Serena Silvi.

References

[1] KaSim: kappa language simulator. http://www.kappalanguage.org.

[2] J. D. Badjić, V. Balzani, A. Credi, S. Silvi, and J. F. Stoddart. A molecular elevator. *Science*, 303(5665):1845–1849, 2004.

[3] V. Balzani, A. Credi, and M. Venturi. *Molecular devices and machines - Concepts and perspectives for the nano world, 2nd Edition*. Wiley-VCH, 2008.

[4] M. Blinov, J. Faeder, B. Goldstein, and W. Hlavacek. Bionetgen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics*, 20(17):3289, 2004.

[5] Y. Cao, H. Li, and L. Petzold. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J Chem Phys*, 121(9):4059–4067, September 2004.

[6] L. Cardelli. Brane Calculi. In Danos and Schächter [15], pages 257–278.

[7] B. Champin, P. Mobian, and J.-P. Sauvage. Transition metal complexes as molecular machine prototypes. *Chemical Society Reviews*, 36(2):358–366, 2006.

[8] F. Ciocchetta, A. Duguid, S. Gilmore, M. L. Guerriero, and J. Hillston. The Bio-PEPA Tool Suite. In *QEST*, pages 309–310. IEEE Computer Society, 2009.

[9] A. Credi, M. Garavelli, C. Laneve, S. Pradalier, S. Silvi, and G. Zavattaro. Modelizations and simulations of nano devices in nanok calculus. In *CMSB 2007*, volume 4695 of *LNCS*, pages 168–183. Springer, 2007.

[10] A. Credi, M. Garavelli, C. Laneve, S. Pradalier, S. Silvi, and G. Zavattaro. nanok: A calculus for the modeling and simulation of nano devices. *Theor. Comput. Sci.*, 408(1):17–30, 2008.

[11] V. Danos, J. Feret, W. Fontana, R. Harmer, and J. Krivine. Rule-based modelling, symmetries, refinements. In *FMSB*, volume 5054 of *Lecture Notes in Computer Science*, pages 103–122. Springer, 2008.

[12] V. Danos, J. Feret, W. Fontana, R. Harmer, and J. Krivine. Abstracting the differential semantics of rule-based models: Exact and automated model reduction. In *LICS*, pages 362–381. IEEE Computer Society, 2010.

[13] V. Danos, J. Feret, W. Fontana, and J. Krivine. Abstract interpretation of cellular signalling networks. In *VMCAI*, volume 4905 of *Lecture Notes in Computer Science*, pages 83–97. Springer, 2008.

[14] V. Danos and C. Laneve. Formal molecular biology. *Theor. Comput. Sci.*, 325(1):69–110, 2004.

[15] V. Danos and V. Schächter, editors. *Computational Methods in Systems Biology, International Conference CMSB 2004, Paris, France, May 26-28, 2004, Revised Selected Papers*, volume 3082 of *Lecture Notes in Computer Science*. Springer, 2005.

[16] P. Degano and R. Gorrieri, editors. *Computational Methods in Systems Biology, 7th International Conference, CMSB 2009, Bologna, Italy, August 31-September 1, 2009. Proceedings*, volume 5688 of *Lecture Notes in Computer Science*. Springer, 2009.

[17] G. Delzanno, C. D. Giusto, M. Gabbrielli, C. Laneve, and G. Zavattaro. The *kappa*-lattice: Decidability boundaries for qualitative analysis in biological languages. In Degano and Gorrieri [16], pages 158–172.

[18] F. Fages and S. Soliman. Formal Cell Biology in Biocham. In *SFM*, volume 5016 of *Lecture Notes in Computer Science*, pages 54–80. Springer, 2008.

[19] M. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *Journal of Physical Chemistry A*, 104(9):1876–1889, 2000.

[20] M. Hucka, A. Finney, H. Sauro, H. Bolouri, J. Doyle, H. Kitano, A. Arkin, B. Bornstein, D. Bray, A. Cornish-Bowden, et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, 2003.

[21] M. Jimenez, C. Dietrich-Buchecker, and J.-P. Sauvage. Towards synthetic molecular muscles: Contraction and stretching of a linear rotaxane dimer. *Angew. Chem. Int. Ed.*, 39(18):3284–3287, 2000.

[22] M. John, C. Lhoussaine, and J. Niehren. Dynamic compartments in the imperative *pi*-calculus. In Degano and Gorrieri [16], pages 235–250.

[23] M. John, C. Lhoussaine, J. Niehren, and A. M. Uhrmacher. The attributed pi calculus. In M. Heiner and A. M. Uhrmacher, editors, *CMSB*, volume 5307 of *Lecture Notes in Computer Science*, pages 83–102. Springer, 2008.

[24] M. John, C. Lhoussaine, J. Niehren, and C. Versari. Biochemical reaction rules with constraints. In G. Barthe, editor, *ESOP*, volume 6602 of *Lecture Notes in Computer Science*, pages 338–357. Springer, 2011.

[25] R. Meyer and R. Gorrieri. On the Relationship between π-Calculus and Finite Place/Transition Petri Nets. In M. Bravetti and G. Zavattaro, editors, *CONCUR*, volume 5710 of *Lecture Notes in Computer Science*, pages 463–480. Springer, 2009.

[26] A. Phillips and L. Cardelli. Efficient, correct simulation of biological processes in the stochastic pi-calculus. In *CMSB*, volume 4695 of *LNCS*, pages 184–199, 2007.

[27] C. Priami and P. Quaglia. Beta binders for biological interactions. In Danos and Schächter [15], pages 20–33.

[28] A. Regev, E. M. Panina, W. Silverman, L. Cardelli, and E. Y. Shapiro. BioAmbients: an abstraction for biological compartments. *Theor. Comput. Sci.*, 325(1):141–167, 2004.

[29] J.-P. Sauvage and C. D.-B. (eds.). *Molecular catenanes, rotaxanes and knots.* Wiley-VCH, 1999.

[30] M. W. Sneddon, J. R. Faeder, and T. Emonet. Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. *Nature Methods*, 8:177183, 2011.

[31] C. Versari and G. Zavattaro. Complex Functional Rates in Rule-based Languages for Biochemistry. *Transactions on Computational Systems Biology*. To appear.