

nano κ : a Calculus for Modelization and Simulation of Nano Devices^{*}

A. Credi¹ M. Garavelli¹ C. Laneve²
S. Pradalier³ S. Silvi¹ G. Zavattaro²

¹*Dipartimento di Chimica “G. Ciamician”, Università di Bologna*

²*Dipartimento di Scienze dell’Informazione, Università di Bologna*

³*Ecole Polytechnique, Paris*

Abstract

We develop a process calculus – the **nano κ** calculus – for modeling, analyzing and predicting the properties of molecular devices. The **nano κ** calculus is equipped with a simple stochastic model, that we use to model and simulate the behavior of a molecular shuttle, a basic nano device currently used for building more complex systems.

1 Introduction

In 2006 the University of Bologna funded an interdisciplinary project of its Departments of Chemistry and Computer Science – the *CompReNDe Project* (Compositional and executable Representations of Nano Devices). The project combines 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 [12,21]. Such expertises are joined together in order to accomplish three main endeavors: (i) deliver a programming model for describing molecular machines that is also amenable to automated simulations and verifications by means of existing algorithms, (ii) apply the model for a formal analysis of real cases of molecular machines to possibly reveal complex behaviors that have not been experimentally observed yet, and (iii)

^{*} Full version of [10].

Research partially supported by the University of Bologna Strategic Project *CompReNDe*: Compositional and Executable Representations of Nano Devices.

use the simulations as tools to assist chemists in the design of novel nano devices.

The CompReNDe research activity started with the initial goal of formalizing a [2]rotaxane [28] into the κ calculus [12] in order to simulate its behavior *in silico* by means of some contemporary stochastic evaluator [15,27,8]. [2]rotaxanes [28] (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*

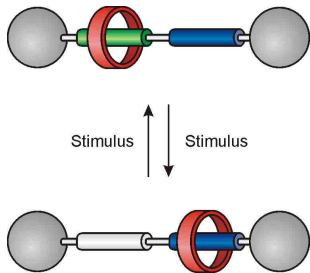


Fig. 1. Schematic representation of a two-station rotaxane and its operation as a controllable molecular shuttle.

shuttles, have been already developed (see [9] and the references therein) and used for building more complex systems [19,18,2].

The κ calculus is a formal language idealizing protein-protein interactions, as a particular restricted kind of graph-rewriting. Bindings are explicit: proteins are nodes with fixed numbers of sites, complexes are connected graphs built over such nodes where bonds are represented by *names*. Biological reactions are modeled by two kinds of rewriting rules: *complexations*, which create bonds, and *decomplexations*, which destroy bonds. Notably, the κ calculus has been compiled into π -calculus [23] introducing a finer-grained concurrent model, the $\mathbf{m}\kappa$ calculus, where reactions have to be at most binary. The significant property of $\mathbf{m}\kappa$ calculus is to be *protein-centred*, rather than reaction-centered as it was the case for κ calculus, thus being amenable to distributed implementations.

We therefore undertook the formalization of a molecular shuttle in $\mathbf{m}\kappa$ calculus and we soon realized that such calculus was inadequate as well. The $\mathbf{m}\kappa$ calculus is much too verbose because it compels designers to reason in terms of bonds and complexations and decomplexations. There are reactions that are neither complexations nor decomplexations, such as the *ion exchanges*. These reactions, used in our molecular shuttle to stimulate the movement of the macrocyclic ring, might be implemented by sequences of complexations

and decomplexations, thus changing the granularity of the chemical semantics. The $\mathfrak{m}\kappa$ calculus model is much too abstract because it overlooks quantitative aspects. Such aspects, in particular reaction rates and the derived stochastic semantics are a must for providing meaningful simulations of molecular machines.

We overcome these inadequacies of the $\mathfrak{m}\kappa$ calculus by defining a new model, the $\mathfrak{nan}\mathfrak{o}\kappa$ calculus, having three types of reactions – *creations*, *destructions*, and *exchanges* – and retaining a stochastic semantics. This stochastic semantics is problematic for the $\mathfrak{nan}\mathfrak{o}\kappa$ calculus because it uses names for representing molecular bonds. In this respect, our model is close to Milner’s π -calculus [23]. However, instead of following the techniques of the stochastic π -calculus [25], we have preferred for $\mathfrak{nan}\mathfrak{o}\kappa$ calculus to extend Cardelli’s language of stochastic interacting processes [7]. In fact, in this way, we get a simple model that may be easily simulated or verified by means of existing well known algorithms [14].

We then apply the $\mathfrak{nan}\mathfrak{o}\kappa$ calculus to describe and analyze an instance of rotaxane, RaH [22,1], for which the dynamic behavior has been experimentally characterized in detail [13]. We have considered two groups of simulations. The first ones are used to validate the model, checking whether the experiments reproduced *in silico* coincides with those already performed *in vitro*. The second ones simulate *in silico* the expected behavior of the rotaxane RaH under conditions not yet observed *in vitro*. Interestingly, we show that under extreme conditions of very low concentration of rotaxane RaH, some of the assumptions, usually taken about the behavior of the rotaxane in standard conditions of concentration, are no longer valid.

Structure of the paper. The next section introduces the $\mathfrak{nan}\mathfrak{o}\kappa$ calculus syntax and semantics. In Section 3 we relate the $\mathfrak{nan}\mathfrak{o}\kappa$ calculus semantics with Interactive Markov Chains and Continuous Time Markov Chains. In Section 4 we present our case study on rotaxane RaH and its modelling into $\mathfrak{nan}\mathfrak{o}\kappa$ calculus, and discuss the results of some simulations. In Section 5 we discuss the related literature and we introduce possible directions for future work.

2 The $\mathfrak{nan}\mathfrak{o}\kappa$ calculus: syntax and semantics

Two disjoint countable sets of names will be used: a set of *species*, ranged over by A, B, C, \dots ; and a totally ordered set of *bonds*, ranged over by x, y, z, \dots . Species are sorted according to the number of *fields* and *sites* they possess. Let $\mathfrak{s}_f(\cdot)$ and $\mathfrak{s}_s(\cdot)$ be two functions returning naturals; the integers

$1, 2, \dots, \mathfrak{s}_f(A)$ and $1, 2, \dots, \mathfrak{s}_s(A)$ are respectively the fields and the sites of A . ($\mathfrak{s}_f(A) = 0$ means there is no field; $\mathfrak{s}_s(A) = 0$ means there is no site). In the following, fields are ranged over by h, i, j, \dots ; sites are ranged over by a, b, c, \dots .

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, \dots . Given a species A , its interfaces are partial functions from $\{1, \dots, \mathfrak{s}_s(A)\}$ to the set of bonds 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. Interfaces, being injective on bonds, cannot express that the endpoints of a bond belong to the same species (cf. *self complexation* in [12]). This design choice simplifies the presentation of $\mathfrak{n}\mathfrak{a}\mathfrak{n}\mathfrak{o}\mathfrak{k}$ calculus.

Fields represent the internal state of a species. The values of fields are defined by maps, called *evaluations*, and ranged over by u, v, \dots . For instance, if A is a species with three fields, $[1 \mapsto 5, 2 \mapsto 0, 3 \mapsto 4]$ is an evaluation of its. As before, 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 into values in $\{0, \dots, n_h\}$. 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.

Definition 1 (Molecules and Solutions) *A molecule $A[u](\sigma)$ is a term where u is a total map on the fields of A . Solutions, ranged over by S, T, \dots , are defined by the following grammar*

$$S ::= A[u](\sigma) \mid S, S$$

The operator “,” is assumed to be associative, so $(S, S'), S''$ is equal to $S, (S', S'')$ (and we always omit parentheses).

Solutions retain the property that bond names always occurs exactly twice. Let \emptyset be the empty map. We use the following shorthand notations: $A(\sigma)$ instead of $A[\emptyset](\sigma)$, $A[u]$ instead of $A[u](\emptyset)$, and simply A instead of $A\emptyset$.

Example 1 *As a running example we consider two typical chemical reactions:*

- $Na + Cl \longleftrightarrow Na^+ + Cl^-$ (sodium chloride) and
- $H + H \longleftrightarrow H_2$ (hydrogen gas) .

In the first reaction, an ion is exchanged between two instances of species Na and Cl. The molecules of the two species can be in two possible states: either they have the extra ion Na^+ and miss an ion Cl^- or they are in the states with all the ions Na and Cl. We model these two possible states using one field ion with values 0 and 1 respectively denoting the absence or the presence of the ion. Formally we can use $\text{Na}[\text{ion}^0]$ and $\text{Na}[\text{ion}^1]$ for Na and Na^+ , and $\text{Cl}[\text{ion}^0]$ and $\text{Cl}[\text{ion}^1]$ for Cl^- and Cl, respectively.

The second chemical reaction represents the creation/destruction of a bond between two hydrogen atoms. This may be described by using a site 1 and bond names. For instance, the solution with H_2 is modelled by $H(1^x), H(1^x)$. An unbound instance of hydrogen is simply represented by H , as its evaluation and interface are both empty.)

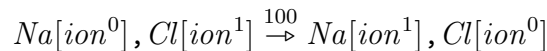
Definition 2 (Reactions) Reactions of nanok calculus are either creations C, or destructions D, or exchanges E. The format of the first two types is $((A, a, u, u', \sigma), (B, b, v, v', \phi), \lambda)$; while the format of exchanges is $((A, u, u', \rho, \rho'), (B, v, v', \psi, \psi'), \lambda)$, such that:

- (1) $\text{dom}(u') = \text{dom}(u)$ and u and u' are partial evaluations of A , $\text{dom}(v') = \text{dom}(v)$ and v and v' are partial evaluations of B ,
- (2) $\text{ran}(\sigma) = \text{ran}(\phi)$ and σ and ϕ are interfaces of A and B , respectively, such that $a \notin \text{dom}(\sigma)$ and $b \notin \text{dom}(\phi)$;
- (3) (for exchanges) ρ, ρ' and ψ, ψ' are interfaces of A and B , respectively, with $\text{ran}(\rho') = \text{ran}(\psi)$ and either $\rho = \rho'$ and $\psi = \psi'$ or $\rho = a^x + \rho''$, $\rho' = a + \rho''$ and $\psi = b + \psi''$, $\psi' = b^x + \psi''$;
- (4) and $\lambda \in \mathbb{R}^+ \cup \{\infty\}$.

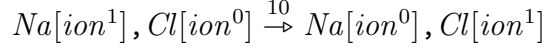
For readability's sake, we write creations as $A[u](a+\sigma), B[v](b+\phi) \xrightarrow{\lambda} A[u'](a^x + \sigma), B[v'](b^x + \phi)$, destructions as $A[u](a^x + \sigma), B[v](b^x + \phi) \xrightarrow{\lambda} A[u'](a + \sigma), B[v'](b + \phi)$, and exchanges as $A[u](\rho), B[v](\psi) \xrightarrow{\lambda} A[u'](\rho'), B[v'](\psi')$.

The difference between the three kinds of rules is concerned with the modification of the interfaces: *creations* produce a new bond between the two unbound sites a and b , *destructions* remove the bond between the sites a and b , while *exchanges* either leave the interfaces unchanged or move one bond from a reactant to the other (bond-flipping exchange).¹

Example 2 The nanok calculus reactions that corresponds to the two reactions of the sodium chloride are



¹ The terms creation and destruction have been preferred to *complexation* and *decomplexation* used in [12,21] because they have a more neutral chemical meaning.



where we have considered a rate 100 for the left to right direction and 10 for the right to left direction.

The **nanok** calculus reactions that corresponds to the two reactions of the hydrogen gas are



where the right direction has been given rate 5 and the left direction has been given rate 0.05.

The formal definition of *reactants* and the corresponding *products* of reactions follows. We use μ to range over ρ_L, ι, x and ρ_R, ι, x and ρ_L, ι and ρ_R, ι and ρ . Let $\bar{\mu}$ be the following operation (notice that $\bar{\bar{\mu}} = \mu$):

$$\bar{\mu} \stackrel{\text{def}}{=} \begin{cases} \rho_R, \iota, x & \text{if } \mu = \rho_L, \iota, x \\ \rho_L, \iota, x & \text{if } \mu = \rho_R, \iota, x \\ \rho_R, \iota & \text{if } \mu = \rho_L, \iota \\ \rho_L, \iota & \text{if } \mu = \rho_R, \iota \\ \rho & \text{if } \mu = \rho \end{cases}$$

Definition 3 (Basic transition relation) *The basic transition relation of solutions, written $\xrightarrow{\mu}_{\ell} \cup \xrightarrow{\mu}_{\ell, \ell'}$, is the least relation that satisfies the following rules (ι are always injective renamings on bonds):*

- (creations) if $\rho = A[u](a + \sigma), B[v](b + \phi) \xrightarrow{\lambda} A[u'](a^x + \sigma), B[v'](b^x + \phi)$ and $\text{dom}(\iota) = \text{ran}(\sigma) (= \text{ran}(\phi))$ and $z \notin \text{ran}(\sigma \circ \iota + \nu)$ then both $A[u + w](a + \sigma \circ \iota + \nu) \xrightarrow{\rho_L, \iota, z}_1 A[u' + w](a^z + \sigma \circ \iota + \nu)$ and $B[v + w](b + \phi \circ \iota + \nu) \xrightarrow{\rho_R, \iota, z}_1 B[v' + w](b^z + \phi \circ \iota + \nu)$;
- (destructions) if $\rho = A[u](a^x + \sigma), B[v](b^x + \phi) \xrightarrow{\lambda} A[u'](a + \sigma), B[v'](b + \phi)$ and $\text{dom}(\iota) = \text{ran}(\sigma) (= \text{ran}(\phi))$ then both $A[u + w](a^x + \sigma \circ \iota + \nu) \xrightarrow{\rho_L, \iota, x}_1 A[u' + w](a + \sigma \circ \iota + \nu)$ and $B[v + w](b^x + \phi \circ \iota + \nu) \xrightarrow{\rho_R, \iota, x}_1 B[v' + w](b + \phi \circ \iota + \nu)$;
- (exchanges) if $\rho = A[u](\sigma), B[v](\phi) \xrightarrow{\lambda} A[u'](\sigma'), B[v'](\phi')$ and $\text{dom}(\iota) = \text{ran}(\sigma) (= \text{ran}(\phi))$ then both $A[u + w](\sigma \circ \iota + \nu) \xrightarrow{\rho_L, \iota, v}_1 A[u' + w](\sigma' \circ \iota + \nu)$ and $B[v + w](\phi \circ \iota + \nu) \xrightarrow{\rho_R, \iota, v}_1 B[v' + w](\phi' \circ \iota + \nu)$, where v is either ε or $\iota(x)$, according to $\text{ran}(\sigma) \setminus \text{ran}(\sigma')$ is \emptyset or $\{x\}$;
- (lifts) if $S \xrightarrow{\mu}_{\ell} S'$ and, when ρ is a creation, $(\text{name}(S') \setminus \text{name}(S)) \cap \text{name}(T) = \emptyset$, then both $S, T \xrightarrow{\mu}_{\ell} S', T$ and $T, S \xrightarrow{\mu}_{\ell + \ell'} T, S'$, where T has ℓ' molecules;

- (communications) if $S \xrightarrow{\mu}_{\ell} S'$ and $T \xrightarrow{\bar{\mu}}_{\ell'} T'$ then $S, T \xrightarrow{\rho}_{\ell, \ell'' + \ell'} S', T'$, where ρ is the rule of μ and S has ℓ'' molecules. If ρ is a creation, then the bond used by the reaction is the least one that is not used in S, T .

The basic transition relation definitely deserves to be spelled out. A reaction, such as $Na[ion^0], Cl[ion^1] \xrightarrow{100} Na[ion^1], Cl[ion^0]$ is a *schema*, namely it only addresses the fields and the sites of the reactants that are useful for the reaction. For example, it may be the case that Na retains a site to be used for other complexes, such as the *sodium peroxide*. In this case, the rule may be applied either to $Na[ion^0]$, where the site is unbound, or to $Na[ion^0](1^x)$. In this latter case, the reaction is instantiated as the transition:

$$Na[ion^0](1^x), Cl[ion^1] \xrightarrow{\rho}_{1,2} Na[ion^1](1^x), Cl[ion^0]$$

The basic transition relation is indexed by numbers. Since the solutions are sequences, these numbers give the exact positions of the reactants in the sequences. In the first three cases, the position is always 1 because the solution consists of one molecule. In the fourth case, the index is increased by the number of the molecules on the left, if any. The last case models a reaction: the solution is split into two parts S and T containing the reactants at positions ℓ and ℓ' , respectively. In the composite solution S, T , the reactants are at ℓ and $\ell'' + \ell'$, where ℓ'' is the number of molecules of S . For example let kM be $\underbrace{M, \dots, M}_{k \text{ times}}$ and let ρ be the hydrogen gas reaction. The following transitions are possible

$$\begin{aligned} 3H(1) &\xrightarrow{\rho}_{1,2} 2H(1^x), H(1) \\ 3H(1) &\xrightarrow{\rho}_{1,3} H(1^x), H(1), H(1^x) \\ 3H(1) &\xrightarrow{\rho}_{2,3} H(1), 2H(1^x) \end{aligned}$$

The basic transition relation is labelled by finite injective renamings. To clarify this point, consider the creation $\varrho = Na(1^x + 2), Na(1^x + 2) \xrightarrow{10} Na(1^x + 2^y), Na(1^x + 2^y)$ (a bond is created between two sodium molecules provided they are already bound). Then take the solution $Na[ion^0](1^z + 2), Na[ion^0](1^v + 2), Na[ion^1](1^z + 2), Na[ion^0](1^v + 2)$. We derive the expected transition

$$\begin{aligned} &Na[ion^0](1^z + 2), Na[ion^0](1^v + 2), Na[ion^1](1^z + 2), Na[ion^0](1^v + 2) \\ &\xrightarrow{\varrho}_{1,3} Na[ion^0](1^z + 2^y), Na[ion^0](1^v + 2), Na[ion^1](1^z + 2^y), Na[ion^0](1^v + 2) \end{aligned}$$

following a structured operational semantics approach [24]. Namely, we focus on the single reactants and lift the transitions to “,”-contextes. This is correct inasmuch as one records the instantiation of bonds in the left-hand sides of reactions with the actual names of the molecules: the two reac-

tants must instantiate bonds in the same way. This is the reason why the first two molecules of the above solution cannot react with ρ . More precisely, $Na[ion^0](1^z + 2) \xrightarrow{\rho_L, \iota, y} Na[ion^0](1^z + 2^y)$, where $\iota = [x \mapsto z]$, and $Na[ion^0](1^v + 2) \not\xrightarrow{\rho_R, \iota, y}$.

Our final remarks regard the fourth and fifth items of Definition 3. Whenever $S \xrightarrow{\rho, \iota, x} T$ and ρ is a creation, the basic transition relation also admits $S \xrightarrow{\rho, \iota, y} T\{y/x\}$, where y is fresh. This nondeterminism is removed when the reaction occurs because the bond has to be the least name not occurring in S . It is also worth to notice that there is no rule lifting a transition $\xrightarrow{\mu} \ell, \ell'$ to a context “,”: we use the associativity of “,” to partition a solution S into S', S'' such that the reactants are in S' and S'' .

The basic transition relation is excessively intensional to be sensible for chemistry. Consider a solution containing hundreds of molecules of the species A and B that could react with ρ . The relation $\xrightarrow{\mu} \ell, \ell'$ distinguishes the two pairs of reactants, and this is not possible in practice. More reasonably, the transition relation should represent *collectively* all the possible combinations of one molecule of species A with one molecule of species B . For instance, the solution A, A, B transits with $\xrightarrow{\rho} 1, 3$ and $\xrightarrow{\rho} 2, 3$. Abstracting out the order of the molecules, we obtain a unique transition whose rate is twice the rate of ρ . However quotienting the solutions with commutativity axioms of “,” does not yield an adequate extensionality. In facts, when ρ is a destruction, between A and B , the solution $A(a^x), A(a^y), B(a^x), B(a^y)$ transits with $\xrightarrow{\rho} 1, 3$ and $\xrightarrow{\rho} 2, 4$ into two solutions that cannot be equated by permutations of the molecules in the solution. In these cases one has to use injective renamings of bonds.

Definition 4 (Structural equivalence) *The structural equivalence between solutions, noted \equiv , is the least equivalence satisfying the following two rules (we remind that solutions are already quotiented by associativity of “,”):*

- (1) $S, T \equiv T, S$;
- (2) $S \equiv T$ if there exists an injective renaming ι on bonds such that $S = \iota(T)$.

Example 3 *Commutativity and injective renaming of the structural equivalence permit to prove the two following equivalences, respectively*

$$Na[h^0], Cl[h^1] \equiv Cl[h^1], Na[h^0] \quad H(b^x), H(b^x) \equiv H(b^y), H(b^y)$$

Combining both commutativity and injective renaming we can prove that

$$H(b^x), H(b^x), H(b^z), H(b^z) \equiv H(b^y), H(b^k), H(b^k), H(b^y)$$

Proposition 1 *Let $S \equiv S'$.*

- (1) *If $S \xrightarrow{\mu} \ell T$ then there is T' and a renaming ι such that $S' \xrightarrow{\iota(\mu)} \ell' T'$ and*

- $T' \equiv S'$;
(2) if $S \xrightarrow{\rho}_{\ell, \ell'} T$ then there is T' such that $S' \xrightarrow{\rho}_{\ell', \ell''} T'$ and $T' \equiv S'$.

The following notations are relevant for the definition of the stochastic transition relation:

- $rate(\rho)$ returns the rate of the reaction ρ ;
- $next(S) = \{(\rho_{\ell, \ell'}, T) \mid S \xrightarrow{\rho}_{\ell, \ell'} T\}$; $next_{\infty}(S) = \{(\rho_{\ell, \ell'}, T) \mid S \xrightarrow{\rho}_{\ell, \ell'} T \text{ and } rate(\rho) = \infty\}$;
- \mathcal{S} has finite rates if, for every $(\rho_{\ell, \ell'}, T) \in \mathcal{S}$, $rate(\rho)$ is not ∞ ;
- let \mathcal{S} be a set of pairs (X, T') (the second element is a solution; the first one is not specified), $[\mathcal{S}]_T$ is the subset of \mathcal{S} of those pairs (X, T') such that $T' \equiv T$;
- $can(\mathcal{S})$ is defined over sets of pairs (X, T) such that the solutions occurring as second element of the pairs are all structurally equivalent. It returns a solution S such that there is X with $(X, S) \in \mathcal{S}$.

Definition 5 (Stochastic transition relation) *The **nanok** calculus stochastic transition relation \mapsto^{λ} , where $\lambda \in \mathbb{R}^+ \cup \{\infty\}$, is the least relation satisfying the following rules:*

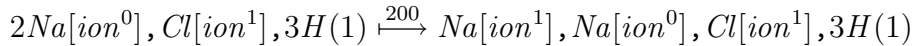
- if $S \xrightarrow{\rho}_{\ell, \ell'} T$ and $rate(\rho) = \infty$ then $S \mapsto^{\infty} can([next_{\infty}(S)]_T)$;
- if $S \xrightarrow{\rho}_{\ell, \ell'} T$ and $next(S)$ has finite rates then $S \mapsto^{\lambda} can([next(S)]_T)$, where

$$\lambda = \sum_{(\rho_{\ell, \ell'}, T') \in [next(S)]_T} rate(\rho)$$

We notice that, by definition, the **nanok** calculus stochastic transition system is such that there is no state with outgoing \mapsto^{∞} and \mapsto^{λ} (λ finite) transitions. Hereafter, the states with \mapsto^{∞} outgoing transitions are called *transient states*, the other ones are called *markovian states*.

The interrelation between basic and stochastic transition relations is as follows: the stochastic one partitions the products of a solution (according to the basic transition relation) into equivalence classes, takes a canonical representative of the class, and defines a transition whose label is the sum of the rates of the reactions in the basic one that yield solutions in the equivalence class.

Example 4 *As examples of stochastic transitions, we consider the reactions of sodium chloride (called ρ) and hydrogen gas (called ρ') of Example 1 for the solution $2Na[ion^0], Cl[ion^1], 3H$. This solution may transit with $\xrightarrow{\rho}_{1,3}$ and $\xrightarrow{\rho'}_{2,3}$ into solutions that are structural equivalent. Therefore we obtain a unique stochastic transition:*



We also observe that there is a unique transition \vdash^{15} outgoing the initial solution and corresponding to $\xrightarrow{\rho'}_{4,5}$, $\xrightarrow{\rho'}_{4,6}$, and $\xrightarrow{\rho'}_{5,6}$.

3 Markov Chains and the nano κ calculus

The stochastic transition relation of nano κ calculus corresponds to an *Interactive Markov Chain (IMC) transition system* with only *silent interactive transitions* [17]. These transitions, which are those labelled ∞ in our model, are executed in the IMC model instantaneously and under the *maximal progress assumption*. That is, the so-called *sojourn time* in a transient state is 0, which amounts to favour silent interactive transitions to those labelled with finite rates (called *markovian* transitions). On the contrary, in a markovian state with n outgoing markovian transitions labelled $\lambda_1, \dots, \lambda_n$, the probability that the sojourn time is less than t is exponentially distributed with rate $\sum_i \lambda_i$, i.e. $\text{Prob}\{\text{delay} < t\} = 1 - e^{-t \sum_i \lambda_i}$, and the probability that the j -th transition is taken is $\lambda_j / (\sum_i \lambda_i)$.

However the models underlying traditional simulation algorithms such as [14] are *Continuous Time Markov transition systems (CTMC)* that do not include interactive transitions. Having a CTMC is therefore primary to run automatic analysis tools for experimenting *in silico* the dynamics of nano-machines specifications in nano κ calculus.

The mismatch between IMC with only silent actions and CTMC systems is due to two main reasons: (i) the nondeterminism and (ii) the persistency of the silent interactive transitions. As regards (i), consider two silent actions that apply to the same reactants and give two different products. If these products have only markovian transitions it is not possible to collect them in a unique solution. As regards (ii), if an infinite sequence of silent interactive transitions exists then the simulation time of the CTMC system will not advance anymore. Therefore collapsing all these transitions, by identifying the initial and final solutions of the sequence, is again not possible.

However, there are cases where the downgrading of an IMC system to a CTMC one is possible without modifying the semantics. This is *when all silent actions may be partitioned into confluent (up-to structural congruence) directed acyclic graphs of finite depth*. In fact, when the silent interactive transitions are partitioned into confluent (up-to structural congruence) directed acyclic graphs, there are no loops (there is no infinite sequence of silent interactive transitions), and all sequences of silent interactive transitions starting from the same state share the same final state (up-to structural congruence), to which the initial state may be safely collapsed. The meaning of this collapse is that we are removing a finite amount of work which is performed in zero

time.

The formal definition of downgrading of IMC to CTMC systems follows. We first introduce the auxiliary function *next markovian state* defined on solutions and yielding sets:

- $nextm(S) = \{((\lambda, T'), T) \mid S \xrightarrow{\lambda} T' \xrightarrow{\infty}^* T \text{ and } \lambda \in \mathbb{R}^+ \text{ and } T \not\xrightarrow{\infty}\}$

We notice that $nextm(S)$ is undefined when S is transient.

Definition 6 (Downgrading of IMC) *An IMC system $(S, \xrightarrow{\lambda})$ is strictly-markovian if*

- (1) *states are either transient or markovian and*
- (2) *every subsystem consisting of silent interactive transitions is a confluent (up-to structural congruence) direct acyclic graph of finite depth.*

Let $(S, \xrightarrow{\lambda})$ be strictly-markovian; the transition relation $\xRightarrow{\nu}$, where $\nu \in \mathbb{R}^+$, is the least one such that:

- *if S is markovian then $S \xRightarrow{\nu} can([nextm(S)]_{\top})$ with*

$$\nu = \sum_{((\lambda, T'), T'') \in [nextm(S)]_{\top}} \lambda$$

It is easy to verify that the relation $\xRightarrow{\nu}$ defines a CTMC system. To assert the soundness of the downgrading process, we show that it maps the classical semantics of IMC, the *markovian bisimulation* [17], to the classical semantics of CTMC, the *lumping equivalence* [20] (see Proposition 2 below).

Notation: Given a strictly-markovian IMC $(S, \xrightarrow{\lambda})$, we denote with \mathcal{S}^m the set of its markovian states and with \mathcal{S}^t the set of its transient states. Given a state $S \in \mathcal{S}$, we write $S \xrightarrow{\tau} S$ if $S \in \mathcal{S}^m$, we write $S \xrightarrow{\tau} T$ if $S \in \mathcal{S}^t$ and $T = can(\{T' \in \mathcal{S}^m \mid S \xrightarrow{\infty}^* T'\})$. Given a markovian state $S \in \mathcal{S}^m$ and a set of states $C \subseteq \mathcal{S}$, we denote with $\mu(S, C)$ the *cumulative rate* obtained as the sum of all rates of the transitions from S to a state in C . This is formally defined as follows:

$$\mu(S, C) = \sum_{(\lambda, T) \in \{(\lambda, T) \mid S \xrightarrow{\lambda} T, T \in C\}} \lambda$$

The cumulative rates can be defined in a similar way also for CTMCs, simply using $\xRightarrow{\lambda}$ instead of $\xrightarrow{\lambda}$. In the following, we overload the notation $\mu(S, C)$ because we use it also for the cumulative rates in CTMCs.

Before presenting the definition of markovian bisimulation and of lumping equivalence, we use the introduced notation to state the following lemma (the

proof is omitted as the statement is a direct consequence of the Definition 6).

Lemma 1 Consider a strictly-markovian IMC $(\mathcal{S}, \mapsto^\lambda)$, two markovian states S and T , and the set of states $C = \{S' | S' \xrightarrow{\tau} T', T' \equiv T\} \cup \{T' | T' \equiv T\}$, we have that:

- $\mu(S, C) > 0$ if and only if there exists one and only one $T'' \equiv T$ such that $S \xrightarrow{\mu(S, C)} T''$.

The above lemma has the following implications:

- the probability distribution of the sojourn time in a markovian state is the same in the IMC and in the downgraded CTMC and
- the probability that one of the paths $S \xrightarrow{\lambda} \xrightarrow{\infty}^* T'$ with $T' \equiv T$ is taken in the IMC corresponds to the probability the unique transition $S \xrightarrow{\lambda'} T''$, with $T'' \equiv T$, is taken in the downgraded CTMC.

In IMCs, the markovian bisimulation is built from the classical concepts of bisimulation corresponding to its interactive and markovian parts. Two bisimilar transient states should have the same outgoing interactive transitions. Two bisimilar markovian states should have same outgoing rates to the bisimulation equivalence classes. It is naturally extended to the notion of *weak markovian bisimulation*. We formally detail this notion for strictly-markovian IMCs as follows.

Definition 7 (Weak markovian bisimulation) Given a strictly-markovian IMC $(\mathcal{S}, \mapsto^\lambda)$, an equivalence relation \mathcal{R} on \mathcal{S} is a weak markovian bisimulation if given SRS' we have that

- if $S \xrightarrow{\tau} T$ then there exists T' such that $S' \xrightarrow{\tau} T'$ and for all \mathcal{R} -equivalence classes C we have that $\mu(T, C) = \mu(T', C)$.

Two states S and S' are bisimilar if SRS' for some weak markovian bisimulation \mathcal{R} . We write $S \approx_I S'$.

Lumping a CTMC consists in agglomerating states that have equivalent behavior, that is from which the outgoing rates to agglomerated states are equals.

Definition 8 (Lumping equivalence) Given a CTMC $(\mathcal{S}, \xrightarrow{\lambda})$, a partitioning \mathcal{P} of \mathcal{S} is a lumping if for every pair of partitions C and C' (i.e. $C, C' \in \mathcal{P}$) we have that

- if $S, T \in C$ then $\mu(S, C') = \mu(T, C')$.

Two states S and T are lumping-equivalent if they are contained in the same partition of a lumping. We write $S \sim_C T$.

The soundness of the downgrading process is formalized as follows: two markovian states in a strictly-markovian IMC are bisimilar if and only if they are lumping-equivalent in the corresponding downgrading.

Proposition 2 *Given a strictly-markovian IMC $(\mathcal{S}, \vdash^\lambda)$, its downgrading $(\mathcal{S}^m, \vdash^\lambda)$, and two markovian states $S, T \in \mathcal{S}^m$, we have that:*

- $S \approx_I T$ if and only if $S \sim_C T$.

Proof We first consider the *only if* part. If $S \approx_I T$ then we have that the partitioning on \mathcal{S}^m composed by the equivalence classes of \approx_I is a lumping, thus $S \sim_C T$ because S and T belong to the same equivalence class. In fact, as $S, T \in \mathcal{S}^m$, we have that $S \xrightarrow{\tau} S'$ implies $S' = S$, and that $T \xrightarrow{\tau} T'$ implies $T' = T$. By definition of weak markovian bisimulation we have also that $\mu(S', C) = \mu(T', C)$ for each \approx_I -equivalence class C , thus also $\mu(S, C) = \mu(T, C)$ as we already proved that $S' = S$ and $T' = T$.

We now consider the *if* part. If $S \sim_C T$ then there exists a partitioning \mathcal{P} such that S and T belong to the same partition. It is not restrictive to assume that all partitions of \mathcal{P} are closed under structural congruence (i.e. if W and Z belong to the same partition, then also W' and Z' belong to the same partition if $W \equiv W'$, and $Z \equiv Z'$). We show how to define an equivalence relation \mathcal{R} on \mathcal{S} which is a weak markovian bisimulation and such that $S \mathcal{R} T$ (thus also $S \approx_I T$). We define \mathcal{R} such that $Q \mathcal{R} Q'$ if and only if:

- $Q = Q'$ or Q and Q' belong to the same partition of \mathcal{P} or
- $Q \xrightarrow{\tau} Q'$ or $Q' \xrightarrow{\tau} Q$ or
- $Q \xrightarrow{\tau} R$ and $Q' \xrightarrow{\tau} R'$ and R and R' belong to the same partition of \mathcal{P} .

It is not difficult to see that, due to the convergence property of strictly markovian IMCs and because the partitions of \mathcal{P} are closed under structural congruence, \mathcal{R} is an equivalence relation. Moreover, each \mathcal{R} -equivalence class C is composed of the markovian states of a partition of \mathcal{P} , that we denote with \overline{C} , plus the transient states Q such that $Q \xrightarrow{\tau} Q'$ for some $Q' \in \overline{C}$. As S and T belong to the same partition of \mathcal{P} , then $S \mathcal{R} T$.

We complete the proof showing that \mathcal{R} is a weak markovian bisimulation. Consider $Q \mathcal{R} Q'$ and an \mathcal{R} -equivalence class C (and its corresponding partition \overline{C} of \mathcal{P}). Let R and R' be such that $Q \xrightarrow{\tau} R$ and $Q' \xrightarrow{\tau} R'$. As we have already observed above, we have that R and R' belong to the same partition of \mathcal{P} . Moreover, R and R' are markovian, thus we can apply Lemma 1 as follows (μ and $\bar{\mu}$ are the cumulative rates on the IMC $(\mathcal{S}, \vdash^\lambda)$ and its downgrading

$(\mathcal{S}^m, \xrightarrow{\lambda})$, respectively):

$$\begin{aligned}
\mu(\mathbf{R}, C) &= \sum_{(\lambda, Z) \in \{(\lambda, Z) | \mathbf{R} \xrightarrow{\lambda} Z, Z \in \overline{C}\}} \lambda && \text{by Lemma 1} \\
&= \bar{\mu}(\mathbf{R}, \overline{C}) \\
&= \bar{\mu}(\mathbf{R}', \overline{C}) && \text{because } \mathbf{R} \text{ and } \mathbf{R}' \text{ belong to the same partition} \\
&= \sum_{(\lambda, Z) \in \{(\lambda, Z) | \mathbf{R}' \xrightarrow{\lambda} Z, Z \in \overline{C}\}} \lambda \\
&= \mu(\mathbf{R}', C) && \text{by Lemma 1} \quad \square
\end{aligned}$$

4 nano κ calculus at work: the rotaxane case study

The investigated rotaxane RaH (Figure 2) [22,1] is made of a stoppered axle containing an ammonium (A) and an electron acceptor bipyridinium (B) stations that can establish hydrogen-bonding and charge-transfer interactions, respectively, with the ring component, which is a crown ether with electron donor properties. Since the hydrogen bonding interactions between the macrocyclic ring and the ammonium center are much stronger than the charge-transfer interactions of the ring with the bipyridinium unit, the rotaxane exists as only one of the two possible translational isomers, denoted as RaH in Figure 2. In solution, addition of a base (e.g., tributylamine) converts the ammonium center into an amine function, giving the transient state Ra that is transformed into the stable state Rb as a consequence of the displacement of the macrocycle onto the B station. The process can be reversed by addition of acid (e.g., trifluoroacetic acid) and the initial state is restored, passing through the transient state denoted as RbH. Nuclear magnetic resonance, absorption and luminescence spectroscopic experiments, together with electrochemical measurements, indicate that the acid-base controlled switching, which is fully reversible and relatively fast, exhibits a clear-cut on-off behavior [1].

The Rotaxane RaH is particularly appropriate to test the modeling approach described in the present paper because it is one of the very few cases wherein not only the thermodynamic properties, but also the dynamic behavior of the system have been experimentally characterized in detail. Specifically, the macrocycle's shuttling process between the ammonium/amine and bipyridinium stations in this rotaxane, driven by the successive addition of base and acid, have been investigated in solution [13]. The rate constants for the "forward" (Ra \rightarrow Rb) and "backward" (RbH \rightarrow RaH) shuttling motions (vertical processes in Figure 2) of the molecular ring, which occur, respectively, upon deprotonation and reprotonation of the ammonium/amine recognition site on the axle (horizontal processes in Figure 2), were found to be $0.72s^{-1}$ and $40s^{-1}$ at $293^\circ K$, respectively.

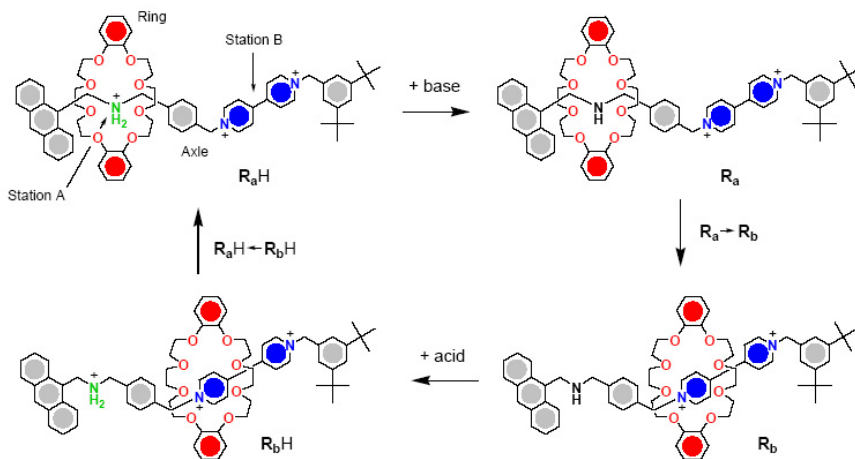


Fig. 2. Schematic representation of the shuttling processes of the molecular ring in the examined rotaxane.

4.1 Modeling the rotaxane RaH in nanok calculus

4.1.1 The nanok calculus molecules.

Figure 3 illustrates the nanok calculus modeling of the rotaxane RaH . We use four species:

- Nh models the ammonium/amine station of the rotaxane: it has one field h and two sites $ring$ and $axle$;
- $Axle$ models the spacer between the two stations: it has two fields s and h and three sites nh , $bipy$, and $ring$;
- $Bipy$ models the bipyridinium station: it has one field h and two sites $ring$ and $axle$;
- $Ring$ models the crown ether ring: it has no field and one site $link$.

The pairs of sites $axle$ of Nh and nh of $Axle$, and $axle$ of $Bipy$ and $bipy$ of $Axle$ are always linked in our modeling. They model the covalent bonds maintaining the structural integrity of the axle. Exactly one site $ring$ of Nh , $Bipy$, and $Axle$ is linked at a given moment at $link$ of $Ring$. The first two cases respectively model the “stable” RaH and Rb states of Figure 2 in which the ring is steadily located around the Nh or the $Bipy$ molecules, respectively. The last case models the “unstable” states; these are the Ra and RbH states of Figure 2 in which the ring is not steadily located. In order to distinguish between the Ra and RbH states, we use the field s of the $Axle$: it holds the value 0 if the ring is around the Nh (Ra state), 1 if it is around the $Bipy$ (RbH state).

Ammonium and amine functions have different chemical nature but can be seen as protonated and deprotonated version of the same species. Thus we

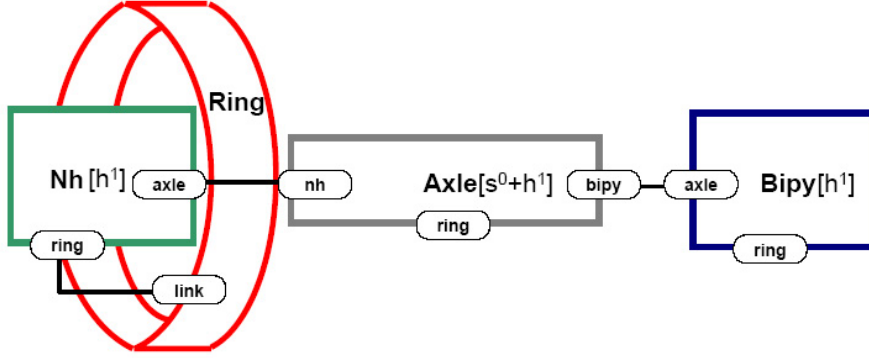


Fig. 3. Initial state of the Rotaxane RaH in `nanok` calculus.

model both by the same `nanok` calculus species Nh . Its field h is used to record the presence or absence of a proton on Nh : its value is 1 if it is protonated, and 0 otherwise.

As $Ring$'s movements are triggered by protonations and deprotonations due to acid-base reactions, we also need to have acid and base molecules in our modeling. We consider the species $Acid$ and $Base$ both with one field h having value 1 in case the acid/base molecule holds the proton to be exchanged, 0 otherwise (for instance $Acid[h^1]$ and $Base[h^0]$ are respectively an acid molecule ready to give a proton and a base molecule ready to receive a proton).

The initial state for rotaxane RaH is thus modeled by the term:

$$Nh[h^1](axle^s + ring^x) , Axle[s^0 + h^1](nh^s + bipy^r + ring) , \\ Bipy[h^1](axle^r + ring) , Ring(link^x)$$

graphically depicted in Figure 3.

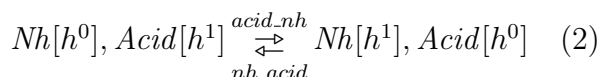
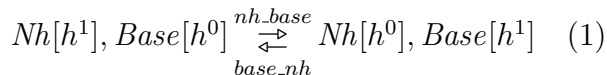
Note that the Nh is initially protonated (and this information is present also in the $Axle$ and the $Bipy$), the $Axle$ is bound to the Nh and the $Bipy$, and the $Ring$ is bound to the Nh .

4.1.2 The `nanok` calculus reactions.

We now present the reactions used in our modeling. Reactions 1, 2, 7 and 8 are presented with a double arrow (that are *reversible* reactions). Formally they correspond to two `nanok` calculus reactions, one achieved reading the reaction from left to right considering the rate over the arrow, and another one achieved reading it from right to left considering the rate below. In this section we do not consider numerical values of rates, this is detailed in part 4.2.

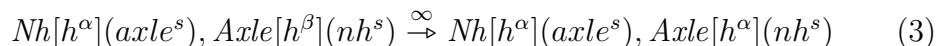
A base can get the proton of a protonated Nh , and a Nh can get a proton

from an acid. These acid-base reactions are reversible. Reactions 1 and 2 model this phenomena. The systems corresponding to the left-hand side and right-hand side coexist, even if one can be much predominant according to the ratio $nh_base/base_nh$ (and $acid_nh/nh_acid$).



The protonation state of the molecule *Nh* needs to be known by *Bipy* because it affects its interaction with *Ring*. Reactions 3 and 4 achieve this by passing information from *Nh* to *Bipy* through *Axle*. These updates are instantaneous because the reactions have infinite rates (this is relevant for the correctness of our simulation, since these reactions have no counterpart in chemistry).

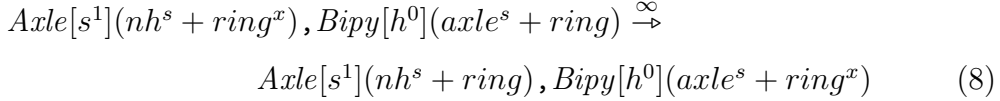
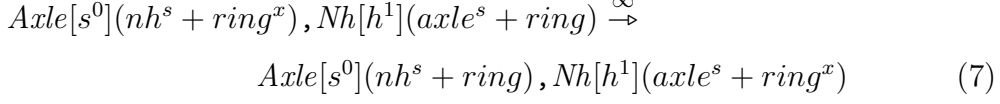
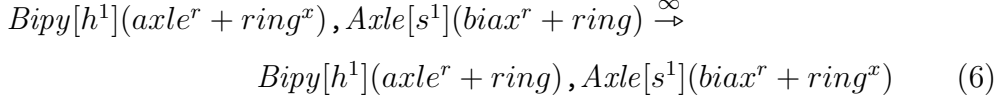
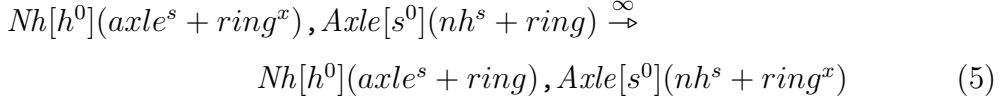
if ($\alpha \neq \beta$)



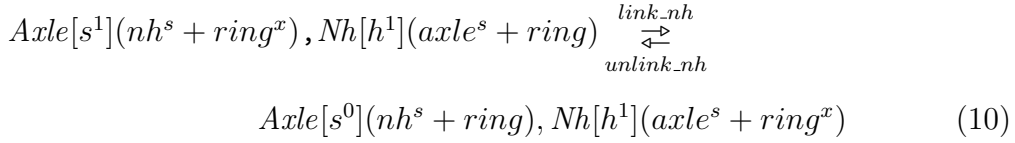
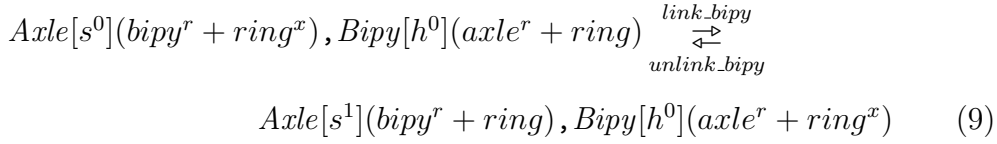
and:



We achieve the modeling of *Ring* movements in two steps. Firstly the instantaneous reactions to deprotonation/reprotonation (reactions 5–8), and secondly the actual *Ring* shuttling (reactions 9 and 10). The reactions (5) and (6) are used to enter in “unstable” states when the *Nh* is deprotonated while the *Ring* is around the *Nh* (reaction (5)), or protonated while the *Ring* is around the *Bipy* (reaction (6)). On the other hand, the reactions (7) and (8) are used to re-enter in a “stable” state in the case the *Nh* returns to its previous (de)protonated state before the *Ring* actually binds to its new station. All these events are immediate consequences of deprotonation or reprotonation of *Nh*; for this reason, they have infinite rates.



We now complete our modeling with reactions 9 and 10 representing the completion of the *Ring* movement. These reactions are reversible because the *Ring* is susceptible to leave its “stable” station due to the Brownian motion.



4.2 Simulation results

It is not difficult to verify that the above modeling of rotaxane RaH in `nanok` calculus yields a strictly markovian IMC system. Therefore we may safely downgrade it to a CTMC system that we use to simulate *in silico* the behavior of the rotaxane RaH.

As previously discussed the rates for the ring movements are respectively $link_bipy = 0.72s^{-1}$ and $link_nh = 40s^{-1}$. On the basis of the estimated equilibrium constants, the rates for the reverse reactions are quantified two orders of magnitude smaller, i.e. $unlink_bipy = 0.0072s^{-1}$ and $unlink_nh = 0.4s^{-1}$.

The aim of the first two simulations depicted in Figure 4 is to check whether the experimentation *in silico* can reproduce the results observed in *in vitro* [13]. The techniques used for the *in vitro* experimentation did not permit to observe

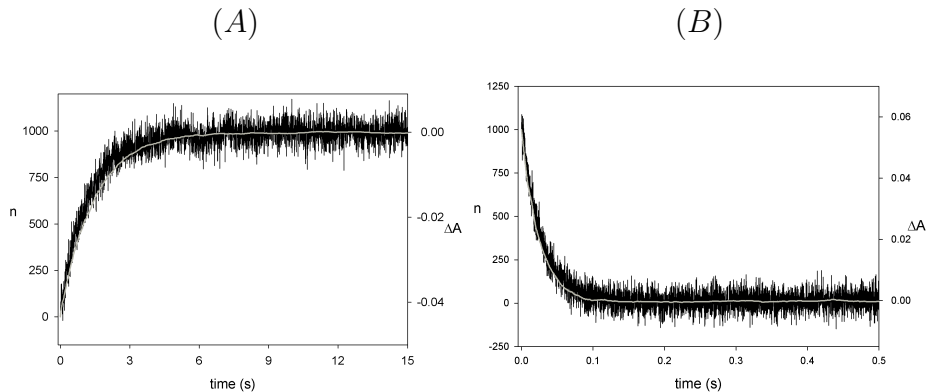


Fig. 4. Comparing the simulations *in silico* with the experiments *in vitro*. Grey traces: number of *Rings* located around *Bipys* during the “forward” Ra→Rb (part A) and the “backward” RbH→RaH (part B). Black traces: UV absorbance changes observed upon the occurrence of the same respective shuttling processes.

and quantify the deprotonation/reprotonation rates (this is not surprising as these are very fast acid-base reactions). Thus, in the simulation we have considered instantaneous deprotonation/reprotonation, i.e. $nh_{base} = acid_{nh} = \infty$ and $base_{nh} = nh_{acid} = 0$. In both simulations, we have considered 1000 rotaxanes: in the first one we have simulated deprotonation and “forward” (Ra→Rb) shuttling, in the second one reprotonation and “backward” (RbH→RaH) shuttling. In the first simulation the shuttling phase is completed in around 6 seconds, while in the second one in 0.1 seconds; this is a consequence of the different rates of the two directions of shuttling. Very remarkably, simulated data are in strike agreement with the experimental results.

After these initial encouraging results, we have decided to use the *in silico* simulation techniques to provide a comprehensive view of the overall reactions depicted in Figure 2, simulating also the deprotonation/reprotonation phases not observed in the *in silico* experimentation. More precisely, the aim of this second group of simulations was to either validate or invalidate the assumption according to which deprotonation/reprotonation can be considered “instantaneous” with respect to the shuttling time. To this aim, we have simulated deprotonation/reprotonation under two different concentrations of rotaxanes. In fact, this is a bimolecular reaction whose rate is influenced by the concentration of the reactants. For instance, at a concentration close to those considered in [13], e.g. $10^{-4}M$, assuming 1000 instances of rotaxane and base/acid, a plausible rate for deprotonation/reprotonation is $2 \times 10^3 s^{-1}$ (with reverse reaction rate on the order of $2 \times 10^{-4} s^{-1}$) while at the concentration $10^{-8}M$ it is $0.2 s^{-1}$ (with reverse reaction on the order of $0.2 \times 10^{-7} s^{-1}$).

We have performed the two simulations, namely deprotonation with subsequent “forward” shuttling and reprotonation with subsequent “backward”

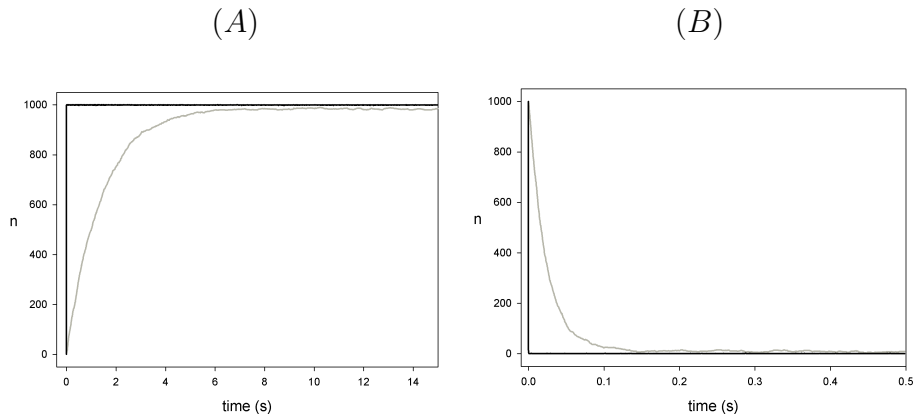


Fig. 5. Number of *Rings* located around *Bipys* (grey trace) and number of deprotonated rotaxanes (black trace) during the “forward” shuttling in the presence of base molecules (part A) and the “backward” shuttling in the presence of acid molecules (part B) at concentration $10^{-4}M$.

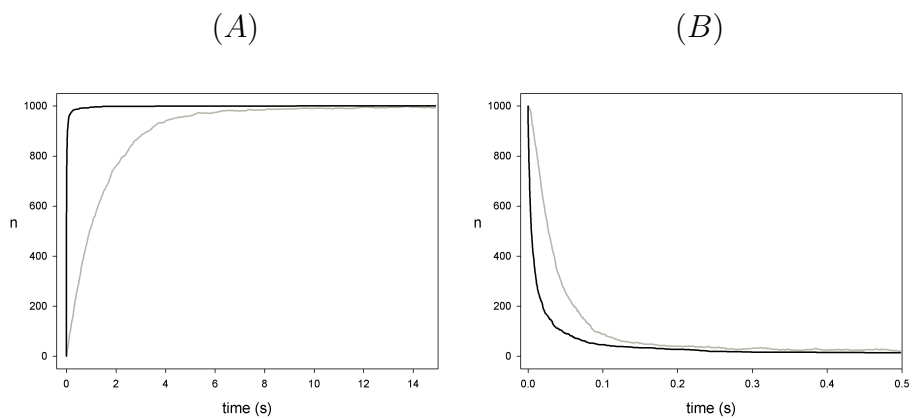


Fig. 6. Number of *Rings* located around *Bipys* (grey trace) and number of deprotonated rotaxanes (black trace) during the “forward” shuttling in the presence of base molecules (part A) and the “backward” shuttling in the presence of acid molecules (part B) at concentration $10^{-8}M$.

shuttling, considering the two different concentrations.

The results at concentration $10^{-4}M$ are reported in Figure 5; they essentially confirm the validity of the “instantaneous” deprotonation/reprotonation assumption at this concentration level. We report in Figure 6 the results for concentration $10^{-8}M$; in this case the rings start moving before the deprotonation/reprotonation phase is over. This proves that in the rotaxane RaH the stimulus and the subsequent shuttling could interplay.

In the light of this observation, we have decided to investigate some additional scenarios not yet considered in the *in vitro* experimentations. In particular, we have decided to analyze the interplay between shuttling and a stimulus given by *weaker* acid/base molecules, that is, for which the ratio between the

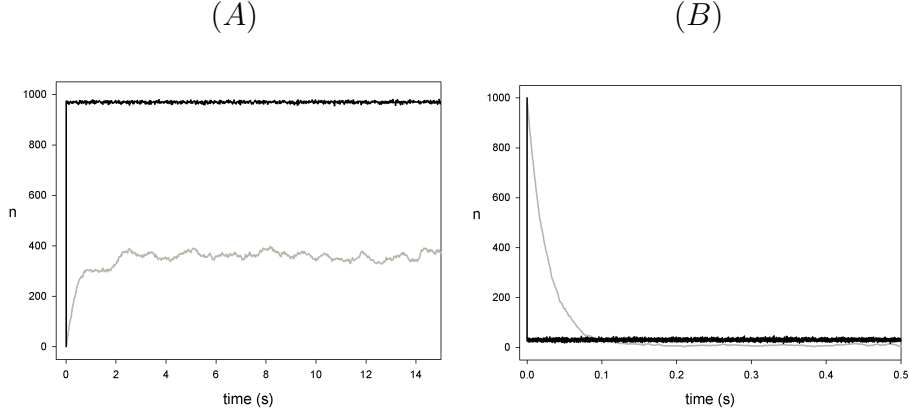


Fig. 7. Number of *Rings* located around *Bipys* (grey trace) and number of deprotonated rotaxanes (black trace) during the “forward” shuttling in the presence of weak base molecules (part A) and the “backward” shuttling in the presence of weak acid molecules (part B) at concentration $10^{-4}M$.

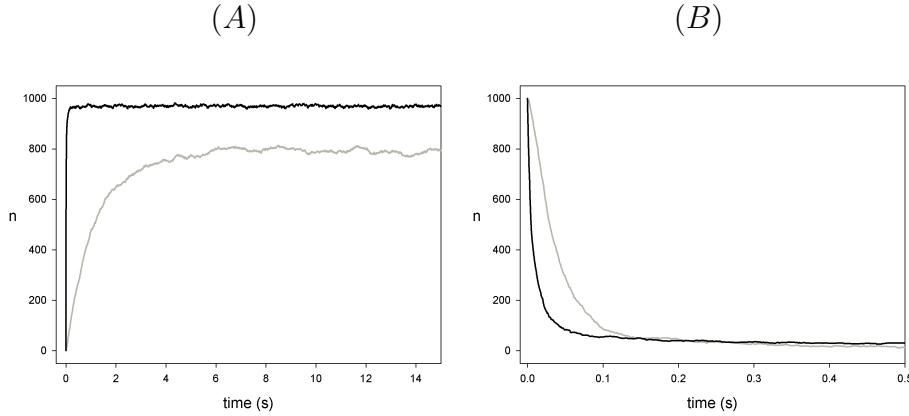


Fig. 8. Number of *Rings* located around *Bipys* (grey trace) and number of deprotonated rotaxanes (black trace) during the “forward” shuttling in the presence of weak base molecules (part A) and the “backward” shuttling in the presence of weak acid molecules (part B) at concentration $10^{-8}M$.

deprotonation/reprotonation rate and the reverse rate is smaller. In fact, the ratio considered in the previously discussed simulations is on the order of 10^7 ; a smaller reasonable ratio could be on the order of 10^3 . Considering this new ratio, assuming 1000 instances of rotaxane and base/acid, at the concentration $10^{-4}M$ the new rates for deprotonation/reprotonation is $2 \times 10^3 s^{-1}$ with reverse reaction rate on the order of $2 s^{-1}$, while at the concentration $10^{-8}M$ it is $0.2 s^{-1}$ with reverse reaction on the order of $0.2 \times 10^{-3} s^{-1}$. Using these new rates, we have simulated the “forward” and “backward” shuttling at both concentrations, $10^{-4}M$ in Figure 7 and $10^{-8}M$ in Figure 8.

Unexpectedly, we found out that the “forward” shuttling is no longer guaranteed. In fact, only in some of the deprotonated rotaxanes the *Ring* actually

moves around the *Bipy*. In other terms, the efficiency of the rotaxane is no longer close to 100% (as it was the case in the *in vitro* experimentations and in the other *in silico* simulations) but it is around 35% for concentration $10^{-4}M$, or 75% for concentration $10^{-8}M$. After a deep analysis of this initially unexpected results, we can conclude that the inefficiency of the rotaxane is justified by the fact that the reverse reaction of deprotonation (i.e. re-protonation) can activate a chain of reactions that allows an already deprotonated rotaxane, with the *Ring* around the *Bipy*, to return in the initial state (protonated with the *Ring* around the *Nh*). This chain of reactions, under these particular circumstances, plays an important role in the equilibrium between the number of deprotonated rotaxanes with the *Ring* around the *Nh* and the number of deprotonated rotaxanes with the *Ring* around the *Bipy*.

5 Conclusion and Related work

We have introduced **nanok**, a calculus designed on purpose for the modelization of nano devices. The calculus is equipped with a stochastic semantics (defined in terms of a CTMC) that can be used to simulate the evolution of the behavior of nano devices using stochastic simulation techniques such as, for instance, the Gillespie algorithm [14]. We have applied the **nanok** calculus to the modelization and simulation of the RaH rotaxane [22,1], a nano device that attracted a lot of attention inside the nano science and technology community, because it revealed very useful for building more complex nano devices [19,18,2]. We have used the **nanok** calculus model of the RaH rotaxane to simulate its behavior under conditions that were not yet considered in the *in vitro* experimentations. We found out that under particular circumstances the nano device is not as efficient as expected. In particular, even if almost all the rotaxanes in a solution are stimulated, only some of them change their internal structure according to the stimulus.

As future work, we intend to use the **nanok** calculus to model and simulate also more complex nano devices, such as the nano elevator presented in [2]. A nano elevator is a system composed of a platform and of three rotaxanes that, once appropriately stimulated, move the platform up or down. We expect to reuse the modeling of the rotaxane presented in this paper. In fact, one of the most important peculiarities of the **nanok** calculus is that it supports compositional modeling: the reactions describing the behavior of the molecules that are part of a nano device, are still valid reactions also when the nano device is itself considered as a part of a more complex system.

We have already discussed in the Introduction the origins of the **nanok** calculus, and its strong relationship with the κ calculus and the **m** κ calculus [12]. Here we simply recall that the **nanok** calculus can be seen as a member of

the family originated by [12]. The κ -calculus presented there beneficiates now from efficient techniques of simulation and analysis [11]. On the contrary of our **nanok** calculus this formalism allows reaction involving an arbitrary number of molecules, but there is no exchanges rules, edges can only be created or destroyed, not moved. These differences are explained by our field of application. Dealing with the behavior of nano-complexes, the relevant reaction we met involve barely more than two molecules, but edges are often exchanged and moved between molecules.

We complete the paper discussing other related work.

The **nanok** calculus has been influenced also by Cardelli's language of stochastic interacting processes [7,6] that has been put in correspondence with Ordinary Differential Equations. The stochastic semantics of the **nanok** calculus, indeed, has been given following these lines. Moreover a subset of the **nanok** calculus expressive enough to describe all the molecular cases we have considered so far is encodable into the stochastic π -calculus. We conjecture that it is also possible to identify the subset of the stochastic π -calculus reached by this encoding, and to encode it into the **nanok** calculus. That would lift the Cardelli's correspondence to the level of the **nanok** calculus and the π -calculus.

The beta-binders [26] are another formalism that can represent complexing molecules. It is based on a π -calculus where the usual communication discipline is relaxed to better represent the complementarity of molecular binding sites. It is achieved by means of a wrapping operator associating an interface to a group of π -processes. Since these operators are not chemical-like, recently a higher level language has been proposed [16] more suitable for the description of molecular process, this language is compiled into the beta-binders.

Finally, in the *calculus of looping sequences* [5,4] a different paradigm is taken. Molecules are represented simply by a name rather than by a π -process and they can be assembled in sequences. Closed chains of molecules are used to represent membranes, while dynamics is governed by rewriting rules on names.

References

- [1] P. R. Ashton, R. Ballardini, V. Balzani, A. Credi I. Baxter, M. C. T. Fyfe, M. T. Gandolfi, M. Gómez-López, M.-V. Martínez-Díaz, A. Piersanti, N. Spencer, J.F. Stoddart, M. Venturi, A. J. P. White, and D. J. Williams. Acid-base controllable molecular shuttles. *Journal of Am. Chem. Soc.*, 120:11932–11942, 1998.
- [2] J.D. Badjic, V. Balzani, A. Credi, S. Silvi, and J.F. Stoddart. A molecular elevator. *Science*, 303:1845–1849, 2004.

- [3] V. Balzani, A. Credi, and M. Venturi. Molecular devices and machines – concepts and perspectives for the nano world. In *Wiley-VCH*, Weinheim, 2007.
- [4] R. Barbuti, A. Maggiolo-Schettini, P. Milazzo, and A. Troina. Bisimulation congruences in the calculus of looping sequences. In *Proceedings of the Third International Colloquium on Theoretical Aspects of Computing (ICTAC'06)*, volume 4281 of *Lecture Notes in Computer Science*, pages 93–107. Springer, 2006.
- [5] R. Barbuti, A. Maggiolo-Schettini, P. Milazzo, and A. Troina. A calculus of looping sequences for modelling microbiological systems. *Fundam. Inf.*, 72(1-3):21–35, 2006.
- [6] L. Cardelli. Artificial Biochemistry, 2007. Available at: <http://lucacardelli.name>.
- [7] L. Cardelli. On Process Rate Semantics. *Theoretical Computer Science*, in press, 2008. Available at <http://dx.doi.org/10.1016/j.tcs.2007.11.012>.
- [8] L. Cardelli and A. Phillips. Spim homepage. At research.microsoft.com/~aphillip/spim/, 2006.
- [9] B. Champin, P. Mobian, and J.-P. Sauvage. Transition metal complexes as molecular machine prototypes. *Chemical Society Reviews*, 36(2):358–366, 2006.
- [10] A. Credi, M. Garavelli, C. Laneve, S. Pradalier, S. Silvi, and G. Zavattaro. Modelization and simulation of nano devices in \mathbf{nanok} calculus. In *Proceedings of Computational Methods in Systems Biology (CMSB'07)*, volume 4695 of *Lecture Notes in Computer Science*, pages 168–183. Springer-Verlag, 2007.
- [11] V. Danos, J. Feret, W. Fontana, and J. Krivine. Scalable simulation of cellular signaling networks. In *Proceedings of 5th Asian Symposium on Programming Languages and Systems (APLAS'07)*, volume 4807 of *Lecture Notes in Computer Science*, pages 139–157. Springer, 2007.
- [12] V. Danos and C. Laneve. Formal molecular biology. *Theoretical Computer Science*, 325(1):69–110, 2004.
- [13] S. Garaudée, S. Silvi, M. Venturi, A. Credi, A.H. Flood, and J.F. Stoddart. Shuttling dynamics in an acid-base-switchable [2]rotaxane. *ChemPhysChem*, 6:2145, 2005.
- [14] D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem*, 81:2340–2361, 1977.
- [15] S. Gilmore and J. Hillston. The PEPA workbench: A Tool to Support a Process Algebra-based Approach to Performance Modelling. In *Proceedings of the 7th int. conference on Computer performance evaluation : modelling techniques and tools*, volume 794 of *Lecture Notes in Computer Science*, pages 353–368. Springer-Verlag, 1994.

- [16] M.L. Guerriero, J.K. Heath, and C. Priami. An automated translation from a narrative language for biological modelling into process algebra. In *Computational Methods in Systems Biology (CMSB'07)*, volume 4695 of *Lecture Notes in Computer Science*, pages 136–151. Springer, 2007.
- [17] H. Hermans. *Interactive Markov Chains: The quest for quantitative quality*, volume 2428 of *Lecture Notes in Computer Science*. Springer-Verlag, Amsterdam, 2002.
- [18] T.-J. Huang, B. Brough, C.-M. Ho, Y. Liu, A.H. Flood, P.A. Bonvallet, H.-R. Tseng, J.F. Stoddart, M. Baller, and S. Magonov. A nanomechanical device based on linear molecular motors. *Applied Physics Letters*, 85(22):5391–5393, 2004.
- [19] M.C. 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.
- [20] J.G. Kemeny and J.L. Snell. *Finite Markov Chains*. D. Van Nostrand Company, 1960.
- [21] C. Laneve and F. Tarissan. A simple calculus of protein and cells. In *Proceedings of the First Workshop on Membrane Computing and Biologically Inspired Process Calculi (MeCBIC'06)*, volume 171(2) of *Electronic Notes in Theoretical Computer Science*, pages 139–154, 2007.
- [22] M.-V. Martínez-Díaz, N. Spencer, and J.F. Stoddart. The self-assembly of a switchable [2]rotaxane. *Angew. Chem. Int. Ed. Engl.*, 36:1904–1907, 1997.
- [23] R. Milner. *Communicating and mobile systems: the π -calculus*. Cambridge University Press, Cambridge, 1999.
- [24] G. D. Plotkin. A Structural Approach to Operational Semantics. Technical Report DAIMI FN-19, University of Aarhus, 1981.
- [25] C. Priami. Stochastic pi-calculus. *Computer Journal*, 38(7):578–589, 1995.
- [26] C. Priami and P. Quaglia. Beta binders for biological interactions. In *Proceedings of Computational Methods in Systems Biology (CMSB'04)*, volume 3082 of *Lecture Notes in Computer Science*, pages 20–33. Springer, 2005.
- [27] C. Priami, A. Regev, E. Shapiro, and W. Silverman. Application of a stochastic name-passing calculus to representation and simulation of molecular processes. *Information Processing Letters*, 80:25–31, 2001.
- [28] J.-P. Sauvage and C. O. Dietrich-Buchecker (eds.). Molecular catenanes, rotaxanes and knots. In *Wiley-VCH*, Weinheim, 1999.