Stochastic Theory for the Kinetics of Migration of Ligands in Biomolecules

P. Hanggi

Department of Physics,
University of Illinois at Urbana-Champaign,
Urbana, Illinois 61801, U.S.A.

(Received 1 November 1977, and in revised form 21 February 1978)

A stochastic approach in terms of master equations with linear and non-linear transition rates for the dynamics of the migration of ligands in biomolecules is presented. Coupling to a bath with constant ligand concentration as well as multiple occupations by ligands of certain sites inside the biomolecule are allowed. Explicit expressions for the fraction of biomolecules that have not bound a ligand at time \( t \) under experimental constraints are found by solving the generating function of the probability obeying the master equation. For highly non-linear systems a computer oriented procedure is presented. The validity of the description with a system of coupled linear deterministic equations is discussed. Relevance to experimental data and applications to other biophysical systems are outlined.

1. Introduction

Most biological processes are stochastic because fluctuations are an inherent consequence of the discrete nature of matter. An interesting problem is then the influence of fluctuations in the description of biological systems. When a large number of particles is involved, the values of the macroscopic quantities will vary closely around their mean values. But the stochastic aspects play a crucial role when the chosen macroscopic variables are subject to fluctuations comparable to their mean values. Even in situations where the fluctuations are small they may trigger a transition to a new macroscopic state, as for instance, in allosteric enzymes. In the present paper we develop the theoretical background for the dynamics of biomolecules with emphasis on the effect of fluctuations rather than confining ourselves to a specific biological system. Much of our knowledge about allosteric interactions and relationships between structure and function can be derived from studies with proteins. Particularly, heme proteins constitute a large group of biomolecules where the effect of fluctuations can be advantageously investigated. These proteins
play primary roles in oxygen storage (myoglobin) and transport (hemoglobin), electron transfer (cytochromes), and detoxification of poisonous substances (glutathione peroxidase). The migration of small ligands in biomembranes is governed by the principles of the classical theory of diffusion.

In the past, the migration process was assumed to be governed by a simple diffusion process. However, recent evidence suggests that the migration process is more complex and involves multiple mechanisms, including active transport, facilitated diffusion, and receptor-mediated endocytosis.

The migration of small ligands in biomembranes is a complex process that involves multiple mechanisms. Understanding the underlying mechanisms of migration is crucial for the development of new therapeutic strategies and the design of novel biomembrane-based technologies.
transition rates in equations (1)-(16) will in general depend in a non- 
linear way on the stochastic variables $n_f(t, \lambda)$. The paper is organized as follows. In section 2 we first deal with the limiting case ('low concentration limit') in which the density in the solvent is so small that at any time $t$ only one ligand sits somewhere inside the biomolecule (no multiple occupancy of different wells). The general case with possible multiple occupations of sites, $L_i$, with ligands and blocking will then be treated in section 3. Diffusion effects of ligands in the solvent and cage effects may also play an important role. In both cases we end up with an analytical expression for the rebinding rate $N_{	ext{reb}}(t, \lambda)$, the fraction of biomolecules that have not bound a ligand at time $t$ under experimental constraints $k$. The analytical solutions for $N_{	ext{reb}}(t, \lambda)$ are made possible by using a coarse-grained description for the migration process of the ligands, i.e., a description intermediate between a fully microscopic approach with a huge number of degrees of freedom and a macroscopic theory which directly describes the (unknown)-non-linear binding rate $N_{	ext{reb}}(t, \lambda)$. The results and other biological applications are discussed in section 2. Some important and useful properties of the generating function are summarized in Appendix A. In Appendix B we give the detailed mathematical development of the solution of the probability function describing the binding process. Appendix C deals with the solution of time-dependent mean values, e.g., $N_{	ext{reb}}(t, \lambda)$, in highly non-linear systems via continued fraction expansions. A convenient numerical, computer oriented procedure is presented.

2. Stochastic Model for Migration of Ligands in the Low Concentration Limit

In the case of a small or even vanishing number of ligands in the solvent we may assume that only one ligand is inside each biomolecule or inside its immediate neighborhood $[L]$ at any time. All the single biomolecule-structures with no ligand are not considered as members of the statistical ensemble for the ligand migration process and the members with more than one ligand are of vanishing influence. Transitions of ligands from state $[L]$, (cage), into other parts of the solvent (diffusion) are assumed to be of minor importance. The site $[L]$ can then be treated in terms of an additional stochastic variable $n_f(t, \lambda)$. In this approximation, only one ligand is occupying a certain state $L_i$ in each biomolecule at any time $t$. In such a "closed system" the number of ligands in a certain site $L_i$ is either 0 or 1. We may assume that the transitions in a given biomolecule occur independently of the transitions in other biomolecules (non-interacting biomolecules). For the sake of simplicity, we may also assume that ligand-transitions in biomolecules occur only between neighboring wells $L_i$.

In the following we introduce a one-dimensional notation. For the stochastic variable $n_f(t, \lambda)$, having at time $t$ and experimental constraints $k$, $n_f$ biomolecules in well $L_i$ at conformational state $i$, we define

$$x_f(t, \lambda) = n_f(t, \lambda)$$

(3)

with

$$K = \sum_{i=1}^{N} n_{	ext{max}}^{(D)} + 1$$

(4)

$K$ will take on values from 1 to $N$. The transition probabilities per unit time in equations (1)-(16) are defined in a similar way. The final binding site will be denoted in the following by $[L, \lambda]$. We use the notation $x(\lambda)$ to denote the stochastic vector process, for the stochastic variables at time $t$ forming the state space and for an actual point in the configuration space. The specific interpretation of $x(\lambda)$ will be understood from the context. Since the behavior of different biomolecules is independent and the configuration changes only by single jumps of the ligand we may assume stationary linear first-order transition rates (McQuarrie, 1968)

$$\ldots, x_1, \ldots, x_{N-1}, 1, \ldots, x_n \ldots$$

(5)

$$\Gamma(x_2 | x_1; \lambda) = \gamma_1 \gamma_2(\lambda)$$

Neglecting fluctuations, the kinetic deterministic rate equations for the variables $x_f(\lambda)$ read

$$\frac{dx_f}{dt} = -\sum_{i=1}^{N} \gamma_{i2} x_i + \sum_{i=1}^{N} \gamma_{2i} x_i, I = 1 \ldots, N$$

(6)

where the total number of all ligands is kept constant in the biological system under consideration (closed system), i.e.,

$$\sum_{i=1}^{N} x_i = N_0$$

(7)

By use of the matrix $A(\lambda)$

$$\Lambda_{ij}(\lambda) = \gamma_{ij}(\lambda), i \neq j$$

$$\Lambda_{ii}(\lambda) = -\sum_{j=1}^{N} \gamma_{ij}(\lambda)$$

(8)

or in matrix form

$$\Lambda(\lambda) = \left( \begin{array}{ccc} \gamma_{12}(\lambda) & \cdots & \gamma_{1N}(\lambda) \\ \vdots & \ddots & \vdots \\ \gamma_{N1}(\lambda) & \cdots & \gamma_{NN}(\lambda) \end{array} \right)$$

(9)

This matrix is the $A$-matrix, $A_{ij} = \gamma_{ij}(\lambda)$.

The structure of the $A$-matrix for $\gamma_{ij} = 4$ is shown in Fig. 1.
Note that $A_{ij}(t)$ for $I \neq J$ is a reaction constant which for ligand-transitions is given usually in terms of a phenomenological Arrhenius equation, a tunnel rate expression, a diffusion controlled expression, or various combinations of all. Assuming the multi-dimensional process $\pi(t) = \{x(t), x'_1(t), \ldots, x'_n(t)\}$ is a Markov process (Hanggi & Thomas, 1974, Hadem, 1975), the master equation, describing the rate of change of the probability $p(x, t; \lambda)$ that the system at time $t$ has the configuration $x$ in state space, reads
\begin{equation}
\frac{\partial p}{\partial t} (x, t; \lambda) = \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} (\lambda) \left[ p(x_{i+1}, x_{i-1}, x'_i; \lambda) - p(x, t; \lambda) \right].
\end{equation}
Here $p(x_{i+1}, x_{i-1}, x'_i; \lambda)$ stands for $p(x_{i+1}, x_{i-1}, x'_i, \ldots; x_{i-1}, \ldots, x_2, x_1; \lambda)$. The experimental data can then be evaluated if the time development of the probability of the system $p(x_i, t; \lambda)$ is known. This master equation can be solved for an arbitrary initial probability, $p(x, 0)$, using the technique of the generating function $G(y_1, \ldots, y_n; t; \lambda)$,
\begin{equation}
G(y_1, \ldots, y_n; t; \lambda) = \sum_{x_1} p(x_1, \ldots, x_n; t; \lambda) \prod_{i=1}^{n} y_i^{x_i}
\end{equation}
yielding
\begin{equation}
\prod_{i=1}^{n} y_i \frac{\partial p(x_1, \ldots, y_n; t; \lambda)}{\partial y_i} = \prod_{i=1}^{n} y_i \frac{\partial G(y_1, \ldots, y_n; t; \lambda)}{\partial y_i} \Big|_{y_i = 0}
\end{equation}
Using some useful and important properties of the generating function $G$, reviewed in Appendix A, the following linear first order differential equation for $G$ is obtained
\begin{equation}
\frac{\partial G(y_1, \ldots, y_n; t; \lambda)}{\partial t} = \sum_{i=1}^{n} \sum_{j=1}^{N} y_j A_{ij} \frac{\partial G(y_1, \ldots, y_n; t; \lambda)}{\partial y_j}.
\end{equation}
For a given initial probability $p(x, 0)$ we obtain with the eigenvalues, $\{\mu_k\}$, and the eigenvectors, $\{v_k\}$, of the matrix
\begin{equation}
\sum_{j=1}^{N} \mu_j v_j v_k = \mu_k v_k,
\end{equation}
and the co-factors, $\delta^k_i$, of the eigenvector $v$ for the solution of $G(x, t; \lambda)$ (see end of Appendix B)
\begin{equation}
G(x, t; \lambda) = \sum_{k=1}^{N} \prod_{i=1}^{n} y_i v_k \sum_{k=1}^{N} \frac{\partial G(x, t; \lambda)}{\partial y_i} v_k^{(i)} \exp(\mu_k t).
\end{equation}
\[\mu_k\] denotes the determinant of $v$.
For the probability $p(x, t; \lambda)$ of having $x_i$ ligands bound at time $t$ and experimental constraints $k$ we find by constructing on the stochastic variable $x_i(t; \lambda)$
\begin{equation}
p(x_i(t_1, t; \lambda)) = \prod_{i=1}^{n} \sum_{x_i} p(x_1, \ldots, x_n; t; \lambda).
\end{equation}
The experimentally monitored function $N_{exp}(t, \lambda)$, the fraction of ligands which have not bound at time $t$ to the binding site, is then given by
\begin{equation}
N_{exp}(t, \lambda) = (N_0 - \sum_{i=1}^{n} x_i)\exp(\mu_k t).
\end{equation}

For the interesting case of an initial multinomial probability $p(x, 0) = N_0! \prod_{i=1}^{n} \frac{p(x_i^0)}{x_i^0!} \prod_{i=1}^{n} x_i\delta^i_i$, with
\begin{equation}
p(x, 0) = e^{-N_0} \prod_{i=1}^{n} \frac{p(x_i^0)}{x_i^0!} \prod_{i=1}^{n} x_i\delta^i_i
\end{equation}
it is known (Saito, 1974) that the probability $p(x, t; \lambda)$ at time $t$ will remain of the multinomial form
\begin{equation}
p(x, t; \lambda) = N_0! \prod_{i=1}^{n} \frac{p(x_i^t)}{x_i^t!} \prod_{i=1}^{n} x_i\delta^i_i.
\end{equation}
The occupation probabilities of the sites, $P_i(A)$, are obtained by solving the equation (22).

$$P_i(A) = A_i \times \rho_i$$

(22)

and coincide exactly for all sites and the same initial condition in the deterministic equation (6) or (10) using the initial transition probability given in equation (10) (or the initial probability obtained from the deterministic equation (23)). Furthermore, in the special case of linear transition probability $P_i(A) = 1 - \lambda_i$, the equation (10) or (15) can be solved directly to obtain the solution. The values of $\lambda_i$ can be determined by solving the equation (23) using a linear regression of the form

$$P_i(A) = A_i \times \lambda_i$$

(23)

and for the covariance $\xi_i^2(A_i)$ of $P_i(A)$ and $P_i(A)$ always the negative result

$$\xi_i^2(A_i) = \sum_i (A_i - \lambda_i)$$

(24)

Finally, we add some comments relating to the choice of the initial probability on the site $i$. The initial probability on the site $i$ is determined from the equilibrium distribution $P_i(A) = A_i \times \rho_i$.

In the general case of biologic distribution, when the initial probability is not known, the determination of the initial probability on the site $i$ is based on the equilibrium distribution $P_i(A) = A_i \times \rho_i$.

On the other hand, when the initial probability is known, the initial probability on the site $i$ is determined from the equilibrium distribution $P_i(A) = A_i \times \rho_i$.
from and into the solvent may occur. [See equations (1)-(3).] The number of all ligands inside a single biomolecule is no longer equal to 1 so that we deal with possible multiple occupancies.

For completeness, we mention that diffusion effects could be described by dividing the state space into cells and considering additional transitions from one cell i to an adjacent cell j with probability d. In the case of a cubic cell system with cell length $l$, the deterministic equations for the concentration of ligands in cell $i$, $c(i, t; \lambda)$, would have the following structure in the limit of a continuous variable, $x$, for the cell index (no fluctuation normalization effects):

$$\frac{\partial c(x, t; \lambda)}{\partial t} = \sum_{j} \frac{d_{ij}}{D} \left( c(x, t; \lambda) - c(x, t; \lambda) \right) + \Delta V c(x, t; \lambda),$$

with

$$D = l^2 d.$$  

(29)

(30)

Here $D$ corresponds to Fick’s diffusion coefficient and the $\{\lambda\}$ denote the rates in cell $i$ or into and from the cell $i$ to the states $\{\lambda\}$ inside the biomolecules in cell $r$.

Treating the solvent as a bath, again we may assume that transitions in one biomolecule do not influence the transitions in other biomolecules. Then the migration process of the ligands can be described by the use of the following set of stochastic variables. We consider a single biomolecule and denote by $c(x, t; \lambda)$ the number of ligands at site $x$ at time $t$ under constraints $\lambda$. Note that in contrast to the case in section 2, $x_k(t; \lambda)$ does pertain to a single biomolecule. Hence, the effect of fluctuations plays a major role. Owing to a multiple occupation and blocking of site $x$, the transition probabilities will in general be non-linear. Blocking is especially important for the binding site which completely stops further transitions when $x_k(t; \lambda) = 1$. As in the previous treatment we assume that the observation changes only by single jumps of ligands. We use then for the stationary non-linear transition probabilities the scaling

$$\left( \ldots, x_k - 1, \ldots, x_k, \ldots \right) \rightarrow \frac{\Gamma(x_k / x_k; \lambda)}{\Gamma(x_k / x_k; \lambda)},$$

(31)

$$\Gamma(x_k / x_k; \lambda) = x_k \Gamma(x_k / x_k; \lambda).$$

(32)

(33)

The additional transitions from and into the bath with a constant ligand concentration $C_0$ are assumed to depend only on the occupation of site $x_k$ and the bath concentration $C_0$ in the form

$$\Gamma(x_k / x_k; \lambda) = x_k \Gamma(x_k / x_k; \lambda)$$

(32)

$$\Gamma(x_k, \lambda) = C_0 \Gamma(x_k, \lambda)$$

(33)

and if $\gamma_0$ can be considered to be independent of $x_k$, we obtain

$$\Gamma(x_k, \lambda) = C_0 \Gamma(x_k, \lambda) = \beta_i(\lambda).$$

With the matrix $M$,

$$M_{ij} = \gamma(x_k, x_k, \lambda) K_{ij} \neq 1,$$

$$M_{ij} = \gamma(x_k, \lambda) - \sum_{k} \gamma(x_k, x_k, \lambda);$$

(35)

(36)

the non-linear deterministic equations for the time evolution of the variables $x$ read

$$\frac{dx(t)}{dt} = M x(t) + \beta,$$

with

$$\beta = (\beta_1, \ldots, \beta_\lambda).$$

(37)

The master equation with the same rates in equations (31)-(34) has the same structure as in section 2 with the sums over $\lambda \Lambda$ extended to include the bath transitions. If the bath couples to one site $K$ only and if the coupling of this site is much stronger than to other sites such that

$$\gamma(x_k, \lambda) \rightarrow 0, \gamma(x_k, \lambda) \rightarrow \gamma(x_k, \lambda) = \rho,$$

(38)

then the average occupation of this site kept constant at the value $\gamma(x_k, \lambda) = \rho C_0$. If in addition, $\gamma(x_k, \lambda)$ is large enough so that fluctuations may be neglected, the site $K$ itself may serve as a bath and we have one less site. The block of the binding site [1] ($x_k = 0$ or 1) is taken into account by

$$\Gamma(x_k / x_k; \lambda) = x_k \Gamma(x_k / x_k; \lambda) \times x_k = 1.$$  

(40)

Equation (40) denotes the usual Kronecker function. The details of the migration process of ligands is then completely given by the solution of the master equation with the non-linear rates. We are interested in the probability

$$p^{(i)}(x_k / x_k; \lambda),$$

(41)

$$p^{(i)}(x_k / x_k; \lambda) = \sum_{x_k = 0, x_k, \ldots, x_k, \lambda, \lambda, \ldots} \gamma(x_k / x_k; \lambda),$$

(42)

which for non-interacting biomolecules is identical with the experimentally monitored function $N(x_k, \lambda)$, the fraction of Biomolecules that have not bound a ligand at time $t$. An analytical solution of the master equation for the non-linear vector process $x_k(t; \lambda)$ is in most cases not available. There may not even be a tractable numerical solution. However, the form of equation (42) forces us to settle for a numerical solution of the mean value $\gamma(x_k, \lambda)$ in

$$\gamma(x_k, \lambda) = \gamma(x_k, \lambda).$$
In order to solve this master equation we use the generating function which, in turn, gives the following evolution equation derived from equation (48) (see also Ref. 3):

\[ \dot{N}(t) = \sum_{a=1}^{n} \sum_{b=1}^{n} \Gamma_{ab}(t) \{ N(a) - N(b) \} \]

where \( \Gamma_{ab}(t) \) is the rate of transition from state \( a \) to state \( b \) at time \( t \).

The continued fraction coefficients \( \xi_{ij}(\alpha, \beta) \) of the matrix \( \mathbf{C}(\alpha, \beta) \) can be evaluated via a recursion scheme, the details of which are given in Appendix B. By using the results of the continued fraction zero set, \( x(\alpha, \beta) \) given in Ref. 3, we can neglect the non-linear terms in the master equation.

Defining further a complete trivial property for the Eigenvalue matrix \( \mathbf{C}(\alpha, \beta) \), the Markovian master equation for the probability \( \rho(\alpha, \beta, t) \) reads then

\[ \frac{\partial \rho}{\partial t} = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} \Gamma_{\alpha\beta}(t) \{ \rho(\alpha, \beta) - \rho(\beta, \alpha) \} \]

where \( \Gamma_{\alpha\beta}(t) \) is the transition rate from state \( \beta \) to state \( \alpha \) at time \( t \).

In situations where the equilibrium conditions are not met, i.e., \( \rho_{eq}(\alpha, \beta) \neq 0 \), the distribution functions will be given by

\[ \rho(t) = \rho_{eq}(t) \exp \left[ -\int_{0}^{t} \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} \Gamma_{\alpha\beta}(t) \{ N(\alpha) - N(\beta) \} dt \right] \]

For the initial probability \( \rho_0(\alpha, \beta, 0) \), we assume a most experimental probability which, for example, in flash photolysis experiments we have after a quench.

\[ \rho_0(\alpha, \beta, 0) = \begin{cases} 0 & \text{if } \alpha \neq \beta \\ \delta(\alpha, \beta) & \text{if } \alpha = \beta \end{cases} \]
preparation procedure (e.g. the flashed-off ligand from the binding site). So we obtain
\[ p(x, 0^+) = (1 - \delta_{x_0}) \prod_{i=1}^{N} \frac{(\delta x_i)^{x_i}}{(x_i)!} \exp(-\langle x_i(0) \rangle), \] (53)
with the initial moments given for the situation in equations (44)-(47):
\[ \langle x_i(0) \rangle = -\sum_{k=1}^{N} \frac{M_i^k}{k} \beta_k; i = 1, \ldots, N. \] (54)
\( M^t \) is the matrix \( M \), defined with constant transition rates \( \gamma_{jk} \), with the first row and column deleted and \( M_{0j} \) defined with \( \gamma_{j1} = 0 \). At an arbitrary time \( t \) the probability \( p(x, t; k) \) will now not remain of the multi-Fokker-Planck form, because of the non-linearity in the transition rates \( \gamma_{jk} \) [equation (45)]. Summation over \( x_1, x_2, \ldots, x_N \) in equation (51) yields for \( N_{x_0} = (x) \) the simple result
\[ N_{x_0} (x, t; k) = \exp(-\langle x_i^* (t, \lambda) \rangle), \] (55)
where (Appendix B)
\[ \langle x_i^* (t, \lambda) \rangle = \langle x_i (t, \lambda) \rangle \] (56)
\[ \langle x_i^* (t, \lambda) \rangle = -\frac{1}{N} \sum_{k=1}^{N} \left[ \frac{1}{\lambda_{jk}} + \frac{1}{\lambda_{kj}} \right] \langle x_j (t, \lambda) \rangle \] (57)
and
\[ \lim_{t \to \infty} \langle x_i^* (t, \lambda) \rangle = -\infty. \] (58)
The present theory allows us to calculate the macroscopic observable \( N_{x_0} (x, k) \) from which we can extract characteristic features of the internal mechanisms such as alternate pathways, restrictions on occupation numbers or volume available in different wells, function of conformational intermediates, or the role of the surrounding solvent. In particular, the simplified model with the assumptions in equations (44)-(47) explains the features of the experimental data of carbon monoxide migration in myoglobin over large ranges of [CO]-ligand concentrations in the solvent, temperature and time (Alberding, Fraunfelder & Hanggi, 1978). In systems more complex than myoglobins multiple occupancies may occur even under biological conditions. For CO-migration to cytochrome-a, of cytochrome oxidase, for instance, Sherrock & Yonetani (1977) have found experimental evidence for a carbon monoxide reservoir that connects to an intermediate well and is occupied by many CO-molecules. The characteristic dynamics for this system can be treated by a straightforward adaptation of the methods presented here.

4. Summary and Conclusions

In this paper we have given a complete description of the kinetics of the migration process of ligands in biomolecules under general experimental situations. Previously, the problem has been studied by solving a set of linear coupled ordinary differential equations (Austin et al., 1975; Alberding et al., 1978). This traditional method of analysis is based upon a deterministic formulation of ligand migration in which reaction constants are viewed as "reaction rates" and the various species concentrations are treated as continuous single-valued functions of time. Although this deterministic formulation is adequate in many cases (see e.g. section 2) there are experimental situations with non-linear transition rates, blocking effects, non-linear cage-solvent effects, history and preparative dependent conformational relaxation, where the non-linear fluctuations play a major role. The influence of fluctuations is also indispensable in noise studies. Experiments under extreme experimental constraints (e.g., high ligand concentrations), to which the present paper is mainly addressed, appear at first sight to have little direct bearing on biological processes. Such experiments, however, help elucidate the internal dynamic features and functions in complex biomolecular systems.

An approach that is more broadly applicable than a deterministic formulation is a stochastic formulation in terms of master equations where reaction constants are not viewed as "reaction rates" but as "reaction probabilities per unit time". From a physical point of view, the stochastic formulation is superior to the deterministic formulation; the stochastic approach is always valid whenever the deterministic approach is valid and is still valid when the deterministic approach is not. The former takes account of fluctuations and time dependent correlations. The stochastic approach, based on a linear evolution equation for the probability \( p(x, t) \) (see in this context also Appendix C), enables us to extract uniquely defined fluctuation renormalized mean value equations, i.e., non-linear renormalized deterministic equations in the sense that we have in equation (37)
\[ M(x|t) \neq M(x,t). \] (59)

In particular, no apologies need be made for the fluctuations. These fluctuations are really present and give rise to macroscopically observable effects in appropriate situations (section 3). In those cases in which fluctuations turn out to be unimportant, that fact too will emerge quite naturally from the formalism presented here. Further, a set of non-linear deterministic equations is often harder to solve than the stochastic properties based on the linear structure of the master equation. (Appendix C)

Because the non-linear biophysical systems are finite in the sense that only
From the definition in equation (A1) it follows that
\[ p \text{G}(y, o) = \mathbb{E}[x|o] = \mathbb{C}(x|o) = (x|o)^{-1} \mathbb{E}[(x|o)^{-1} x], \]
for the moments
\[ p \mathbb{E}(y, o) = (y|o) = (x|o)^{-1} \mathbb{E}[(x|o)^{-1} y], \]
In particular we obtain for the variance \( \sigma_d(y, o) \) of the stochastic variable \( x(t) \)
\[ \sigma_d(y, o) = \mathbb{C}(x|o) = \mathbb{E}[(x|o)^{-1} x], \]
and the covariance \( \sigma_d(y, o) \), \( i \neq j \):
\[ \sigma_d(y, o) = \mathbb{E}[(x|o)^{-1} x] = \left[ \frac{\partial \mathbb{E}(y, o)^{-1}}{\partial y} \right]_{y = x} = \left[ \frac{\partial^2 \mathbb{E}(y, o)^{-1}}{\partial y \partial y} \right]_{y = x}, \]
\[ \text{APPENDIX B} \]
\[ \text{Solution of equation (49) and equation (13)} \]
With help of the conditional probability \( R(x|y, o) \) which is just the solution of the master equation with the initial condition
\[ n(x, o) = \delta(x - x_0), \]
we obtain for the generation function, using a general initial probability \( n(x, o) \)
\[ G(y, o) = \sum_{x, i} R(x|y, o) \left[ \prod_{i=1}^{N} y_i \right] n(x, o) \]
\[ = \sum_{y, x, o} R(x|y, o) \prod_{i=1}^{N} y_i \]
\[ = \sum_{y, x, o} R(x|o) G(y, o, x). \]
Here we have defined \( G(t) \), which is just the generating function solution of equation (49) for the initial probability given in equation (B1) yielding
\[ G(t, y, o) = \prod_{i=1}^{N} y_i. \]
Due to the blocking property of the binding well, we have:
\[ n(x > 1, x_2, \ldots, x_n, 0^*) = 0, \]
giving
\[ G(y, o) = \frac{1}{\sum_{x_1} \sum_{x_2} \ldots \sum_{x_n} \left( \prod_{i=1}^{N} x_i \right) G(y, o, x_1, x_2, \ldots, x_n, 0^*)} \]
\[ \text{STOCHASTICS OF BIOMOLECULAR KINETICS} \]
\[ Finally \ we \ are \ only \ interested \ in \ \mathbb{G}(y_1 = 0, y_2, \ldots, 0; t; x, 0) \]
\[ Setting \ y_1 = 0 \ we \ obtain \ from \ equation (49) \ a \ linear \ first \ order \ differential \ equation \ for \ \mathbb{G}(y_1, y_2, t; x, 0) \]
\[ \frac{\partial \mathbb{G}(y_1, y_2, t; x, 0)}{\partial t} = \sum_{i=2}^{N} \mathbb{G}(y_1, y_2, t; x, 0) \frac{\partial \mathbb{G}(y_1, y_2, t; x, 0)}{\partial y_i} \]
\[ + \sum_{i=1}^{N} \beta_i (y_1 - 1) \mathbb{G}(y_1, y_2, t; x, 0) \]
\[ \sum_{i=1}^{N} \beta_i y_i \mathbb{G}(y_1, y_2, t; x, 0) \]
\[ \text{with} \]
\[ \mathbb{G}(y, o) = \prod_{i=1}^{N} y_i! \]
yielding
\[ dt = \left\{ \sum_{j=1}^{N} b_j d\gamma_j \sum_{j=1}^{N} b_j d\gamma_j \sum_{j=1}^{N} \gamma_j \right\} \frac{dy_j}{d\gamma_j} \exp \left[ \gamma_j \right] \frac{d\gamma_j}{dy_j} - 1 \). \]  
\[ \text{(B13)} \]

Hence, we obtain with equation (B12) and the property
\[ \gamma_{t=0} = -\sum_{j=1}^{N} \gamma_j \varepsilon_j; \]
\[ \mu_j \sum_{j=1}^{N} b_j d\gamma_j \sum_{j=1}^{N} b_j \sum_{j=1}^{N} \gamma_j \exp \left[ \gamma_j \right] - \text{dln} \sum_{j=1}^{N} b_j \exp \left[ \gamma_j \right]. \]
\[ \text{Equation (B13) can now be integrated immediately giving} \]
\[ I_k(t) = \mu_j t + \ln \sum_{j=1}^{N} b_j \exp \left[ \gamma_j \right] - 1 \]  
\[ \text{for} \ k = 1, \ldots, N \]. \]
\[ \text{(B14)} \]

\[ J(t) = \exp \left[ \mu_j t \right] \sum_{j=1}^{N} b_j \sum_{j=1}^{N} \gamma_j \exp \left[ \gamma_j \right]. \]
\[ \text{(B15)} \]

with \([b]^{1}\) meaning the determinant of the matrix of eigenvectors \(b_0, \ldots, b_N\) of \(M\) and \(B^{1}\) the co-factor element defined by
\[ \frac{b_j}{1} = [b_j] B^{1}. \]
\[ \text{(B16)} \]

\(B^{1}\) denotes the transpose of \(B\).

The solution of equation (B9) can be written as
\[ G_j(t) = \psi \left( J_{t=1}^{\infty} I_k(t) \right), \]
\[ \text{where the functional relationship \(\psi\) between the integrals \(I_k\) and \(J\) is determined independent of time. Solving equation (B15) for \(J\), we obtain for \(G_j(0, 0; a)\)} \]
\[ G_j(y, t; a) = \exp \left[ \mu_j t \right] \sum_{j=1}^{N} b_j \sum_{j=1}^{N} \gamma_j \exp \left[ \gamma_j \right] \delta_{y_a, 0} \]  
\[ \cdot \left( 1 + [b]^{-1} \right) \sum_{j=1}^{N} b_j \exp \left[ \gamma_j \right]. \]
\[ \text{(B17)} \]

Because of the absorbing property of the binding site \(\gamma_{a, 0} = 0\), the matrix \(M\) has a vanishing eigenvalue \(\mu_j = 0\). Eliminating \(J\), we have therefore from equation (B16) for \(G_j(y, t; a)\) the final result:

\[ G_j(y; t; a) = \exp \left[ \frac{\gamma_j t}{\sum_{j=1}^{N} \mu_j} \right] \sum_{j=1}^{N} b_j \exp \left[ \gamma_j \right] \delta_{y_a, 0} \]  
\[ \cdot \left( 1 + [b]^{-1} \right) \sum_{j=1}^{N} b_j \exp \left[ \gamma_j \right]. \]
\[ \text{(B20)} \]
The knowledge of the operator $\Gamma$ and the initial probability $p(o)$ enables us to calculate the Taylor expansion of a mean value $\langle x(t) \rangle$ of the stochastic variable $x(t)$

$$
\langle x(t) \rangle = F_x(p(x) \theta) = \langle x(o) \rangle + \sum_{n \geq 1} \frac{F^{(n)}_x}{n!} dt^n
$$
(C3)

The static moments $p_n$ are given by

$$
p_n = \langle x_n \Gamma^n p(o) \rangle, \quad n = 1, \ldots
$$
$$
= \int x^n \Gamma^n p(x) dx.
$$
(C4)

Next we study the Fourier transform of the function $C(t)$

$$
C(t) = \theta \left( (x(o) - (x(t + \infty)) \right)
$$
$$
= \theta \langle x(t) \rangle,
$$
(C5a)

where $\theta(t)$ denotes the step function. For the Fourier transform $C(\omega)$

$$
C(\omega) = \int e^{i \omega \tau} \langle x(\tau) \rangle d\tau
$$
(C6)

we obtain with equation (C3) the sum rule expansion:

$$
C(\omega) = \sum_{n=\infty}^{-\infty} \frac{F^{(n)}}{\omega^{n+1}}
$$
(C7)

where

$$
p_n = \langle x(o) \rangle - \langle x(\omega) \rangle.
$$
(C8)

The series in equation (C7) is in general semi-convergent or asymptotic. Next we construct a continued fraction expansion which serves as an analytical continuation of the series in equation (C7). With $z = -\omega$ for the corresponding continued fractions are given by (Hanggi & Thomas, 1978; Hanggi, 1977):

$$
C(\omega) = \frac{C_o}{z + \frac{C_1}{z + \frac{C_2}{z + \ldots}}}
$$
(C9)

A general evaluation method for the coefficients in equations (C9)-(C9) consists in the requirement that a formal expansion in powers of $1/2$ of the continued fractions equals those appearing in the asymptotic series. A most convenient method for the calculation of the coefficients consists in a recursive calculation scheme (Hanggi, 1977; Gordon, 1968). Here we outline the recursive scheme (Hanggi, 1977) which is usually numerically more stable than that given by Gordon (Gordon, 1968). Starting with

$$
C_1 = D_1, \quad p_1 = \rho
$$
$$
C_2 = -D_2 D_1, \quad D_2 = p_1
$$
$$
C_3 = -D_3 D_2, \quad D_3 = p_2 + p_1 C_2
$$
$$
C_4 = -D_4 D_3, \quad D_4 = p_3 + p_2 (C_1 + C_3)
$$
(C11)

each proceeds from $n = 4$ to the higher terms in the following way: using the auxiliary vector $x$ of dimension $L$

$$
L = 2 \text{ integer } \{(n-1)/2\}
$$
$$
x(2) = x(3) + x(1) = C_1
$$
(C12)

interchange

$$
x(2) \rightarrow x(1), \quad x(1) \rightarrow x(2)
$$
(C13)

with $x(L-1) = 0$ we work upwards:

$$
x(K) = x(K-1) + C_1 x(K-2)
$$
$$
K = L, L-1, \ldots, 4
$$
$$
x(2) = x(1) + C_{n-1}
$$
(C14)

interchange the odd and even component, i.e.

$$
x(2) \rightarrow x(3), \quad x(4) \rightarrow x(2)
$$
$$
x(1) \rightarrow x(2), \quad x(3) \rightarrow x(4)
$$
etc.

The continued fraction coefficient $C_n$ is then given by

$$
C_n = -D_n D_{n-1}
$$
(C15a)

$$
D_n = p_{n+1} + \frac{D_{n+1}}{C_{n+1}} X(2n+1)
$$
(C15b)

With the $C_n$ evaluated by equation (C15), the coefficients of the continued fraction equation (C10) are simply given by the relations

$$
b_1 = C_1
$$
$$
a_1 = -C_2
$$
$$
b_n = -C_n a_{n-1} - b_{n-2}
$$
(C16a)

$$
a_n = -C_{n+1} b_{n-1} - C_{n+2} b_{n-2}
$$
(C16b)

The function $C(\omega)$ (and therefore $\langle x(t) \rangle$ for all times $t$) is then given by the inverse Fourier transform

$$
C(\omega) = \frac{1}{2 \pi} \int C(\omega) e^{-i \omega \tau} d\omega.
$$
(C17)