Contemporary Mathematics
http://ojs.wiserpub.com/index.php/CM/

Research Article

Transcription in Molecular Biochemistry: An Idealized Physical-Mathematical Model

Ramón F. Álvarez-Estrada[1*], Hélia Serrano[2]

[1] Department of Theoretical Physics, Faculty of Physics, Complutense University of Madrid, 28040 Madrid, Spain
[2] Department of Mathematics, Faculty of Chemical Sciences and Technologies, University of Castilla-La Mancha, 13071 Ciudad Real, Spain

E-mail: rfa@ucm.es

Received: 19 January 2026; **Revised:** 16 March 2026; **Accepted:** 31 March 2026

Abstract: Transcription in molecular biochemistry has originated various established physicochemical descriptions: Brownian-ratchet and energy-landscape models for elongation/pausing, kinetic schemes, thermodynamic promoter-binding models, structural-mechanistic approaches and other related ones. A new idealized physical-mathematical model for transcription, complementary to those existing models, is proposed here: it emphasizes a minimal reaction-diffusion/Smoluchowski formulation well suited for spatial-temporal descriptions. The model describes how, in the presence of a molecular chain (the *template*), two initially unbound small molecules (free monomers) become covalently bound to each other, forming a new molecular chain (the transcript) that copies the *template*. The process occurs under the combined action of a promoter and a terminator, attached to the ends of the *template*, and of a moving enzyme in a non-equilibrium state. The enzyme is represented by its active site and is subject to a constant effective force, interpreted as arising from energy released in internal transitions of free monomers and transferred to the enzyme (molecular-motor action). All entities evolve in an aqueous medium at rest, in thermal equilibrium at (approximately) room temperature. The new coarse-grained dynamics is based upon successively simplifying and approximate non-equilibrium non-linear coupled Smoluchowski equations: (i) containing a new quantum correction that enables the bond-forming chemical reaction generating the transcript, and (ii) implementing an effective detachment mechanism for the transcript, the enzyme, and the *template* at termination. Other novelties in the new model are approximate mean first-passage times (τ's) for the three basic processes (transcript formation, its overall motion during binding and enzyme motion). Approximate theoretical formulae are given: τ_b, $\tau_{c,1}$ and $\tau_{E,1}$, respectively, as integrals with integrands dominated by exponentials. For suitable effective interactions, estimates indicate that the integrals yielding τ_b, $\tau_{c,1}$ and $\tau_{E,1}$ be, in orders of magnitude, about 10^{-2} s and, so, about known typical durations of transcription (per pair of new bound monomers in the transcript). A generalization to longer *templates* and transcripts is outlined.

***Keywords*:** transcription, moving enzyme, Smoluchowski equations, chemical reaction, quantum effects, mean first-passage times (MFPT)

MSC: 35K57, 81S30, 92C40

Copyright ©2026 Ramón F. Álvarez-Estrada, et al.
DOI: https://doi.org/10.37256/cm.7420269638
This is an open-access article distributed under a CC BY license
(Creative Commons Attribution 4.0 International License)
https://creativecommons.org/licenses/by/4.0/

Abbreviation

N_c	Number of initially free monomers involved in transcription (and, in Appendix C, also the number of monomers in the active template subchain and in the transcript).
N_c'	Number of bound monomers in the active subchain Te of the template (Appendix C); taken as $N_c' = N_c$.
N_l	Number of monomers in the left spectator subchain of Te_{ext} (Appendix C).
N_r	Number of monomers in the right spectator subchain of Te_{ext} (Appendix C).
Te_{ext}	Extended template chain (template plus promoter P', terminator T' and spectator subchains).
Te	Active subchain of Te_{ext} that is copied into the transcript; for $N_c = 2$ it contains the bound monomers $1'$ and $2'$.
c	Transcript chain formed by covalently binding the free monomers (copy of Te in the toy model).
E	Enzyme (RNA polymerase like motor) that moves along the template and mediates transcript formation.
P'	Promoter region/molecule attached to the template (boundary/initial attractor for E).
T'	Terminator region/molecule attached to the template (repels E at termination).
$1,2,\ldots,N_c$	Labels for the initially free monomers that will form the transcript.
$1',2',\ldots,N_c'$	Labels for the monomers bound in the template active subchain Te (used as the copy template).
$u1$, $u2$, $u3$ (or u_i)	Toy-model internal degrees of freedom representing energy-releasing internal transitions in free monomers (Appendix B).
t	Time variable.
x_1, x_2	3D position vectors (centers of mass) of free monomers 1 and 2.
$x_{1'}$, $x_{2'}$	3D position vectors (centers of mass) of the bound monomers $1'$ and $2'$ in Te.
x_E	3D position vector (center of mass) of the enzyme E.
$x_{P'}$, $x_{T'}$	3D position vectors of promoter P' and terminator T' (integrated out in reduced descriptions).
$x_{u,i}$	Internal position variable for the u_i degree of freedom (Appendix B).
x	One-dimensional spatial coordinate: in Appendix A, the coordinate of a quantum particle; in Section 4.3, the coordinate along the direction selected by the force F for the MFPT of E.
q	Momentum canonically conjugate to x (Appendix A).
q_ρ	Momentum variable canonically conjugate to x_ρ for entity/degree of freedom ρ (Appendix B).
X_c	Overall coordinate of the transcript-forming pair (center-of-mass like coordinate): $X_c = (a_1^{-1}x_1 + a_2^{-1}x_2)/M_c$ (and its N_c-monomer generalization in Appendix C).
x_c	Relative coordinate of monomers 2 and 1: $x_c = x_2 - x_1$.
X_{Te}	Overall coordinate of the two-bead template segment: $X_{Te} = (a_{1'}^{-1}x_{1'} + a_{2'}^{-1}x_{2'})/M_{Te}$ (Appendix B).
x_{Te}	Relative coordinate in the two-bead template segment: $x_{Te} = x_{2'} - x_{1'}$ (Appendix B).
r	Radial separation between free monomers in the relative coordinate: $r = \lvert x_c \rvert$.
$r_{1,0}$, $r_{1,1}$, $r_{1,2}$	Characteristic radii delimiting repulsive/attractive regions of $V_{12}(r)$ (Section 4.1).
r_{ref}	Reflecting boundary radius chosen near the minimum of V_{12} (Section 4.1).
r_{abs}	Absorbing boundary radius (chosen $\simeq r_{1,2}$) used in the MFPT for r (Section 4.1).
r_c	Radial coordinate for the overall motion of the pair/transcript center: $r_c = \lvert X_c \rvert$ (Section 4.2).
$r_{\mathrm{c,min}}$, $r_{c,0}$, $r_{\mathrm{c,max}}$	Inner/transition/outer radii used in the MFPT problem for r_c (Section 4.2).
x_{min}, x_0, x_{max}	Left/transition/right positions used in the one-dimensional MFPT problem for E (Section 4.3).
y_i	Relative coordinate between consecutive free monomers in Appendix C: $y_i = x_{i+1} - x_i$, $i = 1,\ldots,N_c - 1$.

$\nabla_{y_0} \equiv 0, \nabla_{y_{N_c}} \equiv 0$	Boundary conventions for gradients in Appendix C.
k_B	Boltzmann's constant.
T	Absolute temperature of the aqueous medium (heat bath).
η	Viscosity of the aqueous medium.
a_{av}	Average radius of a small molecule (used in Stokes Einstein diffusion estimate).
r_k	Approximate radius of a sphere containing one molecule of type k.
a_k	Dimensionless size ratio $a_k = a_{\text{av}}/r_k$; the diffusion coefficient of type k is taken as Da_k.
D	Reference diffusion coefficient for small molecules in water: $D = k_B T/(6\pi\eta a_{\text{av}})$ (and set to a typical numerical value in the manuscript).
a_{EE}	Dimensionless size ratio for the enzyme: $a_{EE} = a_{\text{av}}/r_E$; its diffusion coefficient is Da_{EE}.
D_u	Diffusion coefficient associated with the internal variables $x_{u,i}$ (Appendix B).
a_{iu}	Dimensionless constants setting the diffusion scale for the internal variable of monomer i (Appendix B).
M_c	Dimensionless total-mass like parameter for monomers 1 and 2: $M_c = a_1^{-1} + a_2^{-1}$ (generalized to $\sum_{i=1}^{N_c} a_i^{-1}$ in Appendix C).
m_c	Dimensionless reduced-mass like parameter for monomers 1 and 2 defined by $m_c^{-1} = a_1 + a_2$.
M_{Te}	Dimensionless total-mass like parameter for the two-bead template segment: $M_{Te} = a_{1'}^{-1} + a_{2'}^{-1}$ (Appendix B).
m_{Te}	Dimensionless reduced-mass like parameter for the two-bead template segment: $m_{Te}^{-1} = a_{1'} + a_{2'}$ (Appendix B).
m	Mass of the quantum particle in Appendix A.
$m_{0,\rho}$	Mass associated with degree of freedom/entity ρ in the extended Wigner description (Appendix B).
$\hbar$	Reduced Planck constant.
β	Inverse thermal energy: $\beta = (k_B T)^{-1}$.
$V(x)$	External potential in the one-particle quantum model of Appendix A (even in x, attractive in a bounded region, vanishing at infinity).
H	Quantum Hamiltonian operator (Appendix A): $H = -(\hbar^2/2m)\, d^2/dx^2 + V$.
V_{ext}	Full interaction potential of the extended transcription system before reductions (Appendix B).
V_{eff}	Effective potential after integrating over P', T' and internal variables (Appendix B).
$V_{\text{eff,tot}}$	Total effective potential including the motor term: $V_{\text{eff,tot}} = V_{\text{eff}} - F \cdot x_E$.
V_{12}	Interaction potential between the two free monomers 1 and 2; depends on $r = \lvert x_c \rvert$.
$V_{1',2'}$	Binding potential between template monomers $1'$ and $2'$ (modeled as a Gaussian-chain/Rouse interaction).
$V_{P',1'}$, $V_{2',T'}$	Potentials describing binding/influence of promoter P' on $1'$ and of terminator T' on $2'$ (Appendix B).
$V_{E,P'}$, $V_{E,T'}$	Potentials describing influence of P' and T' on the enzyme E (Appendix B).
$V_{\text{P}',1',\text{eff}}$, $V_{2',\text{T}',\text{eff}}$	Effective versions of $V_{P',1'}$ and $V_{2',T'}$ after integrating over P' and T' (Appendix B).
$V_{\text{E,P}',\text{eff}}$, $V_{\text{E,T}',\text{eff}}$	Effective versions of $V_{E,P'}$ and $V_{E,T'}$ after integrating over P' and T' (Appendix B).
V_{Eall}	Weak multi-body interaction coupling E with free and template monomers; depends on the set of pairwise distances.
$V_{\text{Eall,eff}}$	Effective potential obtained by averaging V_{Eall} with W_E and W_{Te} in the equation for W_c.
$V_{\text{Eall,eff,re}}$	Effective potential entering the $W_{c,re}$ equation after factorization/averaging.
$V_{\text{Eall,eff,E}}$	Effective contribution of V_{Eall} to the E.
$V_{\text{c,f,eff}}$	Effective potential driving the overall (center-of-mass) coordinate X_c of the pair.
$V_{\text{MFPT,re}}$	Time-independent effective potential used to approximate $V_{\text{Eall,eff,re}}$ in MFPT estimates for r (Section 4.1).

$V_{\mathrm{MFPT,f}}$	Time-independent effective potential used to approximate $V_{\mathrm{c,f,eff}}$ in MFPT estimates for r_c (Section 4.2).
$V_{\mathrm{MFPT,E}}$	Time-independent effective potential used to approximate $V_{\mathrm{Eall,eff,E}} + V_{\mathrm{E,P',eff}} + V_{\mathrm{E,T',eff}}$ in MFPT estimates for E (Section 4.3).
V_{u1}, V_{u2}, V_{u3}	Internal potentials coupling E to the internal variables $u1, u2, u3$ and enabling energy transfer that produces the effective force F (Appendix B).
F	Constant effective force acting on E (molecular-motor action), arising from integrating out internal transition variables; appears via $-F \cdot x_E$.
F_1	Contribution to the effective force from the $u1$ internal transition (Appendix B).
$\rho(t)$, $\rho(x,x',t)$	Quantum density operator and its coordinate representation (Appendix A).
ρ_{in}	Initial density operator at $t = 0$ (Appendix A).
ρ_{eq}	Equilibrium density operator: $\rho_{eq} = \exp(-H/(k_B T))$ (Appendix A).
$W(x,q,t)$	Wigner function of the one-particle quantum model (Appendix A).
W_{in}	Initial Wigner function at $t = 0$ (Appendix A).
$W_{eq}(x,q)$	Equilibrium Wigner function corresponding to ρ_{eq} (Appendix A).
W_{ext}	Extended Wigner function for the many-body transcription system including internal degrees of freedom (Appendix B).
$W_{ext,0}$	Lowest-order (momentum-integrated) moment of W_{ext} (Appendix B).
$W_{ext,0;in}$	Initial condition for $W_{ext,0}$ at $t = 0$ (Appendix B).
W_0	Reduced distribution for $(1,2,1',2',E)$ after integrating out P', T' and internal variables (Appendix B); used in Section 3 as $W_0(x_1,x_2,x_{1'},x_{2'},x_E,t)$.
W_E	Marginal distribution of the enzyme E: $W_E(x_E,t) = \int W_0\, dx_1 dx_2 dx_{1'} dx_{2'}$.
W_{Te}	Marginal distribution of the template segment: $W_{Te}(x_{1'},x_{2'},t) = \int W_0\, dx_1 dx_2 dx_E$; modeled as a two-bead Gaussian chain.
$W_{Te,re}$	Relative-coordinate distribution for the two-bead template segment.
W_c	Distribution for the free-monomer pair / transcript: $W_c(X_c,x_c,t) = \int W_0\, dx_{1'} dx_{2'} dx_E$.
$W_{c,f}$	Marginal of W_c over relative coordinate: $W_{c,f}(X_c,t) = \int W_c\, dx_c$.
$W_{c,re}$	Marginal of W_c over overall coordinate: $W_{c,re}(x_c,t) = \int W_c\, dX_c$.
$W_{c,re,eq}$	Equilibrium (t-independent) solution of the $W_{c,re}$ equation for large t (Section 3.3).
$W_{\mathrm{Te,active}}$, $W_{\mathrm{Te,left}}$, $W_{\mathrm{Te,right}}$	Approximate distributions for the active and spectator subchains of Te_{ext} in Appendix C.
∇	Gradient operator (with subscripts indicating differentiation with respect to the corresponding coordinate).
$\delta^{(3)}$	Three-dimensional Dirac delta distribution (used for point-like initial conditions).
M_Q	Quantum potential operator in the Wigner equation (Appendix A and Appendix B).
M_0	Reduced operator coupling moments in Appendix A and used in the long-time approximation leading to the Smoluchowski form.
$H_{Q,n}(q)$	Orthogonal polynomials (in momentum q) generated by the equilibrium Wigner weight W_{eq} (Appendix A).
$W_n(x,t)$	Non-equilibrium moments of W with respect to $H_{Q,n}$ with $n = 0,1,2,\ldots$ (Appendix A).
j_{EE}	Effective drift field appearing in the Smoluchowski equation for W_E.
J_E, J_{EE}	Probability-current like quantities used in the multi-variable Smoluchowski equations (Appendix B).
θ	Heaviside step function used to split integrals in MFPT approximations (Section 4.1).
$\varphi_b, \varphi_k, \varphi_j$	Quantum eigenfunctions (bound state b, continuum states indexed by k, and generic spectral index j) of the Hamiltonian H (Appendix A).
E_b, E_k, E_j	Energy eigenvalues corresponding to φ_b (bound), φ_k (continuum) and generic φ_j (Appendix A).
k	Wavevector labeling continuum eigenstates φ_k (Appendix A).

$\varepsilon(x)$	Dimensionless quantum correction defined from the equilibrium Wigner function by Eq. (A.4); modifies the effective Smoluchowski operator.
ε	Classical limit of $\varepsilon(x)$: $\varepsilon = -1/2$.
ε_c	Quantum-correction function accounting for the bond-forming reaction between free monomers 1 and 2; depends on $\|x_c\|$ via $\varepsilon_c = \varepsilon_x(\|x_c\|)$.
$\varepsilon_{i'}$, ε_{Te}	Quantum-correction functions associated with template-related degrees of freedom ($i' = P', 1', 2', T'$) in Appendix B.
$\varepsilon_i(y_i)$	Quantum-correction functions for consecutive-bond coordinates y_i in the $N_c > 2$ generalization (Appendix C).
$\tau(r)$	Mean first-passage time (MFPT) associated with the relative separation r of monomers 1 and 2 under the MFPT-approximated dynamics (Section 4.1).
τ_a, τ_b	Contributions to $\tau(r)$ obtained by splitting integrals across different r-regions (Section 4.1).
$\tau_c(r_c)$	MFPT associated with the overall-coordinate radial variable $r_c = \|X_c\|$ (Section 4.2).
$\tau_{c,1}, \tau_{c,2}$	Integral contributions to $\tau_c(r_c)$.
$\tau_E(x)$	MFPT for the enzyme position along the force direction (Section 4.3).
$\tau_{E,1}, \tau_{E,2}$	Integral contributions to $\tau_E(x)$ (Section 4.3).
τ_{free}	Free-diffusion estimate of an MFPT used for order-of-magnitude comparisons (Section 5).
$f(r,t)$	Radial probability density used in the MFPT derivation for $W_{c,re}$ after removing the Jacobian factor (Section 4.1).
$f_c(r_c,t)$	Radial probability density $f_c = r_c^2 W_{c,f}$ used in the MFPT derivation for $W_{c,f}$ (Section 4.2).
$J(r')$	Auxiliary integral function appearing in the formal MFPT solution for $\tau(r)$.
ρ, ρ', ρ'' and σ	Multi-indices used in Appendix B to label the various entities/degrees of freedom in the extended Wigner function (e.g., $\rho \in \{P', 1, 2, 1', 2', E, T', u1, u2, u3\}$).
$[dx]$, $[dq]$, $[dx']$, $[dq']$	Shorthand differential volume elements over the sets of position/momentum variables in Appendix B.
Π	Product symbol used to denote products over sets of indices (Appendix B).
$\mathbb{R}$	Real line (used for continuum integration over k).
i	Imaginary unit, $i^2 = -1$ (Appendix A and Appendix B).
$(\cdot)^*$	Complex conjugation (used for wavefunctions and density-matrix elements).
nm	Nanometer (unit of length) as used in the numerical estimates.
eV	Electronvolt (unit of energy) as used in the numerical estimates.
pN	Piconewton (unit of force) as used in the numerical estimates.
MFPT	Mean First-Passage Time (acronym).
d	Chain length (average bead separation) of the two-bead Gaussian-chain approximation for *Te* (Section 3.3).
r_E	Approximate size (radius) of the enzyme E used to form $a_{EE} = a_{\mathrm{av}}/r_E$.
$r_{P'}$, $r_{T'}$	Effective length parameters summarizing the influence of promoter P' and terminator T' on the overloaded template beads (Appendix B).
$r_{1'}, r_{2'}$	Effective radii of the template beads $1'$ and $2'$ used in $a_{1'} = a_{\mathrm{av}}/(r_{1'} + r_{P'})$ and $a_{2'} = a_{\mathrm{av}}/(r_{2'} + r_{T'})$ (Appendix B).
a_1, a_2	Dimensionless size ratios of free monomers 1 and 2 (special cases of $a_k = a_{\mathrm{av}}/r_k$).
$a_{1'}, a_{2'}$	Dimensionless size ratios of template beads $1'$ and $2'$ overloaded by P' and T' (Appendix B).
a_i	Dimensionless size ratios for the i-th free monomer in the $N_c > 2$ generalization (Appendix C).
Da_{EE}	Diffusion coefficient of E in the Smoluchowski equations, written as the product Da_{EE}.
$W_{E,st}$	Formal stationary solution of the simplified large-$\|x_E\|$ equation (25); noted as not physically acceptable as an equilibrium state.
$\langle\cdot\rangle$	Expectation value with respect to the equilibrium weight (e.g., with respect to W_{eq} in Appendix A).

$\langle \exp f \rangle$	Mean value of $\exp(f)$ defined by $\int a(x)\exp(f(x))dx = \langle \exp f \rangle \int a(x)dx$ (Section 5).
Subscripts $\mathrm{in}, \mathrm{eq}, \mathrm{st}$	Notation for initial ($t = 0$), equilibrium (time-independent thermal equilibrium), and stationary solutions, respectively.
Subscripts eff and ext	Notation for effective (after averaging/integration) and extended (before reductions) quantities, respectively.
Subscripts re and f	Notation for relative-coordinate (re) and overall/free (f) marginals of W_c.
$V_{i,i+1}(y_i)$	Interaction potential between consecutive free monomers i and $i+1$ as a function of their relative coordinate y_i (Appendix C; written as $V_{ii,+1}$ in the manuscript).
$V_E(i, i+1, i', i+1')$	Building-block contribution to V_{Eall} coupling E to a pair of consecutive free monomers $(i, i+1)$ and their template counterparts $(i', i+1')$ (Appendix C).
$X_{Te,in}$	Initial value of the template overall coordinate X_{Te} at $t = 0$ (Section 3.3).
$x_{E,in}$	Initial position of the enzyme (or its reduced coordinate) at $t = 0$ (Section 3.3).
$X_{c,in}$	Initial value of the overall coordinate X_c at $t = 0$ (Section 3.3).
$W_{P'}$	Marginal distribution for the promoter coordinate $x_{P'}$ used in Appendix B factorizations.
$W_{T'}$	Marginal distribution for the terminator coordinate $x_{T'}$ used in Appendix B factorizations.
W_u	Marginal distribution over internal variables $[x_u]$ (Appendix C).
$V_{\mathrm{MFPT,c}}$	Name used in the manuscript for the MFPT-level approximation to $V_{\mathrm{c,f,eff}}$ (implemented as $V_{\mathrm{MFPT,f}}$ in the equations).
a_σ	Dimensionless coefficients used in Appendix B to approximate first momentum moments in terms of $W_{ext,0}$ (generalizing the single-particle coefficient structure in Appendix A).
$[x_\sigma]$, $[q_\sigma]$	Shorthand for the collections of all position and momentum variables in the extended system (Appendix B).
$V_{11'}$, $V_{22'}$	Additional interaction potentials between free monomer 1 and template monomer $1'$ (and similarly 2 with $2'$), mentioned as possible extensions (Appendix B).
$V_{E,PT}$	Notation appearing in Appendix B for the potential describing the influence of promoter/terminator on E (used interchangeably with $V_{E,P'}$ and $V_{E,T'}$ in the toy model).
V_{uj}	Generic notation for the internal transition potentials ($j = 1, 2, 3$) coupling E to internal variables in Appendix B (with explicit cases V_{u1}, V_{u2}, V_{u3}).
x_ρ, q_ρ	Generic notation for the position/momentum variables of the entity labeled by index ρ in Appendix B.

1. Introduction

Transcription is a central process in molecular biochemistry and molecular biology. In its most common biological realization, an RNA polymerase reads a segment of one strand of a double-stranded DNA molecule and synthesizes a complementary RNA chain. Initiation and termination are strongly regulated by promoter and terminator regions and their associated macromolecular complexes. Standard biochemical accounts can be found, for instance, in [1–6]. For wide presentations on physical biochemistry (conformations and dynamics of biological macromolecules, applications to them of statistical thermodynamics, light-matter interaction, a variety of optical techniques, kinetics of enzymes, mathematical modeling,...), see [7–18]. Extensions arise in virology and related settings, where additional actors and regulatory layers may enter; see [19] for an updated account.

From the viewpoint of physics and applied mathematics, transcription combines stochastic transport in a viscous aqueous environment with localized binding/unbinding events and energy-consuming chemical steps that bias motion and drive polymer growth. Single-molecule and sequencing-based measurements have motivated a broad range of quantitative descriptions. Pause-interspersed elongation and backtracking have been discussed within energy-landscape and Brownian-ratchet type frameworks and related stochastic models [20–22]. When several polymerases occupy the same transcription unit, crowding and steric interactions lead naturally to "traffic" and exclusion-type descriptions [23, 24]. On the statistical side, Bayesian and likelihood-based inference methods have been developed to estimate kinetic parameters

from stochastic transcription trajectories and nascent-RNA data [25, 26]. High-resolution structural work has also shown how transcription factors can modulate RNA polymerase conformational equilibria [27]. At larger scales, polymer models connect transcriptional activity with three-dimensional chromatin organization and micro-compartmentalization [28]. More phenomenological generalized master-equation descriptions that couple transcription with other stages of the mRNA life cycle have also been proposed [29]. First-passage concepts also continue to play a role in current work on transcriptional fidelity and termination mechanisms, see, e.g., [30, 31].

The aim of the present paper is different. Upon summarizing how it is planned below in this Section, we shall also be outlining the physical and mathematical aspects in it and the links between the latter two. We propose a new three-dimensional idealized physical-mathematical model designed to be analytically tractable while retaining several ingredients that are essential from a stochastic-process viewpoint: (a) a template chain Te with two distinguished ends and free monomers; (b) a moving enzyme E (treated at the level of an effective active site) that interacts weakly with the template and with free monomers; (c) a promoter P' and a terminator T' attached to the ends of Te and producing, respectively, trapping and repulsive effects on E; (d) a nonequilibrium driving of E represented by an effective constant force, interpreted as the coarse-grained outcome of internal energy-releasing transitions; and (e) a minimal quantum-mechanical remnant that permits the bond-forming chemical reaction that turns initially free monomers into a transcript c. The dynamics of the centers of mass of the interacting entities is described by (overdamped) Smoluchowski equations, extended to incorporate the above reaction and the final detachment stage.

The largest part of the work deals with a model for two initially free monomers. The logical sequence of arguments proceeds through the following successively simplifying and approximate steps for Smoluchowski evolution equations. We summarize here all steps from 0 to 10. Step 0: we start by considering the simple quantum-statistical Smoluchowski equation (A13) in one spatial dimension in Appendix A. Step 1: then, in Appendix B, we consider the evolution equation (B1) for the general Wigner distribution function W_{ext}. Step 2: moreover, in Appendix B, we assume that the moment $W_{ext,0}$ satisfies the evolution equation (B4). Step 3: in Appendix B, by integrating the moment $W_{ext,0}$, we obtain the approximate (irreversible) quantum Smoluchowski equation (B14) for the quantum Wigner function W_0, which, playing a distinguished role, we may expose here:

$$\frac{\partial W_0}{\partial t} = D\frac{1}{M_c}\nabla_{X_c}\left(\nabla_{X_c}W_0 + \frac{W_0}{k_BT}\nabla_{X_c}V_{\text{eff,tot}}\right) + D\frac{1}{m_c}\nabla_{x_c}\left(\nabla_{x_c}(-2\varepsilon_c)W_0 + \frac{W_0}{k_BT}\nabla_{x_c}V_{\text{eff,tot}}\right)$$

$$+ D\sum_{i'=1',2'} a_{i'}\nabla_{x_{i'}}\left(\nabla_{x_{i'}}(-2\varepsilon_{Te})W_0 + \frac{W_0}{k_BT}\nabla_{x_{i'}}V_{\text{eff,tot}}\right) + Da_{EE}\cdot\nabla_{x_E}J_{EE}. \tag{1}$$

Step 4: in Subsection 3.2, we integrate the previous Smoluchowski equation (1), for the quantum Wigner function W_0, and define in (5) the distribution function W_E for E, which satisfies the evolution equation (8). Step 5: besides, in Subsection 3.2, we integrate the Smoluchowski equation (1), and define in (7) the distribution function W_c, which satisfies the evolution equation (10). Step 6: in Subsection 3.3, we integrate W_c in (10), and define the distribution function $W_{c,f}$ in (18), which satisfies the evolution equation (20). Step 7: again, we integrate W_c in (10), and now define the distribution function $W_{c,re}$ in (19), which satisfies the evolution equation (22). Step 8: therefore, the time evolution of the distribution function W_E in (8) is approximated by equations (24) and (25). Step 9: we obtain the approximate evolution equation (28) for the distribution function $W_{c,f}$. Step 10: finally, in Subsection 3.3, we obtain the approximate evolution equations (30) and (31) for the distribution function $W_{c,re}$.

Notice that, since Steps 0, 1, 2 and 3 provide the necessary foundations and a more detailed understanding of our proposed model, they are deferred to Appendix A and B. Appendix B (Steps 1, 2 and 3) extends Appendix A to the seven-body setting underlying the $N_c = 2$ transcription toy model and leads to the effective Smoluchowski description used in Section 3. Appendix C sketches the generalization to longer templates and transcripts ($N_c > 2$). In order to facilitate entering directly into the proposed model, the effective equations in it are given at Step 4 in Section 3. Even with the restriction to $N_c = 2$, the coupled stochastic problem is non-trivial. Afterwards, the presentation follows the logical

sequence up to and including Step 10. In turn, Steps 6–10 can be regarded as the basis for an approximate algorithm. Section 2 discusses qualitative assumptions on E, Te, P', T' and on the initially free N_c monomers and summarizes a simplified physical picture for the formation of c. Section 3 (Steps 4, 5, 6, 7, 8, 9 and 10) specializes to $N_c = 2$ initially free monomers and formulates directly an effective model (its justification being deferred to Appendices A and B). Section 4 deals with Mean First-Passage Times (MFPTs) for the actors in the $N_c = 2$ model. Section 5 outlines representative orders of magnitude, and Section 6 contains conclusions and discussion. Section 7 presents a list of symbols and variables mentioned in the text.

2. Qualitative assumptions and physical picture

In the present Section, we shall describe in more detail the basic constituents and interactions of the proposed model, and comment briefly at the end the physical aspects of evolution in it, thereby preparing for the mathematical formulation and the more detailed link between physics and mathematics in the next Section.

The medium here is an aqueous fluid at equilibrium (at rest) and at thermal equilibrium at absolute temperature T in a range about room temperature. It occupies a volume about the micron scale (interpreted here, mathematically, as the whole three-dimensional space). Free and bound monomers, even if relatively small molecules, are in fact complex structures involving a non small number of degrees of freedom. For instance, let the free monomers 1 and 2 be formed each by a nucleoside triphosphate (containing ribose, a base and a triphosphate, covalently linked among themselves). In transcription, both monomers become covalently bound to each other by means of a phosphodiester bond, with the concomitant release of a molecule of diphosphate from one of the monomers. Due to the number and complexity of the degrees of freedom involved for free monomers (and for bound ones as well), interactions will have to be modeled. As a possibility, one could try a framework known generically as an *optical potential* approach (integrating somehow over unwanted degrees of freedom) employed, for instance, in Nuclear Physics, see [32]. However, as even the *optical potential* approach is not easy to handle, it is far more economical to approximate, more drastically, the actual interactions through suitable effective potentials, as discussed below.

As a first approximation, as we refer to free and bound monomers, we shall be dealing only with their bulk properties, that is, we shall consider only the positions of their centers of mass (concentrating in each the corresponding mass) and disregard their internal structures and quantum features (with one exception below). The subchain Te is formed by N' monomers, bound sequentially and permanently among themselves by strong covalent forces, so as to constitute a large linear molecular chain. The promoter P' (a large linear molecule) and the terminator T' are permanently and covalently bound to the monomers $1'$ and N' of Te, respectively.

Figure 1 displays the initial interaction between a linear molecular chain (formed by the promoter P', the monomers $1'$ and $2'$, and the terminator T'), an enzyme E and two initial free monomers (ie the situation $N_c = 2$). Namely, the enzyme E approaches, thanks to the force F acting on it, and becomes weakly attracted by the promoter P'. Te, P' and T' can also constitute some part of an even larger linear molecular chain (see Appendix C). The larger linear molecular chain, always containing P', Te and T', will be denoted as Te_{ext}. We shall suppose that Te_{ext} diffuses very slowly in the medium.

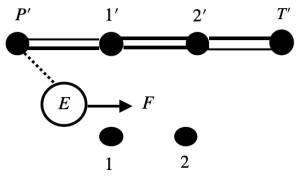


Figure 1. Promoter P', monomers $1'$ and $2'$, and terminator T' are permanently bound among themselves by strong covalent forces, forming a linear molecular chain. The interactions between any of them and enzyme E, and between enzyme E and the free monomers 1 and 2, are weak

The enzyme E is a large and complex molecule with a globular or packed-ball structure (with packed-ball radius a few *nm*) which, as a matter of fact, remains approximately unaltered in transcription. Specifically, in the present simplified analysis, the following approximate properties of E (typical of enzymatic action) will be supposed to hold (and will suffice):

i) its internal properties will not change from the beginning to the end of transcription;

ii) as far as the interactions of E relevant to transcription are concerned, E (disregarding its extension) is approximated by certain reduced part of it (the so-called active site);

iii) just for simplicity, the active site will be supposed to include or coincide with the center of mass of E.

To simplify, E will diffuse in the medium essentially due to the actions of the other actors in transcription and of a constant vector F (having the physical dimension of force), which will play an essential role. The physical generation of F, arising from certain transitions in free monomers, will be discussed below.

As an exception, we consider in Appendix B (exemplified for $N_c = 2$) a short and simplified model of certain decisive transitions affecting internal degrees of freedom of free monomer, which liberate energy to be fed into E and, so, contribute to F (a molecular motor). Such internal degrees of freedom will be regarded as decoupled from the position of the center of mass of the free monomer. Otherwise, we shall disregard the internal structures and quantum effects of the free and bound monomers.

The interactions among different monomers (bound or initially free) will be modeled by local effective (time-independent) potentials depending on the position vectors of their centers of mass. The effective potential between a pair of free monomers (like V_{12} in Subsection 3.1) is assumed to have short range and to be strongly attractive on average (attractive at intermediate distances and repulsive at very short and larger distances, inside its range). So are the internal potentials acting inside Te_{ext}. The various effective potentials between any bound monomer of Te and E, those between E and any free monomer and between E and P' will be supposed to be repulsive at very short distances and, on average, weakly attractive at intermediate and large distances (being allowed to have, possibly, somewhat longer ranges than those between pairs of free monomers). The interaction between the terminator T' and the enzyme E will be supposed to be weakly repulsive and to have similar range. Anyway, all potentials vanish beyond suitably large distances. In short, P', Te (and in some sense, T' as well) and E are determinant for the formation of the transcript: Te_{ext} will not alter the internal state of E and conversely, and none of the latter two will be altered under the action of either the free monomers or c. Te_{ext} is supposed to be always more massive than E which, in turn, will be more massive than any of the free monomers and c. Crucial additional assumptions will be: first, Te_{ext} and E do not form permanent bound states; second, there are no permanent bound states between either any free monomer or c and either Te_{ext} or E.

We now outline the physical evolution and suggest the kind of new mathematical formulation required by it and to be developed, in a direct way, from Section 3.2 onwards. In short: the centers of mass of the free monomers will be essentially regarded as interacting classical Brownian particles (following [33–35], in which quantum features are mostly disregarded) except for one very important quantum remnant here, implementing the possibility of chemical reactions. In fact, the free monomers are subject to chemical reactions among them (approximated by effective chemical reactions among their centers of mass) giving rise, by binding successively among themselves, one by one, to the formation of the transcript, under the combined action of P', Te, T' and E. Figure 2 displays the situation for $N_c = 2$, and Figure C1 for $N_c > 2$. Such chemical reactions among free monomers require to invoke quantum-mechanical effects through a new mathematical recipe baed upon Appendix A, as it will be seen.

3. Model for $N_c = 2$ free monomers

3.1 *Description and interactions*

In this Subsection, we shall discuss in more detail physical aspects of the two free monomers situation ($N_c = 2$), and start the associated mathematical description. Throughout Sections 3, 4 and 5, the extended template chain Te_{ext} will contain only the promoter P', two bound monomers $1'$ and $2'$ in Te, and the terminator T'. A possible qualitative description of transcription for $N_c = 2$ is as follows.

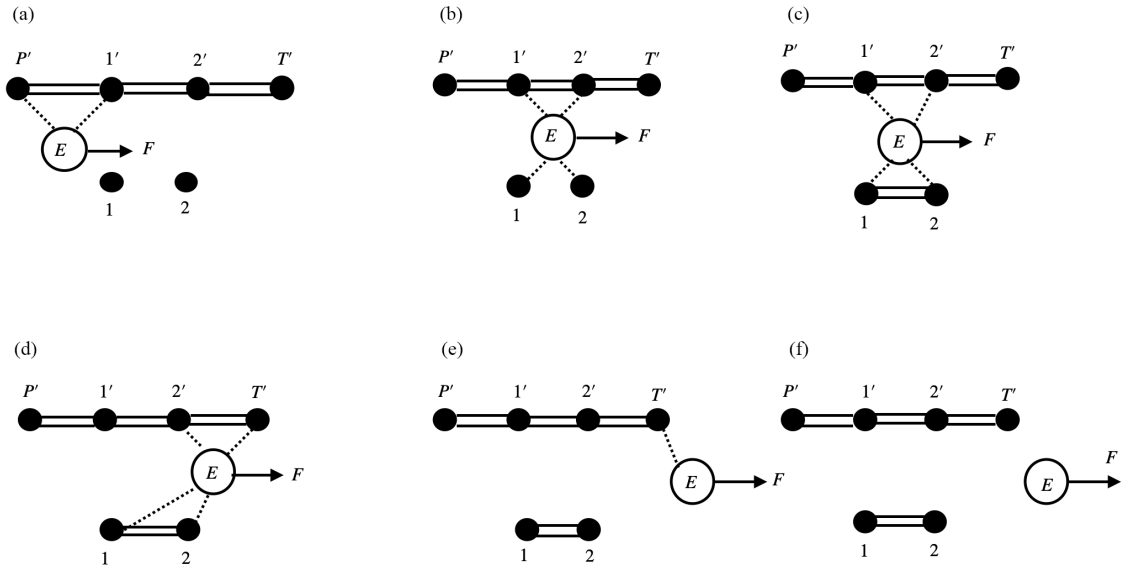


Figure 2. As E moves along the template, it is subject to transient weak attractive interactions with promoter P', monomers $1'$ and $2'$, and to a weak repulsive interaction with terminator T'. The initially free monomers 1 and 2, subject to transient weak attractive interactions with enzyme E, become permanently bound to each other by covalent forces since stage (c)

Figure 2 displays the successive interactions between the extended Te_{ext} (formed only by the promoter P', two bound monomers $1'$ and $2'$, and the terminator T') and the enzyme E and the free monomers 1 and 2, which perform in principle independent Brownian motions in the aqueous medium at the initial time $t = 0$. In particular, Figure 2a displays the enzyme E becoming weakly attracted by the promoter P' as well as by monomer $1'$ in Te. Then, in Figure 2b the enzyme E also approaches and weakly traps the next monomer $2'$ in Te. In the latter process, energy is taken from monomer 1 (which suffers an internal transition) and fed into E, representing a first contribution to the effective F. This would mimic, in the present toy model, the hydrolytic transition of a free monomer containing initially ATP (adenosine triphosphate) into a similar free monomer now containing ADP (adenosine diphosphate) with the liberation of orthophosphoric acid and energy. In our toy model, we identify approximately the centers of mass of the free monomers before and after the transition, focus only on their translation and disregard other products of the transition (say, the orthophosphoric acid). Moreover, the free monomer 2 is also weakly trapped by E, with further liberation of energy transferred from 2 into E (a second contribution to F). Figure 2c displays the region in which E attracts weakly both free monomers 1 and 2, so that the latter experience a short-range reciprocal attraction and become strongly (covalently) bound to each other forming c: the copying of Te into the transcript c is just occurring. In Figure 2d transcription is close to end as Te contains only two bound monomers (the case treated in Sections 3 and 4 and Appendix B). In Figure 2e, the terminator T' certifies the end, by exerting a repulsive action on E (which may continue receiving energy from other free monomers, not participating in transcription, but contributing to the constant force F). Once c has been formed, in Figure 2f, the enzyme E continues its approximate motion: c and Te_{ext} detach from one another. For $N_c > 2$, the process would continue until no more free monomers are added to the transcript c, as described in Figure C1 in Appendix C.

The evolution of those molecular entities, in a medium at thermal equilibrium which induces friction in the former, is irreversible. We accept that, once the various entities reach for adequately long time thermal equilibrium with the medium, their equilibrium states and their (temperature-dependent) probabilities are independent on the friction mechanisms which operated on them off-equilibrium (say, on their history prior to reaching equilibrium).

The evolution will be studied for generic time $t > 0$ which be not too small, such that the velocities of the all entities considered have previously relaxed to equilibrium (see Appendix B). The above global picture will be described based upon Quantum Statistical Mechanics. After suitable long-time approximations, we shall get effective Smoluchowski equations which resemble diffusion equations for interacting Brownian particles, but include quantum effects allowing for

quantum binding of monomers 1 and 2 so as to produce the transcript c (Appendix B, extending Appendix A). Thus, by invoking effective Smoluchowski equations to account for interacting Brownian particles with certain quantum effects, the connection between mathematical formulation and physics is established. The typical time duration for a new pair of covalently bound monomers in the transcript be formed, in the transcription process, is about 10^{-2} s [2]. Then, we would expect, in the context of our new idealized model, that:

1. Te and E attract each other during about some typical 10^{-2} s;
2. The approximate time duration of the evolution of the free monomers 1 and 2 until c is formed is also about 10^{-2} s.

By starting with seven entities (P', monomers $1'$ and $2'$ in Te, T', E and free monomers 1 and 2), as discussed in Appendix B, the three-dimensional position vectors (of the centers of mass) for P' and T' are supposed to have been integrated over, as are the internal degrees of freedom in free monomers giving rise to energy liberation (effectively represented by the constant vector $-F$, with the physical dimension of force). So, the three-dimensional position vectors (of the centers of mass) for the remaining five entities are: x_1 and x_2 for the free monomers 1 and 2, x_E for E, and $x_{1'}$ and $x_{2'}$ for $1'$ and $2'$ (the bound monomers in Te).

In Appendix B, we introduced $V_{\text{eff,tot}}$ (the total effective potential) and V_{eff}, with $V_{\text{eff,tot}} = V_{\text{eff}} - F \cdot x_E$. The effective potential V_{eff}, describing interactions among the bound monomers, the free ones and E, is a sum of (local and t-independent) potentials. In Sections 3, 4 and 5, we shall employ the following contributions in (B12) and (3), which affects the entities indicated by the corresponding position vectors, to V_{eff} (but not the full V_{eff}):

$$V_{12}(|x_2 - x_1|) + V_{\text{E,P}',\text{eff}}(x_E) + V_{\text{E,T}',\text{eff}}(x_E) + V_{\text{Eall}}, \tag{2}$$

where V_{Eall} is defined by

$$V_{\text{Eall}}(x_1, x_2, x_{1'}, x_{2'}, x_E, t) = V_{\text{Eall}}(|x_1 - x_E|, |x_2 - x_E|, |x_{1'} - x_E|, |x_{2'} - x_E|, |x_1 - x_{1'}|, |x_2 - x_{2'}|). \tag{3}$$

The potential V_{12} between 1 and 2 is strongly attractive, on average. All the above potentials approach zero as the absolute values of their arguments tend to $+\infty$. $V_{\text{E,P}',\text{eff}}(x_E)$ and $V_{\text{E,T}',\text{eff}}(x_E)$ are effective potentials describing the influence of P' and T' on E. In agreement with the qualitative discussion in Section 2, V_{Eall} is, on average, weakly attractive and much smaller in magnitude than the largest values of $|V_{12}|$. So, V_{Eall} does not give rise, by itself, to bound states of any pair of the entities associated to its arguments. The physical descriptions so far in Section 2 and the present Subsection have motivated the direct mathematical formulation through Schmoluchowski equation in Subsection 3.2.

3.2 *Approximate smoluchowski equations (Steps 4 and 5)*

Let k_B and T be, respectively, Boltzmann's constant and the absolute temperature in the aqueous medium (approximately at rest and at thermal equilibrium, at room temperature: $k_B T \simeq 0.026\, ev$, electronvolts) in which transcription occurs. η will denote the average viscosity coefficient of that aqueous medium. Let a_{av} be the average radius of a small molecule. We introduce the useful average diffusion coefficient D as $D = k_B T/(6\pi\eta a_{\text{av}})$, see [6]. Here, D is always taken as a value typical for small molecules moving in aqueous media: $D = 10^9$ nm^2/s ($= 10^{-9}$ m^2/s). Let r_k be the approximate radius of a sphere just containing one molecule of type k of interest here (a monomer for c or Te, P', T'or a globular protein for E) and we consider the useful dimensionless numbers: $a_k = a_{\text{av}}/r_k$. The more massive a molecule is, the larger r_k is. For a molecule of type k, the diffusion coefficient is Da_k, to be written below with other notations for a_k, as it suits. So, the more massive a molecule is, the slower is its diffusion (the smaller its diffusion coefficient is).

We shall introduce overall and relative coordinates for the monomers 1 and 2 (either free or forming the transcript c) as $X_c = \frac{a_1^{-1}x_1 + a_2^{-1}x_2}{M_c}$, with $x_c = x_2 - x_1$. The dimensionless $M_c = a_1^{-1} + a_2^{-1}$ and m_c ($m_c^{-1} = a_1 + a_2$) will play roles similar to total and reduced masses of 1 and 2. For E (with approximate size r_E), the dimensionless constant $a_{EE} = a_{\text{av}}/r_E$ is, by assumption, about 10^{-2} to 10^{-3}. ∇ will denote the three-dimensional gradient operator. In terms of X_c and x_c, it holds

$$\nabla_{x_1} = \frac{a_1^{-1}}{M_c}\nabla_{X_c} - \nabla_{x_c} \qquad \text{and} \qquad \nabla_{x_2} = \frac{a_2^{-1}}{M_c}\nabla_{X_c} + \nabla_{x_c}, \tag{4}$$

which enable to express probabilities in terms of (x_1, x_2) or of (X_c, x_c), as it suits (as $dx_1\,dx_2 = dX_c\,dx_c$). We shall consider the distribution function $W_0(x_1, x_2, x_{1'}, x_{2'}, x_E, t)$ for the monomers 1 and 2, *Te* and *E* (Appendix B). It fulfills (B14), the explicit form and justification of which being unnecessary here for the direct formulation that follows. They are justified, through several assumptions and long-time approximations in Appendices A and B. It satisfies $\dfrac{\partial}{\partial t}\displaystyle\int W_0\; dx_1 dx_2\, dx_{1'}\, dx_{2'}\, dx_E = 0$. We introduce separate distribution functions W_E, W_{Te} and W_c for *E*, *Te* and *c*, respectively. Namely, W_E is the approximate distribution for *E* defined by

$$W_E(x_E, t) = \int W_0(x_1, x_2, x_{1'}, x_{2'}, x_E, t)\; dx_1\, dx_2\, dx_{1'}\, dx_{2'}. \tag{5}$$

W_{Te} is the approximate distribution for the bound monomers $1'$ and $2'$ constituting *Te* defined by

$$W_{Te}(x_{1'}, x_{2'}, t) = \int W_0(x_1, x_2, x_{1'}, x_{2'}, x_E, t)\; dx_1\, dx_2\, dx_E. \tag{6}$$

The quantum mechanical distribution function for the monomers 1 and 2 (either unbound to each other or forming *c*) is W_c defined by

$$W_c(X_c, x_c, t) = \int W_0(x_1, x_2, x_{1'}, x_{2'}, x_E, t)\; dx_{1'}\, dx_{2'}\, dx_E, \tag{7}$$

where the use of X_c and x_c is preferred; and $W_c\; dX_c\, dx_c$ is the probability to find X_c and x_c in small volumes $dX_c\, dx_c$ about X_c and x_c at time t.

We shall outline successively the approximations leading from equation (B14) to two coupled non-linear separate Smoluchowski equations for W_E and W_c, through direct integrations and approximate factorizations (understandable through a look at equation (B14), without further details). For W_E, we apply $\displaystyle\int dx_1\, dx_2\, dx_{1'}\, dx_{2'}$ to equation (B14). In the resulting equation, we approximate inside integrals $W_0(x_2, x_1, x_E, x_{1'}, x_{2'}, t) \simeq W_{Te}(x_{1'}, x_{2'}, t) W_E(x_E, t) W_c(x_1, x_2, t)$ for $t \geq 0$. Then, in Step 4, equation (B14) yields

$$\frac{\partial W_E}{\partial t} = Da_{EE} \cdot \nabla_{x_E}\left(\nabla_{x_E} W_E + \frac{W_E}{k_B T}(-F + j_{EE})\right) \tag{8}$$

where Da_{EE} is the diffusion coefficient, and j_{EE} is defined by

$$j_{EE}(x_E, t) = \nabla_{x_E}\left(\int W_{Te}(x_{1'}, x_{2'}, t)\, W_c(X_c, x_c, t)\, \left(V_{\text{Eall}} + V_{\text{E,P}',\text{eff}}(x_E) + V_{\text{E,T}',\text{eff}}(x_E)\right)\; dx_1\, dx_2\, dx_{1'}, dx_{2'}\right). \tag{9}$$

We suppose the initial condition $W_E(x_E, 0) = \delta^{(3)}(x_E - x_{E,in})$, where $x_{E,in}$ is an arbitrary constant initial vector and $\delta^{(3)}$ is the three-dimensional delta function. Since $F \neq 0$, then $W_E(-x_E, t) \neq W_E(x_E, t)$.

In Step 5, we integrate $\int dx_{1'}\,dx_{2'}\,dx_E$ equation (B14) and approximate similarly W_c so that

$$\frac{\partial W_c}{\partial t} = \frac{D}{M_c}\nabla_{X_c}\left(\nabla_{X_c}W_c + \frac{W_c}{k_BT}\nabla_{X_c}V_{\text{Eall,eff}}\right) + \frac{D}{m_c}\nabla_{x_c}\left(\nabla_{x_c}(-2\varepsilon_c)W_c + \frac{W_c}{k_BT}\nabla_{x_c}(V_{12}+V_{\text{Eall,eff}})\right) \tag{10}$$

where $V_{\text{Eall,eff}}$ is defined by

$$V_{\text{Eall,eff}}(X_c,x_c,t) = \int W_E(x_E,t)\,W_{Te}(x_{1'},x_{2'},t)\,V_{\text{Eall}}(X_c,x_c,x_{1'},x_{2'},x_E)\;dx_{1'}\,dx_{2'}\,dx_E \tag{11}$$

with the diffusion coefficients now cast as $\frac{D}{M_c}$ and $\frac{D}{m_c}$ and initial condition $W_{c,in}(X_c,x_c)$. The function

$$\varepsilon_c = \varepsilon_x(|x_c|) \tag{12}$$

accounts for quantum effects between 1 and 2. See Appendix A for its justification in a simpler model.

Notice that $\frac{\partial}{\partial t}\int W_E\;dx_E = \frac{\partial}{\partial t}\int W_c\;dX_c\,dx_c = 0$. Accordingly, those functions can be normalized for $t \geq 0$. By assumption, we impose $\int W_E\;dx_E = \int W_c\;dX_c\;dx_c = 1$, and $\int W_{Te}\;dx_{1'}\;dx_{2'} = 1$, for $t \geq 0$.

Then, j_{EE} becomes

$$j_{EE}(x_E,t) = \nabla_{x_E}\left(V_{\text{E,P}',\text{eff}}(x_E) + V_{\text{E,T}',\text{eff}}(x_E) + V_{\text{Eall,eff,E}}(x_E,t)\right) \tag{13}$$

where $V_{\text{Eall,eff,E}}$ is defined by

$$V_{\text{Eall,eff,E}}(x_E,t) = \int W_{Te}(x_{1'},x_{2'},t)\,W_c(X_c,x_c,t)\,V_{\text{Eall}}(X_c,x_c,x_{1'},x_{2'},x_E)\;dX_c\,dx_c\,dx_{1'}\,dx_{2'}. \tag{14}$$

Another equation could similarly be derived for W_{Te}, coupled to those for W_E and W_c. However, in the spirit of our approximations it will be proper to employ directly an adequate approximation for W_{Te} (see Subsection 3.3). So, equations (8) and (10) constitute (provided a suitable approximation for W_{Te} be assumed) a non linear system of two coupled Smoluchowski equations for W_E and W_c.

3.3 *Further approximations (Steps 6–10)*

Due to the intrinsic difficulty of the two non linear coupled equations for W_c and W_E further approximations, constituting an algorithm formed by the successive Steps 6–10, will be made below. The slowly diffusing Te_{ext} is more massive than E which, in turn, is more massive than the free monomers 1 and 2, and the transcript c.

For Te, since $1'$ and $2'$ are strongly bound to each other (quantum effects being thereby taken care of), we shall suppose, as a zeroth order approximation, that W_{Te} is like a free Gaussian polymer chain formed by two beads (the shortest Rouse chain, see [35, 36]), with some corrections to account for the binding to P' and T' (see below and Appendix B). The monomers $1'$ and $2'$ and the average distance between them are interpreted as two beads and the chain length d (constant and positive), respectively. Therefore,

$$W_{Te}(x_{1'},x_{2'},t) \simeq W_{Te}(X_{Te},t)\, W_{Te,re}(x_{1'}-x_{2'}), \tag{15}$$

where W_{Te} is approximated by

$$W_{Te}(X_{Te},t) = \frac{1}{8\pi^{3/2}} \left(\frac{D}{M_{Te}}\right)^{-3/2} \exp\left(-\frac{M_{Te}(X_{Te}-X_{Te,in})^2}{4Dt}\right), \tag{16}$$

and $W_{Te,re}$ is defined by

$$W_{Te,re}(x_{1'}-x_{2'}) = \left(\frac{3}{2\pi d^2}\right)^{3/2} \exp\left(-\frac{3(x_{1'}-x_{2'})^2}{2d^2}\right), \tag{17}$$

with $W_{Te}(X_{Te},0) = \delta^{(3)}(X_{Te}-X_{Te,in})$, the initial conditions $X_{Te,in}$ and $x_{E,in}$ being close to each other (say, $|X_{Te,in}-x_{E,in}|$ being less than about $2d$), implementing that, at $t=0$, E be close to x'_P. Appendix B gives the definitions of X_{Te} (in terms of $x_{1'}$, $x_{2'}$) and M_{Te}, and explains that the above approximation for W_{Te} and $W_{Te,re}$ in (16) and (17), respectively, includes partially the influences of P' and T' through the large parameter M_{Te}. The spirit of this approximation, affecting only the extreme monomers of Te, would be the more justified the longer Te (say, for the generalization in Appendix C).

We shall carry out a further approximation for finite t, in order to factor out the contributions of the X_c and x_c degrees of freedom. Let $W_{c,f}$ and $W_{c,re}$ be defined by putting

$$W_{c,f}(X_c,t) \;= \int W_c(X_c,x_c,t)\, dx_c, \tag{18}$$

$$W_{c,re}(x_c,t) \;= \int W_c(X_c,x_c,t)\, dX_c, \tag{19}$$

so that $\int W_{c,f}(X_c,t)\, dX_c = \int W_{c,re}(x_c,t)\, dx_c = 1$. We apply $\int dx_c$ to equation (10). In the resulting equation, in Step 6, we approximate inside integrals $W_c \simeq W_{c,f}(X_c,t)\, W_{c,re}(x_c,t)$ as

$$\frac{\partial W_{c,f}}{\partial t} = \frac{D}{M_c}\nabla_{X_c}\left(\nabla_{X_c}W_{c,f} + \frac{W_{c,f}}{k_BT}\,\nabla_{X_c}V_{\mathrm{c,f,eff}}\right) \tag{20}$$

where $V_{\mathrm{c,f,eff}}$ is defined by

$$V_{\mathrm{c,f,eff}}(X_c,t) = \int W_{c,re}(x_c,t)\, dx_c \int W_E(x_E,t)\, W_{Te}(x_{1'},x_{2'},t)\, V_{\mathrm{Eall}}(X_c,x_c,x_{1'},x_{2'},x_E)\, dx_{1'}\, dx_{2'}\, dx_E. \tag{21}$$

We apply $\int dX_c$ to equation (10). In the resulting equation, we proceed similarly. Then, in Step 7, it follows that

$$\frac{\partial W_{c,re}}{\partial t} = \frac{D}{m_c}\nabla_{x_c}\left(\nabla_{x_c}(-2\varepsilon_c)W_{c,re} + \frac{W_{c,re}}{k_BT}\nabla_{x_c}(V_{12}+V_{\mathrm{Eall,eff,re}})\right), \tag{22}$$

where $V_{\text{Eall,eff,re}}$ is defined by

$$V_{\text{Eall,eff,re}}(x_c,t) = \int W_{c,f}(X_c,t)\, dX_C \int W_E(x_E,t)\, W_{Te}(x_{1'},x_{2'},t)\, V_{\text{Eall}}(X_c,x_c,x_{1'},x_{2'},x_E)\, dx_{1'}\, dx_{2'}\, dx_E. \tag{23}$$

The initial conditions are $W_{c,f,in}(X_c)$ and $W_{c,re,in}(x_c)$. By choosing adequately those initial conditions and $X_{c,in}$, one can ensure that at $t=0$, the probability distribution for the free monomer 1 be concentrated near $x_{E,in}$ and $X_{Te,in}$.

We remind that, by assumption, the weak V_{Eall} does not give rise to any bound states. It is necessary that V_{Eall} depends on x_E through the dependences displayed in equation (3) since, otherwise, transcription would be very unlikely or impossible. The dependences of V_{Eall} on $|x_1 - x_{1'}|$ and $|x_2 - x_{2'}|$ (which does not produce bound states) may be not mandatory and could depend on the model.

Equations (22) and (8) have no stationary (t-independent) solutions $W_{c,re,st}(x_c)$ and $W_{E,st}(x_E)$, respectively. In fact, j_{EE} and $V_{\text{Eall,eff,re}}$ depend unavoidably on t, as they depend on W_{Te} which, as a given input function, is t-dependent.

At this level, in equation (8), one should use (15)–(17) and $W_c \simeq W_{c,f}(X_c,t)\, W_{c,re}(x_c,t)$. This, together with the following properties a) and b) constitute Step 8:

a) For sufficiently large t, $V_{\text{Eall,eff,E}}$ vanishes due to $\left(\frac{4D}{M_{Te}}t\right)^{-1/2}$ in W_{Te}, so that

$$\frac{\partial W_E}{\partial t} \simeq Da_{EE}\cdot\nabla_{x_E}\left(\nabla_{x_E}W_E - \frac{W_E}{k_BT}F\right) + D\nabla_{x_E}\left(\frac{W_E}{k_BT}a_{EE}\right)\nabla_{x_E}\left(V_{\text{E,P}',\text{eff}}(x_E) + V_{\text{E,T}',\text{eff}}(x_E)\right). \tag{24}$$

b) If, in addition, $|x_E|$ is large, the short-range potentials $V_{\text{E,P}',\text{eff}}(x_E)$ (<0, attractive) and $V_{\text{E,T}',\text{eff}}(x_E)$ (>0, repulsive) vanishes, so that previous equation (24) becomes

$$\frac{\partial W_E}{\partial t} \simeq Da_{EE}\cdot\nabla_{x_E}\left(\nabla_{x_E}W_E - \frac{W_E}{k_BT}F\right). \tag{25}$$

So, one should study the solution W_E of equation (8) for large t and x_E. For that purpose, we construct the Green's function for equation (25) (as an equality), for any $t>0$ and any x_E, with the initial condition $W_E(x_E,0) = \delta^{(3)}(x_E - x_{E,in})$, $\delta^{(3)}$ being the three dimensional delta function and $x_{E,in}$ being an arbitrary constant initial vector. A standard computation, by performing Fourier transforms in X_c and t in equation (25) (as an equality) and a residue integration, yields for $t>0$ and any x_E:

$$W_E(x_E,t) = \frac{(Da_{EE}t)^{-3/2}}{8\pi^{3/2}}\exp\left(-\frac{(x_E - x_{E,in} - (Da_{EE}t/k_BT)F)^2}{4Da_{EE}t}\right). \tag{26}$$

Then, the previous equation (26) for large $t>0$ and x_E gives an approximation to the solution W_E of equation (8) for large t and x_E. Notice that (25) (as an equality) has a t-independent stationary (st) solution $W_{E,st}(x_E) = \exp\left(\frac{x_E\cdot F}{k_BT}\right)$, which is not acceptable as an equilibrium state.

On the other hand, for sufficiently large t, $V_{\text{Eall,eff}}$ vanishes due to $(4(D/M_{Te})t)^{-3/2}$ in W_{Te}. Then, it holds

$$\frac{\partial W_c}{\partial t} \simeq \frac{D}{M_c}\nabla_{X_c}\left(\nabla_{X_c}W_c\right) + \frac{D}{m_c}\nabla_{\mathbf{x}_c}\left(\nabla_{x_c}(-2\varepsilon_c)W_c + \frac{W_c}{k_BT}\nabla_{x_c}V_{12}\right) \tag{27}$$

with the t-independent V_{12}. In Step 9, at large t, one can factorize exactly $W_c(X_c,x_c,t) = W_{c,f}(X_c,t)W_{c,re}(x_c,t)$, so that

$$\frac{\partial W_{c,f}}{\partial t} = \frac{D}{M_c}\nabla_{X_c}\left(\nabla_{X_c}W_{c,f}\right). \tag{28}$$

It is interesting to study the solution $W_{c,f}$ of (28) for large t and X_c. By proceeding like for W_E, $W_{c,f}$ can be obtained from the Green's function for (28) with the initial condition $X_{c,in}$, at $t = 0$, so that $W_{c,f}(X_c,0) = \delta^{(3)}(X_c - X_{c,in})$ and (28) for any $t > 0$ and X_c yields

$$W_{c,f}(X_c,t) = \frac{1}{8\pi^{3/2}}\left(\frac{D}{M_c}t\right)^{-3/2}\exp\left(-\frac{(X_c - X_{c,in})^2}{4(D/M_c)t}\right) \tag{29}$$

with $\int W_{c,f}(X_c,t)\,dX_c = 1$. Then, for large $t > 0$ and X_c, equation (29) gives an approximation to the desired solution $W_{c,f}$ of (20). On the other hand, in Step 10, we have that

$$\frac{\partial W_{c,re}}{\partial t} = \frac{D}{m_c}\nabla_{x_c}\left(\nabla_{x_c}(-2\varepsilon_c)W_{c,re} + \frac{W_{c,re}}{k_BT}\nabla_{x_c}V_{12}\right). \tag{30}$$

Using (12) and a comparison with Appendix A, the three-dimensional equation (30) is easily seen to have a non-trivial equilibrium (t-independent) solution $W_{c,re,eq}$ fulfilling

$$\nabla_{x_c}(-2\varepsilon_c)W_{c,re,eq} + \frac{W_{c,re,eq}}{k_BT}\nabla_{x_c}V_{12} \;=\; 0. \tag{31}$$

In fact, since ε_c depends on $|x_c|$, equation (31) reduces to a one-dimensional equation for the $|x_c|$-dependence and has a solution: $W_{c,re,eq} = W_{c,re,eq}(|x_c|)$ for $0 \le |x_c| < +\infty$. Thus, the actual three-dimensional case reduces to the one-dimensional one treated in Appendix A.

The equation for $W_{c,re}$ is a direct three-dimensional generalization of equation (A13) in Appendix A. At this stage and a posteriori, we realize the motivation for the structures in (B14) with the contribution $\nabla_{x_c}(-2\varepsilon_c)$'s, but without any similar (-2ε)'s in terms containing ∇_{X_c}'s. That implies that the diffusion for large t described by $W_{c,f}(X_c,t)$, as implied by (28), is a free classical one, uninfluenced by the quantum evolution embodied in $W_{c,re}(x_c,t)$.

Then, for t sufficiently large, $W_{c,re}$ approaches $W_{c,re,eq}$ (equilibrium distribution). Thus, at equilibrium the two free monomers can be either in an unbound state or bound to each other forming the transcript c, with relative probabilities (temperature-dependent), as described by following Appendix A. As the temperature T decreases, the probability for the transcript formation increases, as indicates the analysis in the simplified one-dimensional model in Appendix A. The equilibrium $W_{c,re,eq}$ is independent on the initial condition. The rate at which $W_{c,re,eq}$ is reached does depend on the initial conditions and on the very dynamics, that is, on $V_{\mathrm{Eall,eff,re}}$. Physically interesting V_{12}'s, even if attractive on average, are repulsive in certain limited domain, which may require a long time for the transcript to be formed. A crucial fact is that, due to the action of E (in the presence of Te_{ext}) and if certain conditions are fulfilled, the repulsive domain in V_{12} can be overcome and c can be formed in a time shorter than if E and Te_{ext} were absent. In order to appreciate the essential role of the enzymatic action in the presence of a template and of the quantum effects, we shall estimate an important Mean-First-Passage Time (MFPT) in Subsection 4.1. Other MFPTs are studied in Subsections 4.2 and Subsection 4.3. All MFPTs play an important role in analyzing the approximate consistency of the model.

4. $N_c = 2$ free monomers: MFPTs

It appears physically natural that the three evolutions of $W_{c,re}$, $W_{c,f}$ and W_E have approximately similar time durations: transcription ends, naturally, when the times taken by E and by the overall system of monomers 1 and 2 (as described by $W_{c,f}$) to separate off Te are, in order of magnitude, similar between themselves and to that for 1 and 2 to form the transcript c; that is, about 10^{-2} s.

In order to test whether the present model fulfills those expectations, we shall get, under several simplifying assumptions, some estimates of time durations, by invoking MFPTs for the nonlinear system of three coupled equations (22), (20) and (8) for $W_{c,re}$, $W_{c,f}$ and W_E. For general references about the MFPT formalism, see [37–41]. For other aspects of MFPT approaches, illustrating its versatility, see [42–44]. In the present application of the MFPT, we shall extend [45–49].

In principle, the application of the MFPT formalism to (22), (20) and (8) is very difficult. In fact, among other difficulties: those three Smoluchowski equations are coupled among themselves, the effective potentials in each of them depends on t and on the other W's, …

As an attempt to estimate time durations for the evolutions of W_E, $W_{c,f}$ and $W_{c,re}$ for finite t through the MFPT formalism:

i) First, we approximate the effective t-dependent potential in each equation (22), (20) and (8) by a new effective t-independent potential which, moreover, is independent on the other two W's (and, hopefully, provides an average over them) so as to implement rough separability among the three W's;

ii) Second, we study separately the resulting three MFPTs τ, τ_c and τ_E associated respectively to $W_{c,re}$ (Subsection 4.1), $W_{c,f}$ (Subsection 4.2) and W_E (Subsection 4.3);

iii) Third, angular dependences on position vectors, assumed to be smooth, are smoothed out and, so, disregarded, so that only radial dependences (for τ, τ_c) and a linear one (for τ_E) will be considered.

Specifically, the above approximation i) will be implemented in the MFPT equations and estimates below by replacing: $V_{\mathrm{Eall,eff,re}}$, $V_{\mathrm{c,f,eff}}$ and $V_{\mathrm{Eall,eff,E}} + V_{\mathrm{E,P',eff}} + V_{\mathrm{E,T',eff}}$ by $V_{\mathrm{MFPT,re}}$, $V_{\mathrm{MFPT,f}}$ and $V_{\mathrm{MFPT,E}}$, respectively.

One interesting feature of the MFPTs is that, in spite of the a priori crudeness of the above approximation i), the approximate differential equations satisfied by them bear strong relationships to the system of equations (22), (20) and (8).

4.1 *MFPT for free monomers and c*

We shall study the time partial derivative (22) approximately for t smaller than the large t values at which the approximation in (30) holds and, so, anticipating the approach towards the equilibrium solution. We shall display the crucial roles of ε_c (which allows for the chemical reaction leading to the formation of the transcript), and of $V_{\mathrm{Eall,eff,re}}$, approximated by $V_{\mathrm{MFPT,re}}$.

We assume that the potential V_{12} depends only on $r = |x_c|$. Qualitatively, V_{12} vanishes very quickly beyond a domain with a size about the length d characterizing Te. For simplicity, we suppose that it can support only one bound state (corresponding to the transcript c), with probability distribution concentrated in the interval $(r_{1,0}, r_{1,1})$. The repulsive character of V_{12} means that a tunnelling effect would be required across the interval $(r_{1,1}, r_{1,2})$ in order to achieve a sizable probability distribution concentrated in the interval $(r_{1,0}, r_{1,1})$, but such a tunnelling would make the binding to proceed slowly.

Figure 3 displays the behaviour of the potential V_{12}. Namely, the potential V_{12} is positive, ie repulsive, when the radius is less than $r_{1,0}$ and greater than $r_{1,1}$, and it is negative, ie attractive, when the radius is between $r_{1,0}$ and $r_{1,1}$.

However, the potential $V_{\mathrm{MFPT,re}}$ can make the sum $V_{12} + V_{\mathrm{MFPT,re}}$ (even if still positive) quite appreciably smaller than $|V_{12}|$ in the interval $(r_{1,1}, r_{1,2})$. Consequently, the binding of 1 and 2 to form c could proceed much faster. The influence of $V_{\mathrm{MFPT,re}}$ on V_{12} for $r < r_{1,1}$ is decreasingly small as r approaches r_{ref}.

Figure 4 displays, on the left hand side, the behaviour of potential $V_{\mathrm{MFPT,re}}$. It is positive when the distance is less than $r_{1,0}$, and negative otherwise. On the right hand side, Figure 4 displays the sum $V_{12} + V_{\mathrm{MFPT,re}}$ which is smaller than V_{12} for distances grater than $r_{1,0}$.

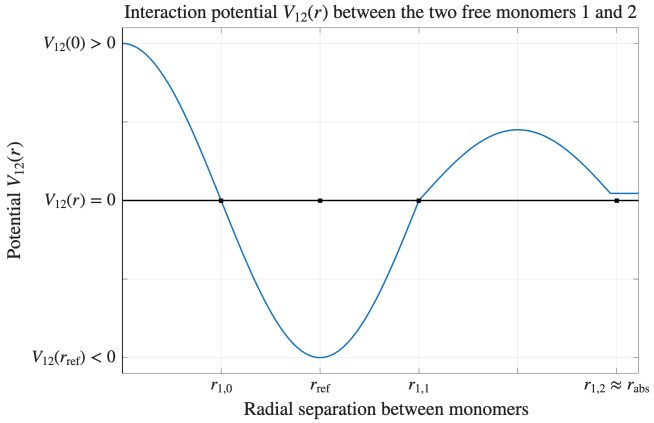


Figure 3. The potential V_{12} is repulsive (ie positive) in the intervals $(0, r_{1,0})$ and $(r_{1,1}, r_{1,2})$. It is strongly attractive (ie negative) in the interval $(r_{1,0}, r_{1,1})$ and reaches its minimum near r_{ref}. It vanishes quickly when the radius is greater than $r_{1,2}$

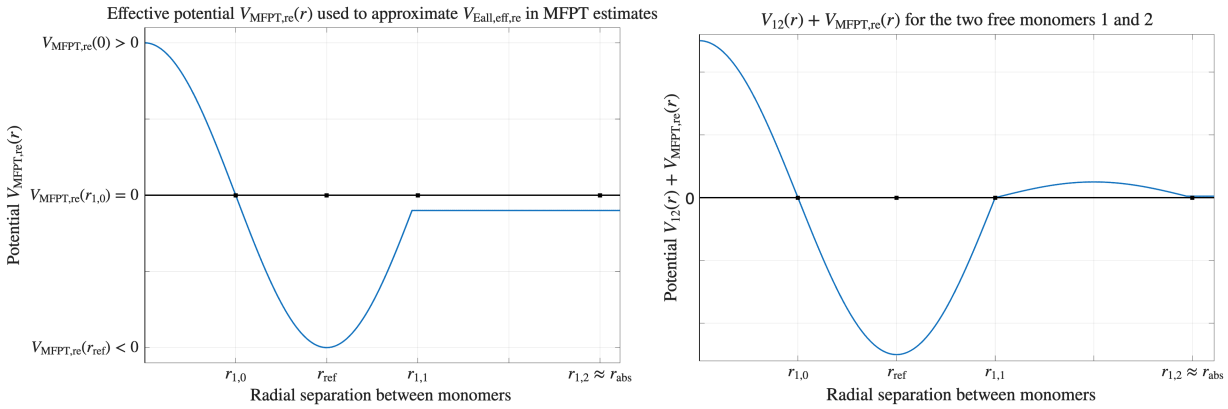


Figure 4. On the left hand side, the potential $V_{\text{MFPT,re}}$ is negative in the interval $(r_{1,1}, r_{1,2})$. On the right hand side, in the interval $(r_{1,1}, r_{1,2})$, the sum $V_{12} + V_{\text{MFPT,re}}$ is positive and appreciably smaller than $|V_{12}|$ in the interval $(r_{1,0}, r_{1,1})$

To estimate approximately the (average) time required for the free monomers 1 and 2 to generate the transcript, under the crucial actions of V_{12} and $V_{\text{MFPT,re}}$, we approximate $W_{c,re}(x_c, t)$ by a new function $W_{c,re}(r,t)$ and (22) by the following radial equation

$$\frac{\partial W_{c,re}}{\partial t} = D\left(\frac{\partial}{\partial r} + \frac{2}{r}\right)\left(\frac{1}{m_c}\frac{\partial}{\partial r}((-2\varepsilon_c)W_{c,re}) + \frac{W_{c,re}}{k_B T}\left(\frac{1}{m_c}\frac{\partial}{\partial r}(V_{12} + V_{\text{MFPT,re}})\right)\right). \tag{32}$$

For convenience, we introduce the function $f(r,t)$ in the previous equation (32) through $f(r,t) = r^2 W_{c,re}(r,t)$, which fulfills

$$\frac{\partial f}{\partial t} = D\frac{\partial}{\partial r}\left(\frac{1}{m_c}\frac{\partial}{\partial r}(-2\varepsilon_c f) - \frac{2}{m_c r}(-2\varepsilon_c)f + \frac{f}{k_B T}\left(\frac{1}{m_c}\frac{\partial}{\partial r}(V_{12} + V_{\text{MFPT,re}}\right)\right). \tag{33}$$

The MFPT function $\tau = \tau(r)$ is the solution of the following and so-called adjoint equation associated with (33)

$$\frac{D}{m_c}(-2\varepsilon_c)\frac{\partial^2\tau}{\partial r^2} - D\left(-\frac{1}{m_c}\frac{2}{r}(-2\varepsilon_c) + \frac{1}{k_BT}\left(\frac{1}{m_c}\frac{\partial}{\partial r}(V_{12}+V_{\text{Eall,eff,re}})\right)\right) + \frac{\partial\tau}{\partial r} = -1 \tag{34}$$

provided that suitable boundary conditions are added. The adequate boundary conditions are $\tau(r_{\text{abs}}) = 0$ (absorption) and $\left.\frac{\partial\tau(r)}{\partial r}\right|_{r=r_{\text{ref}}} = 0$ (reflection). The radial distance r_{ref} is chosen to fulfill $r_{1,0} \leq r_{\text{ref}} \leq r_{1,1}$. V_{12} gets its minimum near r_{ref}. We choose $r_{\text{abs}} \simeq r_{1,2}$. It is supposed that $r_{1,1} - r_{\text{ref}}$ is smaller than $r_{\text{abs}} - r_{1,1}$. The time $\tau(r)$ is interpreted as an estimate of the time required for the individual monomers 1 and 2, separated by a distance r_{abs} and, thus, far from each other, to reach r in $r_{\text{ref}} < r < r_{\text{abs}}$. If $r \simeq r_{\text{ref}}$, then $\tau(r)$ estimates the time required for monomer 1 to become bound to monomer 2, thus giving rise to c.

Then, by direct integration, the solution of equation (34) in the interval $(r_{\text{ref}}, r_{\text{abs}})$, with those boundary conditions, is (with $r_{1,2} = r_{\text{abs}}$) given by

$$\tau(r) = \int_r^{r_{\text{abs}}} \frac{m_c}{D}\frac{J(r')}{r'^2}\,dr', \tag{35}$$

where J is defined by

$$J(r') = \int_{r_{\text{ref}}}^{r'} \frac{r''^2}{-2\varepsilon_c(r'')} \exp\left(\frac{1}{k_BT}\int_{r''}^{r'} \frac{1}{-2\varepsilon_c(r''')}\frac{\partial}{\partial r'''}(V_{12}+V_{\text{MFPT,re}})\,dr'''\right) dr''. \tag{36}$$

Let $V_{12}+V_{\text{MFPT,re}} \neq 0$ and $-2\varepsilon_c \neq 1$. As V_{12} is repulsive in the interval $(r_{1,1}, r_{1,2})$, the chemical reaction for monomers 1 and 2 to form a bound state requires tunnelling towards the interval $(r_{1,0}, r_{1,1})$. In cases of actual biochemical interest (see Section 5), $V_{\text{MFPT,re}}$ is weakly attractive and not sufficiently strong to make $V_{12}+V_{\text{MFPT,re}} < 0$ in the interval $(r_{1,1}, r_{1,2})$. But the contribution of $V_{\text{MFPT,re}}$ may be crucial in decreasing the influence of V_{12} and, so, in accelerating the tunnelling effect.

We insert $1 = \theta(r'-r_{1,1}) + \theta(r_{1,1}-r')$ in equation (35) and $1 = \theta(r-r'') + \theta(r''-r)$ in (36), which gives rise to four contributions. Then, let $r \simeq r_{\text{ref}}$. Only two contributions are immediately seen to survive in principle, namely

$$\tau(r) \simeq \tau_a(r) + \tau_b(r) \tag{37}$$

where τ_a and τ_b are defined by

$$\tau_a(r) = \int_r^{r_{1,1}} \frac{m_c}{Dr'^2}\int_r^{r'} \frac{r''^2}{-2\varepsilon_c(r'')} \exp\left(\frac{1}{k_BT}\int_{r''}^{r'} \frac{1}{-2\varepsilon_c(r''')}\frac{\partial}{\partial r'''}(V_{12}+V_{\text{MFPT,re}})\,dr'''\right) dr''\,dr', \tag{38}$$

$$\tau_b(r) = \int_{r_{1,1}}^{r_{\text{abs}}} \frac{m_c}{Dr'^2}\int_r^{r'} \frac{r''^2}{-2\varepsilon_c(r'')} \exp\left(\frac{1}{k_BT}\int_{r_{1,1}}^{r'} \frac{1}{-2\varepsilon_c(r''')}\frac{\partial}{\partial r'''}(V_{12}+V_{\text{MFPT,re}})\,dr'''\right)$$

$$\times \exp\left(\frac{1}{k_BT}\int_{r''}^{r_{1,1}} \frac{1}{-2\varepsilon_c(r''')}\frac{\partial}{\partial r'''}(V_{12}+V_{\text{MFPT,re}})\,dr'''\right) dr''\,dr'. \tag{39}$$

It will be crucial to apply the properties of $-2\varepsilon_c(r)$ in the quantum regime: it fulfills equation (A12), now with x and $-2\varepsilon(x)$ replaced by $r \geq 0$ and $-2\varepsilon_c(r)$. One invokes the high temperature (close to the classical limit) formal series expansions of equilibrium Wigner distributions (see [50]) and applies them to $W_{c,re,eq}$. Through the actual counterpart for the latter of the exact (A12), one gets formally for that regime: $-2\varepsilon_c(r) \simeq 1 + \frac{\hbar^2}{12m(k_BT)^2}\frac{d^2V_{12}}{dr^2} + ... \simeq \exp\left(\frac{\hbar^2}{12m(k_BT)^2}\frac{d^2V_{12}}{dr^2} + ...\right)$. Compare with [47, 49]. In the previous equation (39), if $r_{1,1} \leq r''' \leq r'$, it is justified to use the above approximation for $-2\varepsilon_c(r''')$ and, moreover, write $-2\varepsilon_c(r''') \simeq 1$ (say, the classical approximation for $-2\varepsilon_c$), since unbound motion dominates and the contribution of bound states is negligible. Using the latter, in τ_b we perform $\int_{r''}^{r_{1,1}} \frac{1}{-2\varepsilon_c(r''')}\frac{\partial}{\partial r'''}(V_{12} + V_{\mathrm{MFPT,re}})\, dr'''$ and notice that $V_{12} + V_{\mathrm{MFPT,re}} \simeq 0$ at $r''' = r_{1,1}$.

For $r''' \leq r_{1,1}$, such r''' belongs to the domain where the bound state of monomers 1 and 2 is concentrated (namely, $r_{1,0} \leq r''' \leq r_{1,1}$) and it would seem open how to approximate reliably $-2\varepsilon_c(r)$. However, it seems natural to expect that in such domain $-2\varepsilon_c(r)$ be > 1, and even large since $\frac{d^2V_{12}}{dr^2} > 0$. That is actually suggested by the (purely formal) expression $-2\varepsilon_c(r) \simeq \exp\left(\frac{\hbar^2}{12m(k_BT)^2}\frac{d^2V_{12}}{dr^2} + ...\right)$. Then, we shall accept that $\frac{1}{-2\varepsilon_c(r''')}$ is sufficiently small to allow to set in τ_b: $\exp\left(\frac{1}{k_BT}\left(\int_{r''}^{r_{1,1}} \frac{1}{-2\varepsilon_c(r''')}\left(\frac{\partial}{\partial r'''}(V_{12} + V_{\mathrm{MFPT,re}})\right) dr'''\right)\right) \simeq 1$ and $\int_r^{r'} \frac{r''^2}{-2\varepsilon_c(r'')}\, dr'' \simeq \int_{r_{1,1}}^{r'} r''^2\, dr''$. Through similar arguments , τ_a is negligible. Then, for r close to r_{ref}, we have

$$\tau(r) \simeq \tau_b \simeq \int_{r_{1,1}}^{r_{\mathrm{abs}}} \left[\frac{m_c}{Dr'^2}\int_{r_{1,1}}^{r'} r''^2 \exp\left(\frac{1}{k_BT}\left(V_{12} + V_{\mathrm{MFPT,re}}\right)(r')\right) dr''\right] dr' \tag{40}$$

where $V_{12} > 0$ and $V_{\mathrm{MFPT,re}} < 0$, in the dominant contributions. If $V_{12} + V_{\mathrm{MFPT,re}} = 0$, then the previous approximation (40) is easily integrated to give (with r close to r_{ref})

$$\tau_b(r) = \frac{m_c}{3D}\left(\frac{r_{\mathrm{abs}}^2}{2} + \frac{r_{1,1}^3}{r_{\mathrm{abs}}} - \frac{3r_{1,1}^3}{2}\right). \tag{41}$$

4.2 *MFPT for $W_{c,f}$*

We approximate $W_{c,f}(X_c,t)$ by a new function $W_{c,f}(r_c,t)$ (with $r_c = |X_c|$), $V_{\mathrm{c,f,eff}}(X_c,t)$ by $V_{\mathrm{MFPT,f}}(r)$ and we replace (20) by the simpler equation for $W_{c,f}(r_c,t)$:

$$\frac{\partial W_{c,f}}{\partial t} = D\left(\frac{\partial}{\partial r_c} + \frac{2}{r_c}\right)\left(\frac{1}{M_c}\frac{\partial W_{c,f}}{\partial r_c} + \frac{W_{c,f}}{k_BT}\frac{1}{M_c}\frac{\partial V_{\mathrm{MFPT,f}}}{\partial r_c}\right). \tag{42}$$

We suppose that $0 < r_{\mathrm{c,min}} \leq r_c \leq r_{\mathrm{c,max}} < +\infty$. For convenience, we introduce the function f_c defined by $f_c(r_c,t) = r_c^2 W_{c,f}(r_c,t)$ and proceed to the MFPT function $\tau_c(r_c)$ as the solution of the following adjoint equation associated with f_c:

$$\frac{D}{M_c}\frac{\partial^2\tau_c}{\partial r_c^2} - D\left(-\frac{2}{M_c r_c} + \frac{1}{k_BTM_c}\frac{\partial V_{\mathrm{MFPT,f}}}{\partial r_c}\right)\frac{\partial\tau_c}{\partial r} = -1 \tag{43}$$

provided that suitable boundary conditions are added. The boundary conditions adequate are $\tau_c(r_{\mathrm{c,min}}) = 0$ and $\left.\frac{\partial\tau_c(r_c)}{\partial r_c}\right|_{r_c = r_{\mathrm{c,max}}} = 0$, where $\tau_c(r_c)$ is interpreted as an estimate of the time required, starting at $r_{\mathrm{c,min}}$, to reach r_c, subject to the approximate $V_{\mathrm{MFPT,f}}$.

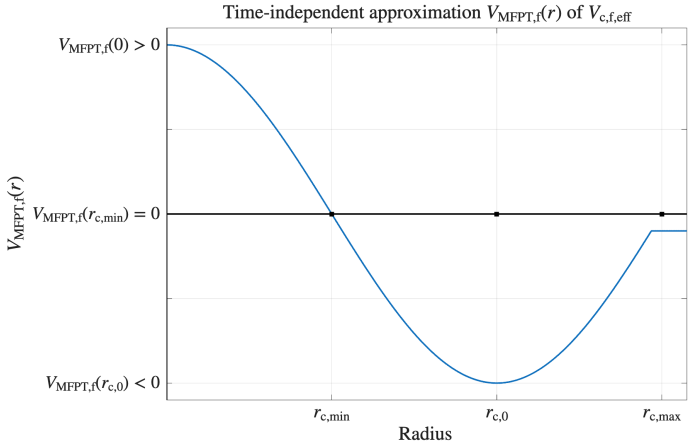


Figure 5. The time-independent effective potential $V_{\mathrm{MFPT,f}}$, used to approximate $V_{\mathrm{c,f,eff}}$, defined in the interval $(0, r_{\mathrm{c,max}})$. $V_{\mathrm{MFPT,f}}$ is repulsive in the interval $(0, r_{\mathrm{c,min}})$, vanishes near $r_{\mathrm{c,min}}$, has a steep decrease in the interval $(r_{\mathrm{c,min}}, r_{c,0})$, and is negative and increases towards 0 in the interval $(r_{c,0}, r_{\mathrm{c,max}})$

Figure 5 displays the time-independent approximation $V_{\mathrm{MFPT,f}}$ of the potential $V_{\mathrm{c,f,eff}}$. $V_{\mathrm{MFPT,f}}$ is repulsive (ie positive) for distances less than $r_{\mathrm{c,min}}$, otherwise is negative and vanishes after $r_{\mathrm{c,max}}$.

Then, by direct integration, the solution of equation (43) with those boundary conditions is

$$\tau_c(r_c) = \tau_{c,1}(r_c) + \tau_{c,2}(r_c), \tag{44}$$

where $\tau_{c,1}$ and $\tau_{c,2}$ are defined by

$$\tau_{c,1}(r_c) = \int_{r_{\mathrm{c,min}}}^{r_c} \left[\frac{M_c}{Dr'^2} \int_{r'}^{r_c} r''^2 \exp\left(\frac{1}{k_B T} (V_{\mathrm{MFPT,f}}(r') - V_{\mathrm{MFPT,f}}(r'')) \right) dr'' \right] dr' \tag{45}$$

$$\tau_{c,2}(r_c) = \int_{r_{\mathrm{c,min}}}^{r_c} \left[\frac{M_c}{Dr'^2} \int_{r_c}^{r_{\mathrm{c,max}}} r''^2 \exp\left(\frac{1}{k_B T} (V_{\mathrm{MFPT,f}}(r') - V_{\mathrm{MFPT,f}}(r'')) \right) dr'' \right] dr'. \tag{46}$$

For $r_{\mathrm{c,min}} < r_c < r_{c,0}$, $\tau_{c,2}(r_c)$ appears to be the dominant contribution (due to $\exp(\frac{-1}{k_B T} V_{\mathrm{MFPT,f}}(r''))$ in the larger interval $r_c < r'' < r_{\mathrm{c,max}}$) compared to $\tau_{c,1}(r_c)$. The closer r_c is to $r_{\mathrm{c,max}}$, the smaller is $\tau_{c,2}(r_c)$ compared to the dominant $\tau_{c,1}(r_c)$ (to be considered in Section 5). The absence of any sort of quantum correction (a counterpart of an ε) in equation (43) makes the structure of $\tau_{c,1}(r_c)$ to differ from that of τ_b above. For $\tau_{c,1}(r_c)$ about $r_{\mathrm{c,max}}$, we search for the contributions which can exponentiate for the typical potentials $V_{\mathrm{MFPT,f}}$ we are considering. Approximate estimates (for instance, performing linear, but different, approximations for $V_{\mathrm{MFPT,f}}$ in $(r_{\mathrm{c,min}}, r_{c,0})$ and $(r_{c,0}, r_{\mathrm{c,max}})$) indicate that, in turn, the terms which produce the dominant exponential approximations in $\tau_{c,1}(r_c)$ for r_c close to $r_{\mathrm{c,max}}$ are

$$\tau_{c,1}(r_c) \simeq \int_{r_{\mathrm{c,min}}}^{r_{c,0}} \left[\frac{M_c}{Dr'^2} \int_{r_{c,0}}^{r_c} r''^2 \exp\left(\frac{1}{k_B T} (V_{\mathrm{MFPT,f}}(r') - V_{\mathrm{MFPT,f}}(r'')) \right) dr'' \right] dr'$$

$$+ \int_{r_{\mathrm{c,min}}}^{r_{c,0}} \left[\frac{M_c}{Dr'^2} \int_{r'}^{r_0} r''^2 \exp\left(\frac{1}{k_B T} (V_{\mathrm{MFPT,f}}(r') - V_{\mathrm{MFPT,f}}(r'')) \right) dr'' \right] dr'. \tag{47}$$

Similar remarks will apply fot $\tau_{E,1}(x)$ in equation (51).

4.3 *MFPT for W_E*

In this case, in order to estimate the MFPT for E, we do not consider radial evolution but a one-dimensional one along the privileged direction x selected by F, subject to an effective t-independent potential $V_{\mathrm{MFPT,E}}$, which approximates $V_{\mathrm{Eall,eff,E}} + V_{\mathrm{E,P',eff}} + V_{\mathrm{E,T',eff}}$ (recall (13) and (14)).

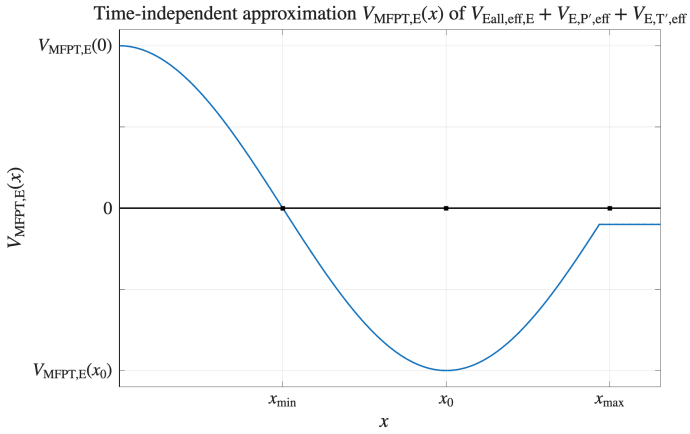


Figure 6. The time-independent potential $V_{\mathrm{MFPT,E}}$ used to approximate $V_{\mathrm{Eall,eff,E}} + V_{\mathrm{E,P',eff}} + V_{\mathrm{E,T',eff}}$. $V_{\mathrm{MFPT,E}}$ is positive when x is less than $x_{\min}$, otherwise is negative, and vanishes when x is greater than $x_{\max}$

Figure 6 displays the time-independent potential $V_{\mathrm{MFPT,E}}$ used to approximate the sum $V_{\mathrm{Eall,eff,E}} + V_{\mathrm{E,P',eff}} + V_{\mathrm{E,T',eff}}$. It can be supposed without essential physical restriction that $V_{\mathrm{MFPT,E}}$ vanishes near $x_{\min}$, has a steep decrease in a limited interval $(x_{\min}, x_0)$, and is negative and increases towards 0 in the interval $(x_0, x_{\max})$.

Then, instead of equation (8), we consider the following equation

$$\frac{\partial W_E}{\partial t} = Da_{EE}\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}W_E + \frac{W_E}{k_B T}\left(-F + \frac{\partial}{\partial x}V_{\mathrm{MFPT,E}}\right)\right) \tag{48}$$

in an interval $(x_{\min}, x_{\max})$. The corresponding adjoint equation for the MFPT $\tau_E(x)$ is given by

$$Da_{EE}\frac{d^2\tau_E}{dx^2} - \frac{Da_{EE}}{k_B T}\left(-F + \frac{dV_{\mathrm{MFPT,E}}}{dx}\right) = -1. \tag{49}$$

The boundary conditions adequate are $\tau_E(x_{\min}) = 0$ and $\left.\dfrac{d\tau_E(x)}{dx}\right|_{x=x_{\max}} = 0$, where $\tau_E(x)$ is interpreted as an estimate of the time required for E, starting at $x_{\min}$ to reach x, subject to both $-F$ and the approximate $V_{\mathrm{MFPT,E}}$.

Like in Subsection 4.2, by direct integration, the solution of the adjoint equation (49) with the previous boundary conditions is

$$\tau_E(x) = \tau_{E,1}(x) + \tau_{E,2}(x), \tag{50}$$

where $\tau_{E,1}$ and $\tau_{E,2}$ are defined by

$$\tau_{E,1}(x) = \int_{x_{\min}}^{x} \frac{1}{Da_{EE}} \left[\int_{x'}^{x} \exp\left(\frac{1}{k_B T} (V_{\mathrm{MFPT,E}}(x') - Fx' - V_{\mathrm{MFPT,E}}(x'') + Fx'') \right) dx'' \right] dx' \tag{51}$$

$$\tau_{E,2}(x) = \int_{x_{\min}}^{x} \frac{1}{Da_{EE}} \left[\int_{x}^{x_{\max}} \exp\left(\frac{1}{k_B T} (V_{\mathrm{MFPT,E}}(x') - Fx' - V'_{\mathrm{MFPT,E}}(x'') + Fx'') \right) dx'' \right] dx'. \tag{52}$$

For $x_{\min} < x < x_0$, $\tau_{E,2}(x)$ appears to be the dominant contribution. The closer x is to $x_{\max}$, the smaller is $\tau_{E,2}(x)$ compared to the dominant $\tau_{E,1}(x)$. The latter will be considered for x close to $x_{\max}$ in Section 5.

5. Some orders of magnitude and MFPTs

We take $D = 10^9$ nm^2/s. By assumption, Te is heavier than E (and E is heavier than the monomers 1 and 2), by a few orders of magnitude. So, by recalling (16) for $W_{Te}(X_{Te}, t)$, for the bulk of the distribution of Te to diffuse appreciably, times appreciably longer than about 10^{-2} s are required. So, in the following estimates, Te is regarded to be effectively at rest. We take $M_c \simeq 10^2$, $m_c \simeq 10$ and $a_{EE} \simeq 10^{-3}$. $|F|$ could range between $|F| \simeq 1$ pN $= 6.4 \times 10^{-3}$ ev/nm $= 0.24 \times k_B$ T/nm at room temperature up to about $|F| \simeq 16$ pN ($\simeq 4k_B$ T/nm).

For τ_b in (40), $\tau_{c,1}$ and τ_E, the MFPTs for free motion (with exponentials replaced by unity, as displayed by (41) for τ_b) bear the structure: $\tau_{free} = (1/D) \times (m) \times (length)^2$, where m stands for m_c, M_c or a_{EE}^{-1} in Subsections 4.1, 4.2 and 4.3, respectively, while $length$ is about 1 nm. The orders of magnitude of τ_{free} for each MFPT are 10^{-8} s (for τ_b) in Subsection 4.1, 10^{-7} s (for $\tau_{c,1}$) in Subsection 4.2, and 10^{-6} s (for $\tau_{E,1}$) in Subsection 4.3.

For two functions $a(x)(\geq 0)$ and $f(x)$ (x denoting one or two variables), let $\int a(x) \exp f(x)\, dx = \langle \exp f \rangle \int a(x)\, dx$ which, by yielding $\langle \exp f \rangle$, could be regarded as the actual definition of the mean value of $\exp f(x)$. Then, the dominant contributions to the MFPTs, namely, τ_b, $\tau_{c,1}$ and $\tau_{E,1}$, can all be recast as the structures $\tau_{\mathrm{free}} \times \langle \exp(potential/k_B T) \rangle$. *potential* should be interpreted as:

1. $V_{12} + V_{\mathrm{MFPT,re}}$ for τ_b;

2. $\ln \frac{c_f}{r'^2} \left[\int_{r_{c,0}}^{r_c} r''^2 \exp\left(\frac{1}{k_B T} (V_{\mathrm{MFPT,f}}(r') - V_{\mathrm{MFPT,f}}(r'')) \right) dr'' + \int_{r'}^{r_0} r''^2 \exp\left(\frac{1}{k_B T} (V_{\mathrm{MFPT,f}}(r') - V_{\mathrm{MFPT,f}}(r'')) \right) dr'' \right]$

for $\tau_{c,1}$, where $c_f > 0$ is a suitable constant with dimensions the inverse of a length and characterizing $V_{\mathrm{MFPT,f}}$, not influencing the corresponding desired exponentiations (recall the end of Subsection 4.2);

3. Another one for $\tau_{E,1}$ similar to the above one for $\tau_{c,1}$ (with the corresponding constant $c_E > 0$ for $V_{\mathrm{MFPT,E}}$, not influencing the corresponding desired exponentiations).

Notice that *potential* also includes the contribution of F in the case of $\tau_{E,1}$.

τ_b). We remind that covalent forces (strong interactions) yielding chemical bonding imply energies which vary between some tens and a few hundred $k_B T$. For instance, the energy required to break a C-N bond is 134.5 $k_B T$ and the dissociation energy of a H-H bond is 169 $k_B T$. A typical V_{12}, being strongly attractive in some domain, is also repulsive (with considerably weaker magnitude) in other domain. Then, back to equation (40), the potential $V_{12}(r')$ is such that

i) is positive in the interval $(r_{1,0}, r_{1,1})$ with variable magnitude $|V_{12}(r')|$ less or equal to the above orders of magnitude times $k_B T$;

ii) is, reasonably, negative in the interval $(r_{1,1}, r_{abs})$ with variable magnitude between 0 and values considerably below the orders of magnitude of $|V_{12}(r')|$ in the interval $(r_{1,0}, r_{1,1})$, considered in i).

The negative potential $V_{\mathrm{MFPT,re}}$ in the interval $(r_{1,1}, r_{abs})$ can be expected to be in magnitude somewhat smaller than V_{12} there. Then, one reasonably expects partial cancellations in $V_{12} + V_{\mathrm{MFPT,re}}$ in the interval $(r_{1,1}, r_{\mathrm{abs}})$, which are responsible for variations in τ by orders of magnitude: for instance, if the positive potential V_{12} and the negative potential $V_{\mathrm{MFPT,re}}$ in the interval $(r_{1,1}, r_{\mathrm{abs}})$ vary in suitable ranges up to about some tens of $k_B T$ (with $V_{12} + V_{\mathrm{MFPT,re}} > 0$). For

$(V_{12}+V_{\mathrm{MFPT,re}})/(k_BT)$ varying about 10 in that range, $\left\langle \exp(\frac{1}{k_BT}(V_{12}+V_{\mathrm{MFPT,re}}))\right\rangle$ can be, on average, near 10^6, so that τ_b be of order 10^{-2} s. This occurs for instance, if $V_{12}/(k_BT)$is about 16 and $V_{\mathrm{MFPT,re}}/(k_BT)$ is about -8 in that range.

$\tau_{c,1}$). It does not seem unreasonable that the orders of magnitude $V_{\mathrm{MFPT,re}}$ and $V_{\mathrm{MFPT,f}}$ (in their contributing ranges to τ_b and $\tau_{c,1}$, respectively) be similar. They are some sort of averages (with somewhat different distribution functions) of V_{Eall}, see (23) and (21). $\tau_{c,1}$ as r_c is close to $r_{\mathrm{c,max}}$, as explained above in this section (for a suitable constant c_f) and follows from the end of Subsection 4.2, can be approximately an integral in some integration domain containing an exponential of a positive quantity about $-V_{\mathrm{MFPT,f}}(r_{c,0})$.Then, for the orders of magnitude of $V_{\mathrm{MFPT,f}}$ involved , it is not unreasonable that $\left\langle \exp(\frac{1}{k_BT}(V_{\mathrm{MFPT,f}}(r')-V_{\mathrm{MFPT,f}}(r''))\right\rangle$ could be, on average, 10^5, so that $\tau_{c,1}$ be of order 10^{-2} s.

$\tau_{E,1}$). The order of magnitude of $V_{\mathrm{MFPT,E}}$, even if it receives contributions from $V_{\mathrm{E,P',eff}}$ (attractive) and $V_{\mathrm{E,T',eff}}$ (repulsive), may plausibly be, in order of magnitude, similar to that of $V_{\mathrm{MFPT,f}}$ (through a similar argument using $V_{\mathrm{Eall,eff,E}}$)): recall (13) and (14)). We consider x close to $x_{\max}$. Through arguments similar as those for $\tau_{c,1}$, the terms which produce the dominant exponential approximations in $\tau_{E,1}(x)$ for x close to $x_{\max}$ are the following

$$\tau_{E,1}(x)\simeq\int_{x_{\min}}^{x_0}\frac{1}{Da_{EE}}\left[\int_{x_0}^{x_{\max}}\exp\left(\frac{1}{k_BT}\exp\left(\frac{1}{k_BT}(V_{\mathrm{MFPT,E}}(x')-Fx'-V_{\mathrm{MFPT,E}}(x'')+Fx'')\right)\right)\,dx''\right]dx'$$

$$+\int_{x_{\min}}^{x_0}\frac{1}{Da_{EE}}\left[\int_{x'}^{x_0}\exp\left(\frac{1}{k_BT}\exp\left(\frac{1}{k_BT}(V_{\mathrm{MFPT,E}}(x')-Fx'-V_{\mathrm{MFPT,E}}(x'')+Fx'')\right)\right)\,dx''\right]dx'. \tag{53}$$

For plausible orders of magnitude of $V_{\mathrm{MFPT,E}}$ and F involved, it is not unreasonable that $\left\langle \exp\left(\frac{1}{k_BT}(V_{\mathrm{MFPT,E}}(x')-Fx'-V_{\mathrm{MFPT,E}}(x'')+Fx'')\right)\right\rangle$ (with the dominant exponential approximations in (53)) be, on average, 10^5, so that $\tau_{E,1}$ be of order about 10^{-2} s.

6. Conclusions and discussions

A new idealized physical-mathematical model for the biochemical transcription process has been presented here. In the simplest case here, transcription occurs in the presence of a molecular chain (*template Te*) formed by two bound monomers $1'$ and $2'$ (all contained in a longer chain Te_{ext}) and under the crucial action of a moving enzyme E subject to an effective force F. With extreme simplifications, the model describes how two unbound and disjoint free monomers (1 and 2) become covalently bound between themselves, so as to form a copy (a transcript c) of Te. See Section 2 and Subsection 3.1 for physical pictures and aspects.

The new model is formulated through a logical sequence of 10 simplifying and approximate Steps successively (indicated in Section 1). The effective model starts in a practical way at Subsection 3.2 (Step 4) in terms of t-dependent distribution functions W_{Te} (for monomers $1'$ and $2'$), W_c (for monomers 1 and 2) and W_E (for enzyme E). (Justifications being deferred to Appendices A and B.) All molecules evolve stochastically in an aqueous medium at rest, at thermal equilibrium at room temperature, in three-dimensional space. Since Te is much more massive than E and the (1,2) pair, by integrating out W_{Te} and performing suitable approximate factorizations, the model boils down to a non-linear system of two coupled non-equilibrium Smoluchowski (diffusion) equations for W_E and W_c, augmented with quantum effects, allowing for the chemical reaction between the unbound monomers 1 and 2 giving rise to c to occur (Steps 4 and 5). At a later stage, W_c is factored out approximately into the distribution function $W_{c,f}$ for the overall evolution of the (1,2) pair times the distribution function $W_{c,re}$ for their relative evolution (influenced by quantum effects) towards forming c. Then, the model yields a non-linear system of three coupled non-equilibrium Smoluchowski equations for W_E, $W_{c,f}$ and $W_{c,re}$ as a result of Steps 6–10, all of which could be regarded as some sort of algorithm. Some properties of the evolutions of W_E, $W_{c,f}$ and $W_{c,re}$ are discussed: in particular, for large t at which the three equations decouple. For finite t, each of the three Smoluchowski equations contains an effective t-dependent potential which depends on the other two distributions and

poses considerable difficulties for approximate solutions (even numerical). Although our main focus is the dynamics of the two free monomers and the formation of c, our analysis also treats the evolutions of $W_{c,f}$ and W_E.

To make detailed numerical simulations of the three-dimensional evolutions of W_E, $W_{c,f}$ and $W_{c,re}$ (requiring specific choices of potentials) lies outside the scope of the present (rather lengthy) work. As an attempt to make plausible the model, time durations for the evolutions of W_E, $W_{c,f}$ and $W_{c,re}$ for finite t have been studied approximately, as another novelty. For that purpose: first, we approximate the t-dependent potential in each Smoluchowski equation by a new effective (t-independent) potential which is independent on the other two distributions (and, hopefully, provides an average over them) and implies rough separability among the three distribution functions; second, we obtain approximate separate formulae for the three Mean First Passage Times (MFPT) τ, τ_c and τ_E associated respectively to $W_{c,re}$ (Subsection 4.1), $W_{c,f}$ (Subsection 4.2) and W_E (Subsection 4.3). In spite of the a priori crudeness of the first approximation, the approximate differential equations satisfied by the MFPTs bear strong relationships to the system of three coupled non-equilibrium Smoluchowski equations. A physically natural condition in transcription seems that the time durations for the processes described by $W_{c,re}$, $W_{c,f}$ and W_E have similar orders of magnitude. Section 5 yields analytically approximations for the leading MFPTs τ_b, $\tau_{c,1}$ and τ_E. The latter provide order of magnitude estimates to those time durations, respectively: for suitable potentials, those three MFPTs can plausibly be of similar order. Such order of magnitude is about 10^{-2} s. That approximate consistency seems possible, in particular, because quantum effects play an important role in suppressing unwanted contributions to τ_b. To make detailed numerical simulations of the MFPTs also lie outside our scope here. However, it is hoped that our arguments and the orders of magnitude of the interactions involved (see Section 5) make both the approximate consistency of the three MFPTs and the present approximate model plausible. The inverse τ_b^{-1} can be interpreted as a rate constant (containing the standard Arrhenius (T-dependent) exponential factor.)

The starting basic justification of the model is summarized below. Appendix A (Step 0) will treat first a simple, but nontrivial, quantum-mechanical statistical model for one particle in one spatial dimension, subject to a potential and a heat bath; then some essential properties at equilibrium, as well as off-equilibrium effective long-time approximations yield an approximate Smoluchowski equation; and finally the model displays crucially both bound and unbound states, allowing for transitions between them and, so, leads to quantum chemical binding (compared to the classical case in which no binding occurs). Appendix B (Steps 1, 2 and 3) outlines the initial formulation involving 10 molecular entities (1, 2, P', $1'$, $2'$, T', E and another three involved in certain internal transitions described by degrees of freedom $u1$, $u2$, and $u3$), in three-dimensional space, and various approximations (in particular, long-time ones). After integrating out variables of P', T' (included in Te_{ext}) and those for the internal transitions $u1$, $u2$, $u3$ (feeding energy into E) and performing suitable approximate factorizations, the model boils down to the effective one in Subsection 3.2, requiring 5 entities (1, 2, $1'$, $2'$ and E) in three-dimensional space.

A generalization to longer Te_{ext} and c is outlined in Appendix C. Te can also be some part of an even larger linear molecular chain. It will be supposed that its curvature is rather small, in some suitable length scale (say, the length of the new chain c). The size of the transcript c will be not greater than that of Te.

The present model is intentionally idealized and differs from established physicochemical descriptions of transcription. For general introductions to ratchets and non-equilibrium stationary states in this context, see [6, 51–53]. Recent literature also includes (i) stochastic models of elongation with explicit pausing/backtracking mechanisms [20–22], (ii) traffic/crowding models when several polymerases share a transcription unit [23, 24], and (iii) inference frameworks aimed at estimating kinetic parameters from single-molecule or sequencing data [25, 26]. The present work is complementary to them: we emphasize a minimal reaction–diffusion/Smoluchowski formulation that, in particular, is suited for MFPT estimates. Standard biochemical and structural aspects are reviewed in [1–5]. Attempts to compare all those models to the one here (which offers complementary descriptions) lie outside our scope.

Acknowledgement

This work was partially supported by Ministerio de Ciencia, Innovación y Universidades (MICIU, Spain), Agencia Estatal de Investigación (AEI, Spain, MCIN/AEI/10.13039/501100011033) and European Regional Development Fund

(ERDF, A way of making Europe) through Grant number PID2022-136374NB-C21. R.F.A.-E. is an associate member of Instituto de Biocomputación y Física de los Sistemas Complejos (BIFI) at the Universidad de Zaragoza in Spain.

Conflict of interest

The authors declare no competing financial interest.

References

[1] Nelson DL, Cox MM. *Lehninger Principles of Biochemistry*. 8th ed. W. H. Freeman; 2021.
[2] Voet D, Voet JG. *Biochemistry*. 4th ed. John Wiley and Sons; 2010.
[3] Watson JD, Baker T, Bell S, Gann A, Levine M, Losick R. *Molecular Biology of the Gene*. 7th ed. Pearson; 2013.
[4] Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K, et al. *Molecular Biology of the Cell*. 6th ed. Garland Publishing; 2014.
[5] Berg JM, Tymoczko JL, Gatto Jr GJ. Stryer L. *Biochemistry*. 9th ed. W. H. Freeman; 2019.
[6] Nelson PC. *Biological Physics: Energy, Information, Life*. W. H. Freeman; 2008.
[7] Murray JD. *Mathematical Biology I: An Introduction*. 3rd ed. Springer; 2002. Available from: https://doi.org/10.1007/b98868.
[8] Mathews CK, van Holde KE and Ahern KG. *Biochemistry*. 3rd ed. Pearson; 1999.
[9] van Holde KE, Johnson WC, Ho PS. *Principles of Physical Biochemistry*. 2nd ed. Pearson College Div; 2005.
[10] Dill KA, Bromberg S. *Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics and Nanoscience*. 2nd ed. Garland Science; 2010.
[11] Cantor CR, Schimmel PR. *Biophysical Chemistry*. Volume 1–3. W. H. Freeman and Co.; 1980.
[12] Wang MD, Schnitzer MJ, Yin H, Landick R, Gelles J, Block SM. Force and velocity measured for single molecules of RNA polymerase. *Science*. 1998; 282(5390): 902–907. Available from: https://doi.org/10.1126/science.282.5390.902.
[13] Larson DR, Zenklusen D, Wu B, Chao JA, Singer RH. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science*. 2011; 332(6028): 475–478. Available from: https://doi.org/10.1126/science.1202142.
[14] Nicholson C, Tao L. Hindered diffusion of high molecular weight compounds in brain extracellular microenvironment measured with integrative optical imaging. *Biophysical Journal*. 1993; 65(6): 2277–2290. Available from: https://doi.org/10.1016/S0006-3495(93)81324-9.
[15] Jing M, Zhang P, Wang G, Feng J, Mesik L, Zeng J, et al. A genetically encoded fluorescent acetylcholine indicator for in vitro and in vivo studies. *Nature Biotechnology*. 2018; 36(8): 726–737. Available from: https://doi.org/10.1038/nbt.4184.
[16] Erban R, Chapman SJ. Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions. *Physical Biology*. 2009; 6(4): 046001. Available from: https://doi.org/10.1088/1478-3975/6/4/046001.
[17] Schöneberg J, Noé F. ReaDDy-a software for particle-based reaction-diffusion dynamics in crowded cellular environments. *PLoS ONE*. 2013; 8(9): e74261. Available from: https://doi.org/10.1371/journal.pone.0074261.
[18] Kabashin AV, Kravets VG, Grigorenko AN. Label-free optical biosensing: going beyond the limits. *Chemical Society Reviews*. 2023; 52: 6554–6585. Available from: https://doi.org/10.1039/D3CS00155E.
[19] Mateu MG, editor. *Structure and Physics of Viruses: An Integrated Guide*. 2nd ed. Springer; 2025.
[20] Qian J, Dunlap D, Finzi L. Basic mechanisms and kinetics of pause-interspersed transcript elongation. *Nucleic Acids Research*. 2020; 49(1): 15–24. Available from: https://doi.org/10.1093/nar/gkaa1182.
[21] Kor R, Mohammad-Rafiee F. Theoretical study of RNA-polymerase behavior considering the backtracking state. *Soft Matter*. 2022; 18(32): 5979–5988. Available from: https://doi.org/10.1039/d2sm00232a.
[22] Yang KB, Rasouly A, Epshtein V, et al. Persistence of backtracking by human RNA polymerase II. *Molecular Cell*. 2024; 84(5): 897–909.e4. Available from: https://doi.org/10.1016/j.molcel.2024.01.019.

[23] Mines RC, Lipniacki T, Shen X. Slow nucleosome dynamics set the transcriptional speed limit and induce RNA polymerase II traffic jams and bursts. *PLoS Computational Biology*. 2022; 18(2): e1009811. Available from: https://doi.org/10.1371/journal.pcbi.1009811.

[24] Zhao Y, Liu L, Hassett R, Siepel A. Model-based characterization of the equilibrium dynamics of transcription initiation and promoter-proximal pausing in human cells. *Nucleic Acids Research*. 2023; 51(21): e106. Available from: https://doi.org/10.1093/nar/gkad843.

[25] Douglas J, Kingston RL, Drummond AJ. Bayesian inference and comparison of stochastic transcription elongation models. *PLoS Computational Biology*. 2020; 16(2): e1006717. Available from: https://doi.org/10.1371/journal.pcbi.1006717.

[26] Stanley JT, Barone GEF, Townsend HA, Sigauke RF, Allen MA, Dowell RD. LIET model: capturing the kinetics of RNA polymerase from loading to termination. *Nucleic Acids Research*. 2025; 53(7). Available from: https://doi.org/10.1093/nar/gkaf246.

[27] Zhu C, Guo X, Dumas P, Takacs M, Abdelkareem M, Broeck AV, et al. Transcription factors modulate RNA polymerase conformational equilibrium. *Nature Communications*. 2022; 13: 1546. Available from: https://doi.org/10.1038/s41467-022-29148-0.

[28] Salari H, Fourel G, Jost D. Transcription regulates the spatio-temporal dynamics of genes through micro-compartmentalization. *Nature Communications*. 2024; 15: 5393. Available from: https://doi.org/10.1038/s41467-024-49727-7.

[29] Shi C, Yang X, Zhang J, Zhou T. Stochastic modeling of the mRNA life process: a generalized master equation. *Biophysical Journal*. 2023; 122(20): 4023–4041. Available from: https://doi.org/10.1016/j.bpj.2023.08.024.

[30] Midha T, Kolomeisky AB, Igoshin OA. Kinetic mechanisms for the sequence dependence of transcriptional errors. *Proceedings of the National Academy of Sciences of the USA*. 2025; 122(28): e2505040122. Available from: https://doi.org/10.1073/pnas.2505040122.

[31] Nielsen S, Zenkin N. Efficient termination of transcription by RNA polymerase I requires a conserved hairpin of the ribosomal RNA precursor. *Science Advances*. 2025; 11(34): eadw2470. Available from: https://doi.org/10.1126/sciadv.adw2470.

[32] Feshbach H. The optical model and its justification. *Annual Review of Nuclear Science*. 1958; 8: 49–104. Available from: https://doi.org/10.1146/annurev.ns.08.120158.000405.

[33] Deutch JM, Oppenheim I. Molecular theory of Brownian motion for several particles. *Journal of Chemical Physics*. 1971; 54(8): 3547–3555. Available from: https://doi.org/10.1063/1.1675379.

[34] Klein R. *Interacting Brownian particle: the dynamics of colloidal suspensions*. IOS Press; 1997. pp. 301–345. Available from: https://doi.org/10.3254/978-1-61499-219-6-301.

[35] Doi M, Edwards SF. *The Theory of Polymer Dynamics*. Oxford University Press; 1986.

[36] Rouse Jr PE. A theory of the linear viscoelastic properties of dilute solutions of coiling polymers. *Journal of Chemical Physics*. 1953; 21: 1272–1280. Available from: https://doi.org/10.1063/1.1699180.

[37] Redner S. *A Guide to First-Passage Processes*. Cambridge University Press; 2007.

[38] Metzler R, Oshanin G, Redner S. *First-Passage Phenomena and Their Applications*. World Scientific; 2014. Available from: https://doi.org/10.1142/9104.

[39] Jacobs K. *Stochastic Processes for Physicists*. Cambridge University Press; 2010.

[40] Van Kampen NG. *Stochastic Processes in Physics and Chemistry*. Elsevier; 2007. Available from: https://doi.org/10.1016/B978-0-444-52965-7.X5000-4.

[41] Risken H. *The Fokker-Planck Equation*. 2nd ed. Springer; 1996. Available from: https://doi.org/10.1007/978-3-642-61544-3.

[42] Serrano H, Álvarez-Estrada RF, Calvo GF. Mean first-passage time of cell migration in confined domains. *Mathematical Methods in the Applied Sciences*. 2023; 46: 7435–7453. Available from: https://doi.org/10.1002/mma.8978.

[43] Serrano H, Álvarez-Estrada RF. Characterization of the mean first-passage time function subject to advection in annular like domains. *Mathematics*. 2023; 11(24): 4998. Available from: https://doi.org/10.3390/math11244998.

[44] Serrano H, Álvarez-Estrada RF. Mean first-passage time in a sphere traversed by a cylinder. *Contemporary Mathematics*. 2025; 6: 8458–8484. Available from: https://doi.org/10.37256/cm.6620258087.

[45] Haengi P, Talkner P, Borkovec M. Reaction-rate theory: fifty years after Kramers. *Reviews of Modern Physics*. 1990; 62: 251–342. Available from: https://doi.org/10.1103/RevModPhys.62.251.

[46] Calvo GF, Alvarez-Estrada RF. The time duration of DNA thermal denaturation. *Journal of Physics: Condensed Matter*. 2008; 20(3). Available from: https://doi.org/10.1088/0953-8984/20/03/035101.

[47] Alvarez-Estrada RF, Calvo GF. Chemical reactions using a non-equilibrium Wigner function approach. *Entropy*. 2016; 18(10): 369–397. Available from: https://doi.org/10.3390/e18100369.

[48] Alvarez-Estrada RF. Quantized constrained molecular chains: vibrations, rotations and polymerization. *Contemporary Mathematics*. 2022; 3(3): 353–385. Available from: https://doi.org/10.37256/cm.3320221586.

[49] Alvarez-Estrada RF. Non-equilibrium Wigner function and application to model catalyzed polymerization. *Entropy*. 2024; 16(2): 104. Available from: https://doi.org/10.3390/e26020104.

[50] Wigner EP. On the quantum correction for thermodynamic equilibrium. *Physical Review*. 1932; 40(5): 749–759. Available from: https://doi.org/10.1103/PhysRev.40.749.

[51] Mallick K. Some recent developments in non-equilibrium statistical physics. *Pramana Journal of Physics*. 2009; 73: 417–451. Available from: https://doi.org/10.1007/s12043-009-0098-4.

[52] Chou T, Mallick K, Zia RKP. Non-equilibrium statistical mechanics: from a paradigmatic model to biological transport. *Reports on Progress in Physics*. 2011; 74. Available from: https://doi.org/10.1088/0034-4885/74/11/116601.

[53] Jarillo J, Ibarra B, Cao-Garcia FJ. DNA replication: in vitro single-molecule manipulation data analysis and models. *Computational and Structural Biotechnology Journal*. 2021; 19: 3765–3778. Available from: https://doi.org/10.1016/j.csbj.2021.06.032.

[54] Hillery M, O'Connell RF, Scully MO, Wigner EP. Distribution functions in physics: fundamentals. *Physics Reports*. 1984; 106(3): 121–167. Available from: https://doi.org/10.1016/0370-1573(84)90160-1.

[55] Chihara TS. *An Introduction to Orthogonal Polynomials*. Gordon and Breach; 1978.

Appendix A. Equilibrium and off-equilibrium quantum distribution functions: a simple model (Step 0)

We consider a microscopic particle with one-dimensional coordinate x ($-\infty < x < +\infty$), subject to an external (t-independent) potential $V(x)$, which is even ($V(x) = V(-x)$), attractive (< 0) in a bounded interval and vanishes quickly as $x \to \pm\infty$. The one-dimensional quantum Hamiltonian of the particle is $H = -(\hbar^2/2m)\frac{d^2}{dx^2} + V$, $\hbar$ being Planck's constant. The particle could have: i) either an unbound motion in $-\infty < x < +\infty$, represented by the wavefunction $\varphi_k = \varphi_k(x)$, with wavevector k, $-\infty < k < +\infty$ and energy $E_k = \hbar^2 k^2/2m$ (≥ 0) ($H\varphi_k = E_k\varphi_k$, continuum spectrum), or ii) only one bound (b) state concentrated about the interval in which V is attractive, with wavefunction $\varphi_b = \varphi_b(x)$, energy $E_b < 0$ and $H\varphi_b = E_b\varphi_b$ (discrete spectrum, $\int \varphi_b\varphi_b^*\, dx = \hbar$).

The particle is also subject to a heat bath (say, a medium) at thermal equilibrium at absolute temperature T, so that it has a stochastic evolution. Let $\rho(t)$ be the statistical density operator of the quantum particle, also represented by the function $\rho(x,x',t)$. It evolves in time ($t > 0$) through $i\hbar\frac{\partial\rho}{\partial t} = H\rho - \rho H$, with initial condition ρ_{in} at $t = 0$. The equilibrium density operator is $\rho_{eq} = \exp(-H/(k_BT))$ (k_B being Boltzmann's constant), expressing that at thermal equilibrium with the heat bath , the particle can be in either φ_b or any φ_k with relative probabilities $\exp(-E_b/(k_BT)) > 1$ or $\exp(-E_k/(k_BT)) < 1$, respectively. Let $W(x,q,t)$ be the general Wigner function, where q is the particle momentum, canonically conjugate of x, so that $W = (\pi\hbar)^{-1}\int_{\mathbb{R}} \rho(x-x',x+x',t)\exp\left(\frac{i2qx'}{\hbar}\right)dx'$. See [50, 54].

For $t > 0$, the exact (t-reversible) dissipation-less quantum master equation for the general off-equilibrium Wigner function W is given by

$$\frac{\partial W(x,q,t)}{\partial t} = -\frac{q}{m}\frac{\partial W(x,q,t)}{\partial x} + M_Q W, \tag{A1}$$

where $M_Q W$ is defined by

$$M_Q W = \int W(x,q',t)\int \frac{i}{\pi\hbar^2}\left(V(x+x') - V(x-x')\right)\exp\left(\frac{i2(q-q')x'}{\hbar}\right)dx'\,dq' \tag{A2}$$

with initial condition W_{in}, determined by ρ_{in}. By integrating the time derivative (A1), one has, for any $t > 0$, $\frac{\partial}{\partial t}\int W\,dq\,dx = 0$. Let W_{eq} be the (t- independent) equilibrium Wigner function, determined by $\rho = \rho_{eq}$.

We introduce (unnormalized and dimensionless) orthogonal polynomials $H_{Q,n} = H_{Q,n}(q)$ in q determined by W_{eq}, which acts as a (in general, non-Gaussian) weight function, n being a non-negative integer. See [55]. In general, the $H_{Q,n}$'s are constructed recurrently as follows. We impose for $n \neq n'$, and any x (left unintegrated), that

$$\int \frac{1}{(2mk_BT)^{1/2}}\, W_{\rm eq}(x,q)H_{Q,n}(q)H_{Q,n'}(q)\, dq = 0, \tag{A3}$$

where $H_{Q,n}(q)$ depends parametrically also on x. As V is even in x, $W_{\rm eq}$ is an even function of q. Then, $H_{Q,n}$ is even or odd in q, for n even or odd, respectively. We choose $H_{Q,0} = H_{Q,0}(q) = 1$ and the following normalizations of $H_{Q,1}$ and $H_{Q,1}$ given by $H_{Q,1}(q) = \frac{q}{(2mk_BT)^{1/2}}$ and $H_{Q,2}(q) = (\frac{q}{(2mk_BT)^{1/2}})^2 + \varepsilon$, where ε (dimensionless), which will be important, is

$$\varepsilon(x) = -\langle \frac{q^2}{2mk_BT} \rangle = -\frac{\int W_{\rm eq}(x,q)\ \frac{q^2}{2mk_BT}\ dq}{\int W_{\rm eq}(x,q)\ dq}. \tag{A4}$$

The explicit structure of the $H_{Q,n}$'s with $n \geq 3$ will not be required here. The classical ($\hbar \to 0$) limit of W_{eq} is proportional to $\exp(-(k_BT)^{-1}((q^2/2m)+V)$ (the classical Boltzmann distribution). In the classical limit, the computation of $\langle \frac{q^2}{2mk_BT} \rangle$ boils down to those of Gaussian integrals. For $\hbar \to 0$, ε is x-independent and

$$\varepsilon = -\frac{1}{2}. \tag{A5}$$

The $H_{Q,n}(q)$'s lead to introduce the non-equilibrium moments, for $n = 0, 1, 2, \ldots$, given by

$$W_n = W_n(x,t) = \int \frac{1}{(2mk_BT)^{1/2}} H_{Q,n}(q) W\ dq \tag{A6}$$

and so on for the equilibrium moments $W_n = W_{eq,n}$, by using W_{eq} (instead of W) as generating function. Then, all $W_{eq,n} = 0$ for $n = 1, 2, 3, \ldots$, while only the lowest one $W_{eq,0}$ is non-vanishing so that

$$W_{eq,0} = \frac{1}{(2mk_BT)^{1/2}} \sum_j \exp(-\beta E_j)\, \varphi_j(x) \varphi_j^*(x) = \frac{1}{(2mk_BT)^{1/2}} \left(\exp(-\beta E_b)\, \varphi_b \varphi_b^* + \int_{\mathbb{R}} \exp(-\beta E_k)\, \varphi_k \varphi_k^*\ dk \right) \tag{A7}$$

$W_{eq,0}$ is dimensionless. An important remark is the following. The classical limit of $W_{eq,0}$ is $\exp(-(k_BT)^{-1}V)$. The latter can be large in the interval in which V is attractive (a classical signal or remnant of the existence of a bound state in the quantum regime). However, $\exp(-(k_BT)^{-1}V)$ has a structure which differs from $W_{eq,0}$, in which the contribution of the bound state is very neat. This corresponds to the different values of ε in the classical and in the quantum regimes: see Section 4. The initial moments $W_{{\rm in},n}$ for W_n are obtained by replacing W by $W_{\rm in}$ in (A6).

The transformation of the one-dimensional equations (A1) and (A2) into an infinite linear hierarchy for the nonequilibrium moments W_n can be performed using the former two and (A6). The lowest equations in that quantum hierarchy (still reversible) are

$$\frac{\partial W_0}{\partial t} = -\ \frac{(2mk_BT)^{1/2}}{m} \frac{\partial W_1}{\partial x}, \tag{A8}$$

$$\frac{\partial W_1}{\partial t} = -\ \frac{(2mk_BT)^{1/2}}{m} \frac{\partial W_2}{\partial x} - M_0 W_0, \tag{A9}$$

where M_0W_0 is defined by

$$M_0W_0 = \ \frac{(2mk_BT)^{1/2}}{2m} \left(\frac{\partial}{\partial x}(-2\varepsilon W_0) + \frac{1}{k_BT} \frac{\partial V}{\partial x} W_0 \right). \tag{A10}$$

The general interest of previous derivatives (A8), (A9) and (A10) here comes from the following. At equilibrium, since $W_{eq,1} = 0$ and $W_{eq,2} = 0$, due to our choice of orthogonal polynomials, the time derivative (A9) implies that

$$M_0 W_{eq,0} = 0, \tag{A11}$$

which holds independently of the way in which W_0 may approach equilibrium. The previous equality (A11) and the expression for M_0 from (A10), with the condition that in the classical ($\hbar \to 0$) limit $-\varepsilon = 1/2$, can be recast as

$$-2\varepsilon(x) W_{eq,0}(x) = \exp\left(\frac{1}{k_B T}\int_x^{+\infty} \frac{1}{-2\varepsilon(x')}\frac{\partial V}{\partial x'}\, dx'\right). \tag{A12}$$

Since $W_{eq,0} \geq 0$, it follows that $-2\varepsilon \geq 0$ for any x. The exact and non-trivial quantum equation (A12) should not be overlooked (by contrast to the classical one, in which $-2\varepsilon = 1$). For the given quantum Hamiltonian H and its unbound and bound eigenfunctions, the natural way to proceed is to get $W_{eq,0}$ from (A7), while ε is given by either (A4) or (A12). See Section 4.

At this stage, one can proceed to the infinite linear hierarchy for the nonequilibrium moments W_n. For simplicity, we shall assume that the moments $W_{\text{in},n} = 0$ for $n > 0$ while only $W_{\text{in},n=0} \neq 0$. Then, W_1 (and all W_n for $n \geq 2$) can be expressed formally, through suitable linear operators, in terms of W_0 and of the initial condition $W_{\text{in},0}$. We shall proceed directly to approximations in that hierarchy for adequately long time t , which do implement irreversibility. For details of the rather lengthy arguments, see [47, 49]. By the very nature of the approximations, they would be expected to be the more unreliable for closer t is to $t = 0$.

In a direct formulation (yielding an effective off-equilibrium time evolution through those approximations): i) in (A9) one drops W_2, and approximates $\frac{\partial W_1}{\partial t}$ by $(mD/k_B T)^{-1} W_1$ thereby introducing an effective diffusion coefficient D (approximating certain involved linear operator), ii) one accepts that, by feeding the resulting W_1 from (A9) into (A8), $W_0 = W_0(x,t)$ fulfills the following Smoluchowski equation for irreversible nonequilibrium evolution:

$$\frac{\partial W_0}{\partial t} = D\left(\frac{2m}{k_B T}\right)^{\frac{1}{2}} \frac{\partial}{\partial x}(M_0 W_0) = D\frac{\partial}{\partial x}\left(\frac{\partial(-2\varepsilon)W_0}{\partial x} + \frac{1}{k_B T}\frac{\partial V}{\partial x} W_0\right), \tag{A13}$$

with an effective initial condition $W_{0,in}$, which needs not coincide with $W_{in,0}$, since the previous equation (A13) is quite possibly unreliable for $t = 0$. Anyway, $W_{in,0}$ may well give hints for $W_{0,in}$. One has $\frac{\partial}{\partial t}\int W_0\, dx = 0$. One can show that W_0 tends to $W_{eq,0}(x)$ as $t \to +\infty$ for any $W_{0,in}$, a non-trivial fact since -2ε is x-dependent. See [49]. This provides an (at least, partial) check of physical consistency of the approximate evolution in (A13). So, whatever be $W_{0,in}$, for very long t the particle will be in either φ_b or any φ_k with relative probabilities $\exp(-E_b/(k_B T)) > 1$, or $\exp(-E_k/(k_B T)) < 1$, respectively: the non-vanishing temperature allows for both possibilities for the particle, each with its own probability. If $-E_b/(k_B T) \gg 1$, the occupancy of φ_b will dominate. However, the very evolution of W_0 towards $W_{eq,0}(x)$ and the speed required to reach the latter depend in principle of $W_{0,in}$. The long-time evolution equation (A13), implementing quantum effects due to ε, is now adequate for chemical reactions. Its structure will provide a basic hint towards (B4), for transcription.

Appendix B. $N_c = 2$ free monomers: towards the approximate smoluchowski equation (Steps 1, 2 and 3)

The system of 5 quantum entities $(1,2,1',2',E)$ involved in transcription in an aqueous medium at rest and thermal equilibrium at T, as considered in Subsection 3.1, will be described by a non-trivial extension of Appendix A. One should start from the extended system of 7 quantum entities $(P',1',2',T',E,1,2)$, augmented with: i) suitable internal degrees of freedom $u1$ and $u2$ in the initially free monomers 1 and 2 respectively and, for possible convenience, ii) degrees of freedom $u3$ for another free monomer 3, not participating directly in transcription but charged with the task of feeding further energy into E to allow it to escape far away (complementing the repulsive action of T'). Those internal variables amount to just toy models enabling to describe simply internal transitions in the free monomers 1, 2 and 3, coupled to E and liberating energy to account for the effective force F. For simplicity, we consider only suitable internal position variables $x_{u,i}$ for the free monomers $i = 1,2,3$. Real transitions in free monomers feeding energy into E, so that the latter behaves as a molecular motor, are far more complex than those in our toy model.

For brevity, the quantum Hamiltonian and density operators will be omitted. Let W_{ext} be the quantum mechanical (reversible) Wigner distribution function extended to the full system of 7 entities $(P',1',2',T',E,1,2)$ plus the internal degrees of freedom uj, $j = 1,2,3$, depending on t and on all positions and momenta (the q's, canonically conjugate to the positions). Thus, $W_{ext}(x_1,x_2,x_{P'},x_{1'},x_{2'},x_{T'},x_E,x_{u,1},x_{u,2},x_{u,3};q_1,q_2,q_{P'},q_{1'},q_{2'},q_{T'},q_E,q_{u,1},q_{u,2},q_{u,3},t) \equiv W([x_\sigma],[q_\sigma],t)$. In Step 1, upon generalizing (A1) and (A2) directly, W_{ext} fulfills the following Wigner dynamical equation, with $[dx'] \equiv \pi_{\rho'=P',1,2,1',2',E,T',(u,1),(u,2),(u,3)} dx'_{\rho'}$ and $[dq'] \equiv \pi_{\rho'=P',1,2,1',2',E,T',(u,1),(u,2)(u,3)} dq'_{\rho'}$, given by

$$\frac{\partial W_{ext}([x_\sigma],[q_\sigma],t)}{\partial t} = - \sum_{\rho=P',1,2,1',2',E,T',(u,1),(u,2),(u,3)} \frac{q_\rho}{m_{0,\rho}} \nabla_{x_\rho} W_{ext}([x_\sigma],[q_\sigma],t) + M_Q W_{ext} \tag{B1}$$

where $M_Q W_{ext}$ is defined by

$$M_Q W_{ext} = \frac{i}{\hbar(\pi\hbar)^{30}} \int\int W_{ext}(([x_\sigma],[q'_\sigma],t) \int (V_{ext}([x_\sigma + x'_\sigma]) - V_{ext}([x_\sigma - x'_\sigma]))$$

$$\exp\left(\frac{2i \sum_{\rho''=P',1,2,1',2',E,T',(u,1),(u,2)(u,3)} (q_{\rho''} - q'_{\rho''}) x'_{\rho''}}{\hbar} \right) [dq'][dx'], \tag{B2}$$

with $\frac{\partial}{\partial t} \int W_{ext}\ [dx][dq] = 0$, for any $t > 0$. V_{ext} will be given below. See [50, 54]. One may introduce the corresponding equilibrium Wigner function, the orthogonal polynomials generated by it and the non-equilibrium moments of W_{ext}, and proceed like in Appendix A. In particular, by integrating W_{ext} over all momenta, we introduce the non-equilibrium moment of lowest order:

$$W_{ext,0}(x_1,x_2,x_{P'},x_{1'},x_{2'},x_{T'},x_E,x_{u,1},x_{u,2},x_{u,3},t) = \int \frac{1}{\pi_{\rho=P',1,2,1',2',E,T',(u,1),(u,2),(u,3)} (2m_{0,\rho} k_B T)^{1/2}} W_{ext}\ [dq] \tag{B3}$$

where $m_{0,\rho}$ denotes the corresponding masses (masses being attributed formally to the uj's). $W_{ext,0}$ can be interpreted as a quantum mechanical probability distribution function for those variables, at time t.

The above Wigner dynamical equation yields an infinite linear hierarchy for all non-equilibrium moments of W_{ext}, thus far without approximation. At that stage, by extending the procedure in Appendix A, one performs suitable long-time approximations and some reasonable assumptions that allow to restrict the hierarchy to the moment of lowest order

$W_{ext,0}$. So, in physical terms, one implements that, in the time evolution, velocities relax much faster than positions, like in the dynamics of interacting Brownian particles, see [34]. Specifically, one accepts that the off-equilibrium σ-th ($\sigma = P', 1, 2, 1', 2', E, T, u1, u2, u3$) first moment can be approximated as the coefficients $\frac{m_{0,\sigma}D}{k_BT}a_\sigma$ times certain linear combinations depending on $W_{ext,0}$ and its first spatial derivatives. Such structures for the σ-th first moment generalize the off-equilibrium approximation $W_1 \simeq (mD/k_BT)\,[M_0W_0]$ in Appendix A. We do not consider the most general structures of that kind: rather we select those adequate for the actual model of transcription.

See Subsection 3.2 and below for the constants a_σ, except for Te. For the latter, we proceed differently, to take care of the fact that, even if at this stage (say, after the integration over its position in this Appendix) P' seems a passive spectator, P' is bound to $1'$ and still plays a role, and so for T' and $2'$. Accordingly, we interpret $a_{1'} = a_{av}/(r_{1'} + r_{P'})$ and $a_{2'} = a_{av}/(r_{2'} + r_{T'})$. Here, $r_{P'}$ and $r_{T'}$ are effective length parameters representing the whole effect of P' and T'. Then, we introduce $X_{Te} = \frac{a_{1'}^{-1}x_{1'} + a_{2'}^{-1}x_{2'}}{M_{Te}}$, $x_{Te} = x_{2'} - x_{1'}$, $M_{Te} = a_{1'}^{-1} + a_{2'}^{-1}$ and $m_{Te}^{-1} = a_{1'} + a_{2'}$, so that m_{Te} and M_{Te} (dimensionless) will play roles for Te similar to effective total and reduced masses of $1'$ and $2'$ (including the effect that $1'$ is "overloaded" with P' and $2'$ with T').

Then, in Step 2, we assume that $W_{ext,0}$ evolves in time according to the following approximate (irreversible) quantum Schmolukowski equation

$$\begin{aligned}\frac{\partial W_{ext,0}}{\partial t} =& D\frac{1}{M_c}\nabla_{X_c}\left(\nabla_{X_c}W_{ext,0} + \frac{W_{ext,0}}{k_BT}\nabla_{X_c}V_{ext}\right) + D\frac{1}{m_c}\nabla_{x_c}\left(\nabla_{x_c}(-2\varepsilon_c)W_{ext,0} + \frac{W_{ext,0}}{k_BT}\nabla_{x_c}V_{ext}\right)\\ &+ D\sum_{i'=P',1',2',T'} a_{i'}\nabla_{x_{i'}}\left(\nabla_{x_{i'}}(-2\varepsilon_{i'})W_{ext,0} + \frac{W_{ext,0}}{k_BT}\nabla_{x_{i'}}V_{ext}\right) + Da_{EE}\cdot\nabla_{x_E}J_E\\ &+ D_u\sum_{i=1,2,3}\frac{1}{a_{iu}}\nabla_{x_{u,i}}\left(\nabla_{x_{u,i}}W_{ext,0} + \frac{W_{ext,0}}{k_BT}\nabla_{x_{u,i}}V_{ext}\right).\end{aligned} \tag{B4}$$

The meaning of the dimensionless constants M_c, m_c and a's is explained in Subsection 3.2. The dimensionless M_c, m_c, M_{Te} and m_{Te} are proportional (but not identical) to the corresponding masses.

$W_{ext,0;in}(x_1, x_2, x_{P'}, x_{1'}, x_{2'}, x_{T'}, x_E, x_{u,1}, x_{u,2}, x_{u,3})$ is the initial distribution at the initial time $t = 0$. The quantum correction $(-2\varepsilon_c)$ for the monomers 1 and 2 generalizes (-2ε) in Appendix A and is discussed in Subsection 3.2. The quantum corrections $(-2\varepsilon_{i'})$ for $i' = P', 1', 2', T'$ plays a similarly necessary role for $1'$, $2'$, P' and T'. It will be unnecessary to explain them in further detail, since P' and T' will be integrated over in this Appendix and so will be $1'$ and $2'$ in Subsection 3.2.

The potential is given by

$$\begin{aligned}V_{ext} =& V_{12}(|x_2 - x_1|) + V_{1',2'}(|x_{2'} - x_{1'}|) + V_{P',1'}(|x_{1'} - x_{P'}|) + V_{2',T'}(|x_{2'} - x_{T'}|)\\ &+ V_{E,P'}(|x_E - x_{P'}|) + V_{E,T'}(|x_E - x_{T'}|) + V_{\text{Eall}} + V_{u1} + V_{u2} + V_{u3}\end{aligned} \tag{B5}$$

$$J_E = \nabla_{x_E}W_{ext,0} + \frac{W_{ext,0}}{k_BT}\nabla_{x_E}V_{ext} \tag{B6}$$

where $V_{1',2'}(|x_{2'} - x_{1'}|)$ is responsible for the permanent binding of $1'$ and $2'$: a simplified model for it is the Gaussian potential generic of the standard Gaussian model and used, in particular, in the Rouse model, see [36, 35]. $V_{P',1'}$ and $V_{2',T'}$ are potentials describing the influence of P' on $1'$ and of T on $2'$, respectively. $V_{E,P'}$ and $V_{E,PT}$ are potentials describing the

influence of P' and T' on E. V_{u1} and V_{u2} account for the internal transitions in the free monomers 1 and 2, in interaction with E, giving rise to energy which will keep E in motion. Similarly, V_{u3} describes similar internal transitions in the free monomer 3, otherwise a spectator regarding the actual transcription process. One has $V_{u1} = V_{u1}(|x_E - x_{u,1}|)$ and so on for the others.

A basic fact is that $W_{ext,0}$ includes quantum effects, that is ε_c accounts for the chemical reaction allowing the monomers 1 and 2, initially free, to become bound to each other. And so on for other ε's.

By performing partial integrations, for any $0 \leq t < +\infty$, it follows

$$\frac{\partial}{\partial t} \int \Pi_{\rho=P',1,2,1',2',E,T',u1,u2,u3} W_{ext,0}\, dx_\rho = 0. \tag{B7}$$

We shall introduce the distribution function W_0 as

$$W_0 = W_0(x_1, x_2, x_{1'}, x_{2'}, x_E, t) = \int W_{ext,0}\, dx_{P'}\, dx_{T'}\, dx_{u,1}\, dx_{u,2}\, dx_{u,3} \tag{B8}$$

and so on for $W_{P'}(x_{P'}, t)$, $W_{T'}(x_{T'}, t)$ and $W_{u1,u2,u3}(x_{u,1}, x_{u,2}, x_{u,3}, t)$, with the adequate integrations. We integrate (B4) for $W_{ext,0}$ over $x_{P'}$, $x_{T'}$, $x_{u,1}$, $x_{u,2}$ and $x_{u,3}$ and, inside integrals we assume, as an approximation, the factorized structure $W_{ext,0} = W_0 W_{P'} W_{T'} W_{u1,u2,u3}$. Those integrations make several contributions (those starting with $\nabla_{x_{i'}}$, $i' = P', T'$, $\nabla_{x_{u,i}}$, $i = 1,2,3$) in (B4) integrate to zero. The terms $V_{u1} + V_{u2} + V_{u3}$ give rise to x_E dependences, as we now explain. Even if the integrations $\int dx_{u,1}\, dx_{u,2}\, dx_{u,3}$ are carried out in principle over their whole three-dimensional spaces, the sizes of the domains where $V_{u1} + V_{u2} + V_{u3}$ contribute are adequately small (for instance, due to V_{uj}, the dominant contribution to the integration over $x_{u,j}$ comes from small $|x_E - x_{u,j}|$). Then, for instance (writing only the relevant contributions), using $V_{u1}(|x_E - x_{u,1}|) \simeq V_{u1}(0) + (\nabla_y V_{u1})_{y=0} \cdot (x_E - x_{u,1})$, reads

$$\nabla_{x_E} \int W_{P'}\, W_{T'}\, W_{u1,u2,u3}\, V_{u1}(|x_E - x_{u,1}|)\, dx_{u,1}\, dx_{u,2}\, dx_{u,3}\, dx_{P'}\, dx_{T'} \simeq \nabla_{x_E}(x_E \cdot (-F_1)) \tag{B9}$$

$$-F_1 = \int W_{P'}\, W_{T'}\, W_{u1,u2,u3}\, \nabla_y V_{u1}(|y|)_{y=0}\, dx_{u,1}\, dx_{u,2}\, dx_{u,3}, \tag{B10}$$

with similar approximations for the integrals of $V_{u2} + V_{u3}$. Then, one approximates the contribution of $V_{u1} + V_{u2} + V_{u3}$ by a global term linear in x_E, $-F \cdot x_E$, with constant F, having physical dimension of force. As those internal transitions feed energy into E, they amount to a molecular motor and, in turn, to a force tending persistently to keep E in motion. Then, on physical grounds (and independently on specific models for those internal transitions), the structure and sign of $-F \cdot x_E$ in comparison with the transcription process (see Figure 1) are reasonable.

After those integrations, and by including $-F \cdot x_E$, V_{ext} gives the total effective potential

$$V_{\text{eff,tot}} = V_{\text{eff}} - F \cdot x_E \tag{B11}$$

where

$$V_{\text{eff}} = V_{12}(|x_2 - x_1|) + V_{1',2'}(|x_{2'} - x_{1'}|) + V_{\text{P}',1',\text{eff}}(x_{1'}) + V_{2',\text{T}',\text{eff}}(x_{2'}) + V_{\text{E,P}',\text{eff}}(x_E) + V_{\text{E,T}',\text{eff}}(x_E) + V_{\text{Eall}} \tag{B12}$$

$$V_{\mathrm{Eall}} = V_{\mathrm{Eall}}(|x_1 - x_E|, |x_2 - x_E|, |x_{1'} - x_E|, |x_{2'} - x_E|, |x_1 - x_{1'}|, |x_2 - x_{2'}|) \equiv V_E(1,2,1',2'), \tag{B13}$$

$V_{\mathrm{P}',1',\mathrm{eff}}(x_{1'})$ and $V_{2',\mathrm{T}',\mathrm{eff}}(x_{2'})$ are effective potentials describing the influence of P' on $1'$ and of T' on $2'$.

V_{eff} describes interactions among the bound monomers, the free ones and E as the sum of (local and t-independent) potentials. Each potential accounts for an interaction affecting the entities indicated by the corresponding subscripts. V_{Eall} approach zero as $|x_2 - x_1|$ and $|x_{2'} - x_{1'}|$ tend to $+\infty$. The terms in $V_{P',1'}(|x_{1'} - x_{P'}|) + V_{2',T'}(|x_{2'} - x_{T'}|) + V_{E,P'}(|x_E - x_{P'}|) + V_{E,T'}(|x_E - x_{T'}|)$ in V_{ext} give rise respectively to those in $V_{\mathrm{P}',1',\mathrm{eff}}(x_{1'}) + V_{2',\mathrm{T}',\mathrm{eff}}(x_{2'}) + V_{\mathrm{E},\mathrm{P}',\mathrm{eff}}(x_E) + V_{\mathrm{E},\mathrm{T}',\mathrm{eff}}(x_E)$ in V_{eff}, when care is taken of integrations and dependences. Thus, $V_{\mathrm{P}',1',\mathrm{eff}}(x_{1'}) + V_{\mathrm{E},\mathrm{P}',\mathrm{eff}}(x_E) = \int W_{P'}(V_{P',1'} + V_{E,P'})\, dx_{P'}$ and so on. In agreement with the qualitative discussion in Section 2, V_{Eall} is, on average, weakly attractive and much smaller in magnitude than the largest values of $|V_{12}|$ and $|V_{1',2'}|$. $V_{1',2'}$ is strongly attractive (< 0) on average.

In Step 3, the result of the above integrations of equation (B4) for $W_{ext,0}$ is the following approximate irreversible quantum Smoluchowski equation for the quantum Wigner function $W_0(X_c, x_c, X_{Te}, x_{Te}, x_E, t)$ for $1, 2, 1', 2'$ and E (provided that the time t be, at least, not short)

$$\begin{aligned}\frac{\partial W_0}{\partial t} =& D\frac{1}{M_c}\nabla_{X_c}\left(\nabla_{X_c}W_0 + \frac{W_0}{k_BT}\nabla_{X_c}V_{\mathrm{eff,tot}}\right) + D\frac{1}{m_c}\nabla_{x_c}\left(\nabla_{x_c}(-2\varepsilon_c)W_0 + \frac{W_0}{k_BT}\nabla_{x_c}V_{\mathrm{eff,tot}}\right)\\ &+ D\sum_{i'=1',2'} a_{i'}\nabla_{x_{i'}}\left(\nabla_{x_{i'}}(-2\varepsilon_{Te})W_0 + \frac{W_0}{k_BT}\nabla_{x_{i'}}V_{\mathrm{eff,tot}}\right) + Da_{EE}\cdot\nabla_{x_E}J_{EE},\end{aligned} \tag{B14}$$

where J_{EE} is defined by

$$J_{EE} = \nabla_{x_E}W_0 + \frac{W_0}{k_BT}\nabla_{x_E}V_{\mathrm{eff,tot}}, \tag{B15}$$

and equation (B6) becomes directly (B15). As a direct justification of (B14), as it stands, it constitutes a variant (adapted to the present model of transcription and the inclusion of some quantum effects for the free monomers) of a well known equation for interacting Brownian particles (specifically, for the dynamics of colloidal suspensions, without quantum effects), see [33–35]. Notice that $\frac{\partial}{\partial t}\int W_0\, dx_1\, dx_2\, dx_{1'}\, dx_{2'}\, dx_E = 0$. The orders of magnitude for the ranges of V_{Eall}, $V_{11'}$ and $V_{22'}$ are chosen about d. One could entertain alternative possibilities for V_{Eall} and include other potentials $V_{11'}$ and $V_{22'}$: these lie outside our scope here.

Appendix C. Longer *Te* and transcript with $N_c > 3$ monomers

We shall give the essentials of the generalizations of the previous physical and mathematical developments in this work for longer transcript c and Te_{ext}.

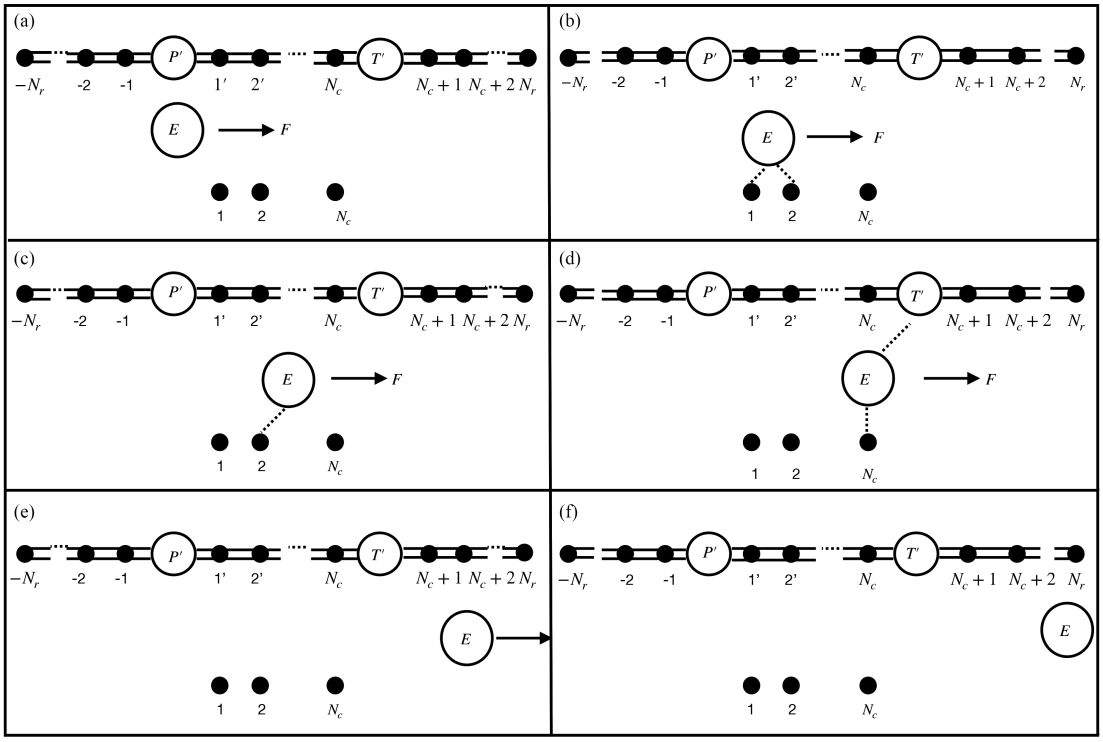


Figure C1. Subfigures (a)–(f) display successive stages in transcription of Te_{ext} (formed by left spectator subchain, promoter P', active subchain, terminator T' and right spectator subchain), enzyme (E) and free monomers (1,2, $\ldots, N_c$). The latter will become, as transcription evolves, successively bound among themselves, so as to form c (although the Subfigures do not display that)

Figure C1 displays the successive stages in transcription of Te_{ext} when there are initially N_c free monomers (namely, ribonucleoside three-phosphate ATP, GTP, CTP or UTP), numbered as $1, 2, \ldots, N_c$. There are also other free monomers, spectator ones, which do not participate directly in transcription but interact with and give energy to E. Te_{ext} is an extended linear chain formed by monomers (individually analogous to the free ones, except for the replacement of U by T), a promoter P' and a terminator T' strongly (covalently) bound sequentially as follows. Te_{ext} is formed by the monomers numbered as $-N_r, -N_r+1, \ldots, -2, -1, P', 1', 2', \ldots, N_c-1, N_c, T', N_c+1, N_c+2, \ldots, N_r$. The monomers $N_c+1, N_c+2, \ldots, N_r$ constitute the right spectator subchain. The monomers $-N_r, -N_r+1, \ldots, -2, -1$ constitute the left spectator subchain. The monomers $1', 2', \ldots, N_c$ are the *active* subchain Te, containing N_c bound monomers and involved in the transcription process through which the free monomers 1, 2,$\ldots, N_c$ will become sequentially bound to one another to form the transcript c. The monomer -1 is connected to P' which, in turn, is connected to $1'$ in the *active* subchain. The monomers N_c is connected to T' which, in turn, is connected to N_c+1 in the right spectator subchain.

Figure C1a displays how the enzyme E approaches, thanks to the force F acting on it, and it becomes weakly attracted by the promoter P' and the monomer $1'$ in Te. Then, in Figure C1b, the enzyme E also approaches and weakly traps the next monomer $2'$ in Te as well as the free monomers 1 and 2. In the latter process, energy is taken from monomer 1 (which suffers an internal transition) and fed into E, representing a first contribution to the effective F. Moreover, the free monomer 2 is also weakly trapped by E, with further liberation of energy transferred from 2 into E (a second

contribution to F). Therefore, the free monomers 1 and 2 experience a short-range reciprocal attraction and become strongly (covalently) bound to each other forming c. In Figure C1c the process follows for the next monomers in Te as well as the free monomers. In Figure C1d transcription is close to end as the terminator T' exerts a repulsive action on E (which may continue receiving, since the beginning of the transcription process, energy from other free monomers, not participating in transcription, but contributing to the constant force F). In Figure C1e, the transcript c has been formed, and in Figure C1f the enzyme E continues its approximate motion.

We shall limit to write the total potential among the free monomers and among E and the free monomers as: $\sum_{i=1}^{N_c-1} V_{ii,+1}(y_i)) + V_{\text{Eall}}$ ($y_i = x_{i+1} - x_i$, $i = 1,...,N_c - 1$). The actual V_{Eall} can be chosen as a sum of successive terms like that in (B13).

$$V_{\text{Eall}} = V_E(1,2,1',2') + V_E(2,3,2',3') + \ldots + V_E(N_c - 1, N_c, N_c' - 1, N_c') \tag{C1}$$

We assume that there are no bound states between: i) any free monomer and any constituent of Te_{ext}, ii) any free monomer and E, iii) E and any constituent of Te_{ext}. Suitable internal transitions are supposed to occur in each free monomer $(1,2,\ldots,N_c)$ involved in transcription (and in the spectator ones, not involved in transcription). All those internal transitions are mediated by interactions coupling all those free monomers to E and liberate energy, to be absorbed by E (molecular motor). The coupling interactions and the transitions are modeled through certain internal variables $[x_u]$), like in Appendix B. The overall effective interaction of all those free monomers with E will be approximated by the interaction potential: $-F \cdot x_E$. It amounts to an effective force F on E. One proceeds successively, by extending Appendix B, from the quantum-mechanical Wigner function. Then, in a practical way (after integrating over all momenta), one starts out from its lowest moment, namely, the quantum mechanical distribution function $W_{ext,0} = W_{ext,0}(x_1,...,x_{N_c},x_E,x_{-N_l},\ldots,x_{-1},x_{1'},\ldots,x_{N_c'},x_{N_c+1},\ldots,x_{N_r},x_{P'},x_{T'},[x_u],t)$. This $W_{ext,0}$ is the natural generalization of the one in Appendix B.

One performs similar long-time approximations and assumptions. At this stage, we suppose that the integration variables $[x_u]$ are somehow integrated over some distribution W_u and their influence approximated by the effective interaction $-F \cdot x_E$, like in Appendix B. Then, let all x's in Te_{ext} (left, "active" and right subchains) as well as $x_{T'}$, $x_{P'}$ be integrated over. Next, we shall proceed directly to the quantum mechanical distribution functions $W_c = W_c(x_1,\ldots,x_{N_c},t)$ for the free N_c monomers (generating eventually the transcript c) and to $W_E = W_E(x_E,t)$ (the distribution associated to E). One performs factorizations inside integrals: $W_0 = W_c \cdot W_E \cdot W_{\text{Te,active}} \cdot W_{\text{Te,left}} \cdot W_{\text{Te,right}}$, with obvious (and not displayed) notations and implicit dependences. One carries out approximations similar to (and generalizing) those for $N_c = 2$ monomers in Subsection 3.3, that is one approximates $W_{\text{Te,active}}$, $W_{\text{Te,left}}$ and $W_{\text{Te,right}}$ by Gaussian freely-jointed chains, by extending the ones in Subsection 3.3.

The most convenient spatial variables for W_c are $y_i = x_{i+1} - x_i$, $i = 1,\ldots,N_c - 1$ and $X_c = M_c^{-1} \sum_{i=1}^{N_c} a_i^{-1} x_i$, with $M_c = \sum_{i=1}^{N_c} a_i^{-1}$, so that one also has $W_c = W_c(y_1,\ldots,y_{N_c-1},X_c,t)$. The dimensionless constants a_i (strictly positive) are the direct generalization of the ones for $N_c = 2$. The non-equilibrium Smoluchowski equation for W_c, with initial condition $W_c = W_{c,in}$ at $t = 0$, generalizing (10), is given by

$$\frac{\partial W_c}{\partial t} = \frac{D}{M_c} \nabla_{X_c} \left(\nabla_{X_c} W_c + \frac{W_c}{k_B T} \nabla_{X_c} V_{\text{Eall,eff}} \right) + D \left(\sum_{i=1}^{N_c} a_i (\nabla_{y_{i-1}} - \nabla_{y_i}) \left(\nabla_{y_{i-1}} (-2\varepsilon_{i-1} W_c) - \nabla_{y_i} (-2\varepsilon_i W_c) \right) \right.$$

$$\left. + \frac{W_c}{k_B T} (\nabla_{y_{i-1}} - \nabla_{y_i}) \left(\sum_{i=1}^{N_c-1} V_{ii,+1}(y_i)) + V_{\text{Eall,eff}} \right) \right) \tag{C2}$$

$$V_{\text{Eall,eff}} = \int W_E \; W_{\text{Te,active}} \cdot W_{\text{Te,left}} \cdot W_{\text{Te,right}} \; V_{\text{Eall}} \; \Pi dx_{Te,E} \tag{C3}$$

with $\nabla_{y_0} \equiv 0$, $\nabla_{y_{N_c}} \equiv 0$ and $\varepsilon_i = \varepsilon_i(y_i)$. Here $\Pi dx_{Te,E}$ stands for the differential volume element for the integrations over x_E and all coordinates in Te_{ext}. Notice that, in part of (C2) use is also made (through V_{Eall}) of the spatial variables x_j, for the sake of shortened expressions. One has a separate non-equilibrium Smoluchowski equation for W_E, omitted for brevity, analogous to (8), with the corresponding new j_{EE}. Equation (C2) and that for W_E constitute a non linear system of two coupled equations for W_c and W_E. A subtlety here is that for very massive Te_{ext}, the counterpart of the approximation in (16) be inadequate (say, the actual counterpart of M_{Te} be so large that the t-dependence in the counterpart of $4(D/M_{Te})t$ be masked, for t values of interest). At this point, it is natural to argue that E could be expected to diffuse for long t and large x_E as shown in (26). Actually, a similar expectation could be entertained for $N_c = 2$, but it was omitted in Subsection 3.3. By invoking the last expectation here, for large t, $V_{\text{Eall,eff}}$ vanishes. Then, in (C2), the structure $\frac{D}{M_c}\nabla_{X_c}(\nabla_{X_c} W_c)$ enables to factor out the diffusive overall evolution of c in the X_c variable from the evolution in all the other variables y_l, with $l = 1,\ldots,N_c - 1$, for large t, $W_c \simeq W_{c,f}(X_c,t)W_{c,re}$, so that

$$\frac{\partial W_{c,re}}{\partial t} = D\sum_{i=1}^{N_c-1} a_i(\nabla_{y_{i-1}} - \nabla_{y_i})\left(\nabla_{y_{i-1}}(-2\varepsilon_{i-1}W_{c,re}) - \nabla_{y_i}(-2\varepsilon_i W_{c,re}) + \frac{W_{c,re}}{k_B T}(\nabla_{y_{i-1}} - \nabla_{y_i})\left(\sum_{i=1}^{N_c-1} V_{ii,+1}(y_i)\right)\right) \quad \text{(C4)}$$

Figure C1 displays the successive stages in transcription. The exact equilibrium solution of previous equation (C4) $\left(\frac{\partial W_{c,re,eq}}{\partial t} = 0\right)$ is $W_{c,re,eq} = W_{c,re,eq}(y_1,\ldots,y_{N_c-1})$, with $\nabla_{y_i}(-2\varepsilon_i)W_{c,re,eq} + \frac{W_{c,re,eq}}{k_B T}\nabla_{y_i} V_{i,i+1} = 0$, $i = 1,\ldots,N_c - 1$. For instance, $W_{c,re,eq} = W_{c,re,eq}(y_1) \times \ldots \times W_{c,re,eq}(y_{N_c-1})$, where each $W_{c,re,eq}(y_i)$ fulfills the counterpart of equation (31) for y_i. Notice that, $W_{c,re,eq}(y_i)$ is also given in terms of wave functions by the direct three-dimensional counterpart of (A7) in Appendix A. In the variables y_l, for large t the set of monomers 1, 2,..., N_c has a non-vanishing probability of being a bound molecular chain (the i-th monomer bound to the $(i-1)$-th and to the $(i+1)$-th ones). That probability increases as temperature decreases.

Thus far, we have disregarded the other spectator or complementary strand of double-stranded DNA. Its inclusion would involve more variables and interactions. However, it appears reasonable that, upon integrating over the new variables in addition to those leading to equation (C2) and performing similar approximate factorizations, Eq. (C2) for W_c associated to c would continue to hold. Now, the corresponding $V_{\text{Eall,eff}}$ would involve the new interactions and integration variables due to the complementary strand. Then, equation (C4) and its consequences would also hold.

Interesting cases are when $N_c = 3$, corresponding to an idealized codon or anticodon, and when N_c is not higher than 20, c corresponding to an idealized *RNA* generated in a typical transcription process occurring in double-stranded DNA.

To estimate time durations (say, through MFPTs) for $N_c > 2$ is far more difficult: however, it would not be unreasonable to expect that the dynamics implemented by (C4), with adequate interactions, would allow for the time duration of the binding of the monomers i and $(i+1)$ $(i = 1,\ldots,N_c - 1)$ be about 10^{-2} s. It is also difficult to estimate time durations for the evolutions of E and $W_{c,f}(X_c,t)$. However, for the mentioned interesting cases, moderate variations, within reasonable ranges, of the various potentials used in Section 5 may allow for not unreasonable estimates of the latter time durations, compatible with the time duration for the binding of the monomers i and $(i+1)$.