PCCP

ROYAL SOCIETY OF CHEMISTRY

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

# Multiscale modeling of reaction rates: application to archetypal $S_N2$ nucleophilic substitutions†

Jonathan Campeggio,[a‡] Marco Bortoli,[a‡] Laura Orian,[a] Mirco Zerbetto*[a] and Antonino Polimeno[a]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

We propose an approach to the evaluation of kinetic rates of elementary chemical reactions within Kramers' theory based on the definition of the reaction coordinate as a linear combination of natural, pseudo Z-matrix, internal coordinates of the system. The element of novelty is the possibility to evaluate the friction along the reaction coordinate, within a hydrodynamic framework developed recently [Campeggio, J. *et al. J. Comput. Chem.* 2019, **40**, 679-705]. This, in turn, allows to keep into account barrier recrossing, i.e. the transmission coefficient that is employed in correcting transition state theory evaluations. To test the capabilities and the flaws of the approach we use as case studies two archetypal $S_N2$ reactions. First, we apply the approach to the standard substitution of chloride ion to bromomethane. The rate constant at 295.15 K is evaluated to $k/c^{\ominus} = 2.7 \cdot 10^{-6}$ $s^{-1}$ (with $c^{\ominus} = 1$ M), which compares well to the experimental value of $3.3 \cdot 10^{-6}$ $s^{-1}$ [R. H. Bathgate, and E. A. Melwyn-Hughes, *J. Chem. Soc* 1959, 2642-2648]. Then, the method is applied to the $S_N2$ reaction of methylthiolate to dimethyl disulfide in water. In biology, such an interconversion of thiols and disulfides is an important metabolic topic still not entirely rationalized. The predicted rate constant is $k/c^{\ominus} = 7.7 \cdot 10^{3}$ $s^{-1}$. No experimental data is available for such a reaction, but it is in accord with the fact that the alkyl thiolates to dialkyl disulfides substitutions in water have been found to be fast reactions [S. M. Bachrach, J. M. Hayes, T. Dao and J. L. Mynar, *Theor. Chem. Acc.* 2002, **107**, 266-271].

## 1 Introduction

The evaluation of rates of elementary chemical reactions is part of a wide class of problems in theoretical/computational chemistry, i.e. the description of activated processes, such as conformation changes in polymers,[1] or folding/unfolding and large amplitude motions in proteins.[2] Activated processes are characterized by two or more relative minima separated by free energy barriers, $\Delta G^{\ddagger}$, large with respect to the thermal energy, i.e. $\Delta G^{\ddagger} >> k_B T$ (here $k_B$ is the Boltzmann constant, $T$ the absolute temperature). In such conditions, the passage from one minimum to another one is a *rare event*. Seen from the perspective of an all-atom molecular dynamics (MD) simulation, activated processes, which are characterized by time scales that can range from few tens of nanoseconds up, occur rarely considering the lengths of the MD trajectories that are nowadays affordable for complex systems. The computational challenge is to link the atomistic description with the kinetics of such rare events.

The straightforward estimate of a reaction rate would require the computation of the mean first passage time (the inverse of which is the reaction rate) across the barrier dividing the reactants state from the products state by means of an MD trajectory long enough to ensure statistical accuracy. However, such a direct route can not often be pursued due to two main issues. Firstly, rare events require the computation of many $\mu$s-long trajectories, carried out with short (of the order of a few femtoseconds) integration time step (notice that the crossing of high free energy barriers is subject also to a numerical problem: due to the finite representation of numbers in a computer, there is a numerical limit to the maximum momentum that a coordinate can acquire, which may prevent the barrier crossing). The second, straightforward, issue, specific to chemical reactions, is that quantum mechanics (QM) must be employed, at least to treat the reacting sub-system. The necessity of accurate QM treatments makes the production of long MD trajectories even more demanding. Still, multiscale approaches are an obvious useful modeling strategy to deal with reaction rates.[3] The basic philosophy of a multiscale method is to

[a] *DIpartimento di Sicenze Chimiche, Universià degli Studi di Padova, Via Marzolo 1, Padova, Italy. Tel: +39 049 5124; E-mail: mirco.zerbetto@unipd.it*

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/cXCP00000x/

‡ These authors contributed equally to this work.

divide the problem into different parts treated at different levels of accuracy: hybrid quantum mechanics / molecular mechanics (QM/MM), MD simulations can be employed to couple the reactive part to its environment, and in each of these sub-systems degrees of freedom can be partitioned into relevant and less relevant coordinates. The well-known challenge of multiscale approaches is to make the computational protocol self-consistent and capable to estimate the wanted observable without resorting to any fitting procedure to free parameters.

In particular, several multiscale approaches have been presented for the study of flexible molecules in the literature.[4–6] In this work, we propose a multiscale computational protocol for evaluating the kinetic rate of an elementary chemical reaction, based on the well-established Kramers[7] description of reactive rare events and coupled with an accurate estimate of the generalized friction tensor of flexible molecules.

A number of simple methods have been developed in the last hundred years to estimate the exponential pre-factor of the Arrhenius equation, trying to model equilibrium, non equilibrium, and even quantum systems.[8–11] Transition state theory (TST), originally developed by Eyring[12] is still the most widely spread and used model, at least among chemists. It is based on the debated concept of "reaction coordinate" as the collective coordinate that distinguishes the reactants state from the products state and passes through the transition state. TST usage, especially in computational software, is widespread because of its simplicity, even if it underestimates (in its original formulation, neglects) any recrossing of the transition state barrier. Thus, TST rates always overestimate the real rates. Many methods have been developed, based on the introduction of a correction factor $\kappa \leq 1$, of the pre-exponential factor, to take into account the fact that the reaction coordinate is coupled to the other coordinates of the system, which are ultimately responsible for recrossing.[13–18]

Kramers theory[7] is the second pillar in the history of reaction rates estimate in condensed phases. In Kramers view, the reaction coordinate is explicitly coupled to a thermal bath, as a generator of fluctuation-dissipation on the coordinate itself. The theory describes naturally recrossing of the transition state barrier, thus providing more accurate estimates of the pre-exponential factor. Improvements of the theory have been proposed, including potential of mean force effects.[19] Kramers theory can be straightforwardly applied also to other rare events processes. For example, it has been satisfactorily employed in interpreting loop formation of semiflexible polymers.[1] More recent approaches, which make also use of short MD simulations compared to the long time dynamics of the rare event, have been developed. We mention here the forward flux sampling (FFS).[20] The approach is based on building a series of hyperplanes between the initial and final states. Then, MD trajectories are computed to estimate the probability of crossing from one hyperplane to the adjacent one. Finally, the total net flux is built from the collected probabilities. The method has been adapted to treat flat energy barriers,[21] and solvent restructuring.[15]

The concepts of reaction coordinate and transition state are so intuitive and rooted into the chemist's view, that they are rarely abandoned even when they are ill-defined. A method that can be considered completely free of the TST-type details in describing the rare event of a chemical reaction is the invariant manifolds approach,[22] which is based on a stochastic (Langevin) description of the relevant dynamics in the phase-space of the reactive system. The method allows one to distinguish among reactive and non-reactive trajectories and then access the flux from reactants to products. Finally, we shall mention the transition path theory[23] as a general purpose method to study rare events. It is based on the analysis of the ensemble of transition paths, which are the "reactive" trajectories, to identify transition mechanisms and compute transition rates.

It is worthy to note that a common denominator among all of the above mentioned methods is the stochastic description of the relevant dynamics with respect to the rare event. In such a description, one important ingredient to be evaluated is the friction along the reaction coordinate, or more in general, the friction tensor along the selected degrees of freedom. Depending on how the reaction coordinates are selected, such stochastic variables may be or not Markovian. In the literature, some authors employ a generalized Langevin equation with the elements of the friction tensor computed from MD simulations as correlation functions of forces.[18] Several authors use a hydrodynamic interpretation of the friction tensor and obtain such a quantity from a Cartesian coordinates description of the relevant coordinates.[1,9,19] However, for a chemist it is more natural to reason in terms of internal degrees of freedom (bonds, bond angles, and torsion angles). For this purpose, we have recently developed a software tool, DiTe2,[24] for the calculation of the diffusion tensor of flexible molecules. As discussed in the next Section, DiTe2 allows to compute diffusive properties along collective coordinates defined as linear combinations of Z-Matrix-like internal coordinates. This makes easy to evaluate the friction tensor along a generalized reaction coordinate in internal coordinates and allows to establish an agile protocol to calculate reaction rates based on a general reaction coordinate.

The purpose of this paper is to present the protocol through two case-studies, namely the evaluation of the reaction rate of the nucleophilic substitution of chloride anion to bromomethane and the nucleophilic substitution of methylthiolate to dimethyl disulfide:

$$\mathrm{Cl^- + CH_3Br \rightarrow CH_3Cl + Br^-} \qquad \textbf{I}$$

$$\mathrm{CH_3S^- + CH_3SSCH_3 \rightarrow CH_3SSCH_3 + CH_3S^-} \qquad \textbf{II}$$

The former process is an archetypal $S_N2$ process reported as example in all organic chemistry textbooks. The latter elementary reaction is of paramount importance in biological systems, in which disulfide bond formation and disruption are essential processes, accompanied by thiol-disulfide exchange pathways in the complex cellular metabolism.[25] The rate and equilibrium constants have been determined for several biochemical reactions, for example for thiol-disulfide exchange reactions involving the oxidized glutathione,[26] for the reduction of disulfide bond in oxytocin and arginine-vasopressin by thiol disulfide interchange with glutathione and cysteine, and the formation of disulfide bonds

by thiol disulfide interchange with oxidized glutathione and cystine,[27] for thiol disulfide exchange in glutaredoxins.[28] The kinetics and thermodynamics of thiol disulfide redox regulation have been critically analyzed by Jensen and co-workers.[29] In addition, the nucleophilic attack of methylthiolate to dimethyl disulfide is an elementary step useful to model the reduction phase of the sulfur mutant of glutathione peroxidase (S-GPx)[30,31] and its mimics.[32–34] More in general, nucleophilic substitutions are archetypal elementary reactions of paramount importance in chemistry. On the basis of the molecularity, i.e. 1 and 2, $S_N1$ and $S_N2$ can be distinguished, which represent the two extreme situations in a broad spectrum of cases where, besides the chemical nature of the reactants, the environment controls the mechanism. Numerous computational studies have appeared in the last decade on model systems for nucleophilic substitutions at carbon,[35] silicon,[35] and phosphrus[36,37] centers, in which the interplay among the nature of the nucleophile, of the substrate, of the leaving group as well as the solvent has been elucidated. More recently, a systematic theoretical work on nucleophilic substitutions involving methylchalcogenolate as nucleophile and a dimethyldichalcogenide as substrate has been published,[38] integrating the pioneering works by Bachrach and co-workers which was mainly focused on sulfides and selenides.[39,40] When considering a thiolate as nucleophile attacking a disulfide substrate, the $S_N2$ mechanism is observed in condensed phase and a symmetric energy profile is calculated using QM methods.[38] Recently, the whole mechanism of $H_2O_2$ reduction mediated by GPx has been reconduced to a sequence of nucleophilic susbtitutions involving calchogenolates and dichalcogenides,[31] similarly to the mechanistic explanation of the oxidation of organoselenides in the initiation of green catalytic processes of paramount importance in organic chemistry.[41–44]

The activation barrier for the case-study reaction (see below) is about 16 in $k_BT$ units at room temperature. Because of such a large barrier, considering the high friction regime and slow dynamics of the reaction coordinate with respect to solvent (water) dynamics, Kramers theory is appropriate to describe the long time dynamics of the reacting system. The two quantities, required to define the Fokker-Planck equation[45] describing the stochastic dynamics of the reaction coordinate, are the free energy and the friction coefficient. Therefore, the computational protocol proposed, described in Section 2, is based on *i*) constructing the reaction path from the reactants to the products via the intrinsic reaction coordinate (IRC) method;[46] *ii*) building the reaction coordinate from the computed path, and *iii*) computing the friction along the reaction coordinate with a hydrodynamics approach. The reaction coordinate is written as a linear combination of those pseudo Z-matrix internal coordinates of the system identified as the relevant ones to describe the reaction mechanism. The coefficients of the combination are obtained following the methodology developed by Parrinello *et al.*[47] based on short, $O(10^0)$ ns, molecular dynamics simulations. Results are commented in Section 3, while a brief discussion is presented in Section 4.

## 2 Theory and Modeling

### 2.1 1D Kramers theory for a generalized reaction coordinate

We assume that a single reaction coordinate exists, and that it can be written as a linear combination of internal, pseudo Z-matrix, coordinates. A Fokker-Planck equation describing the time evolution along the coordinate is postulated on a phenomenological basis. A more rigorous approach would be to employ projection techniques to average "irrelevant" degrees of freedom, but is beyond the scope of this paper. Here we mention only that starting from a complete Hamiltonian in internal coordinates such procedures can be employed profitably.[48,49]

The Fokker-Planck equation[45] for a particle of mass $\mu$ moving along a coordinate $s$, with velocity $\dot{s}$ is

$$\frac{\partial}{\partial t}P(s,\dot{s},t) = -\hat{\Gamma}P(s,\dot{s},t) \tag{1}$$

with

$$\hat{\Gamma} = \dot{s}\frac{\partial}{\partial s} - \frac{1}{\mu}\frac{dU(s)}{ds}\frac{\partial}{\partial \dot{s}} - k_BT\frac{\xi}{\mu^2}\frac{\partial}{\partial \dot{s}}P_{eq}(s,\dot{s})\frac{\partial}{\partial \dot{s}}P_{eq}^{-1}(s,\dot{s}) \tag{2}$$

where $U(s)$ is the potential of mean force acting on $s$, $k_B$ is the Boltzmann constant, $T$ the absolute temperature, $\xi$ the friction opposing to a change in $s$, and $P_{eq}(s,\dot{s})$ the Maxwell-Boltzmann equilibrium probability density.

Following standard vibrational analysis,[50] by roto-translating the coordinates of the system in such a way the Eckart conditions are satisfied with respect to a reference structure, the reduced mass ($\mu$) along the single reaction coordinate ($s$) can be calculated as

$$\mu^{-1} = \sum_{i=1}^{3N}\frac{1}{m_i}\left(\frac{\partial s}{\partial x_i}\right)^2 \tag{3}$$

where $x_i$ is one of the Cartesian coordinates of one of the atoms in the system, and $m_i$ is the inertia (mass) associated to that coordinate. The equilibrium probability density reads

$$P_{eq}(s,\dot{s}) = \frac{\exp\{-[U(s)+T(s,\dot{s})]/k_BT\}}{\langle\exp\{-[U(s)+T(s,\dot{s})]/k_BT\}\rangle} \tag{4}$$

with $\langle\dots\rangle = \int ds \int d\dot{s}\dots$, and the kinetic energy

$$T(s,\dot{s}) = \frac{1}{2}\mu(s)\dot{s}^2 \tag{5}$$

We summarize the approximations introduced up to this point: *i*) only one reaction coordinate is postulated; *ii*) the reaction coordinate is approximately decoupled from all of the other external and internal degrees of freedom, in particular from the overall rotational motion in the Hamiltonian, as well as hydro-dynamically (as discussed in the Appendix); *iii*) the friction coefficient is assumed not to depend upon the reaction coordinate.

It is convenient to introduce the scaled quantities $x = s$, $v = \dot{s}\sqrt{\mu/k_BT}$, $u(x) = U(s)/k_BT$. The Fokker-Planck operator is thus converted into

$$\tilde{\Gamma} = \omega_s\left(v\frac{\partial}{\partial x} - \frac{du}{dx}\frac{\partial}{\partial v}\right) - \omega_c\frac{\partial}{\partial v}P_{eq}(x,v)\frac{\partial}{\partial v}P_{eq}^{-1}(x,v) \tag{6}$$

where the streaming frequency $\omega_s = \sqrt{k_B T/\mu}$, and the collision frequency $\omega_c = \xi/\mu$ have been introduced.
The Fokker-Planck equation is now written in the common form of a massless particle along a single coordinate. Following Kramers' analysis,[7] assuming that the potential can be written as

$$u_i = u_{0,i} \pm \frac{1}{2}\frac{\omega_i^2}{\omega_s^2}(x - x_i)^2 \tag{7}$$

where $i = \mathrm{A}, \mathrm{TS}$, i.e. the point along the reaction coordinate corresponding to the reactants (A) or to the transition state (TS), $u_{0,i}$ is the energy shift at point $i$, $\omega_i = \sqrt{K_i}\omega_s$, and $K_i$ the Hessian of $u$ at point $i$, under the local harmonic approximation. The high barrier/high friction limit gives

$$k \approx \frac{\omega_\mathrm{A}\omega_\mathrm{TS}}{2\pi\omega_c}\exp\left(-E_\mathrm{a}/k_B T\right) \tag{8}$$

where $E_\mathrm{a}/k_B T = u_{0,\mathrm{TS}} - u_{0,\mathrm{A}}$ is the difference between the energy of the transition state and the energy of the reactants, i.e. the activation energy.

### 2.2 QM calculation of the reaction path

The Amsterdam Density Functional software was used in all the calculations based on density functional theory (DFT).[51–54] The scalar zeroth order regular approximation (ZORA) was employed to account for relativistic effects.[55] The OLYP[56–59] density functional was used, in combination with a basis set of large uncontracted Slater-type orbitals (STOs) of triple-ζ quality, augmented with two sets of polarization functions on each atom (TZ2P): 2p and 3d in the case of H, 3d and 4f in the case of C and S. To accurately represent the Coulomb and exchange potentials an auxiliary set of s, p, d, f and g STOs was employed in each SCF cycle. The frozen-core approximation was employed: up to 1s for C, S and Cl and up to 3p for Br. This level of theory has been recently benchmarked and applied with success to study the reactivity of organic halides and dichalcogenides.[35,60,61] Solvent effects (water) have been simulated using the conductor-like screening model (COSMO),[62] as implemented in the ADF program. We have used a solvent-excluding surface with an effective radius for water of 1.93 Å, derived from the macroscopic density, molecular mass, and a relative dielectric constant of 78.39. The empirical parameter in the scaling function in the COSMO equation was chosen equal to 0.0. The radii of the atoms were taken to be MM3 radii,[63] divided by 1.2, giving 1.350 Å for H, 1.700 Å for C, 1.725 Å for Cl, 1.850 Å for Br and 1.792 Å for S. The level of theory is therefore denoted COSMO-OLYP/TZ2P. The transition state for both nucleophilic substitutions was fully optimized at the COSMO-OLYP/TZ2P level of theory and verified through vibrational frequencies calculation. The only imaginary frequency present, associated with the normal mode corresponding to the correct motion, was then followed along a steepest-descent path to arrive at the reactants and products through an IRC calculation.

### 2.3 Reaction coordinate

The reaction coordinate is here defined as a linear combination of internal, pseudo Z-Matrix, coordinates. A heuristic selection of the important descriptors is here pursued, based on the inspection on how the internal coordinates change along the IRC. A similar approach has been followed in a recent paper in the QM/MM investigation of the aminoacylation reaction mechanism of *Thermus thermophilus* glutamyl-tRNA synthetase (GluRS) and its interactions with its cognate tRNA (tRNAGlu).[64] There the authors tested 4 different possible choices of the relevant internal degrees of freedom, and used comparison with the experimental reaction rate to decide the set of internal coordinates that most likely describes the mechanism of the aminoacylation reaction.
The choice of the important internal coordinates for the two reactions studied in this work is discussed in detail in the Results Section. The protocol that has been followed was to collect the selected internal degrees of freedom into an array, $\mathbf{q}(\mathbf{R})$, where $\mathbf{R}$ are the Cartesian coordinates of all of the atoms. We shall stress that internal coordinates are built considering all of the atoms of the species involved in the reaction. Thus, distances and angles may be defined as functions of the coordinates of atoms belonging to different molecules. Then, following the approach proposed by Parrinello *et al.*,[47] the reaction coordinate, $s$, can be built as

$$s = \left(\mathbf{q}_\mathrm{A} - \mathbf{q}_\mathrm{B}\right)^{tr}\left(\boldsymbol{\sigma}_\mathrm{A}^{-1} + \boldsymbol{\sigma}_\mathrm{B}^{-1}\right)\mathbf{q} \tag{9}$$

where the subscripts A and B indicate, respectively, the reactant and the product states. The array $\mathbf{q}_i$ ($i$=A, B) collects the average values of the internal coordinates around the state $i$, while $\boldsymbol{\sigma}_i$ is the covariance matrix describing the fluctuations (as standard deviations) around $\mathbf{q}_i$, and is defined as

$$\boldsymbol{\sigma}_i = \langle \mathbf{q}_i - \langle \mathbf{q}_i \rangle \rangle \otimes \langle \mathbf{q}_i - \langle \mathbf{q}_i \rangle \rangle \tag{10}$$

Following Parrinello's method,[47] $\mathbf{q}_i$ and $\boldsymbol{\sigma}_i$ can be accessed by MD trajectories which are short with respect to the time scale of the reaction.

### 2.4 Molecular dynamics simulations

MD simulations targeted at parametrizing Equation 9 have been calculated with NAMD 2.12,[65] using the CHARMM[66] `par_all36_prot` parameters for the solutes, and TIP3P water.
For reaction **I**, two short MD simulations were carried out in the reactants and products states in order to compute, respectively, $\boldsymbol{\sigma}_A^\mathrm{I}$ and $\boldsymbol{\sigma}_B^\mathrm{I}$. Simulations have been carried out in the canonic ($NVT$) ensemble with $T = 295.15$ K. A cubic simulation box of side length of 20 Å with periodic boundary conditions was employed. A total of 267 water molecules were added to reproduce water density at 295.15 K, while 1 sodium ion was added to neutralize the total charge.
For reaction **II**, instead, due to the symmetry of the problem (i.e., the reactants and the products are the same chemical species) it was sufficient to produce only one single trajectory. Then, in Equation 9 it was set $\boldsymbol{\sigma}_\mathrm{A}^\mathrm{II} = \boldsymbol{\sigma}_\mathrm{B}^\mathrm{II}$. The simulation has been carried out in the canonic ($NVT$) ensemble with $T = 298.15$ K. A cubic simulation box of side length of 20 Å with periodic boundary con-

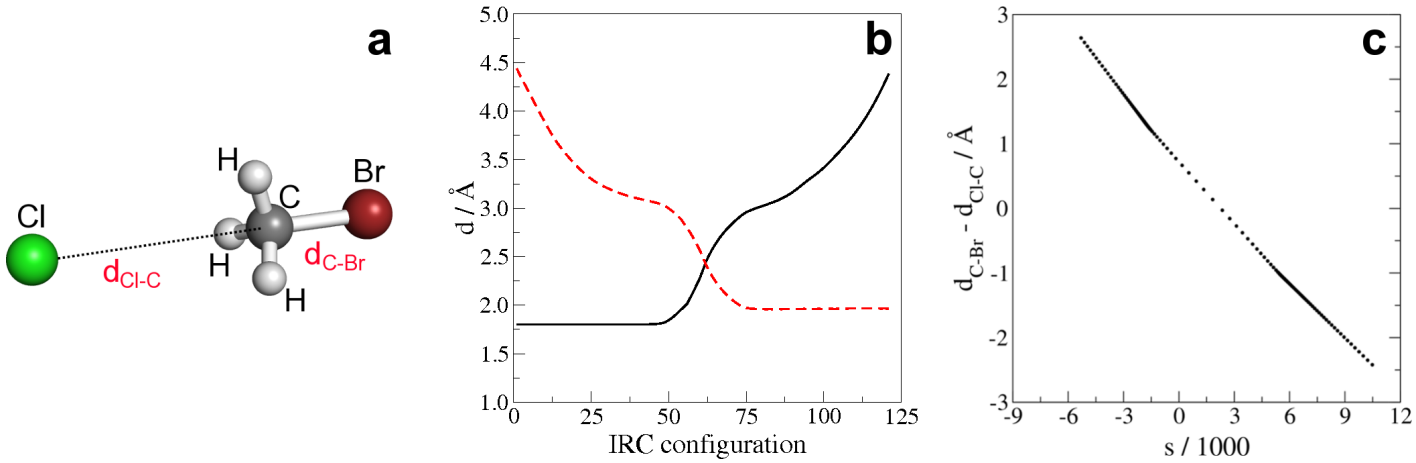


**Fig. 1** (a) Balls & sticks representation of the reactants in reaction I. The two internal degrees of freedom chosen to build the reaction coordinate, namely the two alogen-carbon distances, are highlighted in red. (b) Plot of the change of the $d_{\mathrm{C-Br}}$ (black, solid line), and the $d_{\mathrm{Cl-C}}$ (red, dashed line) distances along the IRC. (c) Plot of the reaction coordinate $s$ determined via the method of Parrinello vs. the intuitive reaction coordinate.

ditions was employed. A total of 260 water molecules were added to reproduce water density at 298.15 K, while 1 sodium ion was added to neutralize the total charge.

Common simulation parameters were: temperature coupling to thermostat the system, pair list distance of 11.5 Å, non-bonded interaction cutoff of 11 Å with smoothing switch at 10.5 Å, and PME electrostatics. Since all bonds with hydrogen atoms were kept rigid, an integration time step of 2 fs was used. The simulation protocol consisted in energy minimization, followed by heating to the target temperature with a ramp increasing temperature by 2 K each 300 MD steps. Then, system equilibration was carried out for 100 ps. Finally, 1 ns of production MD has been computed.

In all of the simulations, the two chemical species were confined to fluctuate around the reactants or products states[47] introducing harmonic constraints. For reaction **I**, the $Cl^- \cdots C$ distance in the reactants state was enforced at 4.4 Å with a harmonic force constant of 920.5 kJ mol$^{-1}$ Å$^{-2}$. Analogously, the $C \cdots Br^-$ distance in the products state was imposed to be 4.4 Å by means of a harmonic force constant of 669.4 kJ mol$^{-1}$ Å$^{-2}$. For reaction **II**, a single harmonic term between the S atom of methylthiolate and the target S atom of dimethyl disulfide was added to the force field to constrain their distance at 4.53 Å, using a force constant of 723.8 kJ mol$^{-1}$ Å$^{-2}$. The force constants were set equal to those of the C-Cl, C-Br, or S-S bonds, while the distances correspond to those in the reactants or products states as in the IRC paths.

### 2.5 Generalized friction tensor

The evaluation of the dissipative properties was based on a hydrodynamic approach.[24] The system, immersed in a homogeneous isotropic fluid of known viscosity, is described as a set of rigid fragments (made of atoms or groups of atoms) connected via different types of joints allowing for stretching, bending and torsional motions. Internal coordinates can be defined as linear combinations of such Z-Matrix like degrees of freedom, and the friction tensor elements along these generalized coordinates can be evaluated.

In our case, the generalized friction tensor $\boldsymbol{\xi}$ is represented by a $7 \times 7$ matrix that can be conveniently partitioned into translation (TT), rotation (RR), internal (ss), and mixed (TR, Ts, Rs) blocks, with s indicating specifically that in this application, the "internal" motion is the reaction coordinate

$$\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\xi}_{\mathrm{TT}} & \boldsymbol{\xi}_{\mathrm{TR}} & \boldsymbol{\xi}_{\mathrm{Ts}} \\ \boldsymbol{\xi}^{tr}_{\mathrm{TR}} & \boldsymbol{\xi}_{\mathrm{RR}} & \boldsymbol{\xi}_{\mathrm{Rs}} \\ \boldsymbol{\xi}^{tr}_{\mathrm{Ts}} & \boldsymbol{\xi}^{tr}_{\mathrm{Rs}} & \xi_{\mathrm{ss}} \end{bmatrix} \tag{11}$$

The DiTe2 software recently developed in our laboratory,[24] allows to compute the friction along generalized internal coordinates defined as linear combinations of internal coordinates, $\boldsymbol{q}$, *i.e.* $s = \sum_j c_j q_j$, where the coefficients of the combination are those provided by Equation 9. In the calculation of the generalized friction tensor, the derivatives $\partial \boldsymbol{R}/\partial s$ are required. Using the map between Cartesian and Z-Matrix coordinates (see ref. 24 for details), it is found that

$$\frac{\partial \boldsymbol{R}}{\partial s} = \sum_j c_j \frac{\partial q_j}{\partial s} \frac{\partial \boldsymbol{R}}{\partial q_j} \tag{12}$$

where the partial derivatives $\partial \boldsymbol{R}/\partial q_j$ are computed analytically in DiTe2,[24] while the partial derivatives $\partial q_j/\partial s$ can be obtained numerically from the $q_j(s)$ plots once the reaction coordinate has been built from the IRC reaction path.

In the Appendix it is shown as a weak hydrodynamic coupling is observed between the rigid-body like motions of the system and $s$. This justifies the employment of the 1D Fokker-Planck operator in Equation 2, where just the internal friction coefficient enters in the expression. Thus, in the following, we shall consider only $\xi = \xi_{ss}$, dropping the subscript to simplify the notation.

## 3 Results

### 3.1 Chloride ion to bromomethane substitution

Reaction **I** has been taken into consideration in this work since, on one hand, experimental data is available. On the second hand, the low dimensionality of the problem allows to easily understand capabilities and flaws of the computational protocol here

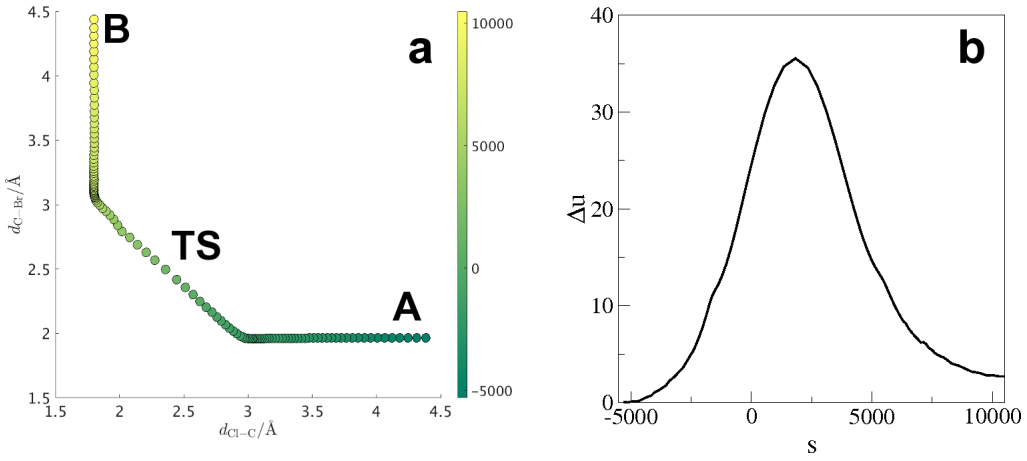


**Fig. 2** (a) Projection of the IRC on the $d_{\mathrm{Cl-C}}d_{\mathrm{C-Br}}$ subspace. The color map shows the corresponding value of the reaction coordinates, permitting to visually identify an IRC configuration as being closer to the reactants (A, yellow), products (B, dark green), or transition state (TS, light green) (b). Energy profile $\delta u(s) = u(s)u_{0,A}$ in $k_BT$ units at 295.15 K as function of the reaction coordinate as obtained from QM calculations in implicit solvent.

proposed to predict rate constants.

As a first step, the reaction coordinate has to be built. In this case, the choice of the set of internal coordinates that constitute the relevant subspace in the substitution reaction is straightforward. In particular, we used the two Cl-C ($d_{\mathrm{Cl-C}}$) and C-Br ($d_{\mathrm{C-Br}}$) distances (shown in Figure 1a) along the straight line containing the 3 heavy atoms of the system. Figure 1b shows how the selected internal coordinates change along the IRC. Initially the $Cl^-$ ion approaches the carbon atom, with $d_{\mathrm{C-Br}}$ remaining constant. Around the transition state, a concerted movement of the two alogens in observed. Finally, the $d_{\mathrm{Cl-C}}$ distance is fixed to its equilibrium value while the $Br^-$ ion leaves the just formed chloromethane.

Usually, the intuitive difference $s_I = (d_{\mathrm{C-Br}} - d_{\mathrm{Cl-C}})$ is used as reaction coordinate. Following the method described by Parrinello,[47] we obtained

$$s = 3575.8\ \text{Å}^{-1} \times (d_{\mathrm{Cl-C}} - 0.74\, d_{\mathrm{C-Br}}) \qquad (13)$$

with the distances expressed in Å.

The resulting definition is coherent with the intuitive construction of the reaction coordinate, with a small asymmetry in the factors multiplying the two alogen-carbon distances that takes into account for the different local librations. The factor 3575.8 is merely a scaling constant that is canceled out at the end in the computation of the rate constant, since the same scaling enters also in the friction tensor (or in the reduced mass, if TST evaluation is to be carried out). Figure 1c plots $s_I$ v.s. $s$ to compare the intuitive reaction coordinate with the one obtained with Parrinello's analysis (Equation 13). A nearly linear relation is observed, confirming that the choices done to parametrize the constraints in the MD simulations allow us to obtain a reliable physical picture.

Figure 2a shows the configurations of the system along the IRC in the ($d_{\mathrm{Cl-C}}$, $d_{\mathrm{C-Br}}$) subspace. The colormap applied to the circles maps between the IRC configuration and the "position" of the system along the reaction coordinate. Circles in yellow are those configurations close to the reactants state, light green around the transition state, while dark green indicates configurations around the products state. The transition state is located at a point where $d_{\mathrm{Cl-C}} = 2.357$ Å and $d_{\mathrm{C-Br}} = 2.495$ Å. The system energy along the reaction coordinate, $\Delta u(s) = u(s) - u_{0,\mathrm{A}}$, is shown in Figure 2b. Transition state is observed at $s = 1829$, and as expected, the final energy of the products ($s = 10482$) is lower than that of the reactants ($s = -5300$). The energy profile reported is the one calculated employing the COSMO model for the solvent effects. As previously reported[67] the contribution of hydration of the halogen ions is underestimated and the use of the experimental value[68] resulted a good approximation. Thus, also in this context, an empirical correction to the free energy of chloride ion employing experimentally determined hydration Gibbs free energy contributions was applied.[68] The computed activation energy for reaction I resulted 95.81 kJ/mol. We assume as further approximation that the curvatures are not affected importantly by such a correction.

For the evaluation of the rate constant, the friction coefficient has been evaluated at the transition state ($s = 1829$), since in deriving Equation 8, $\xi$ is introduced with the flux of the system through the energy barrier. Using the hydrodynamic approach briefly described in Section 2.5, using an effective radius of the beads (heavy, non-H, atoms) of 2 Å, temperature 295.15 K, water viscosity 0.953 cP, and stick boundary conditions. Hydrodynamic interactions among the beads have been computed with the Rotne-Prager model.[69] We estimated $\xi = 5.44 \cdot 10^{-39}$ N m s. The friction tensor is here nothing more than a translational friction of the concerted motion of the two halogen atoms. Converting the friction into a translational diffusion coefficient, one obtains $5.85 \cdot 10^{-10}$ $m^2/s$, which is comparable to experimental translational diffusion coefficients of ions in water, which are in the range $10^{-10} - 10^{-9}$ $m^2/s$.[70] From such a good comparison, we considered acceptable the hydrodynamic interpretation of the friction coefficient. Finally, putting all together, the estimated rate constant is $k/c^{\ominus} = 2.7 \cdot 10^{-6}$ $s^{-1}$, with $c^{\ominus} = 1$ M. The computed value compares very well with the experimental value at 295.15 K of $k_{\mathrm{exp}}/c^{\ominus} = 3.3 \cdot 10^{-6}$ $s^{-1}$.[71]

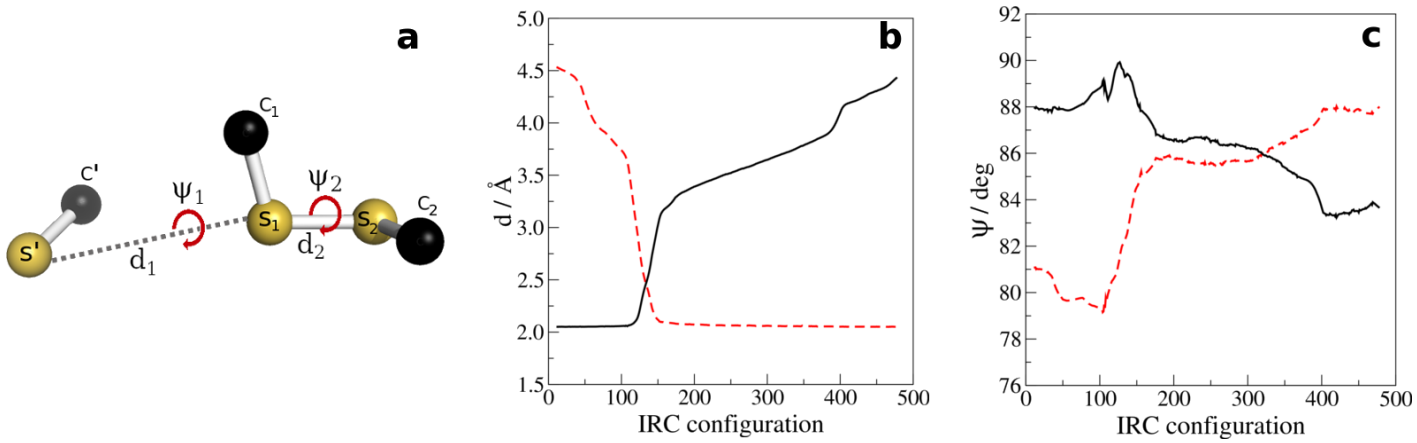


**Fig. 3** (a) Balls and sticks representation of the chemical species involved in the reaction (hydrogen atoms are omitted). The relevant reaction internal degrees of freedom are shown. (b) Variation of the two distances $d_1$ (S'-$S_1$, red dashed line) and $d_2$ ($S_1$-$S_2$, black solid line) along the IRC. (c) Variation of the torsion angles $\psi_1$ (C'-S'-$S_1$-$C_1$, red dashed line) and $\psi_2$ ($C_1$-$S_1$-$S_2$-$C_2$, black solid line) along the IRC.

A comment is to be done here. The activation energy obtained from QM calculations is 7 kJ/mol smaller then the value which is accepted as the reference "experimental" value,[71] which is within the accuracy of DFT methods. With the quotes we want to underline that the real experimental measurement is the rate constant. The activation energy is recovered once an interpretative model for $k$ is employed. The usual value of 102.93 kJ/mol[71] is obtained by considering the pre-exponential factor as temperature independent. In Kramers' theory, however, it depends on temperature mainly because of the friction coefficient at the denominator. Thus, it is expected that the two interpretations give origin to two different partitions into pre-exponential and exponential terms, with consequently different definitions of the activation energy. Once the computational protocol is consistent with the model, the final comparison to be done is on the value of the rate constant.

### 3.2 Methyltiolate anion to dimethyl disulfide substitution

In reaction **II**, four internal degrees of freedom were identified, namely the distances $d_1$ and $d_2$ between, respectively, atoms S'-$S_1$ and $S_1$-$S_2$ (for the labels, see Figure 3a), and the two dihedral angles $\phi_1$ and $\phi_2$ defined, respectively, by atoms C'-S'-$S_1$-$C_1$ and $C_1$-$S_1$-$S_2$-$C_2$. The two S-S distances describe the concerted entrance/exit of a methyl thiolate group, while the two dihedral angles provide the molecular rearrangement necessary to form the final products. Figures 3b,c depict the variation of the selected pseudo Z-Matrix coordinates of the two reactions along the computed IRC.

As highlighted above, for reaction **II** the map in Equation 9 is parametrized with the simplification $\boldsymbol{\sigma}_{\mathrm{A}}^{II} = \boldsymbol{\sigma}_{\mathrm{B}}^{II}$, since reactants and products are the same chemical species. The arrays $\mathbf{q}_{\mathrm{A}}$ and $\mathbf{q}_{\mathrm{B}}$ are built using the two IRC points that are considered the initial and final states. Figures 4a,b show the projections of the IRC configurations on the $d_1 - d_2$ and $\psi_1 - \psi_2$ subspaces. The color provides a visual map between each IRC configuration and its "position" along $s$. Dots in yellow are configurations close to the reactants state (A), the dark green dots are configurations in the products state (B), while light green dots are those around the transition state (TS), for which $s \sim -1025$. Figure 4a shows a very neat and simple transition in the subspace of the distances, with $d_1$ changing in a concerted way with $d_2$, and the TS located at the center of the path between the two extreme points. The torsion angles, in Figure 4b, instead, show a much more complex behavior. Starting from A, $\psi_2$ (the disulfide reagent torsion angle) increases while approaching the TS. The other torsion angle, instead, changes while the thiolate S' atom gets closer to $S_1$. The complex behaviour in the A and B states may be due to the fact that when the reactants (or the products) are further away, the relative orientation of the molecules (which affects the pseudo-torsion angle) is not driven by any important potential energy interaction. In fact, the angles fluctuations found in the $\boldsymbol{\sigma}$ variance-covariance matrix are two orders of magnitude larger than those for the distances.

Figure 4c shows the system internal energy along the reaction coordinate, $\Delta u(s) = u(s) - u_{0,\mathrm{A}}$. As expected, the energy profile is approximately symmetric, with a small asymmetry that may arise from the fact that for simplicity we used $\mathbf{q}_{\mathrm{A}}$ and $\mathbf{q}_{\mathrm{B}}$ taken directly from the IRC and not as averages from the MD trajectory.[47] The energy profile shows a large activation energy $E_{\mathrm{a}}/k_B T = 16.1$. The high friction limit can be invoked, so that Equation 8 can be safely employed.

The friction tensor has been computed on the transition state configuration, with the following parameters: effective radius of the beads (heavy, non-H, atoms) of 2 Å, temperature 298.15 K, water viscosity 0.894 cP, and stick boundary conditions. Hydrodynamic interactions among the beads have been computed with the Rotne-Prager model.[69] The validity of high friction regime approximation is ensured by checking that $\omega_c/2 >> \omega_{\mathrm{TS}}$. To such a purpose, the reduced mass has been computed using the approximate expression given in Equation 3 (we recall that coupling with rotational motion is neglected in the formulation of the kinetic energy). Derivatives of $s$ with respect to Cartesian coordinates of the atoms have been computed analytically using an automatic differentiation algorithm. In the computation, methyl groups have been considered as unique mass points. The com-

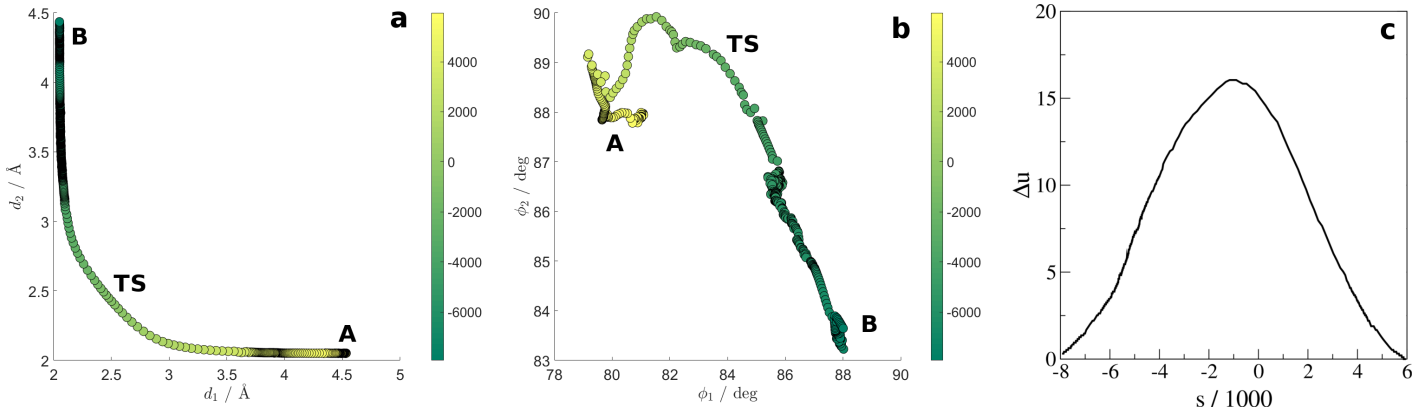


**Fig. 4** (a) Projection of the IRC on the $d_1 - d_2$ subspace. The color map shows the corresponding value of the reaction coordinate $s$, allowing one to visually identify an IRC configuration as being closer to the reactants (A, yellow), products (B, dark green), or transition state (TS, light green). The latter is located around $s = -1025$. (b) Same as panel a, but in the $\psi_1 - \psi_2$ subspace. (c) Energy profile $\Delta u(s) = u(s) - u_{0,\mathrm{A}}$ at 298.15 K as function of the reaction coordinate.

puted reduced mass, which in this case has the units of an inertia tensor since $s$ is adimensional, is practically independent from the value of the reaction coordinate, with changes below 0.1%. The estimated value is $\mu = 1.0975 \cdot 10^{-53}$ kg m$^2$. From a local harmonic approximation of $u(s)$ around the transition state, we find $|K_{\mathrm{TS}}| = 1.5166 \cdot 10^{-6}$. Thus, it is possible to calculate the collision frequency $\omega_c = 5.37 \cdot 10^{14}$ s$^{-1}$, and the frequency $\omega_{TS} = \sqrt{|K_{\mathrm{TS}}| k_B T/\mu} = 2.38 \cdot 10^{13}$ s$^{-1}$. The inequality $\omega_c/2 >> \omega_{\mathrm{TS}}$ can be considered satisfied, allowing to employ Equation 8 to evaluate the reaction rate.

Then, to estimate $k$, we performed a local harmonic approximation of $u(s)$ around A, thus evaluating $K_{\mathrm{A}} = 2.7951 \cdot 10^{-7}$ and $\omega_{\mathrm{A}} = \sqrt{K_{\mathrm{A}} k_B T/\mu} = 1.02 \cdot 10^{13}$ s$^{-1}$. Finally, putting all together, the computed reaction rate is $k/c^{\ominus} = 7.7 \cdot 10^3$ s$^{-1}$. The result obtained for our model reaction is in accord with experimental observations that thiolate to disulfide substitutions in water are fast reactions.[40,72,73]

If Equation 8 is compared with the classical Eyring TST reaction rate formula, with recrossing correction

$$k_{TST} = \kappa \frac{k_B T}{h} \exp\left(-\Delta G^{\ddagger}/k_B T\right) \tag{14}$$

the factor $k_B T/h = 6.2 \cdot 10^{12}$ s$^{-1}$ compares well with $\omega_{\mathrm{A}}/2\pi = 1.6 \cdot 10^{12}$ s$^{-1}$. Thus in simple TST equation, $\omega_{\mathrm{A}}/2\pi$ is the pre-exponential factor[8] counting the flux across the energy barrier. Thus, $\kappa = \omega_{TS}/\omega_c = 0.04$ is the estimated transmission coefficient. Such a low $\kappa$ is expected in the high friction regime, where random "collisions" of the reactive system with the environment makes barrier recrossing a relevant event in determining the reaction rate constant.

A second test has been carried out by considering a different set of internal degrees of freedom, in particular we used the two S-S distances as before, and the two bond angles $\theta_1$ from the C'-S'-S$_1$ atoms and $\theta_2$ from the C$_1$-S$_1$-S$_2$ atoms. Figure 5 shows that, as in the case of the bond distances, also bond angles show a concerted change along the reaction coordinate. Following the computational protocol with such a second set of internal coordinates, the estimation of the reaction rate was $k/c^{\ominus} = 5.8 \cdot 10^3$ s$^{-1}$. There is no qualitative difference between the two estimations of the rate constant with the two choices of the internal coordinates. This result is expected since the analysis of Parrinello is based on employing reasonable descriptors to describe the geometry and fluctuations around the metastable states. In the simple case study reported here, both the choices of distances+dihedral angles, or distances+bond angles are descriptors good enough to locate the relevant collective variable, i.e. the reaction coordinate.

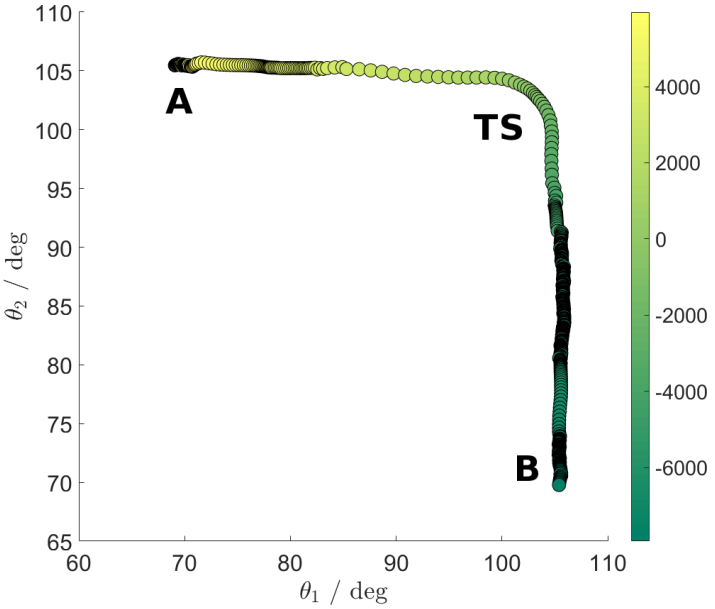


**Fig. 5** Projection of the IRC on the subspace of $\theta_1$ (S'-S$_1$-C$_1$ bond angle) and $\theta_2$ (C$_1$-S$_1$-S$_2$ bond angle). The color map shows the corresponding value of the reaction coordinate $s$, allowing one to visually identify an IRC configuration as being closer to the reactants (A, yellow), products (B, dark green), or transition state (TS, light green). The latter is located around $s = -1025$.

## 4 Conclusions

We reported a multiscale computational protocol to describe the activated dynamics along an IRC defined as a linear combination

of natural Z-matrix internal coordinates within the high friction limit of Kramers theory. The protocol is based on the following tools: *i)* a QM search of the reaction path is performed, producing the corresponding IRC; *ii)* based on the path, the internal coordinates that more likely are important to describe the mechanism of the reaction are considered; *iii)* a reaction coordinate is built following the method proposed by Parrinello,[47] which makes use of very short MD simulations with respect to the length that would be required to describe in a correct statistical way the mean first passage time; *iv)* the generalized friction coefficient along the reaction coordinate is computed with DiTe2;[24] *v)* the parameters computed in the previous points constitute the input to Kramers formula for the reaction rate, which then can be estimated without the need of any fitting procedures to experimental data.

The $S_N2$ substitution of chloride ion to bromomethane has been employed as reference textbook reaction to test the different steps of the computational protocol and the overall predictivity capability. The most important difficulty found is in the ability of estimating the free energy profile along the IRC in presence of important specific solute-solvent interactions. This is a general problem in QM computations and promising multiscale methods are being developed, using free energy perturbation approaches to compute the hydration energy from gas-phase quantum mechanical calculations.[74] In addition, in these reactions, the counterion may also play a role influencing the IRC profile.

The protocol has been then applied to the study of the $S_N2$ reaction of methylthiolate with dimethyl disulfide. The estimated reaction rate is in accord with other similar $S_N2$ reactions of alkylthiolates with dialkyl disulfides in water. Two different choices of the internal degrees of freedom have been tested, both leading to the same (qualitatively) result. This stability of the protocol suggests that once a reasonable set of descriptors of the metastable species are employed, the method introduced by Parrinello to locate the reaction coordinate converges very well.

While the Kramers model has been here employed as a well-established method to treat the activated dynamics, any other comparable approach may be employed. As discussed in the Introduction, most of the methods available are based on a stochastic description of the relevant degrees of freedom. Here we rely, within the hydrodynamic framework, on an accurate evaluation of the friction tensor along generalized coordinates, at least given as linear combinations of the Z-matrix coordinates. This is practically useful for two reasons. On one hand, working in internal coordinates allows one to carry out a more effective complexity reduction if compared with working in Cartesian coordinates. If, for example, one would prefer to use the forward flux sampling method for the reaction here studied, the stochastic trajectories could be computed within the 4-dimensional conformation space of $\mathbf{q}$, instead of considering $3N = 18$ Cartesian coordinates (with methyl groups as single beads). On the other hand, pseudo Z-matrix coordinates are commonly employed by computational chemists to describe molecular geometries. They provide a direct connection between a molecular structure and its changes implied by the reactive event.

Further work is in course along two different routes. On the more practical side, we plan to apply the same computational protocol to a series of analogous reactions involving different chalcogenides as nucleophiles as well as substrates, for which a comprehensive QM study has been carried out by some of the authors. Such an exploration will be used to test whether our protocol is able to reproduce the trend in reactivity that is expected from the nucleophilicity properties of the different chalcogens. After this validation, the same reactions will be studied in the protein environment, where the conditions strongly affect the energetics and might influence the mechanism. For example, in GPx, the regeneration of the initial enzymatic form depends on the selective attack of glutathione to a selenosulfide, which occurs at sulfur, despite the attack at selenium is thermodynamically favored. This peculiarity is controlled by the features of the catalytic pocket, i.e. the arrangement of the conserved residues.

Methodologically, the objective of computing reaction rates and/or understanding reaction mechanisms in complex biomolecular substrates (e.g., proteins), is an important theoretical development. Ideally, it should be based on a multidimensional Fokker-Planck equation in the subspace of relevant coordinates defined from first principles instead of semi phenomenological arguments. Building on some recent work,[48,49] we plan to outline and apply a generalized projection approach to average out non relevant coordinates and recover and numerically solve an effective multidimensional equivalent of Equation 2.

## Appendix. Hydrodynamic decoupling of rotational and internal motions

In formulating Equation 2 we have assumed that the reaction coordinate (that we call the "internal motion") is decoupled from the rotational motion. In the Hamiltonian, such an approximation has been used without any further proof than specifying that internal coordinates should be expressed on a local frame that satisfies the Eckart conditions.[50] In this Appendix, it is shown that it is possible to consider approximately decoupled the two kinds of motion also from the hydrodynamic point of view.

The roto-conformational friction tensors, $\boldsymbol{\xi}$, computed in the reactants (A), products (B), and transition (TS) states, expressed in the frame that diagonalizes the rotational part, read

$$\boldsymbol{\xi}_{\mathrm{A}}/\xi_0 = \begin{bmatrix} 0.570 & 0 & 0 & 4.13 \\ 0 & 1.40 & 0 & 5.02 \\ 0 & 0 & 1.50 & 1.22 \\ 4.13 & 5.02 & 1.22 & 96.7 \end{bmatrix}$$

$$\boldsymbol{\xi}_{\mathrm{B}}/\xi_0 = \begin{bmatrix} 0.579 & 0 & 0 & -4.47 \\ 0 & 1.37 & 0 & 4.39 \\ 0 & 0 & 1.47 & 0.280 \\ -4.47 & 4.39 & 0.280 & 97.2 \end{bmatrix}$$

$$\boldsymbol{\xi}_{\mathrm{TS}}/\xi_0 = \begin{bmatrix} 0.511 & 0 & 0 & 2.88 \\ 0 & 1.13 & 0 & 6.25 \\ 0 & 0 & 1.17 & 0.90 \\ 2.88 & 6.25 & 0.90 & 99.3 \end{bmatrix}$$

where $\xi_0 = 10^{-30}$ kg m$^2$ s$^{-1}$.

Since $\xi_{i,4} << |\xi_{i,i} - \xi_{4,4}|$, with $i = 1, 2, 3$ (the three principal components of the rotational part), it is possible to consider the

hydrodynamic coupling as a perturbation. Equation 2 can be thus seen as a zero-order approximation to the modeling of the relevant dynamics, where the rotational part is projected out since it is not influencing (approximately) the reaction rate. As briefly commented in the main text, further work is planned to obtain rigorously a Fokker-Planck equation for the roto-conformational motion, translating the concept of "reaction coordinate" into that of "relevant internal coordinates" for the description of the reactive event.

## Acknowledgements

Computational work has been carried out on the C3P (Computational Chemistry Community in Padua) HPC facility of the Department of Chemical Sciences of the University of Padua.

## References

1 S. Jun, J. Bechhoefer and B.-Y. Ha, *EPL*, 2003, **64**, 420–426.
2 P. Faccioli and S. a Beccara, *Biophys. Chem.*, 2016, **208**, 62–67.
3 E. E. Santiso and K. E. Gubbins, *Mol. Sim.*, 2004, **30**, 699–748.
4 A. Kovalenko and N. Blinov, *J. Mol. Liq.*, 2011, **164**, 101–112.
5 M. Zerbetto, D. Licari, V. Barone and A. Polimeno, *Mol. Phys.*, 2013, **111**, 2746–2756.
6 M. Zerbetto and A. Polimeno, *Int. J. Quantum Chem.*, 2016, **116**, 1706–1722.
7 H. A. Kramers, *Physica*, 1940, **7**, 284–304.
8 P. Hänggi, P. Talkner and M. Borkovec, *Rev. Mod. Phys.*, 1990, **62**, 251–341.
9 B. J. Berne and M. Borkovec, *J. Chem. Soc., Faraday Trans.*, 1998, **94**, 2717–2723.
10 E. Pollak and P. Talkner, *Chaos*, 2005, **15**, 026116.
11 Y. Xu, K. Song and Q. Shi, *J. Chem. Phys.*, 2018, **148**, 102322.
12 H. Eyring, *J. Chem. Phys.*, 1935, **3**, 107–115.
13 M. V. Basilevsky, G. E. Chudinov and D. V. Napolov, *J. Phys. Chem.*, 1993, **97**, 3270–3277.
14 K. J. Laidler and M. C. King, *J. Phys. Chem.*, 1983, **87**, 2657–2664.
15 R. Hernandez, T. Uzer and T. Bartsch, *Chem. Phys.*, 2010, **370**, 270–276.
16 J. P. Bergsma, B. J. Gertner, K. R. Wilson and J. T. Hynes, *J. Chem. Phys.*, 1987, **86**, 1356–1376.
17 B. J. Gertner, J. P. Bergsma, K. R. Wilson, S. Lee and J. T. Hynes, *J. Chem. Phys.*, 1987, **86**, 1377–1386.
18 B. J. Gertner, K. R. Wilson and J. T. Hynes, *J. Chem. Phys.*, 1989, **90**, 3537–3558.
19 K. Ibuki and M. Ueno, *J. Chem. Phys.*, 1997, **107**, 6594–6602.
20 R. J. Allen, D. Frenkel and P. R. ten Wolde, *J. Chem. Phys.*, 2006, **124**, 024102.
21 H. Mökkönen, T. Ala-Nissila and H. Jónsson, *J. Chem. Phys.*, 2016, **145**, 094901.
22 T. Bartsch, F. Revuelta, R. M. Benito and F. Borondo, *J. Chem. Phys.*, 2012, **136**, 224510.
23 W. E and E. Vanden-Eijnden, *Ann. Rev. Phys. Chem.*, 2010, **61**, 391–420.
24 J. Campeggio, A. Polimeno and M. Zerbetto, *J. Comput. Chem.*, 2019, **40**, 697–705.
25 J. R. Winther and C. Thorpe, *BBA*, 2014, **1840**, 838–846.
26 R. P. Szajewski and G. M. Whitesides, *J. Am. Chem. Soc.*, 1980, **102**, 2011–2026.
27 D. L. Rabenstein and P. L. Yeo, *J. Org. Chem.*, 1994, **59**, 4223–4229.
28 M. M. Gallogly, D. W. Starke and J. J. Mieyal, *Antiox. Redox Sign.*, 2009, **11**, 1059–1081.
29 K. S. Jensen, R. E. Hansen and J. R. Winther, *Antiox. Redox Sign.*, 2009, **11**, 1047–1058.
30 M. Maiorino, V. Bosello-Travain, G. Cozza, G. Miotto, L. Orian, A. Roveri, S. Toppo, M. Zaccarin and F. Ursini, in *Selenium: Its Molecular Biology and Role in Human Health*, ed. D. L. Hatfield, U. Schweizer, P. A. Tsuji and V. N. Gladyshev, Springer International Publishing, Cham, 2016, pp. 223–234.
31 L. Orian, G. Cozza, M. Maiorino, S. Toppo and F. Ursini, in *Glutathione*, ed. L. Flohé, CRC Press, 2018, pp. 53–66.
32 L. Orian and S. Toppo, *Free Radicals Biol. Med.*, 2014, **66**, 65–74.
33 L. P. Wolters and L. Orian, *Curr. Org. Chem.*, 2016, **20**, 189–197.
34 M. Dalla Tiezza, G. Ribaudo and L. Orian, *Curr. Org. Chem.*, 2018, **22**, 1–21.
35 A. P. Bento and F. M. Bickelhaupt, *J. Org. Chem.*, 2008, **73**, 7290–7299.
36 M. A. Van Bochove and F. M. Bickelhaupt, *Eur. J. Org. Chem.*, 2008, **2008**, 649–654.
37 M. A. van Bochove, M. Swart and F. M. Bickelhaupt, *J. Am. Chem. Soc.*, 2006, **128**, 10738–10744.
38 M. Bortoli, L. P. Wolters, L. Orian and F. M. Bickelhaupt, *J. Chem. Theory Comput.*, 2016, **12**, 2752–2761.
39 S. M. Bachrach, D. W. Demoin, M. Luk and J. V. Miller, *J. Phys. Chem. A*, 2004, **108**, 4040–4046.
40 S. M. Bachrach, J. M. Hayes, T. Dao and J. L. Mynar, *Theor. Chem. Acc.*, 2002, **107**, 266–271.
41 C. Santi, B. Battistelli, L. Testaferri and M. Tiecco, *Green Chem.*, 2012, **14**, 1277–1280.
42 L. Sancineto, A. Mariotti, L. Bagnoli, F. Marini, J. Desantis, N. Iraci, C. Santi, C. Pannecouque and O. Tabarrini, *J. Med. Chem.*, 2015, **58**, 9601–9614.
43 C. Santi, L. Capoccia and B. Monti, *Physical Sciences Reviews*, 2018, **3**, year.
44 G. Ribaudo, M. Bellanda, I. Menegazzo, L. P. Wolters, M. Bortoli, G. Ferrer-Sueta, G. Zagotto and L. Orian, *Chem. Eur. J.*, 2017, **23**, 2405–2422.
45 H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, Springer-Verlag.
46 K. Fukui, *Acc. Chem. Res.*, 1981, **14**, 363–368.
47 D. Mendels, G. Piccini and M. Parrinello, *J. Phys. Chem. Lett.*, 2018, **9**, 2776–2781.
48 A. Polimeno, M. Zerbetto and D. Abergel, *J. Chem. Phys.*, 2019, **150**, 184107.

49 A. Polimeno, M. Zerbetto and D. Abergel, *J. Chem. Phys.*, 2019, **150**, 184108.

50 J. Stare, *J. Chem. Inf. Model.*, 2007, **47**, 840–850.

51 E. J. Baerends, D. E. Ellis and P. Ros, *Chem. Phys.*, 1973, **2**, 41–51.

52 G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931–967.

53 C. F. Guerra, J. G. Snijders, G. te Velde and E. J. Baerends, *Theor. Chem. Acc.*, 1998, **99**, 391–403.

54 E. J. Baerends, T. Ziegler, A. J. Atkins, J. Autschbach, D. Bashford, O. Baseggio, A. Bérces, F. M. Bickelhaupt, C. Bo, P. M. Boerritger, L. Cavallo, C. Daul, D. P. Chong, D. V. Chulhai, L. Deng, R. M. Dickson, J. M. Dieterich, D. E. Ellis, M. van Faassen, A. Ghysels, A. Giammona, S. J. A. van Gisbergen, A. Goez, A. W. Götz, S. Gusarov, F. E. Harris, P. van den Hoek, Z. Hu, C. R. Jacob, H. Jacobsen, L. Jensen, L. Joubert, J. W. Kaminski, G. van Kessel, C. König, F. Kootstra, A. Kovalenko, M. Krykunov, E. van Lenthe, D. A. McCormack, A. Michalak, M. Mitoraj, S. M. Morton, J. Neugebauer, V. P. Nicu, L. Noodleman, V. P. Osinga, S. Patchkovskii, M. Pavanello, C. A. Peeples, P. H. T. Philipsen, D. Post, C. C. Pye, H. Ramanantoanina, P. Ramos, W. Ravenek, J. I. Rodríguez, P. Ros, R. Rüger, P. R. T. Schipper, D. Schlüns, H. van Schoot, G. Schreckenbach, J. S. Seldenthuis, M. Seth, J. G. Snijders, M. Solà, S. M., M. Swart, D. Swerhone, G. te Velde, V. Tognetti, P. Vernooijs, L. Versluis, L. Visscher, O. Visser, F. Wang, T. A. Wesolowski, E. M. van Wezenbeek, G. Wiesenekker, S. K. Wolff, T. K. Woo and A. L. Yakovlev, *ADF2017*, `http://www.scm.com`.

55 E. van Lenthe, E. J. Baerends and J. G. Snijders, *J. Chem. Phys.*, 1994, **101**, 9783.

56 N. C. Handy and A. J. Cohen, *Mol. Phys.*, 2001, **99**, 403–412.

57 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

58 B. G. Johnson, P. M. W. Gill and J. A. Pople, *J. Chem. Phys.*, 1993, **98**, 5612–5626.

59 T. V. Russo, R. L. Martin and P. J. Hay, *J. Chem. Phys.*, 1994, **101**, 7729–7737.

60 F. Zaccaria, L. P. Wolters, C. Fonseca Guerra and L. Orian, *J. Comp. Chem.*, 2016, **18**, 1672–1680.

61 M. Bortoli, F. Zaccaria, M. Dalla Tiezza, M. Bruschi, L. P. Wolters, C. Fonseca Guerra, F. M. Bickelhaupt and L. Orian, *Phys. Chem. Chem. Phys.*, 2018, **20**, 20874–20885.

62 A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2*, 1993, 799–805.

63 N. L. Allinger, X. Zhou and J. Bergsma, *J. Mol. Struct. THEOCHEM*, 1994, **312**, 69–83.

64 M. C. R. Melo, R. C. Bernardi, T. Rudack, M. Scheurer, C. Riplinger, J. C. Phillips, J. D. C. Maia, G. B. Rocha, J. a. V. Ribeiro, J. E. Stone, F. Neese, K. Schulten and Z. Luthey-Schulten, *Nature Methods*, 2018, **15**, 351–354.

65 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.

66 B. Brooks, R. Bruccoleri, B. Olafson, D. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 2005, **4**, 187–217.

67 M. Baron, A. Dall'Anese, C. Tubaro, L. Orian, V. Di Marco, S. Bogialli, C. Graiff and M. Basato, *Dalton Trans.*, 2018, **47**, 935–945.

68 Y. Marcus, *J. Chem. Soc., Faraday Trans.*, 1991, **87**, 2995–2999.

69 J. Rotne and S. Prager, *J. Chem. Phys.*, 1969, **50**, 4831–4837.

70 CRC Handbook, *CRC Handbook of Chemistry and Physics, 85th Edition*, CRC Press, 85th edn, 2004.

71 R. H. Bathgate and E. A. Moelwyn-Hughes, *J. Chem. Soc.*, 1959, 2642–2648.

72 S. M. Bachrach and D. C. Mulhearn, *J. Phys. Chem.*, 1996, **100**, 3535–3540.

73 G. M. Whitesides, J. E. Lilburn and R. P. Szajewski, *J. Org. Chem.*, 1977, **42**, 332–338.

74 U. Jug, D. Pregeljc, J. Mavri, R. Vianello and J. Stare, *Comput. Theor. Chem.*, 2017, **1116**, 96–101.