

CHARMMRATE manual, version 2.0, June 26, 2002

CHARMM Element doc/charmmrate.doc

^

File: Polyrate, Node: Top, Up: (doc/charmmrate.doc), Next: Description

```
*****  
*          CHARMM/POLYRATE INTERFACE          *  
*****
```

CHARMMRATE: A Module for Calculating Enzymatic Reaction Rate Constants  
with POLYRATE and CHARMM

CHARMMRATE is an interface of CHARMM and POLYRATE to include quantum mechanical effects in enzyme kinetics. Although CHARMMRATE allows execution of POLYRATE with all existing capabilities, the present implementation is primarily intended for predicting reaction rates in enzyme-catalyzed reactions. CHARMMRATE can be combined with semiempirical combined QM/MM potentials with numerical second derivatives that are computed by the POLYRATE interface programs.

The rate constant for an enzymatic reaction depends on the transition state theory free energy of activation and on an overall transmission coefficient. Quantum effects on the degrees of freedom perpendicular to the reaction coordinate can be incorporated by means of a correction for quantum mechanical vibrational free energy,  $\Delta W_{\text{vib}}$ . As described by M. Garcia-Viloca, C. Alhambra, D. G. Truhlar, and J. Gao, in *J. Chem. Phys.* 114, 9953-9958 (2001), such a correction is calculated by carrying out projected instantaneous normal mode analysis at several configurations along a reaction coordinate as sampled by the umbrella sampling technique (or by any other suitable method) in molecular dynamics simulations with CHARMM. Note that projected instantaneous normal mode analysis involves projecting out the reaction coordinate of the potential of mean force (i.e., the coordinate along which umbrella sampling was carried out); thus it yields different frequencies and modes than would be obtained by ordinary instantaneous normal mode analysis. The correction for quantized vibrational free energy in modes normal to the PMF reaction coordinate is calculated from the average frequencies of the projected instantaneous normal mode analysis and is added to the classical potential of mean force.

The quantum effects on the reaction coordinate are represented by an averaged transmission coefficient obtained by carrying out variational transition state theory (VTST) calculations for individual members (configurations) of the transition state ensemble. These calculations involve a partition of the system into a frozen bath region and a dynamics region that is used in the dynamics calculation. CHARMMRATE has been used to determine the rate constants for the proton transfer reactions catalyzed by enolase and methylamine dehydrogenase and for the hydride transfer reactions catalyzed by alcohol dehydrogenase, xylose isomerase, and dihydrofolate reductase. These studies have demonstrated that inclusion of quantum effects is essential to calculate primary and secondary kinetic isotopic effects (KIEs) for hydrogen transfer reactions. The method used in these studies has evolved to its definitive form that includes free energy simulation to determine the free energy of activation and calculation of the transmission coefficient. Putting all the elements together yields a method that is called ensemble-averaged VTST with

multidimensional tunneling (EA-VTST/MT), the formalism of which is presented in detail in the following recent papers:

C. Alhambra, J. C. Corchado, M. L. Sanchez, M. Garcia-Viloca J. Gao, and D. G. Truhlar J. Phys. Chem. B. 105, 11326-11340 (2001).

D. G. Truhlar, J. Gao, C. Alhambra, M. Garcia-Viloca, J. Corchado, M. L. Sanchez, and Jordi Villa, Acc. Chem. Res. 35, 341-349 (2002).

M. Garcia-Viloca, C. Alhambra, D. G. Truhlar, and J. Gao, J. Comput. Chem. 2002, in press.

This documentation contains a short version and a long, detailed version following the short CHARMM-command description. Users are encouraged to read both parts.

\* Menu:

\* Description:: Description of the POLYRATE driver in CHARMM  
\* Using:: How to run POLYRATE in CHARMM  
\* Installation:: How to install POLYRATE in CHARMM  
\* Status:: Status of the interface code

^  
\_

File: Polyrate, Node: Description, Up: Top, Next: Usage, Previous: Top

```
*****  
* Syntax for the CHARMMRATE Method *  
*****
```

POLYRATE is initiated with the POLYrate command.

[Syntax POLYrate]

```
POLYrate [ atom-selection ] [ RUNIT int ] [ PUNIT int ] [ TSUNIT int ] [ OPUNIT  
int ] [ PMFZpe ] [ ATMA int ] [ ATMB int ] [ ATMC int ]  
[ POLYRATE commands ]  
....  
[*finish]
```

```
atom-spec ::= { residue-number atom-name }  
                  { segid resid atom-name }  
                  { BYNUM atom-number }
```

RUNIT int: Unit specification for input of initial coordinates of the reactant species. The current limitation is that only CHARMM format is allowed for the coordinate file.

PUNIT int: Unit specification for input of initial coordinates of the product species. The current limitation is that only CHARMM format is allowed for the coordinate file.

TSUNIT int: Unit specification for input of initial coordinates of the transition state. The current limitation is that only CHARMM format is allowed for the coordinate file.

OPUNIT int: Unit to write out coordinates of the optimized structures,

of any of the reactant, product, and TS, depending on the species being optimized. The current limitation is that only CHARMM format is used.

PMFZpe ATMA int ATMB int ATMC int: this command switches on the projection operator that is used to project the reaction coordinate out of the Hessian matrix of the system. This is used for projected instantaneous normal mode analysis. The reaction coordinate is defined as the difference in bond distance between the breaking and making bonds.

ATMA int: atom number for the donor atom following the numbering in the general section of POLYRATE commands (see below).

ATMB int: atom number for the transferring atom following the numbering in the general section of POLYRATE commands (see below).

ATMC int: atom number for the acceptor atom following the numbering in the general section of POLYRATE commands (see below).

[POLYRATE commands]

This section contains standard POLYRATE commands. They must follow immediately after the [POLYrate] command in the CHARMM input stream. This section is terminated by the key word [\*finish], lower case with a star in the beginning. For details of the POLYRATE commands, see the POLYRATE documentation.

^\_

File: Polyrate, Node: Usage, Up: Top, Next: Installation,  
Previous: Description

Note: The version number of CHARMMRATE is 2.0/C28b3-P9.0. This means that CHARMMRATE-version 2.0 is based on POLYRATE-version 9.0 and CHARMM-version c28b3. The version number may be abbreviated to 2.0 when no confusion will result.

CHARMMRATE is a module of CHARMM for interfacing it with POLYRATE; the POLYRATE main program becomes a subprogram of CHARMM. POLYRATE can be called to carry out projected instantaneous normal mode analysis and variational transition state theory calculations with semiclassical multidimensional tunneling contributions. When POLYRATE needs the value or gradient of the potential energy surface, it calls a set of interface routines called hooks. The hooks in turn call CHARMM routines for energies and gradients calculated by molecular mechanics or QM/MM methods. The current version has not been parallelized.

Referencing for CHARMMRATE:

"The rate constant (or reaction path or geometry optimization, etc.) calculations were carried out using the CHARMMRATE program[1-3]".

[1] M. Garcia-Viloca, C. Alhambra, J. C. Corchado, M. L. Sanchez, J. Villa, J. Gao, and D. G. Truhlar, CHARMMRATE-version 2.0, University of Minnesota, Minneapolis, 2002, a module of CHARMM (Ref. 2) for interfacing it with POLYRATE (Ref. 3).

- [2] Chemistry at HARvard Macromolecular Mechanics (CHARMM) computer program, as described in B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. 4, 187 (1983).
- [3] J. C. Corchado, Y.-Y. Chuang, P. L. Fast, J. Villa, W.-P. Hu, Y.-P. Liu, G. C. Lynch, K. A. Nguyen, C. F. Jackels, V. S. Melissas, B.J. Lynch, I. Rossi, E. L. Coitino, A. Fernandez-Ramos, J. Pu, and T. V. Albu, R. Steckler, B. C. Garrett, A. D. Isaacson, and D. G. Truhlar, POLYRATE-version 9.0, University of Minnesota, Minneapolis, 2002.

^\_ File: Polyrate, Node: Usage, Up: Top, Next: Contents, Previous: Referencing

```
*****
* Availability of CHARMMRATE                                     *
*****
```

CHARMMRATE-version 2.0/C28b3-P9.0 is a module of CHARMM-version c28b3 for interfacing it with POLYRATE-version 9.0. An earlier version, CHARMMRATE-version 1.0, was distributed as part of the CHARMM program beginning with version 28b1 of CHARMM and was used to interface previous versions of CHARMM and POLYRATE. CHARMMRATE-2.0/C28b3-P9.0 will be distributed beginning with version c28b3 of CHARMM. The user will also require the CRATE utility for modifying POLYRATE to make it compatible with CHARMM. CRATE-version 8.11 corresponds to CHARMMRATE-1.0, and CRATE-version 9.0 corresponds to CHARMMRATE-2.0. CRATE-version 9.0 corresponds to interfacing POLYRATE-version 9.0. The prospective user of CHARMMRATE should obtain a valid license for CHARMM from an authorized CHARMM licensor and valid licenses for POLYRATE and CRATE from the University of Minnesota (<http://comp.chem.umn.edu>).

## 1. INTRODUCTION

CHARMMRATE is an interface of CHARMM and POLYRATE to include quantum mechanical effects in enzyme kinetics. Although CHARMMRATE allows execution of POLYRATE with all existing capabilities for reactions with only one reactant and only one product, the present implementation is primarily intended for prediction of the reaction rates of enzyme-catalyzed reactions. Any CHARMMRATE calculation involves the partition of the system into a primary subsystem (or primary-zone atoms), which contains the subset of atoms involved in the reaction, and the rest of the system (secondary-zone atoms). Only the coordinates of the primary-zone atoms are passed from CHARMM to POLYRATE for both projected instantaneous normal mode analysis and dynamics calculations. Consequently, the quantum mechanical vibrational correction and the dynamics effects are calculated for the primary subsystem in the field of the secondary subsystem.

### 1.A. Capabilities added to CHARMM by CHARMMRATE and references for methods

POLYRATE includes a very large number of options and has multiple capabilities. The user of CHARMMRATE is encouraged to read the POLYRATE manual to learn more about these capabilities. The present section summarizes a few of the capabilities that are liable to be of most interest to CHARMMRATE users.

### 1.A.1. Transition state optimization

Saddle point geometry optimizations for the primary (dynamic) zone in the frozen protein-plus-solvent bath may be performed in various ways; the default option is the Newton-Raphson method with Brent line minimization as described in W. H. Press, S. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1986), p.254. The default option for optimization of the stationary points for reactants and products is to use the BFGS method that has been implemented in POLYRATE. See the POLYRATE manual for further information about the optimization methods available in POLYRATE.

### 1.A.2. Reaction path

In general, reaction paths (RPs) may be defined in various ways. The simplest general method that is reasonably sure to give physically meaningful vibrational frequencies for motions transverse to the reaction path (and hence also physically meaningful free energy of activation profiles) is the steepest descents path in isoinertial coordinates. (An isoinertial coordinates system is one in which the kinetic energy is a sum of square terms and the coordinates are scaled or weighted so that each kinetic energy term has the same reduced mass. All isoinertial coordinate systems are related to each other by orthogonal transformations, and steepest descents paths are invariant under orthogonal transformations.) A steepest descents path is also called a minimum energy path (MEP). The signed distance from the saddle point along the reaction path is called the reaction coordinate, usually denoted  $s$ . (This reaction coordinate,  $s$ , should not be confused with the reaction coordinate used for umbrella sampling, which is called  $z$ .) The isoinertial MEP is sometimes just called the MEP, or it may just be called the RP; other workers prefer to append the word intrinsic, e.g., intrinsic MEP, intrinsic reaction path, intrinsic reaction coordinate, etc.

In CHARMMRATE, the reaction path refers to a multidimensional path for the primary-zone (dynamic) atoms in the presence of the secondary-zone (frozen) atoms.

CHARMMRATE may be used to calculate the isoinertial minimum energy path (MEP) as described in B. C. Garrett, M. J. Redmon, R. Steckler, D. G. Truhlar, K. K. Baldridge, D. Bartol, M. W. Schmidt, M. S. Gordon, *J. Phys. Chem.* 92, 1476-1488 (1988).

### 1.A.3. Free energy of activation profile and variational transition state theory

Vibrational partition functions and generalized free energies of activation (which are free energies of activation for tentative transition states that are not necessarily associated with either a saddle point or with the final variational transition state) are computed along the reaction path by using the quantum mechanical harmonic oscillator approximation in  $3N_1 - 1$  degrees of freedom, where  $N_1$  is the number of atoms in the primary zone, and the reaction coordinate is projected out. This kind of calculation is described in S. E. Wonchoba, and D. G. Truhlar, *J. Chem. Phys.* 99, 9637- 9651 (1993). The generalized free energy of activation as a function of the reaction coordinate (which is the signed distance along the MEP) is called the free energy of activation

profile, and it may be used to calculate reaction rate constants by variational transition state theory (VTST) as described in D. G. Truhlar and B. C. Garrett, *Acc. Chem. Res.* 13, 440-448 (1980). A procedure like this was used in C. Alhambra, J. Gao, J. C. Corchado, J. Villa, and D. G. Truhlar, *J. Am. Chem. Soc.* 121, 2253-2258 (1999), but it is now recommended to use the more complete EA-VTST/MT method, in which this quantity is used to compute a transmission coefficient rather than a rate constant. VTST for a canonical ensemble (i.e., a system at a fixed temperature) is also called canonical variation theory (CVT). In the EA-VTST/MT method (described in Section 2), this step is carried out for several members of the transition state ensemble, and it is used for the quasiclassical part of the ensemble-averaged transmission coefficient.

#### 1.A.4. Transmission coefficient

In CHARMMRATE the EA-VTST/MT transmission coefficient has two parts: a quasiclassical dynamical recrossing part (Section 1.A.3) and a part that accounts for tunneling (transmission through the barrier at energies below the top) and non-classical reflection (reflection caused by diffraction from the barrier top even when the energy is above the barrier); often we just refer to the combination of tunneling and non-classical reflection effects as tunneling (the tunneling is more important than the non-classical reflection because the energies where tunneling occurs have larger Boltzmann factors than the energies where non-classical reflection occurs).

CHARMMRATE can calculate the tunneling part of the transmission coefficient in various ways. The most complete method is the microcanonical optimized multidimensional tunneling (muOMT) approximation as described in Y.-P. Liu, D.-h. Lu, A. Gonzalez-Lafont, D. G. Truhlar, and B. C. Garrett, *J. Am. Chem. Soc.* 115, 7806-7817 (1993). In this calculation, tunneling and non-classical reflection along the reaction path are included by calculating both the large-curvature tunneling (LCT) approximation and the small-curvature tunneling (SCT) approximation and, at each tunneling energy, accepting whichever tunneling approximation yields the larger tunneling probability. This is a poor man's version of a more complete search for the semiclassical tunneling paths that minimize the imaginary action integrals, and it has been extensively validated as summarized by T. C. Allison and D. G. Truhlar, in *Modern Methods for Multidimensional Dynamics Computations in Chemistry*, edited by D. L. Thompson (World Scientific, Singapore, 1998), pp. 618-712.

One may also limit the calculation to just the LCT or SCT approximation or to the zero-curvature tunneling approximation (ZCT) or even the Wigner approximation. The muOMT, LCT, SCT, and ZCT approximations are multidimensional, whereas the Wigner approximation is one-dimensional. The ZCT approximation calculates tunneling along the isoinertial MEP, whereas the muOMT, LCT, and SCT approximations include various amounts of corner cutting, i.e., tunneling on the concave side of the isoinertial MEP, with the amount and nature of the corner cutting depending on the curvature of the reaction path. The computational cost decreases in the following order: muOMT, LCT, SCT, ZCT, Wigner. When tunneling is included, the EA-VTST/MT rate constant is written as

$$k(T) = \gamma(T) k_{\text{TST}}(T)$$

where  $k_{\text{TST}}(T)$  is the TST rate constant that is determined by the free

energy simulation of of stage 1 (including the quantum mechanical correction of step 2 of stage 1), and  $\gamma(T)$  is the transmission coefficient that accounts for classical recrossing (the quasiclassical part of section 1.A.3) and for tunneling and non-classical reflection.

Background for the calculation of KIEs by VTST with multidimensional tunneling approximations is given in D.G. Truhlar, D.-h. Lu, S.C. Tucker, X.G. Zhao, A. Gonzalez-Lafont, T.N. Truong, D. Maurice, Y.-P. Liu, and G.C. Lynch, in *Isotope Effects in Chemical Reactions and Photodissociation Processes*, edited by J. A. Kaye (American Chemical Society Symposium Series 502, Washington, DC, 1992), pp. 16-36.

#### 1.B. CHARMMRATE capabilities that are not included either in POLYRATE or in prior versions of CHARMM: Projected instantaneous normal mode analysis

Two source files of POLYRATE (see the CRATE manual) are modified by the CRATE utility version-9.0 to carry out projected instantaneous normal mode analysis. With these routines quantum mechanical harmonic frequencies of the vibrational modes of the primary subsystem orthogonal to the reaction coordinate may be calculated for a given configuration of the system. This calculation is used in the second step of the first stage of the EA-VTST/MT method (described in Section 2) to include quantum effects on the 3N-7 highest-frequency vibrational modes of the primary zone in a hypersurface orthogonal to the reaction coordinate that is used for umbrella sampling in the first step of stage 1. The constraint that the modes obtained are orthogonal to the reaction coordinate is achieved by a projection operator described in C. Alhambra, J. C. Corchado, M. L. Sanchez, M. Garcia-Viloca, J. Gao, and D. G. Truhlar *J. Phys. Chem. B* 105, 11326-11340 (2001).

#### 1.C. CHARMM options that are of particular interest for use with CHARMMRATE.

CHARMMRATE is of particular interest for calculations of rate constants for enzymatic reactions. Although the program would allow the use of pure molecular mechanics (the CHARMM22 force field) for such calculations, combined quantum mechanical and molecular mechanical (QM/MM) potentials are much more realistic than pure molecular mechanics for chemical reactions. Using CHARMM, QM/MM calculations can now be performed at the ab initio level using GAMESS (B. Brooks and M. Hodoscek, unpublished results), at the density functional level using CADPAC (P. D. Lyne, M. Hodoscek, and M. Karplus, *J. Phys. Chem. A* 103, 3462-3471 (1999)), and at semiempirical molecular orbital levels (AM1 and PM3, with general parameters or with specific reaction parameters) with MOPAC (M. J. Field, P. A. Bash, and M. Karplus, *J. Comput. Chem.* 11 700-733 (1990)). There is more than one choice for joining the QM subsystem to the MM one. The first choice is to use "link atoms" to saturate the valence of the fragment; this requires that certain atomic charges in the MM fragment that are close to the QM region be deleted to avoid artificial polarization of the quantum subsystem. One possibility to avoid these problems is to use the generalized hybrid orbital (GHO) method described in J. Gao, P. Amara, C. Alhambra, and M. Field, *J. Phys. Chem. A* 102, 4714-4721 (1998). The GHO method is currently available for semiempirical calculations with the AM1 and PM3 methods, and it is being extended (work in progress) to ab initio and DFT methods. Another way to correct the link atom artifacts in the original formulation is proposed in C. Alhambra, L.

Wu, Z.-Y. Zhang, and J. Gao, *J. Am. Chem. Soc.* 120, 3858-3866 (1998).

Molecular dynamics simulations of an enzyme-solvent system can be carried out on a QM/MM potential energy surface either using periodic boundary conditions or using stochastic boundary conditions; for the periodic boundary conditions see M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, (Oxford University Press, New York, 1987), Ch. 1, and for the stochastic boundary conditions see C. L. Brooks, A. Brunger, and M. Karplus, *Biopolymers* 24, 843-865 (1985). Free energy perturbation and umbrella sampling techniques can be used to determine the potential of mean force or classical free energy profile for the enzymatic reaction (see C. Alhambra, L. Wu, Z.-Y. Zhang, J. Gao, *J. Am. Chem. Soc.* 120, 3858-3866 (1998)).

## 2. THEORETICAL BACKGROUND: Ensemble-averaged variational transition state theory with multidimensional tunneling (EA-VTST/MT).

This section of the manual summarizes the theoretical framework and the practical procedure for the EA-VTST/MT method developed in C. Alhambra, J. C. Corchado, M. L. Sanchez, M. Garcia-Viloca, J. Gao, and D. G. Truhlar *J. Phys. Chem.* 105, 11326-11340 (2001).

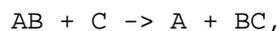
The capabilities added to CHARMM by CHARMMRATE allow the user to calculate the rate constant for an enzymatic reaction with the EA-VTST/MT procedure.

The rate constant for an enzymatic reaction, which is a unimolecular process, is obtained by combining free energy simulations and variational transition state theory (VTST) with microcanonical optimized multidimensional tunneling contributions ( $\mu$ OMT), both in the presence of the protein environment. The potential energy surface (PES) is modeled by a QM/MM method, for example by a semiempirical MO method combined with the CHARMM force field and with the generalized hybrid orbital method (GHO) to treat the boundary between the QM and the MM parts of the system. In addition, a semiempirical term (see quantum.doc LEPS command) or specific reaction parameters (SRP) may be used to improve the accuracy of the PES.

The rate constant is expressed as a function of the free energy of activation,  $\Delta G_{CV}^{\ddagger}$ , calculated by variational transition state theory free energy molecular dynamics simulations, and the transmission coefficient,  $\gamma$ . There are two versions of the method that differ in the approximation used to evaluate  $\gamma$ , in particular a 2-stage version and a three-stage version. The procedure for the former approximation, which has been applied in the study of five hydrogen transfer reactions, has the following two stages:

1) In stage 1 of the calculation, the free energy of activation, including the quantum mechanical vibrational free energy, is computed. This involves the following calculations:

Step 1 - Calculation of the classical mechanical (CM) or transition state theory free energy of activation by computing the potential of mean force (PMF) along a distinguished reaction coordinate. For a reaction



the reaction coordinate,  $z$ , may be defined as:

$$z = r_{AB} - r_{BC}$$

The CM PMF can be evaluated by carrying out classical molecular dynamics on a QM/MM potential energy surface with the umbrella sampling technique implemented in CHARMM (see umbrel.doc) or free energy perturbation theory. Note, for the discussion below that umbrella sampling involves a sequence of overlapping windows (whose centers are separated by about 0.1-0.2 angstroms), each of which is later divided into 50-100 bins. The bins are typically 0.01 angstroms wide. (These specific numerical values are just given as examples; none of these quantities is restricted to lie within those limits.)

Step 2 - Calculation of the quantum mechanical vibrational free energy correction,  $\Delta W_{\text{vib}}$ , which is the difference between the quantal vibrational free energy and the classical vibrational free energy. The addition of  $\Delta W_{\text{vib}}$  to the CM PMF gives the quasi-classical (QC) PMF.  $\Delta W_{\text{vib}}$  may be evaluated by carrying out projected instantaneous normal mode analysis for the primary-zone atoms for many configurations (100-400 per window) obtained in the umbrella sampling step (see Section 1.B). The projected instantaneous normal mode frequencies obtained for the different configurations in a given bin may be averaged and the averaged value used to determine  $\Delta W_{\text{vib}}$ . Strictly speaking, one might argue that one should average the squared frequencies or some Boltzmann factors, but in initial applications it has been found sufficient to average the frequencies themselves.

After these two steps, the value of  $z$  with highest QC free energy is called  $z^*$ , and the bin containing  $z^*$  (or a small set of bins centered on this bin) defines the ensemble of configurations that are representative of the transition state of the enzymatic reaction.

2) Stage 2 has the objective of computing the transmission coefficient,  $\gamma(T)$ , that is the average over transition state configurations  $i$  of the product of two factors,  $\Gamma_i$  and  $\kappa_i$ . The first factor,  $\Gamma_i$ , which is the the quasiclassical transmission factor, corrects the rate constant for classical mechanical dynamical recrossing. The second factor,  $\kappa_i$ , is the semiclassical transmission coefficient that accounts mainly for tunneling, that is, the quantum mechanical effect on the reaction coordinate, which is missing in the calculation of the QC rate constant. The averaged transmission coefficient is:

$$\gamma = \langle \Gamma_i \kappa_i \rangle,$$

where the brackets indicate average over configurations  $i$ . That is, a number of configurations (5-20 for the enzymatic reactions studied so far) within the range  $z = z^* \pm \Delta z$  ( $\Delta z = 0.05-0.01$  angstroms), are chosen as representative of the transition state ensemble. For each of them a CHARMMRATE dynamics calculation is carried out to calculate  $\Gamma_i$  and  $\kappa_i$ . As mentioned above, in such stage-2 calculation the system is divided into a set of primary-zone atoms, which are allowed to move, and the rest of the system, which is fixed at the transition state configurations. For each configuration, the saddle point and the reactant and product structures are optimized. We optimize the saddle point and we calculate an iso-inertial MEP in both directions, i.e., toward the reactant and toward the product. The reactant and the product calculations are used

only to determine the minimum energy at which tunneling is allowed. The effective potential used to calculate the tunneling part,  $\kappa_i$ , of the transmission coefficient (see Section 1.A) involves the zero-point-inclusive energy along this MEP. The user should see test run 2 for an example of how to use a typical solvent configuration.

The approximation described here to obtain the transmission coefficients is called static secondary-zone approximation (SSZ). The product of the averaged transmission coefficient obtained in this way and the quasiclassical rate constant of stage 1 results in the SSZ version of the EA-VTST/MT rate constant. The results obtained in five studies of enzymatic hydrogen transfer reactions demonstrate that the SSZ rate constant is accurate enough to reproduce experimental KIEs.

The SSZ result may be improved by carrying out a further step that has been called stage 3. In stage 3, the free energy of the secondary zone is calculated by free energy perturbation theory along the minimum-energy paths of stage 2. This allows us to include the secondary-zone free energy in the transmission coefficient. This is called the equilibrium-secondary-zone (ESZ) approximation.

The EA-VTST/MT method is described in: C. Alhambra, J. C. Corchado, M. L. Sanchez, M. Garcia-Viloca, J. Gao, and D. G. Truhlar *J. Phys. Chem. B* 105, 11326-11340 (2001), and in D. G. Truhlar, J. Gao, C. Alhambra, M. Garcia-Viloca, J. Corchado, M. L. Sanchez, and J. Villa, *Acc. Chem. Res.* 35, 341-349 (2002). A complete description of an application study is provided in M. Garcia-Viloca, C. Alhambra, D. G. Truhlar, and J. Gao, *J. Comput. Chem.* 2002, in press.

### 3. PROGRAM STRUCTURE

#### 3.A. Overall design

The CHARMMRATE interface for CHARMM and POLYRATE takes advantage of the modular nature of both programs, and, consequently, minimal modifications of CHARMM and POLYRATE were required. The CHARMM program is the main driver of the integrated program, which makes a FORTRAN call to the interface subprogram, CHARMMRATE, to initiate calculations by POLYRATE. The energy and energy gradients for the primary-zone atoms required by POLYRATE are determined by CHARMM through the interface subprogram and are supplied to POLYRATE through a set of subroutines called the POLYRATE hooks.

#### 3.B. Modifications and additions to CHARMM

Only two modifications have been made in the CHARMM program: (1) addition of a one-line keyword processing command in the charmm\_main.src module to initiate the subroutine call to CHARMMRATE; (2) addition of the CHARMMRATE module.

#### 3.C. Modifications and additions to POLYRATE

Specific modifications of the original POLYRATE program have been made primarily for efficient transfer of information between CHARMM and POLYRATE and to eliminate conflicts and other problems during compilation. These modifications are described in the CRATE manual, available at <http://comp.chem.umn.edu/crate>.

## 4. INSTALLATION OF charmmrate AND ITS USE

### 4.A. Program distribution

CHARMMRATE-version 2.0/C28b3-P9.0 is distributed as a module in CHARMM. CHARMM is a copyrighted program distributed by Professor Martin Karplus's research group at Harvard University and by Accelrys, Inc. In addition to CHARMM, which includes the CHARMMRATE module, users also need to obtain the POLYRATE program, which is a copyrighted program distributed by the University of Minnesota (<http://comp.chem.umn.edu>) and the CRATE utility, also available from Minnesota. The CRATE utility will make the changes to the source code of POLYRATE to allow the interface between the two programs. When the CHARMM program (which, beginning with version 28, automatically includes the CHARMMRATE module), the POLYRATE program and the CRATE utility have been obtained, integration of the codes into a single executable file is straightforward as described below.

### 4.B. Installation

The user should carry out the following steps:

1. Install CHARMM.
2. Store the tar file polyrate9.0.tar.Z (obtained from the University of Minnesota, <http://comp.chem.umn.edu>) in the directory chmroot/source/prate and untar it with this command:

```
tar xvf polyrate9.0.tar
```

3. Set an environmental variable, called pr, to the absolute path name of the directory where the polyrate program is stored. Example:

```
C shell
% setenv pr /home/chmroot/source/prate/polyrate9.0
Bourne shell
$ pr = /home/chmroot/source/prate/polyrate9.0
$ export pr
```

4. Store the tar file crate9.0.tar.Z (obtained from the University of Minnesota, <http://comp.chem.umn.edu>) in the directory chmroot/source/prate and untar it with this command:

```
tar xvf crate9.0.tar
```

The directory crate9.0 will then contain the files required to prepare POLYRATE for use with the CHARMMRATE module of CHARMM. These files are described in the CRATE manual. Change the dimensions specified in the param.inc file located in the newly created directory, crate9.0, in order to make them large enough for the system(s) to be studied, but small enough to run in the memory available on the computer chosen to carry out the work. Or use the param.inc file distributed as part of CRATE. See the POLYRATE manual for further discussion of the dimensions in in POLYRATE.

5. Set an environmental variable, `crate`, to the absolute path name of the directory where the CRATE package is stored.

Example:

```
C shell
% setenv crate /home/chmroot/source/prate/crate9.0
 Bourne shell
$ crate = /home/chmroot/source/prate/crate9.0
$ export crate
```

6. Go to the `/build` directory of CHARMM, i.e.  
`cd /home/chmroot/build/'chm_host'`,  
and edit the file `pref.dat`. Add the CHARMMRATE module to the list.
7. If CHARMM has been compiled previously without POLYRATE, remove all the object files in `home/chmroot/lib/'chm_host'`
8. Go to the CHARMM directory `chmroot` and type the command:  
`install.com 'chm_host' (small/medium/large) POLYR > & log &`  
(see `install.doc` in the CHARMM documentation directory). This step will execute the script `install_cr.com`, which will put the CHARMMRATE source code in the directory `/home/chmroot/source/prate`.  
Any modifications desired should be done here.
9. You do not need to run the script `install_cr.com` (described in the CRATE manual) in any further compilations. Therefore it is recommended to comment the following line in the file `/home/chmroot/install.com`:  
`$crate/install_cr.com`
10. After compilation, you will have a new executable in:  
`/home/chmroot/exec/'chm_host'`

NOTE: In order to run projected instantaneous normal mode analysis with CHARMMRATE-version 2.0 it is necessary to do small changes in two of the files located in the directory `/home/chmroot/source/prate` after the compilation process described above. Instructions for these changes are provided in the README file contained in the `crate9.0` directory of the CRATE package.

## 5. DESCRIPTION OF INPUT

CHARMMRATE is run from the CHARMM main input stream. The syntax to execute polyrate from charmm's input stream for a reaction with only one reactant (e.g., an enzyme-substrate precursor complex) and only one product (e.g., an enzyme-substrate successor complex) is:

```
POLYrate SELEction { atom-spec } end [RUNIt int] [PUNIt int]
                [TSUNit int] [OPUNit int]
                [ PMFZpe ] [ATMA int ] [ATMB int] [ATMC int]
_polyrate_input_
*finish
```

We note the use of the CHARMM convention by which one needs to enter

only the first four letters of POLYrate and other words with the first four letters capitalized. Furthermore the parts in brackets are optional. The meanings of the various keywords are:

SELEction { atom-spec } specifies the primary-zone atoms in POLYRATE:

```
atom-spec = { residue-number atom-name }
             { segid resid atom-name }
             { BYNUM atom-number }
```

RUNit int: Unit specification for input of initial coordinates of the reactant species. The current limitation is that only CHARMM format is allowed for the coordinate file.

PUNit int: Unit specification for input of initial coordinates of the product species. The current limitation is that only CHARMM format is allowed for the coordinate file.

TSUNit int: Unit specification for input of initial coordinates of the transition state. The current limitation is that only CHARMM format is allowed for the coordinate file.

OPUNit int: Unit to write out coordinates of the optimized structures of the reactant, product, or TS, depending upon which of these is requested (elsewhere) to be written. The current limitation is that only CHARMM format is used. The coordinate files assigned to these units must be in the CARD format (see CHARMM documentation for details).

PMFZpe ATMA int ATMB int ATMC int: this command switch on the projection of the reaction coordinate out of the Hessian matrix of the system. It is used for projected instantaneous normal mode analysis. The reaction coordinate is defined as the difference in bond distance between the breaking and making bonds.

ATMA int: atom number for the donor atom following the numbering in the general section of POLYRATE commands (see below).

ATMB int: atom number for the transferring atom following the numbering in the general section of POLYRATE commands (see below).

ATMC int: atom number for the acceptor atom following the numbering in the general section of POLYRATE commands (see below).

\_polyrate\_input\_: This section contains a standard POLYRATE fu5 input file. It must follow immediately after the POLYrate command in the CHARMM input stream. For details of POLYRATE input, see the POLYRATE documentation. The initial coordinates have already been setup through the POLYrate command; therefore the GEOM record in the POLYRATE fu5 input file may be omitted. If, however, the GEOM record is present, the Cartesian coordinates given in this record will replace the data set up through the POLYrate command. This is not recommended.

\*finish The last record to be read by POLYRATE from the CHARMM main input stream. This will terminate I/O operations from unit 5

by POLYRATE, and POLYRATE calculations will proceed.

## 6. TEST RUNS

This section describes two test runs. Each test job includes a full input file, initial coordinates and parameter files. They are located in:

/home/chmroot/test/cquantumtest.

### 6.1. Test Job 1 - Direct dynamics of chorismate to prephenate in the gas phase

This test job reads in three initial guess coordinates for the reactant state, product state, and transition state, optimizes their geometries, and performs a CVT calculation to yield the predicted rate constants at various temperatures. This test job takes roughly 3 hours on an SGI Octane 2 computer running under the Irix 6.5 operating system. (Test Job 2 is a shorter test run.)

#### 6.1.A. Input files

The cr01.inp file contains the CHARMM input stream for a direct dynamics calculation of the chorismate to prephenate rearrangement reaction. Similar calculations can be carried out for the substrate in the enzyme active site, provided that appropriate boundary conditions are set up in the CHARMM input.

The charmm22.top and charmm22.par files are the CHARMM topology and parameter files. They are required for all CHARMM calculations.

Three coordinate files are provided for this test job, corresponding to the initial guess coordinates for the reactant (gs.crd), product (prod.crd), and transition state (ts.crd) for the dynamics calculation with POLYRATE.

#### 6.1.B. Description of the CHARMM input stream

The majority of the CHARMM commands are straightforward. The three initial guess coordinate files must be opened as formatted files in the CHARMM input stream before the POLYrate command is initiated. Certain FORTRAN unit numbers are default file choices in POLYRATE. Therefore, these numbers should not be used in the CHARMM input file unless the POLYRATE defaults are changed. Please consult the POLYRATE documentation for a full list and description of these files.

Five (5) FORTRAN files will be used by POLYRATE to write out the computational results. They are files with unit numbers 14, 25, 26, 27, and 61, which should be opened in the CHARMM input stream before CHARMMRATE calculations.

Section 6.1.C summarizes the contents of these files.

All input instructions immediately following the POLYrate command are those of POLYRATE. A full description of these commands can be found in the POLYRATE documentation.

#### 6.1.C. Description of CHARMMRATE output

cr0114.out - computed reaction rates at various temperatures using the TST and CVT methods.

cr0125.out - potential energy along the reaction coordinate  $s$ , which measures distance along the minimum energy path (MEP), and computed transmission coefficient, if requested. Since the test run is for a CVT calculation, no multidimensional tunneling is included in the test run.

cr0126.out - computed vibrational frequencies that have been requested for printing out along  $s$ .

cr0127.out - coordinates along  $s$ .

cr0161.out - optimized geometries, energies, vibrational frequencies and the Hessian for the reactant, product, and transition state.

## 6.2. Test Job 2 - Geometry optimization of chorismate in a water bath

This test job performs geometry optimization of chorismate in the presence of a frozen water bath, arbitrarily taken from the trajectory of a molecular dynamics simulation of chorismate in water.

The cr02.inp file is the CHARMM input stream command file. In addition to the CHARMM topology and parameter files, the cr02.crd file is required; it contains the instantaneous (initial) coordinates of chorismate in water from a molecular dynamics simulation.

The file optcr02.crd contains the optimized coordinates of chorismate in water.