

## Supporting Information

### Direct Dynamics Study of Deprotonated Guanine-Cytosine Base Pair: Intra-Base Pair Proton Transfer, Thermal Dissociation vs. Collision-Induced Dissociation, and Comparison with Experiment

Jianbo Liu\*

Department of Chemistry and Biochemistry, Queens College and the Graduate Center of the City University of New York, 65-30 Kissena Blvd., Queens, NY 11367, USA

**Computational details** Trajectories were calculated using the chemical dynamics program VENUS 99 of Hase *et al.*<sup>1,2</sup> to set up initial conditions, and the Hessian-based predictor-corrector algorithm<sup>3</sup> implemented in Gaussian 09<sup>4</sup> to integrate classical equations of motions, with the Hessian matrix recalculated every 5 steps. Trajectories were propagated with a step size of 0.25 amu<sup>1/2</sup> Bohr (~ 0.5 fsec), which was small enough for SCF convergence as well as to keep the total energy constant. Initial guess for each quantum chemistry calculation was obtained from the previous trajectory step. Energy was checked during the simulation to ensure the total energy was conserved to better than 10<sup>-4</sup> Hartree. A quadratically convergent SCF procedure<sup>5</sup> was used (*i.e.* SCF = XQC) in case the usual, but much faster, first-order SCF method failed to converge within the allotted number of cycles. Because millions of gradients and Hessian evaluations were required, the level of theory used for simulations had to be modest. B3LYP/6-31G\* was chosen as it represented a good compromise between computational demand and chemistry accuracy.

Batches of trajectories (100 trajectories for each batch) were calculated for the thermal excitation of G·[C – H]<sup>-</sup> and G·[C – H]<sup>-</sup>\_PT. Base pairs had zero-point energies (ZPEs) in all vibrational modes (with a total of 5.55 eV), and were excited to a vibrational temperature of 960 K which equals 3.0 eV of excitation energy in average. The excitation energy was randomly distributed among vibrational modes using the quantum Boltzmann sampling scheme,<sup>6</sup>

$$P(n_i) = \exp\left(-\frac{n_i h\nu_i}{k_B T_{vib}}\right) \left[1 - \exp\left(-\frac{h\nu_i}{k_B T_{vib}}\right)\right] \quad (1)$$

---

\* E-mail: jianbo.liu@qc.cuny.edu. Tel: 1-718-997-3271.

where  $\nu_i$  and  $n_i$  are the vibrational frequency and quantum number of the  $i^{\text{th}}$  mode, respectively, and  $T_{vib}$  is the vibrational temperature. Quasi-classical initial vibrational states were prepared by giving individual reactant atoms displacements from equilibrium and momenta that are appropriate to the initial rovibrational state, with random phases for different modes. Rotational energy was added by sampling classical Boltzmann distribution at 300 K.

Trajectories for barrier excitation were initiated at the TS and directed randomly towards  $G\cdot[C-H]^-$  and  $G\cdot[C-H]^-_{PT}$ . Reaction coordinate momentum was sampled using the cumulative distribution function with the average reaction coordinate translational energy of  $k_B T_{rc}$ ,<sup>6</sup>

$$P(s) = \pm[-2k_B T_{rc} \ln(1 - R)]^{1/2} \quad (2)$$

where  $R$  is a random number, and  $T_{rc}$  is the temperature chosen for the reaction coordinate. In this work  $T_{rc}$  was set equal to  $T_{vib}$ . Two activation temperatures, 960 and 1330 K, were used for sampling Boltzmann distributions of TS's vibrational levels and reaction coordinate energy, which correspond to total excitation energies of 3.0 and 5.0 eV, respectively. Initial rotational temperature for the TS was set at 300 K. A batch of 100 trajectories was calculated at each activation temperature.

Finally, bimolecular collisions of  $G\cdot[C-H]^-$  and  $G\cdot[C-H]^-_{PT}$  with Ar were simulated at  $E_{col}$  of 3.0 and 5.0 eV. Because in the experiment deprotonated G·C was generated by electrospray ionization and thermalized to ambient temperature within a guided-ion-beam apparatus, a 300 K Boltzmann distribution was sampled to generate base pair initial vibrational (ZPE included) and rotational energies.

Trajectories were started with center-of-mass separation of 7.0 Å between the randomly-oriented base pair and Ar, at which point the interaction between reactants is relatively small. Relative velocities were added to the base pair and Ar corresponding to the desired center-of-mass  $E_{col}$ . Approximately 100 trajectories were calculated for each combination of base pair starting conformation and  $E_{col}$ . The purpose of scattering dynamics study was to probe gross features of collisional activation and intra-base pair proton transfer, thus all trajectories were calculated at an impact parameter  $b = 0$  Å rather than randomly sampling a range of reactive  $b$ .

The nature of the base pair reactions and the energy range of interest are ideal for QCT calculations. The electronic excited states of these systems are high enough to rule out the likelihood of non-adiabatic behavior, allowing use of a single reference quantum chemistry method. Thermal excitation and collision energies are sufficiently large that tunneling should be negligible. One issue with using the QCT method is that vibrational energy ( $E_{vib}$ ) is not quantized in molecules. Lack of quantization presumably has an effect on how energy is distributed among vibrational modes.<sup>7,8</sup> It is possible to have dissociation trajectories where the product  $E_{vib}$  is below ZPE. But such unphysical collisions turn out to be rare in the present study presumably because trajectories were calculated at high excitation energies, reducing the errors associated with treating vibrational motion classically.

The random number generator seed used in setting up trajectory initial conditions was identical. Each batch of thermal excitation and collision trajectories, therefore, used the same pseudo-random sequence to sample reactant parameters (orientations, rotational and vibrational phases, etc.). As a result, it is easy to compare trajectories of different starting structures, temperatures or  $E_{col}$ , because corresponding trajectories from different batches have identical initial conditions, apart from the parameters being varied. The pseudo-random sampling procedure has minimized a potentially serious problem that may arise when reactivity is sharply dependent on one or more reactant parameters. By using the same pseudo-random sequence for each batch, the error from inadequate sampling of reactant parameter space is the same for all batches, and tends to cancel when comparing different batches.

Trajectories were terminated when the center-of-mass separation between guanine and cystine exceeded 8.0 Å, or a maximum integration time (4 ps) was reached. A total of 780 trajectories were calculated in the work, each taking at least 50 days of CPU time on a 64-bit Linux computation cluster consisting of six Intel dual Xeon (16-core, 2.30 GHz) nodes. gOpenMol<sup>9</sup> was used for trajectory visualization. Analysis of individual trajectories and ensemble averages were done with programs written for these purposes. Kinetic energies of relative motion (*i.e.*, translation energy  $E_{trans}$ , reactant collision energy  $E_{col}$ , and product recoil energy  $E_{recoil}$ ) were calculated straightforwardly from the center-of-mass velocities of reactants and products. Rotational angular momenta and rotational energies

( $E_{rot}$ ) were calculated as  $J = r \times p$  and  $E_{rot} = (J \cdot \omega)/2$ , where  $\omega$  is the angular velocity, determined as  $\omega = J \cdot I^{-1}$ , and  $I$  is the moment of inertia tensor. For non-dissociation trajectories,  $E_{rot}$  was calculated for the intact base pair; and for dissociation trajectories, product  $E_{rot}$  was calculated as the sum of two separate rotators.

Upon the suggestion by one of the reviewers, the three representative collision trajectories demonstrated in Fig. 5 were recalculated at the B3LYP/6-31G\*\* level of theory. The 6-31G\*\* basis set adds  $p$  functions to hydrogen atoms in addition to the  $d$  functions on heavy atoms, and could be used to test how a basis set with explicit polarization term for hydrogen atoms affects intra-base pair proton transfer and base pair dissociation. The comparison of B3LYP/6-31G\* and B3LYP/6-31G\*\* trajectories confirms that the trajectories follow identical reaction pathways and produce nearly the same product energy distributions in the two sets of calculations, with the only difference being that the duration for intra-base pair proton transfer becomes shorter at B3LYP/6-31G\*\*.

## References

- 1 X. Hu, W. L. Hase and T. Pirraglia, *J. Comput. Chem.*, 1991, **12**, 1014-1024.
- 2 W. L. Hase, K. Bolton, P. de Sainte Claire, R. J. Duchovic, X. Hu, A. Komornicki, G. Li, K. Lim, D. Lu, G. H. Peslherbe, K. Song, K. N. Swamy, S. R. Vande Linde, A. Varandas, H. Wang and R. J. Wolf, VENUS 99: A General Chemical Dynamics Computer Program, Texas Tech Univeristy Lubbock, TX, 1999.
- 3 V. Bakken, J. M. Millam and H. B. Schlegel, *J. Chem. Phys.*, 1999, **111**, 8773-8777.
- 4 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian 09, Revision D.01, Gaussian, Inc, Wallingford, CT, 2013.
- 5 G. B. Bacskay, *Chem. Phys.*, 1981, **61**, 385-404.
- 6 G. H. Peslherbe, H. Wang and W. L. Hase, *Adv. Chem. Phys.*, 1999, **105**, 171-201.
- 7 A. Untch, R. Schinke, R. Cotting and J. R. Huber, *J. Chem. Phys.*, 1993, **99**, 9553-9566.
- 8 W. H. Miller, W. L. Hase and C. L. Darling, *J. Chem. Phys.*, 1989, **91**, 2863-2868.
- 9 L. Laaksonen, gOpenMol, Center for Scientific Computing, Espoo, Finland, 3.0 edn., 2005, available at [www.csc.fi/gopenmol/](http://www.csc.fi/gopenmol/).