Supporting Information

Photodissociation Dynamics of CO-Forming Channels on the Ground State Surface of Methyl Formate at 248 nm: Direct Dynamics Study and Assessment of Generalized Multicenter Impulsive Model

Federico Palazzetti¹ and Po-Yu Tsai^{2*}

¹ Dipartimento di Chimica, Biologia e Biotecnologie, Università degli Studi di Perugia, Perugia 06123, Italy

²Department of Chemistry, National Chung Hsing University, Taichung 402, Taiwan

^{*}To whom correspondence should be addressed.

Email: pytsai@nchu.edu.tw

More about electronic structure calculation and direct dynamics

ediculated via various busis sets. [Kedi/inol]									
	TS1	TS2	TS4	TS5	RTS				
	$\Delta E(B3LYP)$	$\Delta E(B3LYP)$	$\Delta E(B3LYP)$	$\Delta E(B3LYP)$	ΔE(M06-2X)				
6-31G(d,p)	67.63	-	74.35	-	-				
6-31+G(d,p)	66.62	-	73.59	-	-				
6-31++G(d,p)	66.53	71.09	73.50	-	-				
6-311G(d,p)	65.20	68.81	73.82	84.71	96.15				
6-311+G(d,p)	65.02	69.18	74.45	83.74	95.60				
6-311++G(d,p)	65.01	69.13	74.43	83.75	95.61				
aug-cc-pVTZ	65.14	69.28	74.38	84.82	96.94				

Table S1. The energies (including the ZPE) of the stationary points relative to cis-HCOOCH₃ calculated via various basis sets. [kcal/mol]

Table S2 Information for each dissociation pathway [unit of energy is in kcal/mol]

	TS1	TS2	TS4	TS5	RTS ^(a)
No. of trajectory	1000	1000	1000	1000	1245
No. of trajectory to product	971	983	977	976	973
No. of trajectory to reactant	29	17	23	24	170
No. of trajectory to radicals	0	0	0	0	102
Final distance between products [Å]	5.5	7.0	5.5	5.5	5.5
mean dissociation time [fs] ^(b)	88.07±23.22	105.63±33.02	68.00±11.00	103.39±22.82	168±56
endergonicity of products	17.2	43.2	17.2	17.2	16.5
energy of saddle point	69.8	75.8	78.7	89.0	102.8
exit barrier height	52.6	32.6	61.5	71.8	86.3
total excess energy	154.11	154.11	154.11	154.11	154.8
excess energy of products	136.91	110.91	136.91	136.91	138.35
excess energy above exit barrier	84.31	78.31	75.41	65.11	52.05

^(a) RTS was calculated by UM06-2X while other SPs were calculated by B3LYP functional.

^(b) The dissociation time depends on the final distance between the products in each pathway.

Total excess energy = 248nm photon energy + ZPE of the parent molecule.

Endergonicity of products: energy relative to the energy at equilibrium geometry of cis-form methyl formate.

Excess energy of products=total excess energy-endergonicity of product.

Exit barrier height = energy of the saddle point–endergonicity of product.

Excess energy above the exit barrier = total excess energy – exit barrier height.



B3LYP & UM06-2X /6-311G(d,p)

Figure S1. The energy level diagram of the five dissociation channels. The energy values of each saddle point and products relative to the *cis* form of methyl formate are reported with including the zero-point energy of each stationary point.



Figure S2 The structures of (a) roaming saddle point (RTS) and (b) the configuration of the IRC at $0.75(\text{amu})^{1/2}$ -Bohr downhill from RTS("sp-15"), both calculated at level of UM06-2X/6-311G(d,p).

More about GMCIM: Initial conditions

The available energy is set to be the excess energy above the starting structure (the saddle points, except for roaming pathway we choose the configuration the same as in direct dynamics) including the ZPE. There are 3N-7 (N is the number of atoms) transverse normal mode in the parent molecule if the structure is a saddle point on potential energy surface. For the i-th normal mode with harmonic frequency vi, the probability P(Ei) to have the i-th mode with mode energy $E_i = n_i h v_i$ can be expressed by the vibrational density of states. The density of states is calculated by the Beyer-Swinehardt algorithm. Once the probability $P(E_i)$ of the i-th mode at all the possible E_i is determined, the actual value of E_i is randomly sampled with aid of cumulative distribution function. The procedure for energy sampling of the next mode ((i+1)-th) is similar, except that the energy of the i-th mode, obtained from previous sampling, is subtracted from the available energy. When the energies of all the 3N-7 modes have been determined, the residual of the available energy is assigned to the reaction coordinate, that is the direction of the imaginary normal mode at the saddle point. The randomly- chosen energies of the remaining 3N-6 degrees of freedom are individually converted to Cartesian coordinates and momenta of classical harmonic oscillator with aid of the Cartesian eigenvectors of normal modes. These eigenvectors can be taken from the normal mode analysis of electronic structure calculations. The converted Cartesian coordinates and momenta of all the normal modes are added to constitute the initial mass-weighted Cartesian coordinate and velocity required by GMCIM. The above steps are repeated many times to obtain a large amount of different initial conditions with constant available energy. The velocity vectors of each initial condition are checked to ensure that there is no component corresponding to end-over-end rotation of parent molecule. If there is any of such component, an extra angular velocity of the opposite rotating direction is added to the velocity vector to compensate the rotation. Once the rotating

motion has been eliminated from the velocity vector, the velocity vector is further rescaled to ensure that the energy conservation is respected. In this study, about 1000 different initial conditions are generated for non-rotating parent molecules in each pathway.

Identification of conserved and transitional modes

For the mass-weighted Cartesian eigenvector of each transverse mode, $\mathbf{C}^{(i)}$, a column vector, pick out the component of each fragment. If the components of the k-th fragment picked from the i-th eigenvector is $\mathbf{C}_{k}^{(i)}$, then record the relative contribution of the fragment by comparing the norm square of the eigenvector and the component: $\|\mathbf{C}_{k}^{(i)}\|^{2} / \|\mathbf{C}^{(i)}\|^{2}$. Then normalize the component of the k-th fragment, construct the projection operator for the k-th fragment, \mathbf{P}_{k} , that contains the translation and rotation of the fragment.

$$\mathbf{U}_{\mathbf{k},\mathbf{vib}}^{(i)} = \left(\mathbf{I} - \mathbf{P}_{\mathbf{k}}\right)^{\mathrm{T}} \mathbf{U}_{\mathbf{k}}^{(i)} \quad , \quad \mathbf{U}_{\mathbf{k}}^{(i)} \equiv \mathbf{C}_{\mathbf{k}}^{(i)} / \left\|\mathbf{C}_{\mathbf{k}}^{(i)}\right\|$$

The relative contribution of vibration in the k-th fragment can be related to the norm-square ratio again: $\|\mathbf{U}_{\mathbf{k},\mathbf{vib}}^{(i)}\|^2 / \|\mathbf{U}_{\mathbf{k}}^{(i)}\|^2$ To identify a transverse mode is either conserved or a transitional one, examine the total contribution from the vibrations of all the fragments, $N_{\text{vib}}^{(i)}$:

$$N_{vib}^{(i)} = \frac{1}{\left\|\mathbf{C}^{(i)}\right\|^{2}} \sum_{k} \frac{\left\|\mathbf{C}_{k}^{(i)}\right\|^{2} \left\|\mathbf{U}_{k,vib}^{(i)}\right\|^{2}}{\left\|\mathbf{U}_{k}^{(i)}\right\|^{2}}$$

If $N_{vib}^{(i)}$ is close to 1, then the i-th transverse mode is a conserved mode; if $N_{vib}^{(i)}$ is close to 0, the mode is the transitional one. According to our examination in the five dissociation pathways considered in this studies, over 90% of the transverse modes give the contribution very close to either 1 or 0. In these 10% exceptions, few of them give values around 0.7 in average, assigned to conserved modes; the others are still below 0.2 in average, assigned to transitional modes. The number of conserved modes are found to equal to number of the total vibrational modes in the

products. In principle, there are 13 vibrational modes in the products of two-body channel and 8 vibrational modes in the products of three-body channel of methyl formate.

Product energy disposal and potential energy of transitional modes.

The potential energy of transitional modes is released by considering the following scheme of coordinates. For the k-th fragment, take the center of mass (COM) position vectors of the last and the first IRC points (Send & so). well as the difference between them, as $\Delta \mathbf{r}_{\mathbf{cm},\mathbf{k}} \equiv \mathbf{r}_{\mathbf{cm},\mathbf{k}} \left(s_{end} \right) - \mathbf{r}_{\mathbf{cm},\mathbf{k}} \left(s_0 \right)$. (The displacements of transverse vibrations are already included in the position vectors)For an example of triple fragmentation, see Figure S3. Note that $\mathbf{r_{cm}}(s_{end})$ and $\mathbf{r}_{cm}(s_0)$ of the fragments may not be parallel if the IRC has curvature.



Figure S3. The position vectors of fragments in a concerted triple fragmentation.

According to the reasons which we provided in Method section, the potential energy of transitional mode can release in the direction of dissociation coordinate as long as the sudden approximation is valid. The dissociation coordinate can be well-defined by a single relative coordinate of the two bodies' center of mass in two-body dissociation, thus the released

potential energy can be distributed radially to fragments' translation, subject to the constraint of momentum conservation. There is no unique dissociation coordinate for three-body channel, the transitional energy can distribute to the translation of product with arbitrary amount and direction in each fragment, without violating the constraint of momentum conservation. For a process of k-body dissociation, direction of translational energy release for each fragment can be decided by $\mathbf{r}_{cm}(s_{end})$ and $\Delta \mathbf{r}_{cm}$. We assume that the transitional potential is released to fragments' translation without generate extra orbiting angular momentum, therefore the direction of energy release is set to be $\mathbf{r}_{cm}(s_{end})$ vector for each fragment. The relative amount of the transitional potential acquired by each fragment is decided via the relative norm of $\Delta \mathbf{r}_{cm}$. The following procedure can be adopted for the k-body dissociation.

1. Calculate the energy partition of transitional potential, V_{tt} , which is sum of the potential energies of all the transitional modes, to the k-th fragment by considering the norm-square of all the Δr_{cm}

$$\Delta E_{T,k} = V_{tt} \frac{\left\| \Delta \mathbf{r}_{\mathbf{cm,k}} \right\|^2}{\sum_k \left\| \Delta \mathbf{r}_{\mathbf{cm,k}} \right\|^2}$$

2. If the previously determined COM velocity of k-th fragment is $v_{cm,k}$, the unit vector for the direction of potential energy release ($\Delta E_{tt,k}$) is the direction of $r_{cm}(s_{end})$:

$$\hat{\mathbf{u}}_{\mathbf{k}} \equiv \frac{\mathbf{r}_{\mathrm{cm,k}}\left(s_{end}\right)}{\left\|\mathbf{r}_{\mathrm{cm,k}}\left(s_{end}\right)\right\|} \qquad \hat{\mathbf{u}}_{\mathbf{v},\mathbf{k}} \equiv \frac{\mathbf{v}_{\mathrm{cm,k}}\left(s_{end}\right)}{\left\|\mathbf{v}_{\mathrm{cm,k}}\left(s_{end}\right)\right\|}$$

The direction of the COM velocity-gain vector of the k-th fragment (defined as $\Delta v_{cm,k}$) due to translational energy gain ($\Delta E_{T,k}$) is the direction of \mathbf{u}_k .

$$\Delta \mathbf{v}_{\mathrm{cm},\mathrm{k}} = \left\| \Delta \mathbf{v}_{\mathrm{cm},\mathrm{k}} \right\| \hat{\mathbf{u}}_{\mathrm{k}}$$

The norm of the velocity-gain vector can be obtained by solving the second-order algebraic equation:

$$\Delta E_{T,k} = \frac{m_k}{2} \left(\mathbf{v}_{\mathrm{cm,k}} + \left\| \Delta \mathbf{v}_{\mathrm{cm,k}} \right\| \hat{\mathbf{u}}_k \right)^2 - \frac{m_k}{2} \left(\mathbf{v}_{\mathrm{cm,k}} \right)^2$$
$$\Rightarrow \left\| \Delta \mathbf{v}_{\mathrm{cm,k}} \right\| = -\left\langle \hat{\mathbf{u}}_k, \mathbf{v}_{\mathrm{cm,k}} \right\rangle + \left[\left\langle \hat{\mathbf{u}}_k, \mathbf{v}_{\mathrm{cm,k}} \right\rangle^2 + \frac{2\Delta E_{T,k}}{m_k} \right]^{1/2}$$

The performance and advantages of GMCIM

GMCIM only requires the following information provided by electronic structure calculations: the relative energy of saddle points and products from the reactants, the equilibrium structures of saddle points and products, the IRC of the reaction path, and the hessian matrices of the saddle points. Typically, the acquirement of these information is fulfilled as long as the IRC calculation has been done, therefore no extra electronic structure calculation is necessary. Under the same number of initial conditions, a simulation of the post-transition state dynamics by GMCIM is much faster than the one performed by direct dynamics. A computer program is written in MATLAB language to implement GMCIM in this work. All the key equations of GMCIM reported in this study and previous work (J. Chem. Phys. 148, 234101 (2018)) are simple enough to take the advantage of array programming. In practice, the program implementing GMCIM is executed by running all the initial conditions in parallel. For example, the elapsed time to obtain the 10^3 trajectories of TS1 in direct dynamic (b3lyp/6-311g(d,p)) is over 5 days on a Linux workstation with dual Intel Xeon Silver 4116 CPU (24 physical cores), but it only take 25 seconds for a set of 10³ initial conditions in simulation of GMCIM on a laptop computer with Intel i5 8250U CPU (4 physical cores). It took 140 seconds to finish a set of 10^4 conditions in GMCIM, and one can predict naturally that the elapsed time for 10^4 trajectories in direct dynamics will be more than 50 days, as long as the trajectories are calculated sequentially in direct dynamics. Hence for a study of post TS dynamics, GMCIM can easily provide a statistically significant number of results. Considering that the quality of the result it generates within short elapsed time, GMCIM to obtain approximate solutions to the post TS dynamics with limited computational resources, especially for theoretical interpretations of the experimental observations on the basis of IRC and sudden approximation. In summary:

(i) GMCIM minimizes the requirement of electronic structure calculations: when a saddle point has been computationally investigated until the IRC is obtained, the information required by a simulation of GMCIM is also fulfilled;

(ii) the post-transition state dynamics of polyatomic molecule is reduced by GMCIM to become an effective one-dimensional (IRC) problem. The shape of the PES along the remaining 3N-7 coordinates is approximated to a quadratic one at the saddle point, followed by simplifying the dynamics of the 3N-7 degrees of freedom via sudden approximation; (iii) due to the reasons described in (i) and (ii), GMCIM can be implemented in array programming with high performances to generate large amount of results.

Additional Figures:



Figure S4 The rotational distribution of CO(v=1,j) from all the pathways as well as the two sets of experimental result (480K and 1080K/470K bimodal, Boltzmann distribution assumed).



Figure S5 The rotational Boltzmann fit of CO(v=1,j) from TS4. The data point at far left of the figure (CO(v=1,j=0)) is excluded in the fitting.



Figure S6 The rotational Boltzmann fit of CO(v=1,j) from TS5. The data point at far left of the figure (CO(v=1,j=0)) is excluded in the fitting.



Figure S7 The rotational Boltzmann fit of CO(v=1,j) from RTS.



Figure S8 The two-dimensional correlation diagram of translational energy/rotational angular momentum of CO products, dissociated from various pathways.



Figure S9 The two-dimensional correlation diagram of rotational angular momentum/ translational energy of methanol and formaldehyde products, dissociated from various pathways.



Figure S10 The two-dimensional correlation diagram of rotational angular momentum of methanol or formaldehyde products with the rotational quantum number j of CO, dissociated from various pathways.



Figure S11 The two-dimensional correlation diagram of vibrational energy of polyatomic products with the translational energy of CO, dissociated from various pathways.