Supporting Information

An AIMD study of CPD repair mechanism in water: role of solvent in ring splitting

Ali A. Hassanali[a], Dongping Zhong[a,b,c] and Sherwin J. Singer[a,c]

[a] Biophysics Program, Ohio State University

[b] Department of Physics, Ohio State University

[c] Department of Chemistry, Ohio State University

March 11, 2011

1 Charge transfer to solvent

As described in the main body of this manuscript, we observed a significant amount of charge delocalization from the CPD anion to the solvent. It is well known that DFT methods suffer from the self-interaction error which can lead to spurious charge transfer states. We tested the validity of the charge delocalization observed in our simulations by examining calculations using several Gaussian basis sets and levels of theory, and for the dimer clustered with different numbers of water molecules. We chose a configuration at the shallow free energy minimum where both C5-C5′ and C6-C6′ bonds are intact (position of nascent anion) and included a variable number of waters from the full simulation configuration in the test calculations. Inclusion of all 32 waters was not possible for the more expensive methods. In order to make sure that the delocalization effect was not specific to the configuration where the C5-C5′ and C6-C6′ were intact, we repeated the CP2K/Gaussian comparison for the Hartree-Fock calculations, with one configuration chosen from the partially cleaved minima (point 1 of Fig. 4 in preceding paper[1] of this series) and for another configuration chosen from the split products well (point 4 of Fig. 4 in preceding paper[1] of this series). Gaussian calculations were performed using the Gaussian 03/09 packages.[2,3]

Table S1 shows the total amount of excess charge on the water for configurations of the thymine dimer anion surrounded by either 6 or 10 waters. The charges from CP2K are the density derived DDAP charges[4] while those from Gaussian are obtained by using the CHELPG[5,6] utility. The results show that moving to more flexible basis sets still results in a substantial amount of charge delocalization. Furthermore, using Hartree-Fock methods, which are not plagued by SIE, the charge delocalization is also quite substantial. Hence, significant charge delocalization onto water molecules appears not be be an artifact of using density functional theory for this anionic system. We have also computed the extent of charge delocalization onto the water using the long range corrected density functional LC-BLYP developed by Hirao and co-workers[7] that is available in Gaussian 09. In the LC-BLYP density functional, the long range part of the exchange interaction is treated with full Hartree-Fock exchange and has been shown to improve the description of charge transfer states. The results in Table S1 show that the LC-BLYP functional still exhibits significant charge delocalization onto the solvent for all fragments. The comparison of the Hartree-Fock

methods to the CP2K results for two different fragment configurations along the splitting process shown in the last 6 rows of Table S1, demonstrate the the charge delocalization occurs throughout the splitting process of the the C5-C5′ and C6-C6′ bonds. Table 1 shows that the charge transfer to solvent is quite sensitive to the size of the periodic box. We have verified that the charge delocalization onto the solvent also occurs in simulations performed with a larger system size. (See Supporting Information for calculations performed with larger box size.)

Table S1: Total T<>T anion charge delocalized on water molecules according to DDAP[4] (CP2K) or CHELP[5,6] (Gaussian) density-based charge partitioning. The number of water molecules from the original simulation configuration is indicated. The CP2K calculations were performed in a cubic periodic cell of side length 9.8Å except for two values, indicated with a dagger (†), in which a cell dimension of 14.8Å was used and double dagger (‡) in which a cell dimension of 20.8Å was used. The last 6 rows are CP2K and Gaussian (ROHF and LC-BLYP) calculations performed with different configurations as described in the text.

method	software	basis	6 water	10 water
BLYP	CP2K	DZVP	−0.288, −0.154†, −0.159‡	−0.706, −0.499†, −0.501‡
BLYP	CP2K	TZV2P	−0.337, −0.184†, −0.194‡	−0.813, −0.577†, −0.580‡
ROHF	Gaussian	6-311++G**	−0.412	−0.351
LC-BLYP	Gaussian	6-311++G**	−0.122	−0.395
BLYP	Gaussian	6-311++G**	−0.311	−0.499
B3LYP	Gaussian	6-311++G**	−0.193	−0.441
BLYP	Gaussian	aug-cc-PVDZ	−0.345	−0.572
BLYP	CP2K	DZVP	−0.310, −0.333†, −0.332‡	−0.317, −0.338†, −0.342‡
ROHF	Gaussian	6-311++G**	−0.219	−0.384
LC-BLYP	Gaussian	6-311++G**	−0.242	−0.411
BLYP	CP2K	DZVP	−0.304, −0.398†, −0.398‡,	−0.423, −0.593†, −0.598‡
ROHF	Gaussian	6-311++G**	−0.291	−0.589
LC-BLYP	Gaussian	6-311++G**	−0.345	−0.646

The scheme that is used to partition the total charge among atomic and molecular species is not unique. Design of an optimal, universal charge scheme is still an unresolved issue in the literature, and the best choice of method will vary from system to system. In order to test the sensitivity of our results to the type of charge scheme applied we have also calculated MK (Merz-Kollman),[8,9] NPA (natural population analysis)[10,11] and Mulliken charges[12] for the 6 and 10 water thymine fragment of the nascent anion. We have also compared different charge schemes for the two other fragments that were chosen at different points along the splitting process. Those results are provided in Table S2. There is a qualitative difference in the charge scheme assignments between electrostatic potential derived (CHELPG and MK), NPA and Mulliken charges. The extent of charge delocalization is much less with NPA charges, although still significant. As expected, the Mulliken charges are very sensitive to the level of theory and basis set. Szefczyk and co-workers[13] have shown that at least for Lewis acid/base systems, CHELPG charges work much better than NPA charges in predicting the extent of charge transfer in these systems. Furthermore, NPA charges are not fit to any

underlying electronic density. We therefore place the most confidence in our charge distributions from density-based charge schemes, although qualitatively NPA also indicates significant charge delocalization. The density-based charges are reported in the results presented in the main text.

Table S2: Total T<>T anion charge delocalized on water molecules according to several charge partitioning schemes.

waters	method	basis	CHELPG	MK	NPA	Mulliken
6	BLYP	6-311++G**	−0.311	−0.313	−0.233	−0.138
10	BLYP	6-311++G**	−0.499	−0.550	−0.186	0.043
10	UHF	6-311++G**	−0.351	−0.390	−0.107	0.228
6	ROHF	6-311++G**	−0.412	−0.418	−0.377	0.207
10	ROHF	6-311++G**	−0.351	−0.389	−0.108	0.227
10	ROHF	6-311++G**	−0.384	−0.537	−0.128	0.381
10	ROHF	6-311++G**	−0.589	−0.698	−0.154	0.260

2 Role of solvent in C6-C6′ bond splitting: 32 Waters

In this manuscript we demonstrated that the splitting of the C6-C6′ bond is accompanied by an increase in electron density of the C6 and C6′ carbon atoms and the movement of solvent closer to those atoms. The main body of this manuscript illustrated these mechanisms for one trajectory. Below we show these processes for three other trajectories.

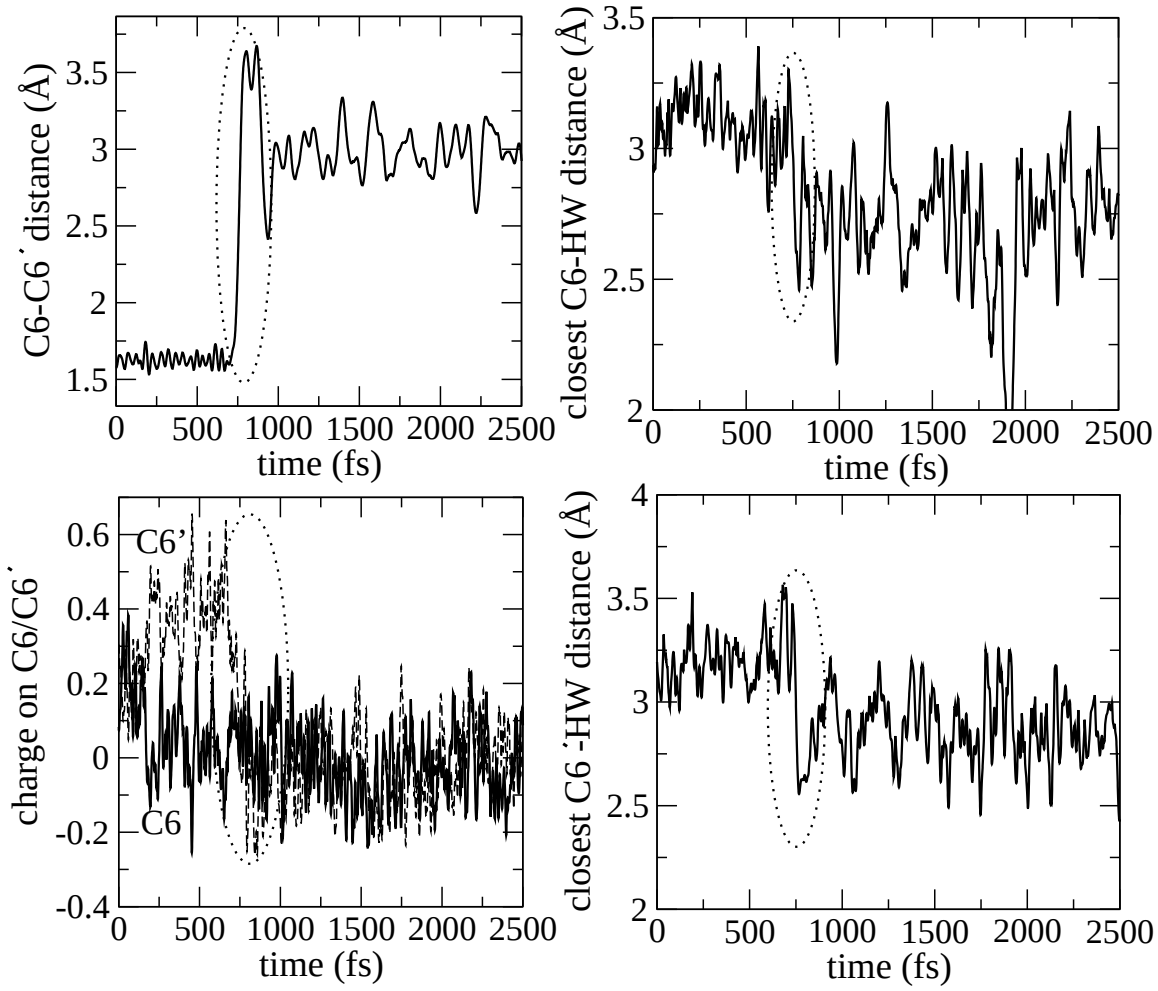

Figure S1: Top left panel shows C6-C6′ bond splitting over time, top right panel shows closest water distance to C6, bottom left shows charge on C6 and C6′ and bottom right shows closest water distance to C6′. As the C6-C6′ bond splits at about 2ps a water molecule moves closer toward the C6(C6′) carbon atom. For this trajectory we find that the charge on both the C6 and C6′ carbon atoms transitions from less negative to more negative at the same time, when the C6-C6′ bond splits.

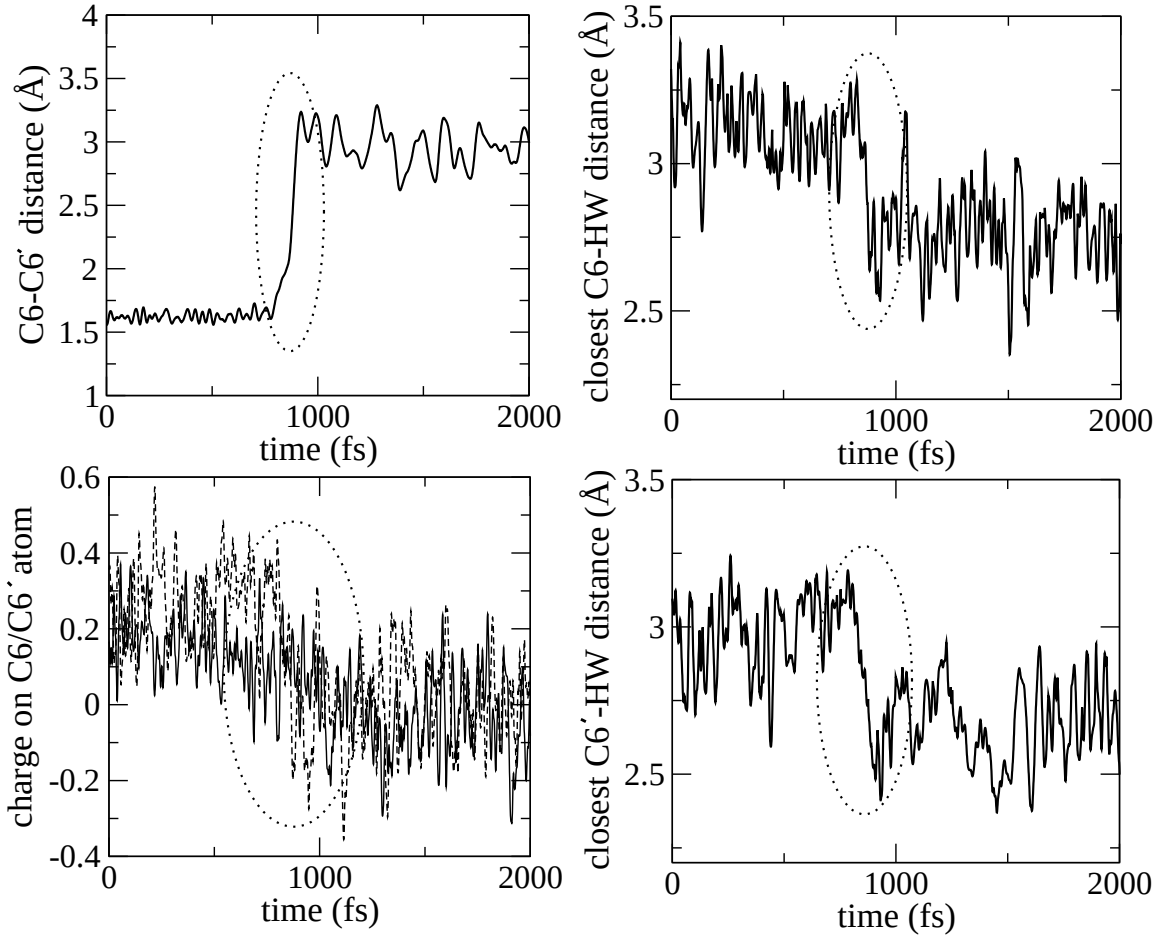

Figure S2: Top left panel shows C6-C6′ bond splitting over time, top right panel shows closest water distance to C6, bottom left shows charge on C6 and C6′ and bottom right shows closest water distance to C6′. The splitting of the C6-C6′ bond for this trajectory occurs at about 900fs. However for this trajectory we find that the charge on the C6 carbon atom gradually decreases between 400-1000fs (the decrease begins before the split of the C6-C6′ bond) while the charge on the C6′ carbon atom decreases over a shorter time interval at about 900fs.

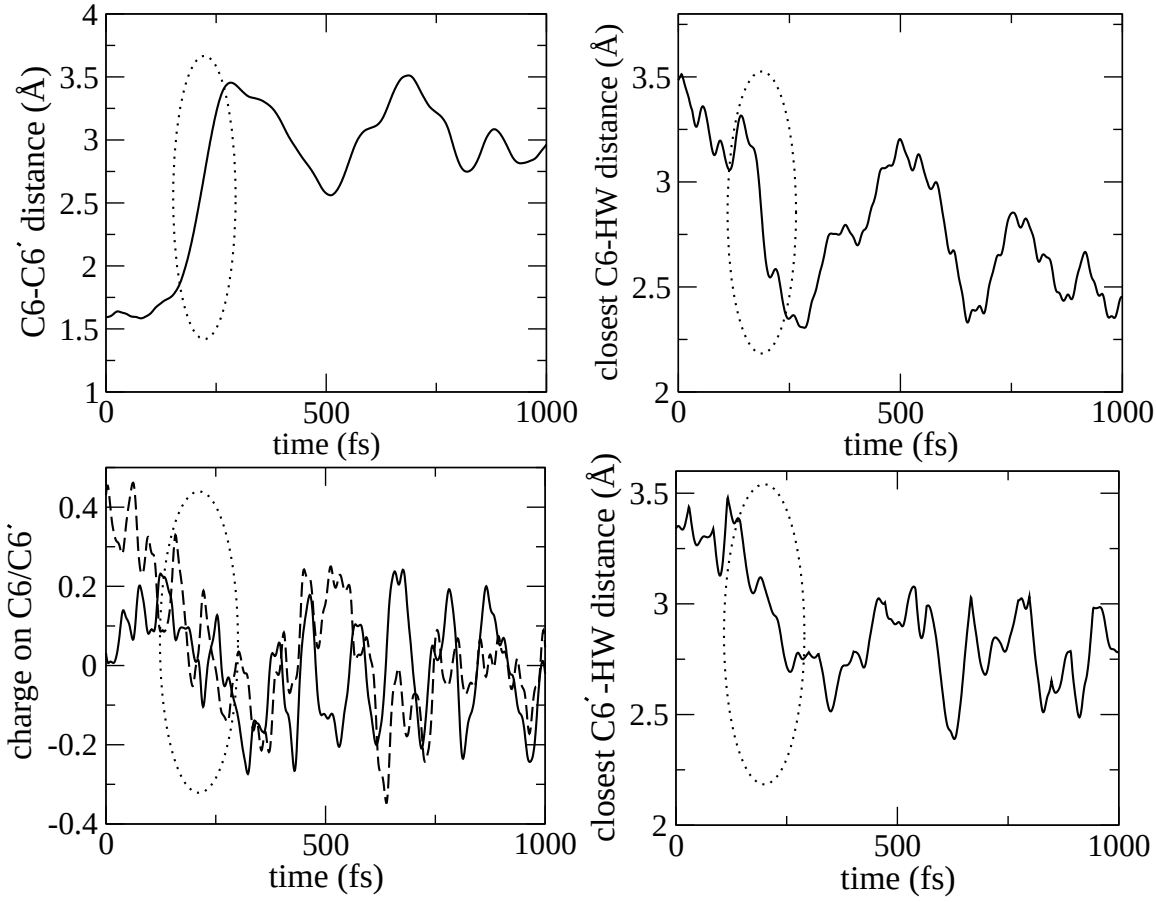

Figure S3: Top left panel shows C6-C6′ bond splitting over time, top right panel shows closest water distance to C6, bottom left shows charge on C6 and C6′ and bottom right shows closest water distance to C6′. For this trajectory, the splitting of the C6-C6′ bond occurs at 250fs. The charge on the C6 and C6′ carbon atoms and the movement of a water molecule closer to these atoms for this trajectory, transitions on a similar timescale as the splitting of the C6-C6′ bond.

3 Non-equilibrium effects, splitting times: 64 Waters

Errors due to finite box size effects are always a concern in *ab initio* simulations. While repeating our umbrella sampling calculations for a larger box size is not computationally feasible, we have attempted to check some of the important features of the splitting process for a larger system size. The system consists of the thymine dimer surrounded by a hydration pocket of 64 waters in a cubic box of side length 12.35Å. The non-equilibrium calculation documented in this work, where an electron is injected into an ensemble of neutral configurations, was analyzed for the larger system size, although for significantly fewer trajectories owing to the computational cost. The neutral thymine dimer surrounded by 64 waters was equilibrated for $3ps$ after which 30 initial configurations for non-equilibrium trajectories were taken from an ensuing $1.5ps$ of simulation. The initial velocities for the non-equilibrium simulations were chosen from a Maxwell-Boltzmann distribution at 300K. The individual non-equilibrium trajectories were run for $0.5ps$.

The most significant change in our results going from 32 to 64 waters was a lessening of non-equilibrium effects associated with cleavage of the C5-C5′ bond, as discussed below. The changes in other properties that we were able to test with a larger system were minor.

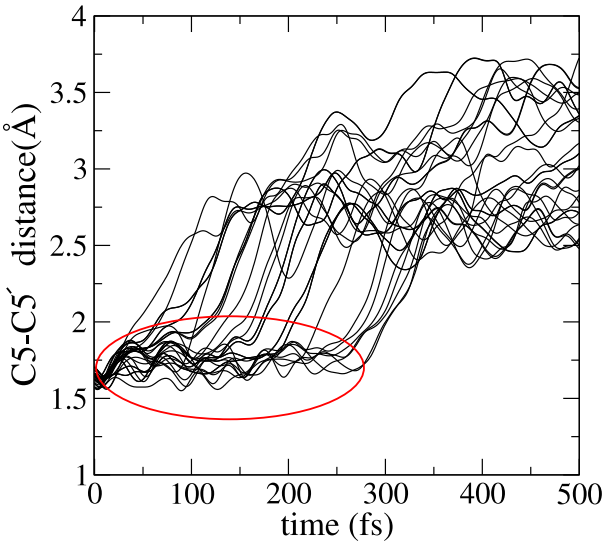

Figure S4: Non-equilibrium dynamics of C5-C5′ bond after electron injection for larger system size. Highlighted in red is the partial trap exhibited by several trajectories.

Fig. S4 shows the evolution of the C5-C5′ bond for the 30 non-equilibrium trajectories. There are several trajectories highlighted within the red circle, that remain trapped at a fixed C5-C5′ bond length for a certain period before splitting. The splitting times of the C5-C5′ bond range from less than $50fs$ to $300fs$ and the C5-C5′ bond for all trajectories split within the first $500fs$. The presence of a partial trap as indicated by the red circle in Fig. S4, is in qualitative agreement with our analysis of the non-equilibrium trajectories for the smaller box system. However, the delay time associated with this trapping is $\sim$ 3 times shorter for the larger system. For the 32 water system, we found that the C5-C5′ bond remained uncleaved up to $500fs$ for about half of the

population that was begun from the nascent anion.

The reasons for the decrease in trapping time in the larger system, other than possible lack of equilibration of the neutral dimer prior to adding an excess electron, are not clear. In the smaller box with 32 waters, the solvent configuration near the C5 carbon atom plays an important role in the splitting mechanism of the C5-C5′ bond. For the neutral dimer, we have not found any obvious or significant differences in the local water configuration around the C5 carbon atom between the 32 and 64 water systems. However, the average C5-C6-C6′-C5′ dihedral angle in the smaller system is close to zero degrees (4.8 ± 3.9), while in the larger system, based on the limited sampling available, the average dihedral angle is somewhat larger (17.6 ± 2.7). The extent of puckering in the dihedral angle might affect the solvent configuration near the C5 atoms, or manner in which the C5-C5′ σ^* antibonding orbital is populated. These differences in the initial conditions could potentially modify the subsequent dynamics of the anion.

With regard to splitting of the C6-C6′ bond, we find that the C6-C6′ bond for 15 of the 30 trajectories split within 0.5ps. This is consistent with our results for the 32 water system, and suggests that finite box size errors do not have significant effects on our TST estimate of the splitting time of the C6-C6′ bond.

Role of solvent in C6-C6′ bond splitting

In the manuscript, we have shown that the splitting of the C6-C6′ bond is accompanied by an increase of the solvent density in the vicinity of the C6 and C6′ carbon atoms. Figs. S5-S6 on the following pages show the time evolution of the splitting of the C6-C6′ bond and the closest water hydrogen to the C6 carbon atom in the left and right panels respectively, for 8 of the trajectories where the C6-C6′ bond split within $0.5 ps$. Each figure illustrates the splitting process for four trajectories. The data shows that for all of the trajectories, C6-C6′ bond splitting is accompanied by the movement of a water molecule closer to the C6 carbon atom, as found for simulations using a smaller box size.

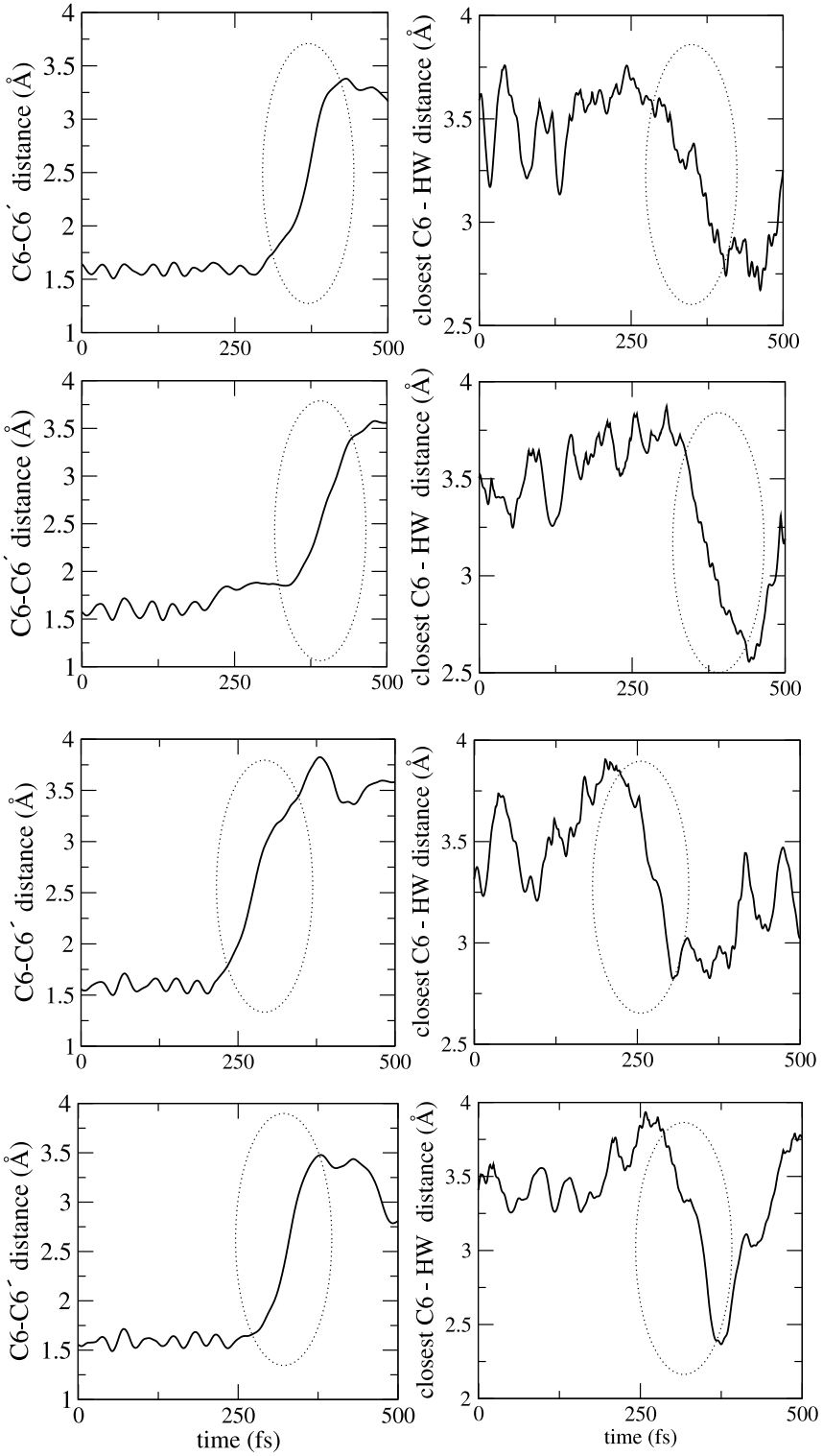

Figure S5: Non-equilibrium dynamics of C6-C6′ bond after electron injection for the larger system size containing 64 waters. Four different trajectories are shown. The evolution of the C6-C6′ bond length is given in the left panels, and the distance from the closest water hydrogen (HW) to C6 carbon atom appears in the right panels.

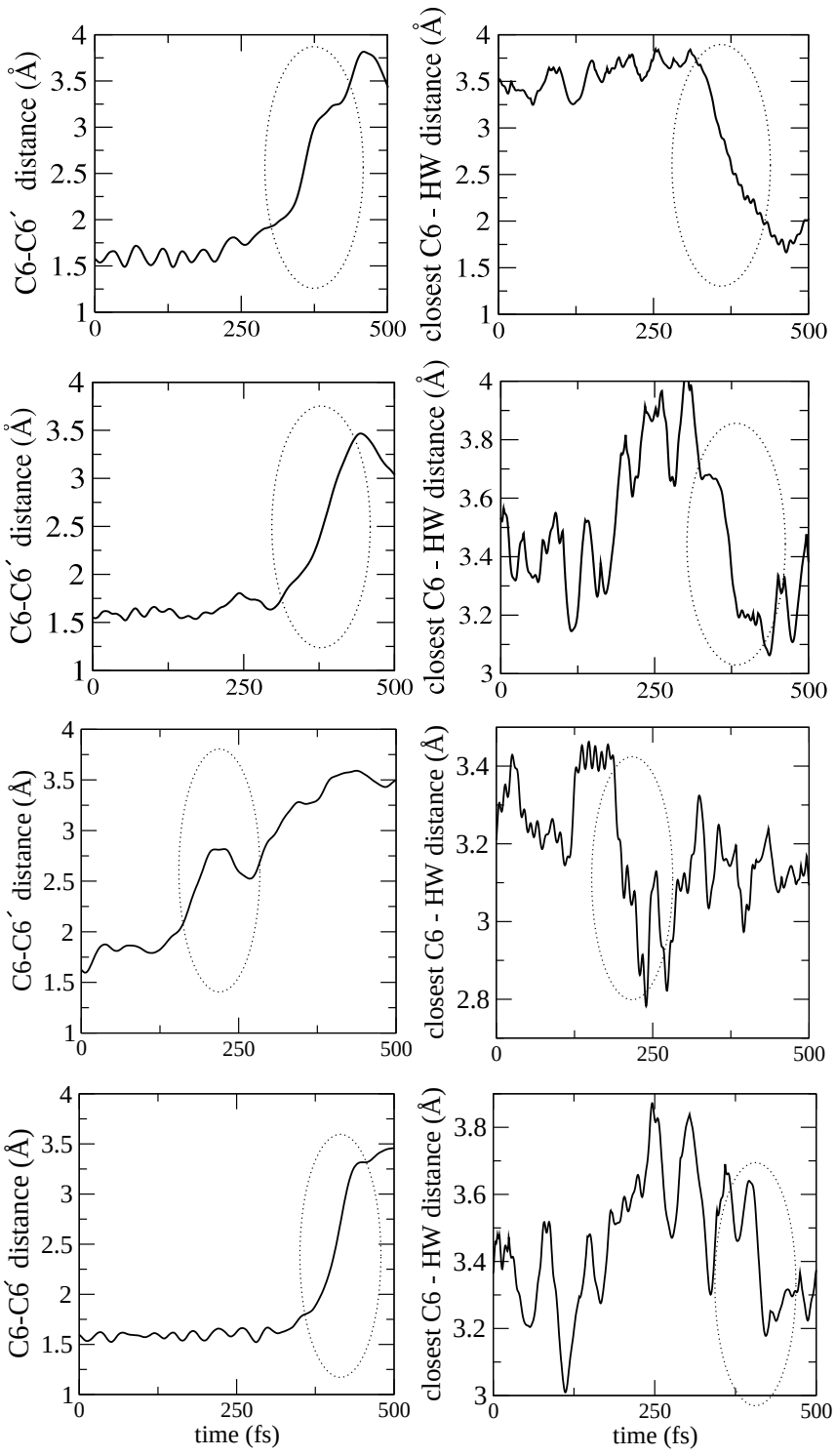

Figure S6: Non-equilibrium dynamics of C6-C6′ bond after electron injection for the larger system size containing 64 waters. Four different trajectories are shown. The evolution of the C6-C6′ bond length is given in the left panels, and the distance from the closest water hydrogen (HW) to C6 carbon atom appears in the right panels.

4 Charge transfer to solvent and simulation system size

We have documented a substantial amount of charge transfer of the excess electron from the thymine dimer to the solvent. This result has also been verified by a comparison with quantum chemistry cluster calculations. In Table S1 of the manuscript we showed that the amount of charge delocalized onto the solvent is quite sensitive to the size of the periodic box. We find a significant amount of charge transfer to solvent with the thymine dimer surrounded by 64 waters in a larger box as seen in Fig. S7 which shows the evolution of the total charge on the dimer in two trajectories from the larger system size calculations. In both trajectories we find a substantial amount of charge delocalization onto the solvent. These results confirm that the charge delocalization onto the solvent is not an artifact of finite box size effects.

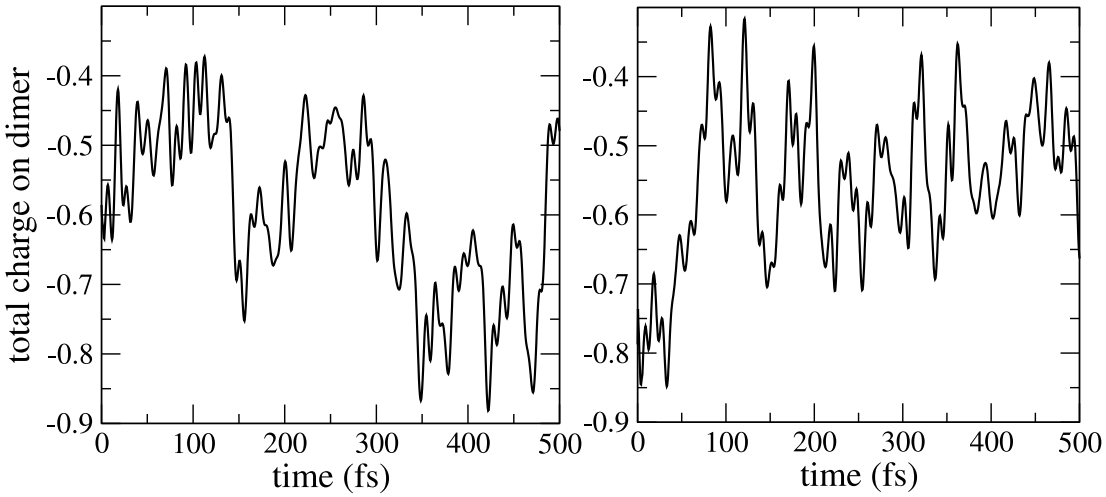

Figure S7: Non-equilibrium dynamics of the evolution of the total charge on the thymine dimer after electron injection. Shown are two trajectories. A substantial amount of charge is delocalized onto the solvent during the splitting process.

References

[1] Hassanali, A. A.; Zhong, D.; Singer, S. J. *J. Phys. Chem.* **2011**, *B* (preceding manuscript).

[2] Gaussian 03, Revision C.02. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.

[3] Gaussian 09. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. **2009**.

[4] Blöchl, P. E. *J. Chem. Phys.* **1995**, *103*(17), 7422–7428.

[5] Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*(6), 894–905.

[6] Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*(3), 361–373.

[7] Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2004**, *120*(18), 8425–8433.

[8] Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*(2), 129–145.

[9] Besler, B. H.; Merz Jr., K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*(4), 431–439.

[10] Reed, A. E.; Weinhold, F. *J. Chem. Phys.* **1983**, *78*(6), 4066–4073.

[11] Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*(2), 735–746.

[12] Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*(10), 1833–1840.

[13] Szefczyk, B.; Sokalski, W. A.; Leszczynski, J. *J. Chem. Phys.* **2002**, *117*(15), 6952–6958.