



A fragment-based approach towards *ab-initio* treatment of polymeric materials

RESHMA S PINGALE¹, SUBHASH S PINGALE^{2,*} and ANJALI KSHIRSAGAR^{1,3}

¹Interdisciplinary School of Scientific Computing, Savitribai Phule Pune University, Pune 411 007, India

²Department of Chemistry, Savitribai Phule Pune University, Pune 411 007, India

³Department of Physics, Savitribai Phule Pune University, Pune 411 007, India

*Corresponding author. E-mail: sspingale@chem.unipune.ac.in

Published online 20 June 2017

Abstract. The broad range of applications of π -conjugated polymeric materials in industries such as automobiles, textiles, packaging, medical etc. have led to their extensive studies in both academic and industrial fields. Predicting the structure of these polymers is important for the study of their properties. The present work uses a ‘divide and conquer’-type approach for the *ab-initio* studies of these polymeric systems. The method employs a fragmentation technique with independent fragment optimization for obtaining optimized geometries of the oligomers of various polymeric materials such as polyfuran, polypyrrole, polythiophene and other such π -conjugated polymers. A few test calculations performed in the study provide fair concurrence between the energies and the HOMO–LUMO energy gaps obtained using the fragmentation-based approach with those obtained using the full optimization of the whole oligomer. Also, a significant reduction in time complexity occurs for the present fragment-based approach compared to the parent system optimization. The results are encouraging and prompt for studies of large polymeric materials.

Keywords. π -Conjugated polymer; divide and conquer; *ab-initio*; fragmentation.

PACS Nos 31.15.A–; 36.20.Kd

1. Introduction

The wide-ranging applications of π -conjugated polymeric materials, due to their attractive electrical and optical properties, have led to the extensive studies of these materials in commercial as well as research-based industries [1]. A systematic theoretical prediction of the electronic structures of these π -conjugated polymers is important for studying their electrical, optical, magnetic and structural properties. The need for a reliable and highly accurate study of the properties necessitates the use of *ab-initio* quantum chemical methods for predicting an environmentally stable structure of these systems. However, the steep scaling behaviour of the *ab-initio* methods with increasing number of constituent atoms have always been a formidable challenge in their applications for predicting accurate energies, geometries and properties of large polymeric systems with the available computational resources [2]. For a system with N basis functions, the Hartree–Fock (HF) theory has a scaling factor of approximately $O(N^3)$ whereas, Møller–Plesset second-order perturbation theory (MP2)

has a scaling factor of $O(N^5)$ and coupled cluster (CC) even worse [3]. Thus, applications of these methods is restricted to a few tens of atoms. Many attempts have been made all over the globe to devise alternate linear scaling methods to reduce formidable complexity of *ab-initio* methods. Of these, ‘divide and conquer’ (DC)-type linear scaling algorithms have shown considerable success [4,5]. One such algorithm is the molecular tailoring approach (MTA) developed by Gadre and coworkers during the early nineties for one-electron property calculations of large molecules [6]. In their approach, they cut the parent system into a set of overlapping small fragments and perform *ab-initio* calculations on these fragments. In analogy with the way a tailor stitches pieces of cloths to get the desired pattern, the results were stitched to get the desired properties of the whole system. It was later extended to evaluate energy, gradients, Hessian matrix and geometry optimization of large systems including π -conjugated systems [7–11]. Further, one such similar technique is also developed by our group which fragments the parent system as the MTA does. However, the approach differs in the way

the fragments are optimized and the results are stitched together. The fragments are optimized independent of each other and parameters of these optimized fragments are stitched together to get the resultant geometry of the whole system [12]. The work here reports the use of this full fragment optimization-based algorithm for large π -conjugated polymeric materials. A comparison of timings, electronic energies and highest occupied molecular orbital (HOMO)–lowest unoccupied molecular orbital (LUMO) energy gaps (E_{HL}) obtained using the full optimized fragments approach with the corresponding quantities calculated with optimization of the full parent system is presented here.

2. Methodology

The π -conjugated test cases used in the present work are oligomers of various polymeric systems with $n = 5$ and $n = 7.5$ as shown in figure 1. All the calculations are done at the B3LYP/6-31G(d,p), B3LYP/6-31++G(d,p) and HF/6-31++G(d,p) level of theory using the package GAUSSIAN09 [13] on a HPC compute node@2.66 GHz with number of processors as 16 and memory of 32 GB. Oligomers of 10 monomer length (10-mer) units of π -conjugated systems, viz., polyacetylene (PA), polypyrrole (PPy), polycyclopentadiene (PCp), polyfuran (PFu), poly-*p*-phenylene (PPP), polythiophene (PTh), and one non- π -conjugated system, viz., polyethylene (PE) [14] were studied using the fragment optimization-based approach. Further, these oligomers were also studied for a chain length of 15 monomer (15-mer) units at the B3LYP/6-31G(d,p) level of theory. For the purpose of benchmarking these results, we also have done full geometry optimization on the whole parent system.

3. Results and discussions

A comparison of the energies obtained using both the fragment optimization-based approach (E_{frag}) and full parent system optimization (E_{full}) for the B3LYP/6-31G(d,p) level of theory for the oligomers studied here of chain length 10-mer units is given in table 1. The table also lists the time required for the fragment-based optimization (T_{frag}) and time for the full optimization (T_{full}). The last column of the table gives the ratio T_{full}/T_{frag} .

As can be seen from the DFT investigation at the B3LYP/6-31G(d,p) level of theory presented in table 1, for the case of PA oligomer, performing a full optimization on this system, the energy (E_{full}) obtained is -775.28656 a.u. and the required time (T_{full}) for this calculation is 12.14 min. When we perform fragment optimization-based calculations, the energy (E_{frag}) obtained is -775.28565 a.u. which is accurate upto three decimal places when compared with the full optimized energy and the time taken is almost three times less than that taken for the full parent system optimization. Similar results are observed for the PE oligomer. For the PPy, PCp, PFu, PPP and PTh oligomers, the energies obtained using the fragment-based optimization show a good concurrence with the corresponding energies obtained using the actual *ab-initio* full optimization. The errors in energies of the fragment-based optimization compared to full optimization for these oligomers were observed to be well within the normal acceptable limit of accuracy of 1 mHartree [11]. But when we compared the time complexities of the fragment-based optimization with those of the actual full optimization, the fragment-based method required more time than the latter method.

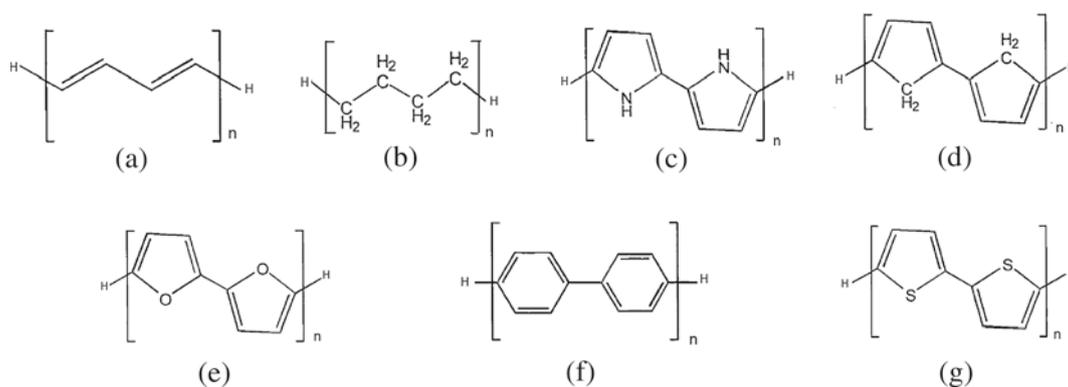


Figure 1. Oligomers with n number of dimer units (as shown in square brackets of various polymeric systems). This is one pair. Here, $n = 5$ (10-mer) and $n = 7.5$ (15-mer). Such pairs constitute the actual systems studied. (a) Polyacetylene (PA), (b) polyethylene (PE), (c) polypyrrole (PPy), (d) polycyclopentadiene (PCp), (e) polyfuran (PFu), (f) poly-*p*-phenylene (PPP), (g) polythiophene (PTh).

Table 1. Geometry optimization performed for the 10-mer oligomers at B3LYP/6-31G(d,p) level of theory.

System	No. of atoms	E_{init} (a.u.) (1)	E_{full} (a.u.) (2)	E_{frag} (a.u.) (3)	T_{full} (min)	T_{frag} (min)	$T_{\text{full}}/T_{\text{frag}}$
PA	42	-775.23223	-775.28656	-775.28565	12.14	4.33	2.80
PE	62	-787.50911	-787.53622	-787.53622	4.05	3.98	1.02
PPy	82	-2091.01694	-2091.08284	-2091.08280	14.51	27.98	0.52
PCp	92	-1930.40590	-1930.47593	-1930.47554	17.66	37.58	0.47
PFu	72	-2289.53791	-2289.61985	-2289.61979	17.23	23.25	0.74
PPP	102	-2311.80803	-2311.83421	-2311.83417	23.48	25.01	0.94
PTh	72	-5519.27094	-5519.37490	-5519.37480	17.98	27.14	0.66

(1) E_{init} – Initial single-point energy before optimization, (2) E_{full} – full optimization energy, (3) E_{frag} – energy calculated using fragment-optimized method. T_{full} – Time taken for (2) and T_{frag} – time taken for (3).

Table 2. Geometry optimization performed for the 10-mer oligomers at B3LYP/6-31++G(d,p) level of theory.

System	No. of atoms	E_{init} (a.u.) (1)	E_{full} (a.u.) (2)	E_{frag} (a.u.) (3)	T_{full} (min)	T_{frag} (min)	$T_{\text{full}}/T_{\text{frag}}$
PA	42	-775.26634	-775.31976	-775.31890	20.89	9.2	2.27
PE	62	-787.52773	-787.55481	-787.55481	58.37	11.0	5.31
PPy	82	-2091.10030	-2091.16245	-2091.16229	187.37	64.39	2.91
PCp	92	-1930.46775	-1930.53529	-1930.53488	379.24	185.02	2.05
PFu	72	-2289.61654	-2289.69357	-2289.69351	95.6	39.61	2.41
PPP	102	-2311.88093	–	-2311.90691	–	139.53	–
PTh	72	-5519.32986	–	-5519.43055	–	50.74	–

(1) E_{init} – Initial single point energy before optimization, (2) E_{full} – full optimization energy, (3) E_{frag} – energy calculated using fragment optimized method. T_{full} – Time taken for (2) and T_{frag} – time taken for (3). Dash (–) mark indicates that geometry optimization could not be done with available resources.

To investigate the reason as to why the fragment-based calculations take more time compared to the actual full parent system optimization in the case of some oligomers, we approached the problem by increasing the calculation load in two ways. In the first way, we calculated these 10-mer oligomers at the higher basis sets of B3LYP/6-31++G(d,p) and HF/6-31++G(d,p). The results for these calculations are given in tables 2 and 3 respectively.

For the B3LYP/6-31++G(d,p) level of theory, significantly different results are seen (see table 2) compared to the previous calculations at the B3LYP/6-31G(d,p) level with respect to the time complexities of the two calculations. For the test cases of PCp, PFu, PA, PE and PPy, the energies obtained using the fragment-based approach were found to be fairly matching with the full optimized energies with an accuracy of 3 to 4 decimal places. However, the thing of significant interest we observed here was the reduction in the time taken for the fragment-based optimization method compared to the full parent optimization. The time complexity results obtained for these 10-mer oligomers at the B3LYP/6-31++G(d,p) level of theory were found to be improved with bulky calculations and observed in contrast to those obtained at the B3LYP/6-31G(d,p) level of theory. It could be seen that for these oligomers, the

time taken for full parent optimization calculations is about twice the time taken for the fragment-based optimization calculations. For the PTh and PPP oligomers, a comparison between E_{full} and E_{frag} could not be done as full optimization was not possible for these systems at such a higher level of theory with the available computational resources. So we have done a comparison of the fragment-based optimized energies with that of the initial single-point unoptimized energies (E_{init}) for the PPP and PTh input systems. We can see that the energy is lower for the PPP from -2311.88093 a.u. to -2311.90691 a.u. and for the PTh from -5519.32986 a.u. to -5519.43055 a.u., thus resulting in a more stable structure for these oligomers. The fact to be noted is that the fragment-based optimization for these systems was completed within a feasible and considerable short time duration. Similar results are also obtained for the HF/6-31++G(d,p) level of theory calculations. From table 3, we can see that the E_{frag} values are in good agreement with the E_{full} values for all the test cases except for PPP and PTh. For these two oligomers, full parent system optimization could not be done, and hence we compared the fragment-optimized energies with that of the initial input system. The comparisons show that the fragment optimization method gave more stable

Table 3. Geometry optimization performed for the 10-mer oligomers at HF/6-31++G(d,p) level of theory.

System	No. of atoms	E_{init} (a.u.) (1)	E_{full} (a.u.) (2)	E_{frag} (a.u.) (3)	T_{full} (min)	T_{frag} (min)	$T_{\text{full}}/T_{\text{frag}}$
PA	42	-770.05271	-770.09551	-770.09540	6.81	4.33	1.57
PE	62	-781.88410	-781.92286	-781.92286	9.8	7.21	1.36
PPy	82	-2077.82439	-2077.91658	-2077.91650	89.9	42.45	2.12
PCp	92	-1917.68514	-1917.75561	-1917.75560	180.1	50.41	3.57
PFu	72	-2275.93715	-2276.07164	-2276.07163	78.1	46.22	1.69
PPP	102	-2296.75303	–	-2296.79359	–	123.88	–
PTh	72	-5502.53986	–	-5502.63361	–	198.38	–

(1) E_{init} – Initial single-point energy before optimization, (2) E_{full} – full optimization energy, (3) E_{frag} – energy calculated using fragment optimized method. T_{full} – Time taken for (2) and T_{frag} – time taken for (3). Dash (–) mark indicates that geometry optimization could not be done with available resources.

structures than the initial input ones for the PPP and PTh oligomers.

Thus, we observed that doing the calculations at higher basis sets, the fragment-based method showed a considerable advantage in terms of the time complexities in addition to the good concurrence of energies when compared with the full parent system optimization. Furthermore, a comparison of the energies obtained using the B3LYP/6-31++G(d,p) and HF/6-31++G(d,p) level of theories, shows that the HF method results in higher energies. This can be due to the fact that the HF method does not fully incorporate the electron correlation and this results in its higher energies. We have mentioned this HF and DFT comparison here so as to point out that our further studies are mainly focussed only on the DFT calculations.

In the second approach to analyse as to why the fragment optimization-based method calculations takes more time than the full parent system optimization at the B3LYP/6-31G(d,p) level of theory, we decided to increase the size of the oligomers under consideration. Here, we increased the chain length of the oligomers from 10-mer units to 15-mer units and performed the calculations at the same lower basis set of B3LYP/6-31G(d,p). The results for these 15-mer oligomers of PA, PE, PPy, PCp, PFu, PPP and PTh are given in table 4. In terms of energy comparison between the full optimization calculations and the fragment-based optimization, similar results were obtained as in the case of 10-mer oligomers. For all the 15-mer oligomers, the fragment-based method showed a good agreement giving an accuracy of upto 3 to 4 decimal places for the energies E_{frag} calculated when compared to E_{full} . Moreover, when we compare the time taken for the full and fragment-based optimization, considerably good results are seen. For the 15-mer oligomers of PTh, PCp and PFu, a time advantage of about two times more was seen in the fragment-based optimization calculations than the full parent

optimization calculations, whereas, for the PPP and PPy, the fragment-based optimization took a little more time compared to the full parent optimization calculations. In the case of PA and PE, however, both the methods took almost same time for the calculations which is in contrast to the results we obtained for the 10-mer PA and PE oligomers. This could be attributed to the smaller size of these systems and very short time taken for the calculations of these two oligomers. Further analysis needs to be done in these cases.

Furthermore, to test the applicability of the fragment-based optimization approach for other properties of the oligomeric systems, we applied the method for calculating the HOMO–LUMO energy gap (E_{HL}) of the oligomer test cases. The optimized geometries obtained after applying the fragment optimization-based method were used for calculating the HOMO and LUMO energies of the tested oligomers. The E_{HL} for these oligomers using the fragment optimization-based method ($E_{\text{HL-frag}}$) were then compared with the corresponding E_{HL} obtained using the full parent system optimization method ($E_{\text{HL-full}}$).

Table 5 summarizes the $E_{\text{HL-full}}$ and $E_{\text{HL-frag}}$ values in atomic units for the 10-mer oligomers at the B3LYP/6-31G(d,p) and B3LYP/6-31++G(d,p) level of theory and for the 15-mer oligomers at the B3LYP/6-31G(d,p) level of theory. As can be seen from table 5, $E_{\text{HL-frag}}$ for the fragment optimized method are in very good agreement with the corresponding $E_{\text{HL-full}}$ obtained using the full parent system optimization. We further plotted the graphs (see figure 2) of $E_{\text{HL-full}}$ vs. $E_{\text{HL-frag}}$ for the 10-mer oligomers and 15-mer oligomers at the B3LYP/6-31G(d,p) level of theory. A linear fit to these graphs resulted in a correlation coefficient of 1.00 for both the 10-mer and 15-mer oligomers which shows the accuracy of the HOMO–LUMO energies obtained by using the fragment-based method. Further, a correlation coefficient of 1.00 was also obtained

Table 4. Geometry optimization performed for the 15-mer oligomers at B3LYP/6-31G(d,p) level of theory.

System	No. of atoms	E_{init} (a.u.) (1)	E_{full} (a.u.) (2)	E_{frag} (a.u.) (3)	T_{full} (min)	T_{frag} (min)	T_{full}/T_{frag}
PA	62	-1162.25437	-1162.34146	-1162.34023	6.9	7.74	0.89
PE	92	-1180.66026	-1180.70137	-1180.70137	7.63	7.72	0.99
PPy	122	-3135.92990	-3136.03133	-3136.03129	23.28	21.58	1.08
PCp	137	-2895.01553	-2895.12517	-2895.12464	68.07	33.82	2.01
PFu	107	-3433.71205	-3433.83857	-3433.83851	36.9	19.0	1.94
PPP	152	-3467.11381	-3467.15437	-3467.15434	55.71	45.38	1.23
PTh	107	-8277.31012	-8278.46774	-8278.46765	76.87	32.72	2.35

(1) E_{init} – Initial single-point energy before optimization, (2) E_{full} – full optimization energy, (3) E_{frag} – energy calculated using fragment optimized method. T_{full} – time taken for (2) and T_{frag} – time taken for (3).

Table 5. HOMO–LUMO energy gap (E_{HL}) for the 10-mer oligomers at B3LYP/6-31G(d,p) and B3LYP/6-31++G(d,p) and for the 15-mer oligomers at B3LYP/6-31G(d,p) level of theory.

System	10-mer				15-mer	
	6-31G(d,p)		6-31++G(d,p)		6-31G(d,p)	
	E_{HL} -full	E_{HL} -frag	E_{HL} -full	E_{HL} -frag	E_{HL} -full	E_{HL} -frag
PA	0.082	0.089	0.081	0.088	0.067	0.075
PE	0.370	0.370	0.280	0.280	0.367	0.367
PPy	0.113	0.114	0.112	0.113	0.109	0.110
PCp	0.059	0.064	0.056	0.061	0.053	0.057
PFu	0.098	0.099	0.097	0.098	0.093	0.094
PPP	0.145	0.146	–	0.145	0.142	0.143
PTh	0.085	0.086	–	0.085	0.081	0.081

E_{HL} -full: HOMO–LUMO energy gap for full optimization, E_{HL} -frag: HOMO–LUMO energy gap for fragment optimization. All energies are in a.u. Dash (–) mark indicates that E_{HL} -full could not be calculated with available resources.

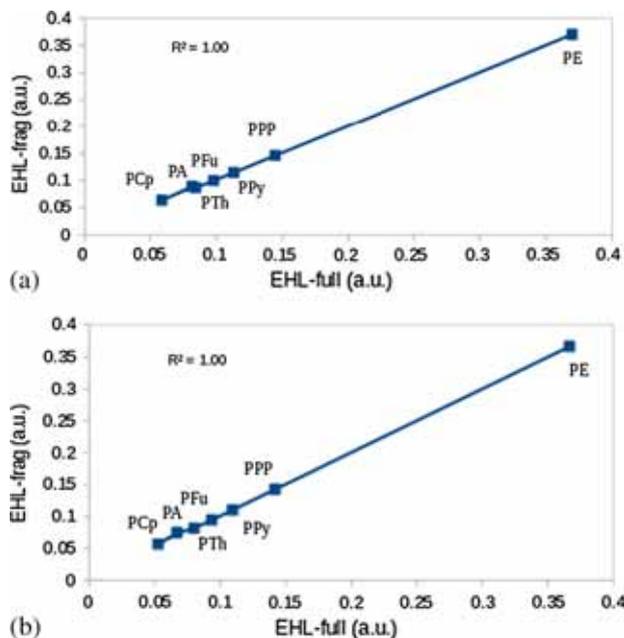


Figure 2. Plot of E_{HL} -full vs. E_{HL} -frag for (a) 10-mer and (b) 15-mer systems at the B3LYP/6-31G(d,p) level of theory.

for the 10-mer oligomers at the B3LYP/6-31++G(d,p) for the PA, PE, PFu, PPy and PCp oligomers. In the case of the PPP and PTh oligomers, E_{HL} -full could not be calculated as optimized geometries for these systems using the full optimization was not possible at the B3LYP/6-31++G(d,p) level of theory. However, using the fragment optimization-based method we obtained the values of E_{HL} -frag for PPP and PTh as 0.145 a.u. and 0.085 a.u. respectively. Thus, the fragment-based method proved to be advantageous in cases where property calculations is not feasible using full optimization.

4. Conclusions

Here we tested the applicability of the fragment optimization-based approach for geometry optimization and electronic property calculations of large polymeric materials. The method presently is tested for optimization of oligomers of length 10 and 15 monomer units at different levels of theory. Analysis of the energies

obtained and time taken for the fragment optimization method was done with those obtained from the full parent system optimization calculations. A fairly good concurrence is observed for the energies obtained using fragment optimization-based method with those of the full parent system optimized energies. Although the time complexity observed for the oligomers studied here are not very good at the B3LYP/6-31G(d,p) level of theory for the fragment optimization-based approach compared to the full parent system optimization method, considerably good results are observed at a higher level of theory, viz., B3LYP/6-31++G(d,p) and HF/6-31++G(d,p). Further, we also observed that at these higher basis sets, actual calculation are not possible for the oligomers of PTh and PPP polymeric systems. Thus, we can see the potential applicability of the fragment optimization method wherein large systems which are difficult to optimize otherwise, could be easily treated with the fragment optimization-based approach. Also by increasing the length of the oligomers, a good time advantage was also obtained at the lower level of theory for the fragment optimization method. Applicability of the method for calculating the HOMO–LUMO energy (E_{HL}) also showed very good results when compared with the E_{HL} values of the full parent system calculations. With the accuracy and time advantage in case of the results obtained from the present fragment-based study, we can further utilize the method for studies of larger systems at higher level of theories and also explore the algorithm further not only for energy and E_{HL} calculations but also for other electronic property studies which will be useful for the design of such π -conjugated polymeric materials for better applications.

Acknowledgements

Financial support of DST, Govt. of India, New Delhi, in the form of Women Scientist Fellowship SR/WOS-A/PS-19/2013 to RSP is thankfully acknowledged.

References

- [1] J Casanovas and C Alemán, *J. Phys. Chem. C* **111**, 4823 (2007)
- [2] H-A Le, H-J Tan, J F Ouyang and R P A Bettens, *J. Chem. Theory Comput.* **8**, 469 (2012)
- [3] S R Gadre, K V J Jose and A P Rahalkar, *J. Chem. Sci.* **122**, 47 (2010)
- [4] Q Zhao and W Yang, *J. Chem. Phys.* **102**, 9598 (1995)
- [5] S Li, W Li and T Fang, *J. Am. Chem. Soc.* **127**, 7215 (2005)
- [6] S R Gadre, R N Shirsat and A C Limaye, *J. Phys. Chem.* **98**, 9165 (1994)
- [7] S R Gadre, A P Rahalkar and V Ganesh, *Indian Assoc. Nucl. Chem. Alli. Sci.* **V**, 267 (2007)
- [8] V Ganesh, R K Dongare, P Balanarayan and S R Gadre, *J. Chem. Phys.* **125**, 104109 (2006)
- [9] S R Gadre and V Ganesh, *J. Theor. Comp. Chem.* **5**, 835 (2006)
- [10] S D Yeole and S R Gadre, *J. Chem. Phys.* **132**, 094102 (2006)
- [11] V Ganesh, *J. Comput. Chem.* **30**, 661 (2009)
- [12] R S Pingale, A Kshirsagar, S R Gadre and S S Pingale, (details of the approach will be published elsewhere)
- [13] M J Frisch, G W Trucks, H B Schlegel, G E Scuseria *et al*, Gaussian09, Revision A02 (Gaussian, Inc, Wallingford, CT, 2009)
- [14] J Ma, S Li and Y Jiang, *Macromolecules* **35**, 1109 (2002)