Theor Chem Acc (2011) 130:449–453
DOI 10.1007/s00214-011-1021-x

REGULAR ARTICLE

Application of resolution of identity approximation of second-order Møller–Plesset perturbation theory to three-body fragment molecular orbital method

Michio Katouda

Received: 25 April 2011 / Accepted: 10 August 2011 / Published online: 25 August 2011

Abstract The resolution of identity (RI) approximation of second-order Møller–Plesset perturbation (MP2) theory, termed as RI-MP2, is applied to three-body fragment molecular orbital (FMO3) method. New implementation of FMO3 RI-MP2 is developed based on an efficient parallel RI-MP2 code developed recently in our group. Using this new implementation, the accuracy and computational time of FMO3 RI-MP2 calculations are assessed for water clusters, polyalanines, and proteins. The errors arising from RI-MP2 are sufficiently small in the FMO3 MP2 calculations that they give quantitative accuracy for practical chemical applications. Considerable time savings are attained in the FMO3 MP2 calculations with the application of RI-MP2.

Keywords Three-body fragment molecular orbital method · Resolution of identity approximation · Second-order Møller–Plesset perturbation theory · RI-MP2 · Density fitting MP2

1 Introduction

The fragment molecular orbital (FMO) method [1–10] is an efficient approach for the rapid *ab initio* quantum chemical calculations of large biological molecules and nanomolecules. FMO reduces the computational costs considerably by dividing the target system into fragmented systems and performing *ab initio* quantum chemical calculations for each monomer and dimer (and trimer, if necessary) of fragments. The three-body FMO method (FMO3) [7–10], which is based on three-body expansion of total energy, is a promising tool for treating biological molecules in quantitative accuracies.

The dispersion interactions play an important role in biological and nano-scale molecules. Second-order Møller–Plesset perturbation (MP2) theory [11–15] is the simplest electron correlation method at the *ab initio* level to account for dispersion interactions, which widely used density functional theory (DFT) methods fail to describe. MP2 is applied successfully to FMO [10, 16–19]. This opens up many practical chemical applications of FMO to biological molecules such as protein–ligand bindings [1, 19–21].

Despite the computational efficiency of the FMO scheme, practical FMO3 MP2 calculations of large biological and nano-scale molecules are still computationally demanding. For two-body FMO (FMO2) MP2 calculations, the resolution of identity (RI) MP2 [22], which is often termed as RI-MP2 or density fitting MP2, and Cholesky decomposition (CD) MP2 [23] were applied to reduce these computational costs by introducing the approximation of ERIs. For FMO3, however, no attempt to apply RI-MP2 or CD MP2 has been performed to date, and the application is strongly desired. In this article, the RI-MP2 method is applied to FMO3 calculations. The accuracy and computational times of FMO3 with RI-MP2 are assessed by performing test calculations.

2 Computational details

RI-MP2 [24–30] is based on the RI approximation of ERIs [31, 32] in which the four-center two-electron repulsion

---

Dedicated to Professor Shigeru Nagase on the occasion of his 65th birthday and published as part of the Nagase Festschrift Issue.

M. Katouda (✉)
Department of Theoretical and Computational Molecular Science, Institute for Molecular Science,
Okazaki 444-8585, Japan
e-mail: katouda@ims.ac.jp

Table 1 Number of amino acid residues (N_{res}), atoms (N_{atom}), AOs (N_{AO}), and auxiliary basis functions (N_{aux}) for test molecules

Molecule	N_{res}	N_{atom}	N_{AO}		N_{aux}^{a}
			6-31G*	6-311G*	
$(H_2O)_{16}$		48	288	384	2,256
$(H_2O)_{32}$		96	576	768	4,512
$(H_2O)_{64}$		192	1,152	1,536	9,024
$(Ala)_{10}$	10	112	884	1,161	6,165
$(Ala)_{20}$	20	212	1,684	2,211	11,715
$(Ala)_{40}$	40	412	3,284	4,311	22,815
1L2Y	20	304	2,456	3,222	16,974

[a] The cc-pVTZ auxiliary basis set was used for RI-MP2 calculations

integral (ERI) is approximated by the product sum of two-center and three-center ERIs. RI-MP2 reduces the computational costs as well as the required memory and disk sizes while maintaining reliable accuracies for practical chemical applications. Recently, we have developed an efficient parallel RI-MP2 energy code and tested it with molecules consisting of up to 300 atoms and 4,000 atomic orbitals (AOs) [28, 29]. For the FMO3 RI-MP2 calculations, this code was modified to interface with the FMO code in the GAMESS-US [33, 34] program.

Using this code, the assessment of FMO3 and FMO2 with RI-MP2 was performed for water clusters $(H2O)_n$ ($n = 16, 32, 64$), α-helices of alanine α-$(Ala)_n$ ($n = 10, 20, 40$), β-strands of alanine β-$(Ala)_n$ ($n = 10, 20, 40$), and a trp-cage mini protein (PDB ID: 1L2Y). These molecules were used for the FMO test calculations in earlier studies [7, 8, 10, 16]. α-$(Ala)_n$ and β-$(Ala)_n$ are capped respectively with the acetyl and—$NHCH_3$ groups at the N-terminus and C-terminus. $(H2O)_n$ contains hydrogen bonds where three-body effects are significant. Three-body effects are also significant for α-$(Ala)_n$ because they possess large dipole moments and multiple hydrogen bonds. β-$(Ala)_n$ is good example of a linear polymer for which inter-chain interactions are slight. 1L2Y protein was chosen as an example of small proteins. Each system was divided into fragments, with one or two molecules or residues per fragment. In the following discussion, shorthand notation of FMO*n*/*m* is used, where *n* and *m* denote *n*-body FMO expansion based on dividing the total system by *m* molecules (residues) per fragment.

6-31G* [35, 36] and 6-311G* [37] basis sets were used. Because the RI-MP2 auxiliary basis sets optimized for 6-31G* and 6-311G* are not available, Weigend's auxiliary basis set for cc-pVTZ [27] was used instead. The numbers of amino acid residues (N_{res}), atoms (N_{atom}), AOs (N_{AO}), and auxiliary basis functions (N_{aux}) are presented in Table 1. Spherical harmonic Gaussian functions were used throughout. The core orbitals were frozen. Each system was computed without symmetry. No electrostatic potential (ESP) approximation was applied for FMO3 calculations. ESP approximation was applied for FMO2 calculations. The same settings for cutoff of distant fragments assessed in Ref. [10] were used. For the comparison of computational times of FMO3, FMO2, and full RI-MP2 results, FMO3, FMO2, and full MP2 calculations were also performed using the efficient parallel semi-direct MP2 code developed in our group [14].

All calculations were performed on 32 nodes of a single-core processor personal computer (PC) cluster connected with gigabit Ethernet, where each node has a 3.2 GHz processor (Pentium 4 640; Intel Corp.), a 4 GB dual-PC3200 DDR2 memory, and a 400 GB SATA2 hard disk with 8 MB cache.

3 Results and discussion

We first investigated the accuracy of RI-MP2 in FMO3 calculations, which is important for practical chemical applications. The errors of MP2 correlation energy with FMO3 and FMO2 from results of full MP2 calculations are presented in Table 2. These results demonstrate that the errors from RI-MP2 approximation are small enough to perform practical chemical applications with quantitatively reliable accuracies. The maximum absolute error of RI-MP2 is 1.207 mHartree for the FMO3/1 calculations of β-$(Ala)_{40}$ with 6-311G*. The RI-MP2 errors for 6-311G* are larger than those for 6-31G* because the same auxiliary basis set is used for both cases. To improve the RI-MP2 errors for 6-311G* more extended and optimized auxiliary basis functions are required.

For molecules such as $(H2O)_n$ and α-$(Ala)_n$, which have large dipole moments and multiple hydrogen bonds, FMO3/2 is important to obtain the quantitative accurate results. For example, the error of FMO3/2 RI-MP2/6-31G* calculation of α-$(Ala)_{40}$ is -0.915 mH while the errors of FMO3/1 RI-MP2/6-31G* and FMO2/2 RI-MP2/6-31G* are -4.511 and -1.531, respectively. Therefore, the application of RI-MP2 to FMO3/2 is essential and robust tool for the practical calculations because of the small errors from the RI-MP2 approximation.

We next investigate the computational speed of FMO3 RI-MP2. Results of the wall-clock time for RI-MP2 and

Table 2 MP2 correlation energy errors (mHartree) for FMO3/*m* and FMO2/*m* ($m = 1, 2$) calculations from full MP2 results (Hartree)

Molecules	FMO2/1		FMO2/2		FMO3/1		FMO3/2		Full
	MP2	RI-MP2	MP2	RI-MP2	MP2	RI-MP2	MP2	RI-MP2	MP2
6-31G*									
$(H_2O)_{16}$	−1.290	−1.278	−0.724	−0.710	0.022	0.044	0.020	0.038	−3.005 613[a]
$(H_2O)_{32}$	−2.520	−2.505	−1.984	−1.961	0.098	0.151	−0.085	−0.048	−6.025 657[a]
$(H_2O)_{64}$	−4.719	−4.700	−2.001	−1.966	[c]	[c]	0.171	0.248	−12.080 094[a]
α-(Ala)$_{10}$	1.037	1.097	−0.305	−0.235	−0.628	−0.550	−0.026	0.046	−7.944 869[a]
α-(Ala)$_{20}$	1.710	1.814	−0.824	−0.696	−1.924	−1.772	−0.339	−0.204	−15.175 683[a]
α-(Ala)$_{40}$	3.275	3.469	−1.777	−1.531	−4.810	−4.511	−1.175	−0.915	−29.634 999[a]
β-(Ala)$_{10}$	0.027	0.110	0.200	0.285	−0.081	0.005	0.013	0.098	−7.891 064[a]
β-(Ala)$_{20}$	0.471	0.626	0.621	0.782	−0.157	0.004	0.051	0.211	−15.066 664[a]
β-(Ala)$_{40}$	1.407	1.706	1.493	1.804	−0.303	0.009	0.130	0.441	−29.417 902[a]
1L2Y	2.564	2.760	3.325	3.540	−1.811	−1.574	−0.164	0.060	−22.051 981
6-311G*									
$(H_2O)_{16}$	−5.983	−5.733	−3.348	−3.093	0.941	1.208	0.523	0.784	−3.306 166[a]
$(H_2O)_{32}$	−14.190	−13.700	−8.061	−7.558	3.777	4.326	1.312	1.835	−6.628 391[a]
$(H_2O)_{64}$	−33.167	−32.207	−17.019	−16.030	[c]	[c]	5.031	6.089	−13.288 258[a]
α-(Ala)$_{10}$	−17.877	−17.589	−4.073	−3.782	−4.807	−4.511	−0.066	0.224	−8.551 858[a]
α-(Ala)$_{20}$	−42.920	−42.380	−11.519	−10.976	−12.485	−11.925	−0.378	0.165	−16.332 256[a]
α-(Ala)$_{40}$	−93.743	−92.695	−25.853	−24.803	−27.253	−26.166	0.000[b]	1.048	(−31.891 538)[b]
β-(Ala)$_{10}$	−8.324	−7.994	0.348	0.676	−0.395	−0.067	0.036	0.364	−8.494 006[a]
β-(Ala)$_{20}$	−18.027	−17.401	1.052	1.673	−0.855	−0.234	0.143	0.763	−16.217 459[a]
β-(Ala)$_{40}$	−37.742	−36.525	2.131	3.337	−2.127	−0.920	0.000[b]	1.206	(−31.664 037)[b]

[a] From Ref. [10]

[b] FMO3/2 results are taken as the reference

[c] FMO3 Hartree-Fock self-consistent field (SCF) calculations were not converged

MP2 calculations with FMO3, FMO2, and full (non-FMO) are presented in Table 3. The application of RI-MP2 always hastened calculations. In fact, 1.4–14.1 times of acceleration is achieved for FMO3 calculations. The wall-clock times for FMO3/2 RI-MP2 calculations are less than those for FMO3/1 MP2 calculations except for the calculations of 1L2Y with 6-31G*. Considerable reduction of computational time was accomplished for FMO3/2 calculation of 1L2Y protein with 6-31G*: FMO3/2 RI-MP2 required 8.2 h while FMO3/2 MP2 required 23.9 h. The FMO3 MP2 calculations of large proteins are considerably time consuming, but the RI-MP2 method is a promising tool for the practical chemical applications of FMO3 with reasonable computational costs.

It is also notable that the results of FMO RI-MP2 scales almost linearly with the size of α-(Ala)$_n$ and β-(Ala)$_n$ whereas the those of full RI-MP2 do not scale linearly. For small molecules such as α-(Ala)$_{10}$ and β-(Ala)$_{10}$, the results of the wall-clock time of full RI-MP2 calculations are comparable with those of FMO3/2 RI-MP2. However, the results of FMO3/2 RI-MP2 of large molecules such as α-(Ala)$_{20}$, β-(Ala)$_{20}$, and 1L2Y are more than four times faster than those of full RI-MP2. This demonstrates the computational advantage of FMO RI-MP2 over full RI-MP2.

4 Concluding remarks

In this study, RI-MP2 was applied to speed up FMO3 MP2 calculations. Test calculations demonstrate the quantitative accuracy and computational efficiency of application of RI-MP2 to FMO3. The maximum absolute error of FMO3 RI-MP2 is at most 1.207 mHartree for every molecule tested in this study. Considerable reductions of computational times were observed in FMO3 calculations by the application of RI-MP2. Especially, FMO3/2 RI-MP2/6-31G* calculation of 1L2Y protein was finished within 8.2 h whereas FMO3/2 MP2/6-31G* took 23.9 h.

Acknowledgments The author thanks Prof. Shigeru Nagase for reading the manuscript carefully and giving many suggestive comments. The author also thanks Prof. Kazuo Kitaura, Dr. Hiroaki Umeda, and Dr. Dmitri G. Fedorov for fruitful discussions. This work was supported by the Nanoscience Program in the Next Generation

Table 3 Wall-clock time (minutes) for RI-MP2 and MP2 calculations with FMOn/m (n = 2, 3 and m = 1, 2) and full molecules (non-FMO)

Molecule	FMO2/1			FMO2/2			FMO3/1			FMO3/2			Full		
	MP2	RI-MP2	Speedup	MP2	RI-MP2	Speedup	MP2	RI-MP2	Speedup	MP2	RI-MP2	Speedup	MP2	RI-MP2	Speedup
6-31G*															
$(H_2O)_{64}$	[a]	[a]		2.5	0.1	18.2	[b]	[b]		6.3	1.4	4.4	76.4	49.2	1.6
α-$(Ala)_{10}$	1.9	1.5	1.2	5.3	1.9	2.7	20.3	3.1	6.6	27.0	9.8	2.7	23.0	11.4	2.0
α-$(Ala)_{20}$	4.0	0.8	4.8	10.0	2.4	4.1	45.9	4.0	11.3	84.7	24.3	3.5	465.7	260.2	1.8
α-$(Ala)_{40}$	7.6	0.7	10.5	25.1	6.8	3.7	102.0	16.4	6.2	176.3	37.9	4.7			
β-$(Ala)_{10}$	1.7	0.4	4.3	2.0	1.2	1.8	6.2	2.3	2.6	9.8	6.9	1.4	12.8	11.1	1.2
β-$(Ala)_{20}$	3.6	0.4	10.4	3.0	1.5	2.0	11.9	2.7	4.4	19.8	9.7	2.0	363.1	257.2	1.4
β-$(Ala)_{40}$	3.4	0.3	11.2	9.5	1.6	5.8	24.5	5.3	4.6	46.3	18.2	2.5			
1L2Y	24.5	8.6	2.8	68.6	28.0	2.4	409.6	131.4	3.1	1,434.8	491.1	2.9	3,549.8	2,028.8	1.7
6-311G*															
$(H_2O)_{64}$	0.4	0.3	1.2	3.6	0.2	16.3	[b]	[b]		18.0	7.3	2.5	193.9	116.5	1.7
α-$(Ala)_{10}$	4.4	0.2	28.5	14.9	4.9	3.0	46.7	9.0	5.2	68.4	12.6	5.4	55.8	19.5	2.9
α-$(Ala)_{20}$	10.1	2.3	4.4	27.4	8.9	3.1	109.1	10.7	10.2	198.8	47.1	4.2	815.0	463.0	1.8
α-$(Ala)_{40}$	15.8	0.2	103.1	92.7	20.6	4.5	235.0	26.9	8.7	438.6	102.4	4.3			
β-$(Ala)_{10}$	3.9	0.5	8.3	4.4	1.0	4.4	14.8	2.5	6.0	22.9	11.6	2.0	34.3	18.8	1.8
β-$(Ala)_{20}$	2.9	0.2	13.5	8.5	1.5	5.6	31.2	2.2	14.1	47.9	13.5	3.6	596.7	462.0	1.3
β-$(Ala)_{40}$	8.3	0.5	17.8	27.2	9.6	2.8	63.2	9.1	6.9	86.9	27.3	3.2			

[a] Wall-Clock times are too short to observe

[b] FMO3 Hartree-Fock self-consistent field (SCF) calculations were not converged

Super Computing Project of the MEXT. Some preliminary calculations were performed at the Research Center for Computational Science, Okazaki, Japan.

References

1. Fedorov DG, Kitaura K (eds) (2009) The fragment molecular orbital method: practical applications to large molecular systems. CRC Press, Boca Raton
2. Fedorov DG, Kitaura K (2006) In: Starikov EB, Lewis JP, Tanaka S (eds) Modern methods for theoretical physical chemistry of biopolymers. Elsevier, Amsterdam, pp 3–38
3. Nakano T, Mochizuki Y, Fukuzawa K, Amari S, Tanaka S (2006) In: Starikov EB, Lewis JP, Tanaka S (eds) Modern methods for theoretical physical chemistry of biopolymers. Elsevier, Amsterdam, pp 39–52
4. Fedorov DG, Kitaura K (2007) Extending the power of quantum chemistry to large systems with the fragment molecular orbital method. J Phys Chem A 111:6904–6914
5. Kitaura K, Ikeo E, Asada T, Nakano T, Uebayasi M (1999) Fragment molecular orbital method: an approximate computational method for large molecules. Chem Phys Lett 313:701–706
6. Nakano T, Kaminuma T, Sato T, Fukuzawa K, Akiyama Y, Uebayasi M, Kitaura K (2002) Fragment molecular orbital method: use of approximate electrostatic potential. Chem Phys Lett 351:475–480
7. Fedorov DG, Kitaura K (2004) The importance of three-body terms in the fragment molecular orbital method. J Chem Phys 120:6832–6840
8. Fedorov DG, Kitaura K (2006) The three-body fragment molecular orbital method for accurate calculations of large systems. Chem Phys Lett 433:182–187
9. Fedorov DG, Kitaura K (2005) Coupled-cluster theory based upon the fragment molecular-orbital method. J Chem Phys 123:134103
10. Fedorov DG, Ishimura K, Ishida T, Kitaura K, Pulay P, Nagase S (2007) Accuracy of the three-body fragment molecular orbital method applied to Møller–Plesset perturbation theory. J Comput Chem 28:1476–1484
11. Møller C, Plesset MS (1934) Note on an approximation treatment for many-electron systems. Phys Rev 46:618–622
12. Cremer D (1998) In: Schleyer PvR, Aliinger NL, Clark T, Gasteiger J, Kollman PA, Schaefer HF III, Shreiner PR (eds) Encyclopedia of computational chemistry. Wiley, Chichester, pp 1706–1735
13. Pople JA, Binkley JS, Seeger R (1976) Theoretical models incorporating electron correlation. Int J Quantum Chem Symp 10:1–19
14. Ishimura K, Pulay P, Nagase S (2006) A new parallel algorithm of MP2 energy calculations. J Comput Chem 27:407–413
15. Ishimura K, Pulay P, Nagase S (2007) New parallel algorithm for MP2 energy gradient calculations. J Comput Chem 28:2034–2042
16. Fedorov DG, Kitaura K (2004) Second order Møller–Plesset perturbation theory based upon the fragment molecular orbital method. J Chem Phys 121:2483–2490
17. Mochizuki Y, Nakano T, Koikegami S, Tanimori S, Abe Y, Nagashima U, Kitaura K (2004) A parallelized integral-direct second-order Møller–Plesset perturbation theory method with a fragment molecular orbital scheme. Theor Chem Acc 112:442–452
18. Mochizuki Y, Koikegami S, Nakano T, Amari S, Kitaura K (2004) Large scale MP2 calculations with fragment molecular orbital scheme. Chem Phys Lett 396:473–479
19. Mochizuki Y, Yamashita K, Murase T, Nakano T, Fukuzawa K, Takematsu K, Watanabe H, Tanaka S (2008) Large scale FMO-MP2 calculations on a massively parallel-vector computer. Chem Phys Lett 457:396–403
20. Fukuzawa K, Mochizuki Y, Tanaka S, Kitaura K, Nakano T (2006) Molecular interactions between estrogen receptor and its ligand studied by the ab initio fragment molecular orbital method. J Phys Chem B 110:16102–16110
21. Nakanishi I, Fedorov DG, Kitaura K (2007) Molecular recognition mechanism of FK506 binding protein: an all-electron fragment molecular orbital study. Proteins Struct Funct Bioinform 68:145–158
22. Ishikawa T, Kuwata K (2009) Fragment molecular orbital calculation using the RI-MP2 method. Chem Phys Lett 474:195–198
23. Okiyama Y, Nakano T, Yamashita K, Mochizuki Y, Taguchi N, Tanaka S (2010) Acceleration of fragment molecular orbital calculations with Cholesky decomposition approach. Chem Phys Lett 490:84–89
24. Feyereisen M, Fitzgerald G, Komornicki A (1993) Use of approximate integrals in ab initio theory. an application in MP2 energy calculations. Chem Phys Lett 208:359–363
25. Bernholdt DE, Harrison RJ (1996) Large-scale correlated electronic structure calculations: the RI-MP2 method on parallel computers. Chem Phys Lett 250:477–484
26. Weigend F, Häser M (1997) RI-MP2: first derivatives and global consistency. Theor Chem Acc 97:331–340
27. Weigend F, Köhn A, Hättig C (2002) Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. J Chem Phys 116:3175–3183
28. Katouda M, Nagase S (2009) Efficient parallel algorithm of second-order Møller–Plesset perturbation theory with resolution-of-identity approximation (RI-MP2). Int J Quant Chem 109:2121–2130
29. Rahalkar AP, Katouda M, Gadre SR, Nagase S (2010) Molecular tailoring approach in conjunction with MP2 and RI-MP2 codes: a comparison with fragment molecular orbital method. J Comput Chem 31:2405–2418
30. Katouda M, Nagase S (2010) Application of second-order Møller–Plesset perturbation theory with resolution-of-identity approximation to periodic systems. J Chem Phys 133:184103
31. Vahtras O, Almlöf J, Feyereisen MW (1993) Integral approximations for LCAO-SCF calculations. Chem Phys Lett 213:514–518
32. Kendall RA, Früchtl HA (1997) The impact of the resolution of the identity approximate integral method on modern ab initio algorithm development. Theor Chem Acc 97:158–163
33. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) General atomic and molecular electronic structure system. J Comput Chem 14:1347–1363
34. Gordon MS, Schmidt MW (2005) In: Dykstra CE, Frenking G, Kim KS, Scuseria GE (eds) Theory and applications of computational chemistry: the first forty years. Elsevier, Amsterdam, pp 1167–1189
35. Hehre WJ, Ditchfield R, Pople JA (1972) Self—consistent molecular orbital methods. XII. Further extensions of gaussian—type basis sets for use in molecular orbital studies of organic molecules. J Chem Phys 56:2257–2261
36. Hariharan PC, Pople JA (1973) The influence of polarization functions on molecular orbital hydrogenation energies. Theor Chim Acta 28:213–222
37. Krishnan R, Binkley JS, Seeger R, Pople JA (1980) Self—consistent molecular orbital methods. XX. A basis set for correlated wave functions. J Chem Phys 72:650–654