Geometry Optimization of the Active Site of a Large System with the Fragment Molecular Orbital Method

ARTICLE *in* JOURNAL OF PHYSICAL CHEMISTRY LETTERS · JANUARY 2011

Impact Factor: 7.46 · DOI: 10.1021/jz1016894

3 AUTHORS, INCLUDING:

Yuri Alexeev
Argonne National Laboratory
29 PUBLICATIONS **357** CITATIONS

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

pubs.acs.org/JPCL

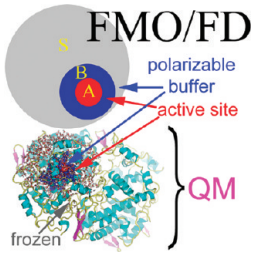
Geometry Optimization of the Active Site of a Large System with the Fragment Molecular Orbital Method

Dmitri G. Fedorov,*[,†] Yuri Alexeev,[‡] and Kazuo Kitaura[†,§]

[†]NRI, National Institute of Advanced Industrial Science and Technology (AIST), Central 2, Umezono 1-1-1, Tsukuba, 305-8568, Japan, [‡]Institute of Food Research, Norwich Research Park, Colney NR4 7UA, U.K., and [§]Graduate School of Pharmaceutical Sciences, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan

ABSTRACT An efficient formulation of the fragment molecular orbital method is introduced based on dividing the system into frozen and polarizable domains. The former is computed once taking into account the many-body polarization of the whole system, while the latter is recalculated for each step of a geometry optimization. We performed ligand docking and calibrated the method on the complexes of the Trp−cage miniprotein construct (PDB: 1L2Y) with neutral and charged ligands and applied it to optimize a partially solvated structure of prostaglandin H(2) synthase-1 in complex with the reversible competitive inhibitor ibuprofen (PDB: 1EQG) containing 19471 atoms at the B3LYP-D/6-31G* and RHF/STO-3G levels of theory for the polarizable and frozen domains, respectively. The optimization took 32 h on six dual CPU quad-core 2.83 GHz Xeon nodes. Our method requires no fitted parameters and allows optimizations of large systems based solely on quantum mechanics.

SECTION Biophysical Chemistry



Driven by the increasing computational power, quantum mechanical (QM) methods are becoming applicable to large systems. There are methods considering the whole system,[1−3] numerous fragment-based approaches,[4−17] as well as other schemes.[18−23] Quantum refinement has been applied to improve X-ray structures.[24,25]

The fragment molecular orbital (FMO) method has been proposed by Kitaura et al.[26−29] Geometry optimizations[30−32] and molecular dynamics[33] have been performed, and it has been applied to protein−ligand binding,[34,35] protein folding,[36] and drug design.[37] The distinctive feature of FMO is the inclusion of the electrostatic (ES) potential (ESP) from the whole system into each individual fragment calculation and in using the many-body expansion[38] to account for the inter-fragment interactions. Although FMO has been routinely used to calculate single-point energies, its applications to geometry optimizations have been scarce.[30−32,34,39] A new approach for an efficient optimization of a part of the large system is necessary. The first attempt is given by the partial energy gradient method.[32] The importance of treating a sufficiently large active site has also been stressed for other methods.[23]

We propose an approach based on the multilayer FMO and the concept of the polarizable buffer. In the total system S, we define the active domain A, which contains atoms whose positions are to be optimized, the polarizable buffer domain B, which includes all atoms of A and a part of S surrounding A; the rest of the system is the frozen domain F. The relation between domains can be shown as $A \subseteq B \subset S$, $S = F \cup B$; see Figure 1. Note that neither A nor B have to be continuous in space; they can be geometrically in several separated regions. F, B, and A are defined fragment-wise.

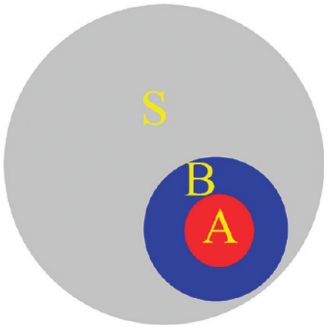


Figure 1. Schematic representation of the active (A) and polarizable buffer (B) and frozen F domains within the total system S. B includes A; $A \subseteq B \subset S$.

Next, we make the approximation that the electronic state of F is unchanged under geometry relaxation of A. This means that the electronic state of S is computed once, and during the geometry optimization of A, only the electronic densities of the polarizable buffer B are relaxed. We note that "frozen" here means that these densities are computed by considering their full mutual many-body polarization in the initial geometry, and consequently, they are not changed.

Received Date: December 15, 2010
Accepted Date: January 11, 2011
Published on Web Date: January 24, 2011

DOI: 10.1021/jz1016894 | *J. Phys. Chem. Lett.* **2011**, *2*, 282–288

The regular FMO equation is

$$E = \sum_I E'_I + \sum_{I>J} \Delta E_{IJ} \tag{1}$$

where $\Delta E'_{IJ} = E'_{IJ} - E'_I - E'_J$, $\Delta E_{IJ} = \Delta E'_{IJ} + \mathrm{Tr}(\Delta \mathbf{D}^{IJ}\mathbf{V}^{IJ})$, and E'_I and E'_{IJ} are the internal monomer and dimer energies, respectively; $\Delta \mathbf{D}^{IJ} \equiv \mathbf{D}^{IJ} - (\mathbf{D}^I \oplus \mathbf{D}^J)$ and $\mathbf{V}^{IJ}$ are the dimer density difference and ESP matrices, respectively ($\mathbf{D}$ denotes electron densities).[40] Equation 1 can be rewritten as

$$\begin{aligned} E &= E^{\mathrm{F}} + E^{\mathrm{B}} + \Delta E^{\mathrm{BF}} \\ &= \left[\sum_{I\in \mathrm{F}} E'_I + \sum_{\substack{I>J \\ I,J\in \mathrm{F}}} \Delta E_{IJ}\right] + \left[\sum_{I\in \mathrm{B}} E'_I + \sum_{\substack{I>J \\ I,J\in \mathrm{B}}} \Delta E_{IJ}\right] \\ &\quad + \left[\sum_{I\in \mathrm{B}, J\in \mathrm{F}} \Delta E_{IJ}\right] \end{aligned} \tag{2}$$

If the electronic state of F is frozen, then its energy can be approximately considered constant. $E^{\mathrm{F}} \approx$ const and $\nabla E^{\mathrm{F}} \approx 0$ during geometry optimizations (it is an approximation because fragments in F should be polarized by the fragments in B during optimization). Another approximation is to consider constant the energy of cross dimers BF not involving A, $\Delta E^{(\mathrm{B}\wedge\mathrm{A})\mathrm{F}}$, where B ∧ A denotes the part of B excluding A.

$$\begin{aligned} \Delta E^{\mathrm{BF}} &= \sum_{I\in \mathrm{B}, J\in \mathrm{F}} \Delta E_{IJ} = \Delta E^{\mathrm{AF}} + \Delta E^{(\mathrm{B}\wedge\mathrm{A})\mathrm{F}} \\ &= \sum_{I\in \mathrm{A}, J\in \mathrm{F}} \Delta E_{IJ} + \sum_{I\in \mathrm{B}\wedge\mathrm{A}, J\in \mathrm{F}} \Delta E_{IJ} \end{aligned} \tag{3}$$

Ignoring the second term in this equation as well as E^{F}, the energy of FMO/FD is defined as

$$\begin{aligned} E^{\mathrm{FMO/FD}} &= E^{\mathrm{B}} + \Delta E^{\mathrm{AF}} \\ &= \sum_{I\in \mathrm{B}} E'_I + \sum_{\substack{I>J \\ I,J\in \mathrm{B}}} \Delta E_{IJ} + \sum_{I\in \mathrm{A}, J\in \mathrm{F}} \Delta E_{IJ} \end{aligned} \tag{4}$$

We developed a general multilayer FMO (MFMO)[27] gradient in this work and used it to improve the efficiency even further. Assign F to layer L_1 and B to layer L_2. A possible choice is RHF/STO-3G for L_1 and DFT/6-31G* for L_2. Then

$$E^{\mathrm{FMO/FD}} = \sum_{I\in \mathrm{B}} E'^{\mathrm{L}_2}_I + \sum_{\substack{I>J \\ I,J\in \mathrm{B}}} \Delta E^{\mathrm{L}_2}_{IJ} + \sum_{I\in \mathrm{A}, J\in \mathrm{F}} \Delta E'^{\mathrm{L}_2,\mathrm{L}_1}_{IJ} \tag{5}$$

where we enforced the condition that in the third sum, all dimers are computed with the ES approximation[40] among the nuclear charges and densities of $\mathbf{D}^{I,\mathrm{L}_2}$ and $\mathbf{D}^{J,\mathrm{L}_1}$; $\Delta E_{IJ} = \Delta E'_{IJ}$ (no charge transfer, $\Delta \mathbf{D}^{IJ} = 0$)) at the mixed layer level. Contrary to the general MFMO scheme, the self-consistent charge (SCC) loop of L_1 need not to be repeated after the initial calculation (if $\Delta E'^{\mathrm{L}_2,\mathrm{L}_1}_{IJ}$ is replaced by $\Delta E'^{\mathrm{L}_1,\mathrm{L}_1}_{IJ}$ in the spirit of MFMO, then one has to repeat SCC of L_1 for fragments in B for each geometry).

A further speed-up can be achieved by assuming that the energy of the inactive dimers in B not involving A fragments (I ∈ B ∧ A, J ∈ B) is constant (in the second term of eq 5). This approximation is referred to as frozen domain and dimers (FDD).

$$\begin{aligned} E^{\mathrm{FMO/FDD}} &= \sum_{I\in \mathrm{B}} E'^{\mathrm{L}_2}_I + \sum_{\substack{I>J \\ I\in \mathrm{A}, J\in \mathrm{B}}} \Delta E^{\mathrm{L}_2}_{IJ} \\ &\quad + \sum_{I\in \mathrm{A}, J\in \mathrm{F}} \Delta E'^{\mathrm{L}_2,\mathrm{L}_1}_{IJ} \end{aligned} \tag{6}$$

One can ask, what is the role of these inactive dimers, which are neglected in FDD? They do not contain active atoms. However, they describe the energy relaxation of buffer B, whose electronic state changes, and they make the gradient and the energy consistent. Without these dimers, the energy gradient error increases, making it impossible to optimize to a stricter threshold.

The FD and FDD methods have the following features. (a) The long-range electrostatic interactions are described via the third term in eqs 5 and 6. We note that they are evaluated using the electron densities[40] without any multipole (i.e., point charge) approximation. (b) Full many-body polarization effects are included within the buffer B, and the initial many-body polarization is considered in the whole system S (term 1, denoting the internal energy of B computed under the total electrostatic field of S). (c) Quantum effects are incorporated for all fragments in B (FD, term 2) or for all fragments in B, including at least one fragment in A (FDD). (d) Computational efficiency is achieved because only a part of the system (B) is recomputed for each point. The larger B is, the more savings FDD offers compared to FD.

The difference between FMO/FD and molecular mechanics (MM)-based QM/MM is as follows: the former has a polarizable electrostatic environment (represented by densities or charges in ESP, depending on the distance) versus typically fixed standard charges in the latter; the quantum region is fragmented in FMO, and thus, the computations are more efficient; also, there is no need for atom typing in FMO, which can be advantageous for treating ligands without standard parameters. The active region in FMO/FD can be more easily resized because all atoms are treated quantum mechanically.

During geometry optimizations, some fragments can move, changing the distance R_{AF} between the active and frozen domains (defined based on the interfragment distances R_{IJ}). If it becomes too small, one will have to redefine B and restart geometry optimization.

$$R_{\mathrm{AF}} \equiv \min_{I\in \mathrm{A}, J\in \mathrm{F}} \{R_{IJ}\} \tag{7}$$

In FMO/FD, we relaxed the approximation thresholds used in FMO (the electrostatic dimer approximation threshold was set to 1.5). In addition, because the total FMO/FD energy is not large, the integral and SCF convergence thresholds were lowered and set to 10^{-9} for integrals and 10^{-6} for SCF/SCC convergence. All fragmentations were performed with one residue per fragment. For the FMO/FD gradient, we differentiated eqs 5 and 6 and used the hybrid orbital projector

DOI: 10.1021/jz1016894 | J. Phys. Chem. Lett. 2011, 2, 282–288

gradient by Nagata et al.[41] The standard geometry optimizer based on numeric Hessian updates in GAMESS was employed, with the important modification of using only atoms in domain B to construct the Hessian. In the accuracy tests, we employed RHF with the STO-3G and 3-21G basis sets in order to run many optimizations in reasonable time. For the final application, we used 6-31G* and the dispersion correction.[42] FMO/FD was parallelized with the general distributed data interface similar to other FMO calculations.[43] All calculations were performed on the Soroban cluster consisting of six dual CPU quad-core 2.83 GHz Xeon nodes equipped with 16 GB of RAM and connected by Gigabit Ethernet. Geometry convergence thresholds were 10^{-4} and $1/3 \times 10^{-4}$ au/bohr for the maximum and root-mean-square (RMS) gradients, respectively.

The RMS deviations (RMSD) between two sets of coordinates were computed only for the atoms active during geometry optimization. In other words, the normalizing factor in RMSD is given by the number of active atoms (e.g., 16 for L1, rather than 320 in the whole system). The RMS gradient used as a criterion in verifying the optimization convergence was computed also only for the active atoms.

For tests, we chose the Trp−cage miniprotein construct (PDB: 1L2Y)[44] and *o*-phenolic acid (neutral and deprotonated) as the ligands, denoted by L1 and L2, respectively (Figure 2). Ligands were docked to the protein using docking software Autodock Vina.[45] The protein structure was fixed during docking, while ligands were fully flexible. For the initial structure of the protein, we took the experimental one. It is important to have a charged representative test system; 1L2Y in our model had five charged residues, and we tried both a neutral and a negatively charged ligand. The influence of the B domain was studied as a function of radius R_B (fragments for which the interatomic distance to atoms in domain A is smaller than R_B are included in B). We chose the set of R_B based in the van der Waals radii of the typical contact atoms, O and H, so that their sum $1.2 + 1.4 = 2.6$ Å was multiplied by 1, 1.5, 2, and 2.5 (roughly corresponding to the FMO distance of these values).

For domain A, we tried taking (a) only the ligand (L1 or L2) and (b) the ligand and the most important part (P) of the protein based on the docking results, which for L1 and L2 were Pro-18 and Gln-5, respectively (the fragmentation in FMO is at Cα, and the peptide bond in −Cα−CO−NH− is not divided, so that two residue fragments have[38] −Cα and −CO−NH−). In the case of L1 + P and L2 + P, we defined domain B based on R_B for the full domain A = L1 + P or L2 + P. To improve FMO/FDD, we tried to include two residues (Pro-17 and Pro-19) into A, in addition to Pro-18, and denoted this by L1 + P′ (in which case, the coordinates of the atoms of these two additional residues were not optimized).

As shown in Figure 3 and Table 1, a reasonable choice of the size of B (R_B) is 3.9 and 6.5 Å for neutral and charged ligands, respectively (depending on the desired accuracy, 5.2 Å can be also used for the latter). Below, we focus mainly on the 3-21G results. When comparing FD and FDD, the general trend is that the latter apparently has a systematic error, which cannot be reduced by using a larger R_B. An

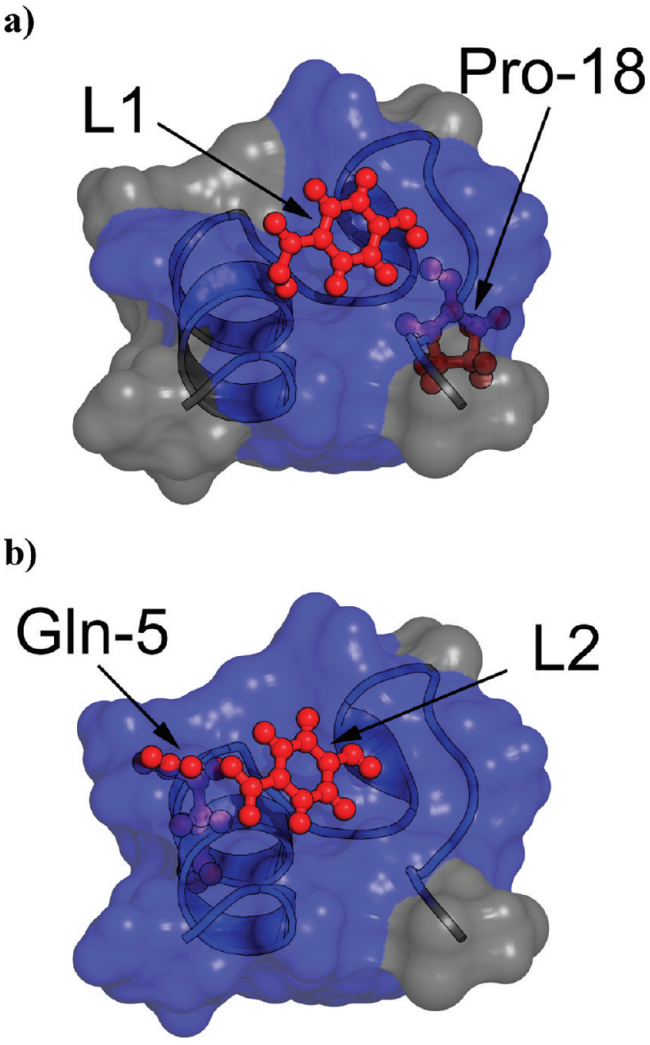


Figure 2. Structures at the FMO-RHF/3-21G level (L + P; atoms are optimized in the ligand and one shown residue) for the complexes of Trp−cage with (a) neutral (bound mostly to Tyr-3, Gln-5, and Pro-18) and (b) deprotonated *o*-phenolic acid (bound mostly to Gln-5 and Arg-16). Active (A), polarizable buffer (B), and frozen (F) domains are colored red, blue, and gray, respectively.

expansion of A from P to P′ (Figure 3d) does not show an improvement. The convergence of the error with respect to R_B for the charged ligand L2 is slow and is driven by the $1/r$ decay of the Coulomb interaction.

It is interesting to see how much of the error is due to F, whose electronic state is frozen at the initial geometry. To test this, we performed a two-step optimization (denoted by FD2), where we first did a regular FD optimization to full convergence and then restarted the optimization, whereby the electronic state of the domain F was recomputed at the initial minimum geometry. For L2, the second optimization did not take any steps for $R_B \geq 5.2$ Å (i.e., the initial gradient was below the threshold), whereas for the other two values, about 30 steps were required. It can be seen in Figure 3e that the error was considerably reduced for $R_B \leq 3.9$ Å; however, it is still somewhat large. For L1, on the other hand, for all tested R_B, the initial geometry of the second optimization

DOI: 10.1021/jz1016894 | J. Phys. Chem. Lett. 2011, 2, 282–288

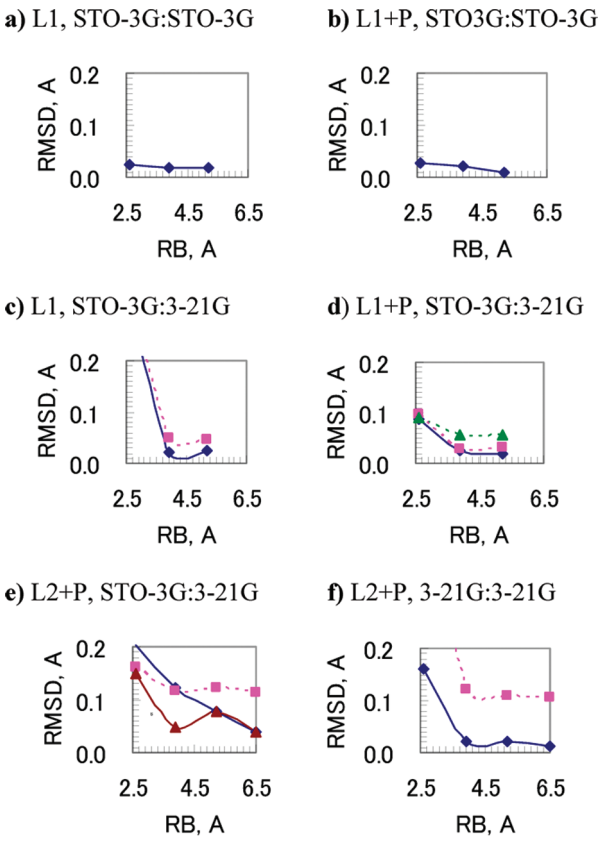


Figure 3. RMSD between full FMO and FMO/FD as a function of the size of the polarizable domain B (R_B), all in Å. L1 (a−d) and L2 (e−f) are ligands, neutral, and deprotonated *o*-phenolic acid, respectively. P denotes that a part of the protein is included in the geometry optimization. The basis sets used in domains F and B are listed as F:B. FD and FDD results are shown as solid (diamond) and dashed (square) lines, respectively. Dashed lines with triangles show FDD results for L1 + P′. Triangles with solid lines depict FD2 results.

was already a minimum. In general, we conclude that assuming that the active domain A does not get close to F (the criterion in eq 7), the effect of freezing the domain B is almost negligible, and therefore, the error observed for the charged ligand is mostly due to the use of a smaller basis set for domain F.

To verify this last supposition, we performed optimizations for L2 with F described by 3-21G. It can be expected that charged ligands require a basis set in F which properly describes its polarization. Indeed, the FD errors are reduced roughly by a factor of 3−5 (except for R_B = 2.6 Å) when F is described with 3-21G rather than STO-3G. It is instructive to observe that FDD errors, on the other hand, are reduced very little by using 3-21G in F. This means they have a different origin (in the neglect of the inactive dimers and their derivatives). On the other hand, the neutral L1 ligand had reasonable errors with the small basis set in F.

FMO/FD and FMO/FDD can be used not only to optimize structures but also to quickly compute the complex formation energy. At the optimized (L1 + P and L2 + P) geometry of the complex, these values (kcal/mol) for L1 were −38.0, −37.2 (−37.9), and −43.6 (−37.8); for L2, they were −117.0, −113.0 (−117.1), and −124.0 (−117.6) for regular FMO, FMO/FD and FMO/FDD, respectively, at their corresponding minima (the values in parentheses were obtained from single-point FMO energies). We note that the above binding energies in vacuum can be corrected for the desolvation effects.[34,35,46]

Table 1. RMSD (Å) between the Full and Frozen Domain FMO for the Complexes of Trp−Cage (1L2Y) and Its Ligands, L1 = HOPh-COOH (R_B = 3.9 Å) and L2 = $HOPhCOO^-$ (R_B = 6.5 Å)

ligand	protein[a]	basis set	level	RMSD
L1	none	STO-3G[b]	FD	0.017
L1	Pro-18	STO-3G[b]	FD	0.022
L1	none	3-21G[b]	FD	0.022
L1	none	3-21G[b]	FDD	0.050
L1	Pro-18	3-21G[b]	FD	0.024
L1	Pro-18	3-21G[b]	FDD	0.027
L1	Pro-18	3-21G[b]	FDD[c]	0.057
L2	Gln-5	3-21G[b]	FD	0.040
L2	Gln-5	3-21G[b]	FDD	0.115
L2	Gln-5	3-21G[d]	FD	0.012
L2	Gln-5	3-21G[d]	FDD	0.105

[a] The part of the protein whose geometry was optimized along with the ligand. [b] Shown for domain B; STO-3G for F. [c] With two extra fragments included in domain A, denoted by P′ in the main text. [d] Shown for domain B; 3-21G for F.

Next, we briefly mention the timings. For L1, the FMO, FMO/FD, and FMO/FDD (R_B = 3.9 Å) optimizations with 3-21G took 9.4, 2.8, and 1.9 h, respectively, on six dual CPU quad-core 2.83 GHz Xeon nodes (48 cores). The exact amount of the time savings depends strongly upon the system size, R_B, and the resultant size of domains A and B, and it is clearly not very meaningful to pay too much attention to the exact cost for each calculation presented here. The generation of the electron densities for buffer F takes a small time; therefore, the actual cost of FMO/FD optimizations only weakly depends on the system size and is mainly determined by the size of domain B (independent of the system size) and the number and flexibility of the optimized atoms in A. Therefore, FMO/FD can be efficiently applied to large-scale calculations at the cost of computing the polarizable buffer B in the Coulomb bath of F.

Nonsteroidal anti-inflammatory drugs block prostanoid biosynthesis by inhibiting prostaglandin H(2) synthase. As an example of a ligand optimization in a large system, we chose the complex of prostaglandin H2 synthase (COX-1) and the active enantiomer of ibuprofen (PDB: 1EQG).[47] The ibuprofen's carboxylate group forms hydrogen bonds with Arg-120 and Tyr-355, while its distal end lies in the hydrophobic cavity. The X-ray COX-1 structure is a dimer which consists of chains A and B. All crystallographic water and ligands except for ibuprofen in chain A were removed. Ibuprofen in chain A was solvated by a 25 Å water shell (567 water molecules; Figure 4). After solvation, the structure was protonated to pH 7 and optimized in CHARMM[48] program version C34B2 (CHARMM force field version 27, gradient RMS criterion of 0.1 kcal·mol^{-1}·Å^{-1}).

DOI: 10.1021/jz1016894 | J. Phys. Chem. Lett. 2011, 2, 282–288

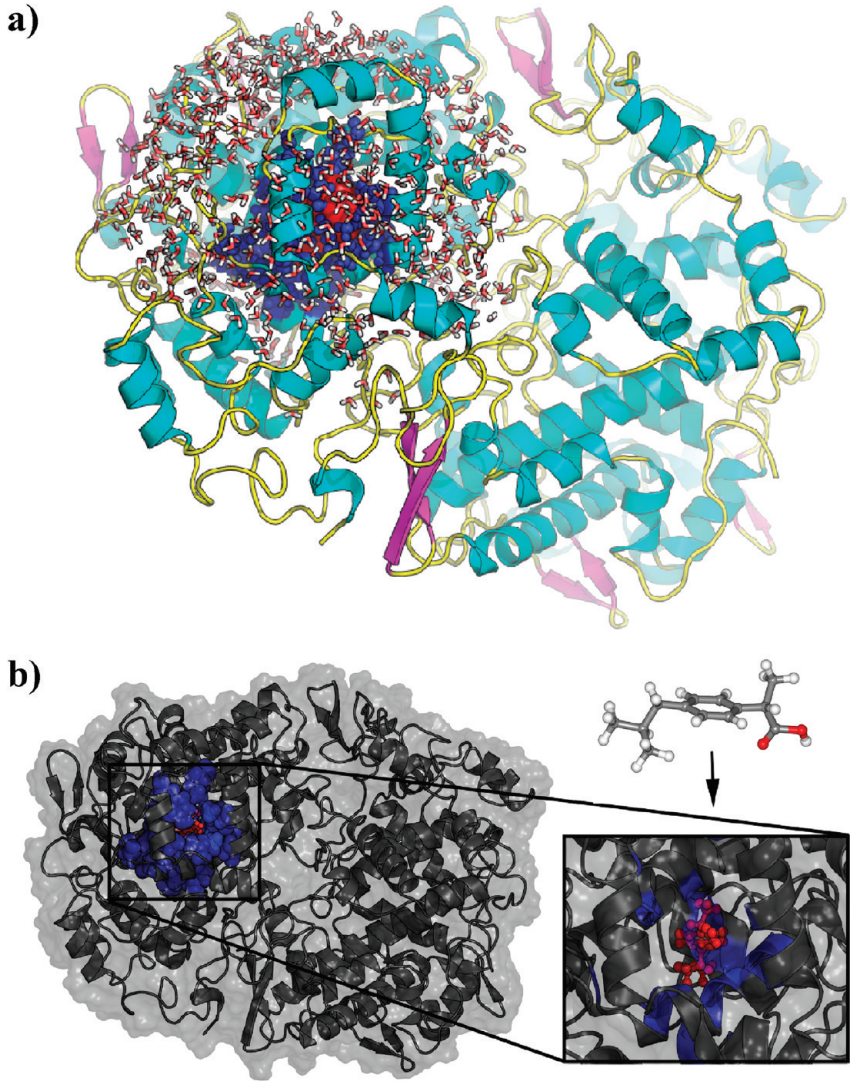


Figure 4. (a) Solvated structure of prostaglandin H(2) synthase-1 (COX-1) in complex with ibuprofen (PDB: 1EQG), optimized with FMO/FDD/RHF/STO-3G:B3LYP-D/6-31G*. (b) Domain definition: active (A), polarizable buffer (B), and frozen (F) domains are colored red, blue, and gray, respectively.

Next, we refined this structure with 19471 atoms by an FMO/FDD/RHF/STO-3G:B3LYP-D/6-31G* optimization (D indicates the dispersion correction[42]), where domain B was treated at the B3LYP-D/6-31G* level with the SG-1 grid, while domain F was described by RHF/STO-3G. Domain A included only the neutral ligand, all atoms of which were optimized to 2×10^{-4} au/bohr. In this calculation, we used R_B = 3.9 Å, which resulted in domain B having 25 fragments, including 19 protein residues, a ligand, and 5 water molecules). The FMO/FDD optimization of the system containing 19471 atoms took 31 steps and 32 h on six dual CPU quad-core 2.83 GHz Xeon nodes, of which 98 min were spent on domain F, and each step of buffer B took about 59 min.

FMO/FD and FMO/FDD have been implemented into GAMESS[49] and will be made available to the public in future releases. The input file generation for them is a minor addition to the regular multilayer FMO calculations, with the specification of domain A as a list of active fragments. The computational cost of the new methods is reasonably low so that practical geometry optimizations are possible using a modern single-node PC. We expect FMO/FD and FMO/FDD to become efficient tools for drug design and other studies. Because of the consistent energy and gradient, these methods can also be used for molecular dynamics, if only a part of the system is relaxed, for example, to evaluate the entropic contribution to the free energy of binding.[35]

AUTHOR INFORMATION

Corresponding Author:

*To whom correspondence should be addressed. E-mail: d.g.fedorov@aist.go.jp.

ACKNOWLEDGMENT The authors would like to thank Dr. Michael Schaefer for helpful discussions during the initial phase of this project. D.G.F. and K.K. were partially supported by the Next Generation SuperComputing Project, Nanoscience Program (MEXT, Japan). Y.A.

DOI: 10.1021/jz1016894 | J. Phys. Chem. Lett. 2011, 2, 282–288

was supported by a BBSRC competitive strategic grant to the Institute of Food Research (U.K.).

REFERENCES

(1) Scuseria, G. E. Linear Scaling Density Functional Calculations with Gaussian Orbitals. *J. Phys. Chem. A* **1999,** *103*, 4782–4790.
(2) Nikitina, E.; Sulimov, V.; Zayets, V.; Zaitseva, N. Semiempirical Calculations of Binding Enthalpy for Protein−Ligand Complexes. *Int. J. Quantum Chem.* **2004,** *97*, 747–763.
(3) Stewart, J. J. P. Application of the PM6 Method to Modeling Proteins. *J. Mol. Model.* **2009,** *15*, 765–805.
(4) Mitani, M.; Aoki, Y.; Imamura, A. Geometry Optimization of Polymers by the Elongation Method. *Int. J. Quantum Chem.* **1997,** *64*, 301–323.
(5) Gao, J. Toward a Molecular Orbital Derived Empirical Potential for Liquid Simulations. *J. Phys. Chem. B* **1997,** *101*, 657–663.
(6) Xiang, Y.; Zhang, D. W.; Zhang, J. Z. H. Fully Quantum Mechanical Energy Optimization for Protein−Ligand Structure. *J. Comput. Chem.* **2004,** *25*, 1431–1437.
(7) Ganesh, V.; Dongare, R. K.; Balanarayan, P.; Gadre, S. R. Molecular Tailoring Approach for Geometry Optimization of Large Molecules: Energy Evaluation and Parallelization Strategies. *J. Chem. Phys.* **2006,** *125*, 104109.
(8) Sorkin, A.; Dahlke, E. E.; Truhlar, D. G. Application of the Electrostatically Embedded Many-Body Expansion to Microsolvation of Ammonia in Water Clusters. *J. Chem. Theory Comput.* **2008,** *4*, 683–688.
(9) Hua, W. J.; Fang, T.; Li, W.; Yu, J. G.; Li, S. H. Geometry Optimizations and Vibrational Spectra of Large Molecules from a Generalized Energy-Based Fragmentation Approach. *J. Phys. Chem. A* **2008,** *112*, 10864–10872.
(10) Söderhjelm, P.; Ryde, U. How Accurate Can a Force Field Become? A Polarizable Multipole Model Combined with Fragment-wise Quantum-Mechanical Calculations. *J. Phys. Chem. A* **2009,** *113*, 617–627.
(11) Gordon, M. S.; Mullin, J. M.; Pruitt, S. R.; Roskop, L. B.; Slipchenko, L. V.; Boatz, J. A. Accurate Methods for Large Molecular Systems. *J. Phys. Chem. B* **2009,** *113*, 9646–9663.
(12) Mata, R. A.; Stoll, H.; Cabral, B. J. C. A Simple One-Body Approach to the Calculation of the First Electronic Absorption Band of Water. *J. Chem. Theory Comput.* **2009,** *5*, 1829–1837.
(13) Huang, L.; Massa, L.; Karle, I.; Karle, J. Calculation of strong and weak interactions in TDA1 and RangDP52 by the kernel energy method. *Proc. Natl. Acad. Sciences U.S.A.* **2009,** *106*, 3664–3669.
(14) Touma, T.; Kobayashi, M.; Nakai, H. Time-Dependent Hartree−Fock Frequency-Dependent Polarizability Calculation Applied to Divide-and-Conquer Electronic Structure Method. *Chem. Phys. Lett.* **2010,** *485*, 247–252.
(15) He, X.; Merz, K. M., Jr. Divide and Conquer Hartree−Fock Calculations on Proteins. *J. Chem. Theory Comput.* **2010,** *6*, 405–411.
(16) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. The Effective Fragment Potential Method: A QM-Based MM Approach to Modeling Environmental Effects in Chemistry. *J. Phys. Chem. A* **2001,** *105*, 293–307.
(17) Steinmann, C.; Fedorov, D. G.; Jensen, J. H. Effective Fragment Molecular Orbital Method: A Merger of the Effective Fragment Potential and Fragment Molecular Orbital Methods. *J. Phys. Chem. A* **2010,** *114*, 8705–8712.
(18) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM Implementation of the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) Method. *J. Phys. Chem. B* **2001,** *105*, 569–585.
(19) Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. Geometry Optimization with QM/MM, ONIOM, and Other Combined Methods. I. Microiterations and Constraints. *J. Comput. Chem.* **2003,** *24*, 760–769.
(20) Lu, Y.; Mei, Y.; Zhang, J. Z. H.; Zhang, D. Electron Polarization Critically Stabilizes the Mg^{2+} Complex in the Catalytic Core Domain of HIV-1 Integrase. *J. Chem. Phys.* **2010,** *132*, 131101.
(21) Mayhall, N. J.; Raghavachari, K.; Hratchian, H. P. ONIOM-Based QM:QM Electronic Embedding Method Using Löwdin Atomic Charges: Energies and Analytic Gradients. *J. Chem. Phys.* **2010,** *132*, 114107.
(22) *Combined Quantum Mechanical and Molecular Mechanical Methods*; Gao, J., Thompson, M. A., Eds.; ACS Symposium Series 712; American Chemical Society: Washington, DC, 1998.
(23) Hu, L.; Eliasson, J.; Heimdal, J.; Ryde, U. Do Quantum Mechanical Energies Calculated for Small Models of Protein-Active Sites Converge? *J. Phys. Chem. A* **2009,** *113*, 11793–11800.
(24) Canfield, P.; Dahlbom, M. G.; Hush, N. S.; Reimers, J. R. Density-Functional Geometry Optimization of the 150000-Atom Photosystem-I Trimer. *J. Chem. Phys.* **2006,** *124*, 024301.
(25) Ryde, U.; Greco, C.; De Gioia, L. Quantum Refinement of [FeFe] Hydrogenase Indicates a Dithiomethylamine Ligand *J. Am. Chem. Soc.* **2010,** *132*, 4512–4513.
(26) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment Molecular Orbital Method: An Approximate Computational Method for Large Molecules. *Chem. Phys. Lett.* **1999,** *313*, 701–706.
(27) Fedorov, D. G.; Ishida, T.; Kitaura, K. Multilayer Formulation of the Fragment Molecular Orbital Method (FMO). *J. Phys. Chem. A* **2005,** *109*, 2638–2646.
(28) Fedorov, D. G.; Kitaura, K. Extending the Power of Quantum Chemistry to Large Systems with the Fragment Molecular Orbital Method. *J. Phys. Chem. A* **2007,** *111*, 6904–6914.
(29) *The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems*. Fedorov, D. G., Kitaura, K., Eds.; CRC Press: Boca Raton, FL, 2009.
(30) Fedorov, D. G.; Ishida, T.; Uebayasi, M.; Kitaura, K. The Fragment Molecular Orbital Method for Geometry Optimizations of Polypeptides and Proteins. *J. Phys. Chem. A* **2007,** *111*, 2722–2732.
(31) Li, H.; Fedorov, D. G.; Nagata, T.; Kitaura, K.; Jensen, J. H.; Gordon, M. S. Energy Gradients in Combined Fragment Molecular Orbital and Polarizable Continuum Model (FMO/PCM) Calculation. *J. Comput. Chem.* **2010,** *31*, 778–790.
(32) Ishikawa, T.; Yamamoto, N.; Kuwata, K. Partial Energy Gradient Based on the Fragment Molecular Orbital Method: Application to Geometry Optimization. *Chem. Phys. Lett.* **2010,** *500*, 149–154.
(33) Komeiji, Y.; Mochizuki, Y.; Nakano, T.; Fedorov, D. G. Fragment Molecular Orbital-Based Molecular Dynamics (FMO-MD), A Quantum Simulation Tool for Large Molecular Systems. *J. Mol. Struct.: THEOCHEM* **2009,** *898*, 2–7.
(34) Nakanishi, I.; Fedorov, D. G.; Kitaura, K. Molecular Recognition Mechanism of FK506 Binding Protein: An All-Electron Fragment Molecular Orbital Study. *Proteins: Struct., Funct., Bioinf.* **2007,** *68*, 145–158.
(35) Sawada, T.; Fedorov, D. G.; Kitaura, K. Role of the Key Mutation in the Selective Binding of Avian and Human

DOI: 10.1021/jz1016894 | *J. Phys. Chem. Lett.* **2011,** *2*, 282–288

Influenza Hemagglutinin to Sialosides Revealed by Quantum-Mechanical Calculations. *J. Am. Chem. Soc.* **2010,** *132*, 1686–16872.
(36) He, X.; Fusti-Molnar, L.; Cui, G.; Merz, K. M., Jr. Importance of Dispersion and Electron Correlation in Ab Initio Protein Folding. *J. Phys. Chem. B* **2009,** *113*, 5290–5300.
(37) Mazanetz, M. P.; Ichihara, O.; Law, R. J.; Whittaker, M. Prediction of Cyclin-Dependent Kinase 2 Inhibitor Potency Using the Fragment Molecular Orbital Method. *J. Cheminf.* **2011,** *3*, 2.
(38) Fedorov, D. G.; Kitaura, K. The Importance of Three-Body TeRMS in the Fragment Molecular Orbital Method. *J. Chem. Phys.* **2004,** *120*, 6832–6840.
(39) Fujimura, K.; Sasabuchi, Y. The Role of Fluorine Atoms in a Fluorinated Prostaglandin Agonist. *ChemMedChem.* **2010,** *5*, 1254–1257.
(40) Nakano, T.; Kaminuma, T.; Sato, T.; Fukuzawa, K.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. Fragment Molecular Orbital Method: Use of Approximate Electrostatic Potential. *Chem. Phys. Lett.* **2002,** *351*, 475–480.
(41) Nagata, T.; Fedorov, D. G.; Kitaura, K. Importance of the Hybrid Orbital Operator Derivative Term for the Energy Gradient in the Fragment Molecular Orbital Method. *Chem. Phys. Lett.* **2010,** *492*, 302–308.
(42) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H−Pu. *J. Chem. Phys.* **2010,** *132*, 154104.
(43) Fedorov, D. G.; Olson, R. M.; Kitaura, K.; Gordon, M. S.; Koseki, S. A New Hierarchical Parallelization Scheme: Generalized Distributed Data Interface (GDDI), and an Application to the Fragment Molecular Orbital Method (FMO). *J. Comput. Chem.* **2004,** *25*, 872–880.
(44) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. Designing a 20-Residue Protein. *Nat. Struct. Biol.* **2002,** *9*, 425–430.
(45) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, And Multithreading. *J. Comput. Chem.* **2010,** *31*, 455–461.
(46) Murata, K.; Fedorov, D. G.; Nakanishi, I.; Kitaura, K. Cluster Hydration Model for Binding Energy Calculations of Protein−Ligand Complexes. *J. Phys. Chem. B* **2009,** *113*, 809–817.
(47) Selinsky, B. S.; Gupta, K.; Sharkey, C. T.; Loll, P. J. Structural Analysis of NSAID Binding by Prostaglandin H2 Synthase: Time-Dependent and Time-Independent Inhibitors Elicit Identical Enzyme Conformations. *Biochemistry* **2001,** *40*, 5172–5180.
(48) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, And Dynamics Calculations. *J. Comput. Chem.* **1983,** *4*, 187–217.
(49) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; et al. General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.* **1993,** *14*, 1347–1363.

DOI: 10.1021/jz1016894 | *J. Phys. Chem. Lett.* **2011,** *2*, 282–288