Pure Appl. Chem., Vol. 75, Nos. 11–12, pp. 2405–2410, 2003.

Workshop 1.5

Fragment molecular orbital study of the binding energy of ligands to the estrogen receptor*

Kaori Fukuzawa[1], Kazuo Kitaura[2], Kotoko Nakata[3], Tsuguchika Kaminuma[4], and Tatsuya Nakano[3,‡]

[1]Fuji Research Institute Corporation, 2-3 Kanda Nishiki-cho, Chiyoda-ku, Tokyo 101-8443, Japan; [2]National Institute of Advanced Industrial Science and Technology, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan; [3]National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan; [4]Chem-Bio Informatics Society, 4-3-16 Yoga #301, Setagaya-ku, Tokyo 158-0097, Japan

Abstract: We examined the published data for the binding affinity of typical ligands to the α-subtype of the human estrogen receptor with use of an approximate molecular orbital method applicable to interacting molecular clusters. An ab initio procedure for "molecular fragments" proposed recently to deal with such macromolecules as proteins was applied to the molecular orbital calculations. The receptor protein was primarily modeled using 50 amino acid residues surrounding the ligand. For a few ligand-receptor complexes, the binding energy was also calculated with use of 241 amino acid residues contained in the entire binding domain. No significant difference was found in the calculated binding energy between the complex modeled with ligand-surrounding 50 amino acids and that with residues of the entire domain. The calculated binding energy was correlated very well with the published relative binding affinity for typical ligands.

INTRODUCTION

The effect of estrogenic ligands is induced by their binding to the estrogen receptors (ERs) [1–3]. Since a variety of unknown compounds could bind to the ligand-binding domain (LBD) of the ER and exert hormone-like effects on human and wildlife health, the ER is an important research target for the development of therapeutic agents [3,4] as well as the screening of endocrine disruptors [5]. A number of experimental and theoretical efforts have been carried out for the mechanism of the interaction of ligands with the ER LBD. Most of the theoretical works, however, have stood on empirical force field approximations [6–8]. Although they are suited for calculating macromolecules in terms of the computational time, empirical approaches may not be accurate enough theoretically. Hoping to establish a time-saving and versatile computational procedure for biomacromolecules, we recently proposed the fragment molecular orbital (FMO) method [9]. Here, we report the result of our FMO study for the interaction of ligands with the α-subtype of ER carried out to elucidate its submolecular mechanism theoretically and accurately.

*Report from a SCOPE/IUPAC project: Implication of Endocrine Active Substances for Human and Wildlife (J. Miyamoto and J. Burger, editors). Other reports are published in this issue, *Pure Appl. Chem.* **75**, 1617–2615 (2003).

‡Corresponding author

METHODS

In the FMO method [9], a single molecule or a molecular cluster (a group of molecules interacting to each other noncovalently) is dealt with after being divided into fragments to which electron pairs are assigned according to certain rules. The molecular orbitals (MOs) for fragments and fragment pairs (combinations of two fragments) are calculated under conditions under which the orbitals are forced to localize as the closed shell within the corresponding region. For fragments to which no electron pair is allocated from the bond when detached in the fragmentation, the MO is built from usual atomic basis functions of the constituent atoms according to the conventional linear combination of atomic orbitals to yield molecular orbitals (LCAO-MO) framework. For fragments in which bonding electron pair is left, the atomic valence basis function of the partner atom, with which the fragment is connected originally, is used additionally in the LCAO-MO model. The initial calculation for each fragment MO yields the initial electron density distribution.

The Hamiltonian for each fragment is composed to include the terms for the electrostatic potential governed by electrons in the surrounding fragments and all nuclei in the molecule. Since the electrostatic potential of each fragment depends on the electron distribution of surrounding fragments, the electron density distribution of each fragment is calculated first using the initial electron distribution calculated in a manner described above. A set of "Schrödinger" equations for every fragment with the initial electron density is solved iteratively until the electron density distribution for all fragments converges self-consistently. Likewise, the Hamiltonian of each fragment pair has the terms for the potential arising from electrons in the surrounding fragments and the terms from every nuclear charge in the molecule. The set of equations for fragment pairs is solved using the electrostatic potential from the converged electron density distribution of the surrounding fragments. The potential energy of fragments and fragment pairs at the HF/STO-3G level is calculated to estimate the energy of the total system.

The ligand molecules examined here are shown in Fig. 1. The coordinates of heavy atoms in the ER complex of EST, RAL, DES, and OHT were fixed as being equivalent to those of the PDB files, entries 3ERE, 1ERR, 3ERD, and 3ERT, of the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB), respectively [10–12]. For ligands such as ESTA, GEN, TAM, BISA, BISF, CLO, and OHC, the PDB files for the ERα complex are not available. Thus, the binding geometry of the first two ligands was approximated first by superimposing the "phenoxy" substructure of the phenol moiety on that of EST in the 3ERE, while that of the others was by superimposing their "phenoxy" substructure or corresponding phenyl group on that of OHT in the 3ERT file. Then, the geometry of GEN was approximated by that in the ERβ-GEN complex taken from the PDB 1QKM file. TAM, CLO, and OHC were modeled with the Insight II system [13] based on the geometry of OHT, and the others were optimized using the HF/6-31G(d) method. The geometry of hydrogen atoms was modeled with the Insight II system [13] and the CHARMm force field calculations [14].

Hydrogen bonds, occurring between the ligand and surrounding residues directly as well as through the mediation of a single water molecule, have been shown to stabilize the ER ligand binding [15]. In this study, the most stable geometry of the hydrogen bond network was calculated at the HF/6-31G(d) level [16] with use of a model molecular cluster consisting of such hydrogen-bonding residues in the LBD as Glu 353, Leu 387, Arg 394, and His 524, each of the ligands and the single water molecule (Model 3).

The entire LBD of the receptor protein containing 241 amino acid residues (Model 1) was used for the calculation only for some ligands. The binding domain was, however, primarily modeled with use of 50 amino acid residues "directly" surrounding the ligand (Model 2) as displayed in Fig. 2. To make the fragmentation of the receptor protein, the peptide chain was divided at the Cα atom into blocks of every two residues in a manner as shown in Fig. 3. The ligand as well as the hydrogen-bonding water molecule was treated as a single fragment.

All the FMO calculations were carried out with an FMO program package, ABINIT-MP [17], mostly on dual Pentium III 1-GHz clusters equipped with 32 processor units. The time required for cal-

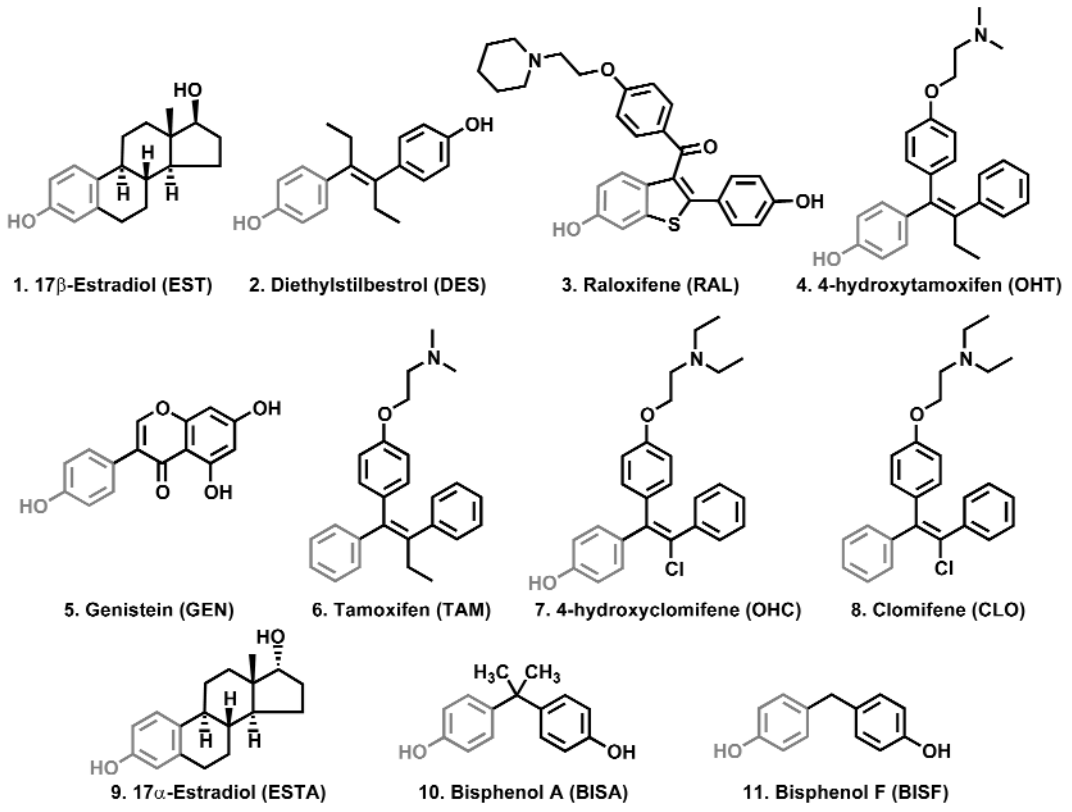

Fig. 1 Ligands used for the calculation of the binding energy. Light black substructures represent the moiety to be superimposed with the corresponding moiety in reference compounds.

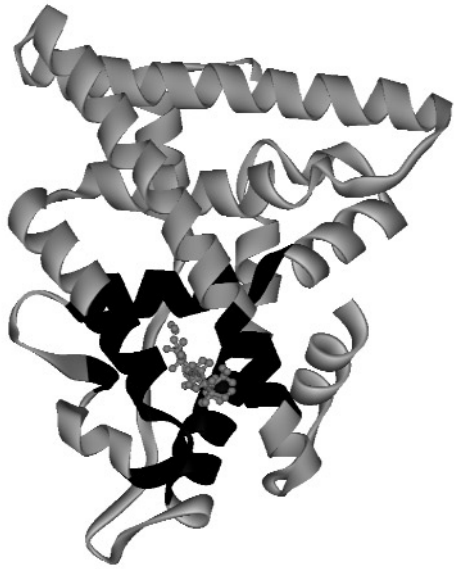

Fig. 2 The ribbon display of the ERα LBD complexed with 17β-estradiol (1, EST). Model 1 including 241 residues is shown as the entire picture. Fifty residues surrounding "directly" the ligand for Model 2 are dark-colored. The ligand and the water molecule are displayed inside the matrix using ball and stick.

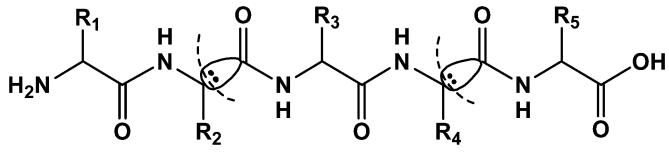

Fig. 3 Fragmentation of peptides indicated as broken arcs.

culating entire ERα LBD containing 241 residues with ca. 4000 atoms was about 14 h. The accuracy of the FMO method has been examined using crambin, a protein series with 46 residues [9]. The ab initio total energy values calculated at the HF/STO-3G level for [Pro^{22}, Leu^{25}]crambin with and without the FMO approximation are –17779.5030 and –17779.5024 a.u., respectively, corresponding to a difference below 0.5 kcal/mol. The computational time is "drastically" reduced with the FMO procedure compared to that without the FMO approximation.

RESULTS AND DISCUSSION

The energy of each of the three systems, i.e., the receptor, $E_{receptor}$, ligand E_{ligand}, and the ER ligand complex, $E_{complex}$, can be calculated from the sum of energy values of fragments and the counterpart for fragment pairs within each system under certain conditions [9]. In the calculation of the $E_{receptor}$ value, the hydrogen-bonding water molecule was included as a fragment along with "dipeptide" fragments. The binding energy for a given ligand (ΔE_{ligand}) can be expressed in eq. 1 as the difference in the energy between complex and components.

$$\Delta E_{ligand} = E_{complex} - (E_{receptor} + E_{ligand}) \quad (1)$$

The binding energy relative to that of 17β-estradiol (EST), $\Delta\Delta E_{ligand}$, in eq. 2 is the value to be compared with the experimental relative binding affinity (RBA) value. The RBA value of 17β-estradiol is defined as 100.

$$\Delta\Delta E_{ligand} = -(\Delta E_{ligand} - \Delta E_{EST}) \quad (2)$$

The $\Delta\Delta E_{ligand}$ values estimated using Model 2 are plotted against the published values of log (RBA/100) in Fig. 4.

The ligands 1–6, 9, and 10, of which the experimental RBA value is known, are shown as a circle in Fig. 4. For these 8 compounds, the correlation between $\Delta\Delta E$ and log (RBA/100) seems to be promising, the correlation coefficient r being 0.837. In particular, there is a very good correlation (r = 0.931) for the 7 ligands omitting TAM (6). From the correlation equation (n = 8), the log (RBA/100) value of ligands 7, 8, and 11, of which the RBA value is unknown, can be estimated with use of the calculated $\Delta\Delta E$ value. These 3 compounds are shown as a square in the plot.

The $\Delta\Delta E$ value was also calculated according to Model 1 for the complex of ligands 1~4. The result was almost identical with that calculated with Model 2. The difference in the $\Delta\Delta E$ value between two models was mostly below 3 kcal/mol, suggesting that the binding between ER and ligand is local. Another interesting finding was a difference in the charge distribution between complexed and individual component molecules. The total charge of ligands was changed to be negative with the values –0.00 ~ –0.18 when complexed with ER. The greatest negative charge influx occurs from Glu 353 to ligands, and a slight efflux is observed into Arg 394 and His 524. Such charge transfer is highly related with the binding energy. In fact, the ΔE tends to be greater with the increase in the difference of the charge distribution. Thus, most of the stabilization in the ER–ligand docking arises from the ligand-Glu 353 interaction. This observation seems to indicate that the charge is variable in the ER–ligand interaction, and therefore atomic charges should be calculated dynamically instead of using fixed charges as in classical calculations.

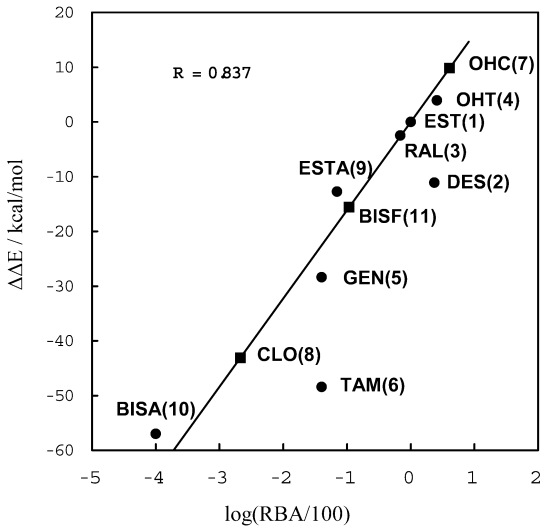

Fig. 4 Relationship between calculated relative binding energy ($\Delta\Delta E$) and experimental relative binding affinity [log (RBA/100)] of eight ligands (●), and the estimation of log (RBA/100) for three ligands (■). The regression line is drawn so that it is forced to pass the origin of coordinates.

To summarize, we have applied the ab initio FMO method to ER ligand binding which allows us to accurately predict the relative binding energy of xenoestrogenic ligand molecules from a "single" energy calculation. Given a variety of compounds, some of which could bind to the ER, such methods as we have proposed may provide a powerful tool for assessing the affinity of putative xenoestrogens in silico prior to biological studies. For further improvements, it is necessary to optimize not only the hydrogen bond, but also the geometry of the ligand and surrounding residues to estimate possible effects, in particular, those according to induced-fit in the ER ligand binding. Such functions are under development in our group.

REFERENCES

1. K. Paech, P. Webb, G. G. J. M Kuiper, S. Nilsson, J.-Å. Gustafsson, P. J. Kushner, T. S. Scanlan. *Science* **277**, 1508–1510 (1997).
2. G. G. J. M. Kuiper, J. G. Lemmen, B. Carlsson, J. C. Corton, S. H. Safe, P. T. van der Saag, B. van der Burg, J.-Å. Gustafsson. *Endocrinology* **139**, 4252–4263 (1998).
3. T. Barkhem, B. Carlsson, Y. Nilsson, E. Enmark, J.-Å. Gustafsson S. Nilsson. *Mol. Pharmacol.* **54**, 105–112 (1998).
4. S. Nilsson, G. Kuiper, J.-Å. Gustafsson. *Trends Endocrinol. Metab.* **9**, 387–395 (1998).
5. C. Sonnenschein and A. M. J. Soto. *J. Steroid Biochem. Molec. Biol.* **65**, 143–150 (1998).
6. S. P. Bradbury, O. G. Mekenyan, G. T. Ankley. *Environ. Toxicol. Chem.* **17**, 15–25 (1998).
7. B. C. Oostenbrink, J. W. Pitera, M. M. H. van Lipzig, J. H. N. Meerman, W. F. van Gunsteren. *J. Med. Chem.* **43**, 4594–4605 (2000).
8. P. D. Kirchhoff, R. Brown, S. Kahn, M. Waldman, C. M. Venkatachalam. *J. Comput. Chem.* **22**, 993–1003 (2001).
9. T. Nakano, T. Kaminuma, T. Sato, K. Fukuzawa, Y. Akiyama, M. Uebayasi, K. Kitaura. *Chem. Phys. Lett.* **351**, 475–480 (2002).
10. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne. *Nucleic Acids Res.* **28**, 235–242 (2000); <http://www.rcsb.org/pdb/>.

11. A. M. Brzozowski, A. C. W. Pike, Z. Dauter, R. E. Hubbard, T. Bonn, O. Engström, L. Öhman, G. L. Greene, J.-Å. Gustafsson, M. Carlquist. *Nature* **389**, 753–758 (1997).
12. A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard, G. L. Greene. *Cell* **95**, 927–937 (1998).
13. InsightII Version 98.0, Molecular Simulations Inc., San Diego, CA (1998).
14. CHARMm, Version 25.2, Revision: 98.0731.
15. D. M. Tanenbaum, Y. Wang, S. P. Williams, P. B. Sigler. *Proc. Nat. Acad. Sci. USA* **95**, 5998–6003 (1998).
16. Gaussian 98, Revision A.7, M. J. Frisch et.al., Gaussian, Inc., Pittsburgh PA (1998).
17. ABINIT-MP: <http://moldb.nihs.go.jp/abinitmp/>.