Letter


Manabu SUGIMOTO a,b,c* and Takafumi INOUE b

a Faculty of Advanced Science and Technology, Kumamoto University, 2-39-1 Kurokami, Chuo-ku, Kumamoto 860-8555, Japan
b Department of Applied Chemistry and Biochemistry, Faculty of Engineering, Kumamoto University, 2-39-1 Kurokami, Chuo-ku, Kumamoto 860-8555, Japan
c Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan
* e-mail: sugimoto@kumamoto-u.ac.jp

(Received: November 28, 2016; Accepted for publication: December 26, 2016; Online publication: February 25, 2017)

A numerical model, called the interaction energy projection method (IEPM), is suggested to evaluate electronic similarity among protein-ligand complexes. Herein we apply the method by referring to the "inter-fragment interaction energy (IFIE)," calculated using the fragment molecular orbital (FMO) method, in two human estrogen receptor complexes.

Keywords: Electronic similarity, Protein-ligand docking, Fragment molecular orbital method, Inter-fragment interaction energy, Estrogen receptor, Interaction energy projection method

1 Introduction

Search for functional molecules is one of the most important and challenging issues in applied chemistry. It is also true in pharmaceutical science which demands the discovery of molecules bound to target proteins [1]. For systematic and efficient molecular search, development of quantitative measures of protein-ligand complexes seems demanded.

In this paper, we propose a method describing the amino acid residue (AAR)–ligand interaction, which is called "Interaction Energy Projection Method (IEPM)." By using this method, we can quantitatively compare similarity/dissimilarity among protein-ligand complexes. In the present implementation, we refer to the "inter-fragment interaction energy (IFIE)" between AAR-ligand pairs calculated using the fragment molecular orbital (FMO) method developed by Kitaura et al. [2]. Therefore, the similarity evaluated in this paper can be called "electronic similarity" reflecting the electronic interaction in the complexes.

Herein we use the IFIE data at the MP2/6-31G* level recently calculated by Anzaki et al. on the human estrogen receptor α with the Y537S mutation (hERα/Y537S in Figure 1(a)) [3]. We evaluate the electronic similarity between two complexes of hERα, which are called 2QA6 [4] and 3UU7 [5]. Their ligands (KN2 and 2OH ligands, respectively) are shown in Figure 1.

2 3D Description of Protein-Ligand Interactions

The IEPM method is decomposed into three steps: (A) three-dimensional description of the reference and target complexes, (B) orientation matching, and (C) similarity measurement. The atomic coordinates of the target protein-ligand complexes are taken from the Protein Data Bank [4,5].
The algorithm in Step (A) mentioned above is as follows:
(A1) the origin of the coordinate is taken as the geometric center of the ligand in the ER-ligand complex.
(A2) The IFIE is projected onto a sphere with a radius of R (5000 Å in the present study).
(A3) The position of each AAR is represented using the polar coordinate \((r, \theta, \phi)\) of its \(C_\alpha\) atom.
(A4) A Gaussian function in the \((\theta, \phi)\) plane is defined to represent the AAR-ligand interaction. Overlap of the Gaussian functions defined to all the ligand-amino-acid residue pairs gives the projection sphere. We call this sphere "projected interaction pattern (PIP)." The resultant function \(P = P(\theta, \phi)\) is given as
\[
P = \sum_{\text{all AAR}} E_{LI} \exp\left[-\alpha \left(\theta-\theta_j\right)^2 + \left(\phi-\phi_j\right)^2\right]
\]
where \(E_{LI}\) is the IFIE value for the ligand L and the summation is taken over all AARs. \(\alpha\) is defined so that the half-width of the Gaussian is equal to 0.01r_{la} where r_{la} is the AAR-ligand distance. The topologies of the PIPs for 2QA6 and 3UU7 are shown in Figure 2 (a).

3 Orientation Adjustment for Similarity Analysis

In order to have maximum overlap between two PIPs, we adjust their mutual orientations using the following algorithm:
(B1) Initially, the \(10^4\) sampling points are randomly generated on the mapped surface.
(B2) The number of sampling points near a peak is determined to reflect its height: when the IFIE value is \(x\), we determine the number \(n\) as ceiling\(|x|\) (Iverson's ceiling function for the absolute value of \(x\)). The \(n\) points near the peak axis in the \((\theta, \phi)\) space are chosen among the \(10^4\) points. The points selected in this process are called "geometric-feature points" (GFPs).
(B3) Rotation of the target complex is carried out by minimizing the averaged distance between the pairs of the GFPs of the reference and target complexes shown in Figure 2. In this process one GFP of the target is chosen as the nearest neighbor of each GFP of the reference complex. It is noted that translational matching is not necessary because the origin of the coordinate is taken as the geometric center of the ligand.

We show the overlap of GFPs of 2QA6 and 3UU7 after orientation adjustment in Figure 2 (d). The reoriented PIP of
3UU7 is shown in Figure 2 (e). These two figures indicate that the present procedure is reasonably good.

4 Electronic Similarity

The electronic similarity among the protein-ligand complexes is evaluated using the following algorithm:

(C1) The IFIE value is plotted with respect to θ and ϕ where the reoriented PIP is referred to. This process generates figures as shown in Figure 3.

(C2) The electronic similarity S is calculated by using the following integral:

\[
S = \frac{1}{\int P_r^2 \cos \theta d\theta \cos \phi d\phi} \int P_t^2 \cos \theta d\theta \cos \phi d\phi
\]

where the denominator is to normalize both P_r and P_t of the reference and target complexes, respectively.

The calculated value of S between 2QA6 and 3UU7 was 63.0% in the present method. This is not so high although the complexes are known to exist experimentally. The unexpectedly small value may be due to the fact that, although the AAR giving the most intense peak in 2QA6 and 3UU7 is common, the IFIE in 3UU7 is about 1/2 of that in 2QA6. This implies that, although a sufficiently strong binding site is necessary, the multi-site interaction should also play an important role.

5 Summary and Conclusions

We have developed a numerical method called the interaction energy projection method (IEPM) in order to compactly describe electronic AAR-ligand interactions in protein-ligand complexes. The present method facilitates numerical evaluation of electronic similarity in the complexes. It is expected that important patterns in the complexes will be found through systematic applications of the present method.

A part of this work was financially supported by MEXT KAKENHI Grant Number 26102015. This research was also done in activities of the FMO drug design consortium (FMODD). We are grateful to Prof. K. Fukuzawa, Prof. S. Tanaka, and Mr. S. Anzaki for providing the IFIE data used in this paper.

References


Figure 3. Mapping of the IFIE data onto the (θ, ϕ) plane: (a) 2QA6 and (b) 3UU7.