Random matrix theory for an inter-fragment interaction energy matrix in fragment molecular orbital method

Masanori Yamanaka

Department of Physics, College of Science and Technology, Nihon University
1-8 Kanda-Surugadai, Chiyoda-ku, Tokyo 101-8308, Japan
*E-mail: yamanaka@phys.cst.nihon-u.ac.jp

(Received June 29, 2015; accepted January 22; published online November 22, 2018)

Abstract

The statistical properties of the inter-fragment interaction energy matrix of the fragment molecular orbital method are analyzed using the random matrix theory. The eigenvalue and eigenvector distributions, the inverse participation ratio, and the unfolded eigenvalue statistics are compared with the corresponding random matrix ensemble. Cluster analysis of the fragments with strong correlations is presented using the inverse participation ratio of the random matrix theory.

Key Words: fragment molecular orbital, random matrix, principal component analysis

Area of Interest: Molecular recognition and molecular modeling

1. Introduction

The random matrix theory (RMT) was introduced in physics in the 1950s to describe nuclear energy spectra. Because the first principal calculation of the spectrum of such a complex interacting many-body system is difficult, we attempted to consider an ensemble of the Hamiltonian, which is the mathematical representation of the hermitian matrix, with random components instead of a single Hamiltonian with definite components [1–5]. Many findings and applications have been established, both in Mathematics and Physics, which correspond to RMT. A prominent application is the nonperturbative description of strings and D-branes, which is known as the matrix model [6].

Recently, RMT has been widely applied to multivariate techniques. One example is in principal component analysis (PCA) [7–10], which is a multivariate statistical technique that decomposes a data set into several inter-correlated quantitatively dependent variables called principal components. In the past two decades, RMT has been applied to time-series analysis, particularly in the field of econophysics [11–13]. In particular, if the cross-correlation matrix of stocks is regarded as a random matrix, then the matrix can be described by a corresponding ensemble of RMT [1–4]. Several studies confirmed the assumption that correlation among stocks would be detected as a
deviation from the random matrix [11–15]. One of the advantages of RMT is that it provides criteria to distinguish eigenvectors whose components contain correlated elements from those whose components do not contain correlated elements. This has shed some light on the problem, unsolved for a long time, of the number of eigenvectors deemed to be meaningful in PCA where only some of the eigenvectors (i.e., the one with the largest eigenvalues) have been typically analyzed as the principal component. The application of RMT has been adopted in other areas as well [16]. For example, in biophysics, RMT was used to analyze the NMR spectra of biomolecule [17, 18], to classify RNA [10–21], to investigate the optical properties of proteins [22, 23], to extract a specific network topology in proteome [24, 25], to describe the displacement covariance matrix of the elastic network model [26], and to perform sequence analysis of protein residues [27, 28]; moreover, RMT was used in molecular dynamics [29–31].

The fragment molecular orbital (FMO) method [33–36] is an *ab initio* computational method that can be used to determine the electronic states of a complex molecule, particularly a huge biomolecule, very efficiently. In this method, structurally complex molecules are divided into small fragments and *ab initio* or density functional quantum-mechanical calculations are performed for each fragment and fragment pair. During the calculation, a quantity of the pair interaction energies is obtained - also called as the inter-fragment interaction energies (IFIEs) - that is associated with the two fragments and interpreted as a two-body interaction between them. The IFIEs contain useful information on intra- and inter-molecular interactions, which we cannot be obtained by using any other method. Several methods have been proposed to analyze the pair interactions in detail. Configuration analysis for fragment interaction (CAFI) allowed extracting the stable components of the polarization and charge transfer for IFIEs [37]. In the pair interaction energy decomposition analysis (PIEDA), pair interactions were decomposed into the electrostatic, exchange-repulsion, charge transfer, and dispersion components [38]. A fragment interaction analysis based on local MP2 (FILM) was implemented to extract the electron correlation component of the interaction energy in each orbital [39]. A theoretical scheme to evaluate effective, screened interactions between fragments was proposed [40]. Recently a new attempt was made for the cluster analysis of IFIE matrix [41], where advanced clustering analyses based on a self-organizing map and multi-dimensional scaling were carried out, and a novel method for virtual ligand screening to explore drug candidates binding to target proteins was proposed. The interaction of protein complexes was extensively investigated [41–89]; in particular, numerous important findings were obtained by the FMO method on the influenza virus [50, 55, 58, 68, 69, 73, 75, 76, 79, 85]. The matrix representation of IFIE values is called the IFIE matrix, and an IFIE map was introduced for the comprehensive analysis of the entire set of IFIEs [53], in order to extract the information regarding the interaction among fragments through the colored map. The visualized cluster analysis of the protein-ligand interaction (VISCANA) was also introduced for the virtual ligand screening based on the FMO [46]. The VISCANA classifies the structurally similar ligands in terms of the interaction patterns of a ligand with amino acids. Because the IFIE matrix is a type of a cross-correlation matrix, investigating it by the RMT is quite intriguing.

Herein, we study the inter-fragment interaction energy matrix of the FMO method by the random matrix theory. We mainly present the statistical properties of the eigenvalues and eigenvectors and the inverse participation ratio (IPR). The probability density of the eigenvalues is not agreement with Wigner's semicircle law [90]. There are four eigenvalues located outside the Wigner's semicircle distribution. The nearest neighbor and the next-to-nearest neighbor unfolded eigenvalue statistics that are consistent with those of the Gaussian orthogonal and symplectic ensembles, respectively. All IPRs are above the RMT values; this implies that the correlation sectors of fragments are rather localized. The distribution of the components of the eigenvector with the largest eigenvalue is a unique feature. It is a binary distribution; this means that the
fragments belonging to the corresponding sector interact with the same weight. This may have profound implications for the interaction between the protein and DNA. Thus, we propose a cluster analysis based on IPR and RMT.

2. Random matrix

For an \( N \times N \) real symmetric random matrix, we denote the eigenvalues \( \lambda_k \), where \( k = 1, 2, \ldots, N \) and set \( \lambda_1 > \lambda_2 > \ldots > \lambda_N \), and also denote the i-th component of the k-th eigenvector associated with the k-th eigenvalue \( \lambda_k \) by \( u_{ki} \), where \( i = 1, 2, \ldots, N \). In the limit \( N \to \infty \), the probability density, \( \rho \), of the eigenvalues is analytically given [90] by

\[
\rho(\lambda) = \frac{1}{2\pi\sigma} \sqrt{4\sigma^2 - \lambda^2},
\]

where \( \sigma \) is the variance of the matrix elements. Equation (1) is known as Wigner's semicircle law. The probability density is bounded by

\[-2 \sqrt{\sigma} < \lambda < 2 \sqrt{\sigma}.\]

Some of the universal quantities in the RMT are the nearest- and the next-to-nearest-neighbor spacings, \( P_{nn}(s) \) and \( P_{nnn}(s) \), respectively.

\[
P_{nn}(s) = \frac{\pi s}{2} \exp\left(\frac{-\pi s^2}{4}\right),
\]

\[
P_{nnn}(s) = \frac{2^{18} s^4}{3^6 \pi^3} \exp\left(-\frac{64}{9\pi} s^2\right),
\]

where a spacing \( s \) is defined by \( s = \xi_{i+1} - \xi_i \) and \( \xi_i \) is the unfolded eigenvalue [91,92]. In the limit \( N \to \infty \), the components of the eigenvectors \( u_{ki} \) are described by the Gaussian distribution with a mean of zero and unit variance [91,92] as follows:

\[
u_k(u) = \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{u^2}{2}\right).
\]

The IPR is a useful quantity in characterizing the distribution of the components [5,93–100], which is defined for each eigenvector, \( u_k \), by

\[
I_k = \sum_{i=1}^{N} u_{ki}^4.
\]

IPR measures the reciprocal of the number of eigenvector components that contribute significantly. A delocalized eigenvector with identical components \( u_{ki} = 1/\sqrt{N} \) has \( I_k = 1/N \), and a completely localized eigenvector with a single nonzero component has \( u_{ki} = 1 \) for particular \( i \), and the remaining zeros have \( I_k = 1 \). For the Gaussian orthogonal ensemble, the average is \( <I_k> = 3/N \). Therefore, a
deviation from $3/N$ implies a correlation among the fragments. In particular, a deviation greater than $3/N$ implies a localized correlation and that less than $3/N$ implies a global correlation. In RMT, the eigenstate whose eigenvalue deviates from the RMT value (1), i.e., goes outside of the RMT boundary (2), is considered as a correlated sector. The correlation magnitude is evaluated from the deviation magnitude. The eigenvectors with especially small eigenvalues have large IPR values in most cases, implying a strong correlation between the fragments inside the corresponding sector. This is in contrast with PCA, where only some of their largest eigenvalues and the corresponding eigenvectors are typically analyzed.

To analyze the correlated clusters in more detail, we study IPRs further, using the method of Utsugi et al. [15], where the z-index is defined in each fragment as follows:

$$z_i(\delta) = \sum_{k: I_k \geq \delta} u_{ki}^2,$$

(7)

where $i$ represents the fragment number and $\delta$ is the threshold for IPR when extracting the eigenstates which deviate from RMT prediction. The summation in Eq. (7) is over the components that satisfy the IPR condition $I_k \geq \delta$. The i-th fragment is regarded as significantly correlated if $z_i$ is larger than a certain threshold $\Delta$.

In FMO, the RMT applied as follows. If all fragments are correlated at random, the IFIE matrix is a random matrix. Because IFIE matrix is a real symmetric matrix, it becomes a real symmetric random matrix. The universality class of the real symmetric random matrix is the Gaussian orthogonal ensemble. Therefore, the eigenvalue distribution of the random IFIE matrix follows (1); the eigenvalues are bounded by (2); the nearest- and next-to-nearest-neighbor eigenvalue spacings, are (3) and (4), respectively, after unfolding the eigenvalue; the distribution of the eigenvalue components follows (5); and the IPR is $< I_k > = 3/N$. Let us assume that there is a single cluster composed of correlated fragments. In that case, the components of the correlated fragment cluster are embedded in a random IFIE matrix. The eigenvalue distribution of the IFIE matrix slightly deviates from (1), where one of the eigenvalues is usually pushed outside of the upper limit of (2); the IPR of the corresponding eigenvalue deviates from $3/N$, and the eigenvector components that correspond to the correlated fragment shift from (5), where it is usually the largest or the second largest eigenvalue. The nearest- and next-to-nearest-neighbor eigenvalue spacings, (3) and (4), are not subject to change. When there are several clusters of correlated fragments, each fragment interacts with the other in the same cluster. In that case, the components of the correlated fragment cluster are also embedded in a random IFIE matrix. Some of the eigenvalues are usually pushed outside of the upper limit of (2), and the eigenvalue distribution of the IFIE matrix slightly deviates from (1); sometimes, some of the eigenvalues are also pushed outside of the lower limit of (2) to satisfy the sum rule of the eigenvalues. The IPR of the corresponding eigenvalues deviates from $3/N$. The eigenvector components that correspond to the correlated fragment shift from (5) indicate a different eigenvector for each correlated cluster. Therefore, among the eigenvalues obtained by diagonalizing the IFIE matrix, each eigenvalue that deviates from (2) is a cluster of fragments. In addition, examining the ingredients of the eigenvector, the components whose distribution deviates from (5) are fragments constituting the cluster. This is in principle the same as PCA; however RMT utilizes the quantitative criterions, (1), (2), and (5) and sometimes (3) and (4) as judgment conditions for the procedure. The RMT is a taxonomy of the random matrix, and a statistical property is analytically derived for every category of the random matrix [5, 16]. For example the IFIE matrix examined here corresponds to the Gaussian orthogonal ensemble whose eigenvalue distribution follows Wigner's semicircle law (1), and in the case of the time series analysis [11–15, 26, 29–32], it is a kind of the symplectic matrix, so the eigenvalue distribution follows the
Marcenko-Pastur law [101].

3. IFIE matrix and Dataset

We analyze the IFIE matrix [47, 53, 102] of the cyclic AMP receptor protein complex ((CRP)-cAMP-DNA), PDBID:1O3Q [103]. The IFIE matrix is available on URL [102]. Previous studies [47, 53] elucidate the sequence-specific binding of the CRP complexed with a cAMP and DNA duplex and the stability using the FMO calculations. In these studies, the IFIEs were analyzed considering the interactions of CRP-cAMP with each base pair, the DNA duplex with each amino acid residue, and each base pair with each residue. In the calculation [47, 53], they adopted the fragments in FMO as the amino acid residue, DNA bases, DNA backbone, and cAMP. The quantum-mechanical calculation is at MP2/6-31G level. The IFIE is a $245 \times 245$ matrix. The consistency with the IFIE matrix which is used here is checked with the figures for the same IFIE matrix in Refs. [47, 53].

4. Results and statistical properties

The probability density of the components of the IFIE matrix, normalized to unit variance, i.e., $\sigma = 1$, is shown in Fig. 1. The function is characterized by a sharp peak around the origin and two broad low peaks on either side. We diagonalize the IFIE matrix numerically and obtain the eigenvalues $\lambda_k$. The probability density of the eigenvalues is shown in Fig. 2. The four eigenvalues outside Wigner's semicircle distribution are $\lambda_1 = 11.82$, $\lambda_2 = 3.784$, $\lambda_3 = 2.834$, and $\lambda_4 = 2.581$. Before the normalization of the variance, the respective values were $\lambda_1 = 185.0$, $\lambda_2 = 59.23$, $\lambda_3 = 44.36$, and $\lambda_4 = 40.40$. To test for universal properties, we calculate their distributions (Fig. 3); these are consistent with the predictions of the Gaussian orthogonal ensemble in RMT. Because of the small size of the IFIE matrix, the convergence is poor and we cannot make a definite conclusion. Fig. 4 represents the IPRs. All IPRs exceed the value predicted by the RMT, $3/245$. This means that a cluster of correlated fragments tends to be localized. The IPRs of a large number of eigenvectors considerably deviate from the values of the RMT.

Figs. 5–10 show the components of the eigenvectors and their probability densities, for $u_1$ to $u_8$, $u_{50}$ to $u_{57}$, $u_{100}$ to $u_{107}$, $u_{150}$ to $u_{157}$, $u_{200}$ to $u_{206}$, $u_{239}$ to $u_{245}$, and the eight eigenvectors with the largest IPR values. On the horizontal axes of the eigenvector components shown in Figs. 5–10, we set components 1-200 as amino acid residues, components 201-244 as the DNA, and component 245 as the cAMP in the matrix in order provide a graphical representation. The numerals in the other figures, tables, and caption notations in the PDB file. For example, the corresponding residue numbers in PDB are 8-207. Figs. 11 and 12 represent the image of the eigenvectors from $u_1$ to $u_{12}$ and twelve eigenvectors of largest IPR, respectively.
Figure 1. Probability density of the IFIE matrix components, which is normalized to unit variance, $\sigma=1$.

Figure 2. Probability density function of the eigenvalues $\lambda$ of the IFIE matrix. The green curve represents the RMT result, Wigner's semicircle law, for a real symmetric matrix Eq. (1). The variance of the IFIE matrix is normalized to $\sigma=1$ and the upper and lower bounds of Wigner's semicircle law are $\pm 2$.

Figure 3. Nearest-neighbor (upper panel) and the next-nearest-neighbor (lower panel) spacing distributions of the unfolded eigenvalues $\xi$. The green curves represent the prediction of RMT, i.e., $P_{nn}(s)$ of Eq. (3) (upper panel) and $P_{nnn}(s)$ of Eq. (4) (lower panel). The blue curves are the Poisson distributions, which show no correlation among the eigenvalues.

Figure 4. Inverse participation ratio (IPR), defined by Eq. (6), as a function of eigenvalue $\lambda$ for the IFIE matrix. The thick lines are predictions of RMT, which is $3/245$. 
Figure 5. Continued on the following page.
Figure 5. Eigenvectors (left column) and the probability density functions (right column) of eigenvectors $u_k$ of the IFIE matrix.
Eigenvectors associated with (a) the largest eigenvalue $\lambda_1$, (b) $\lambda_2$, (c) $\lambda_3$, (d) $\lambda_4$, (e) $\lambda_5$, (f) $\lambda_6$, (g) $\lambda_7$, and (h) $\lambda_8$. In each plot of the probability density functions, the green curve represents the distribution by RMT.
Figure 6. Continued on the following page.
Figure 6. Eigenvectors (left column) and the probability density functions (right column) of eigenvectors $u_k$ of the IFIE matrix.

Eigenvectors associated with (a) $\lambda_{50}$, (b) $\lambda_{51}$, (c) $\lambda_{52}$, (d) $\lambda_{53}$, (e) $\lambda_{54}$, (f) $\lambda_{55}$, (g) $\lambda_{56}$, and (h) $\lambda_{57}$. In each plot of the probability density functions, the green curve represents the distribution by RMT.
Figure 7. Continued on the following page.
Figure 7. Eigenvectors (left column) and the probability density functions (right column) of eigenvectors $u_k$ of the IFIE matrix. Eigenvectors associated with (a) $\lambda_{100}$, (b) $\lambda_{101}$, (c) $\lambda_{102}$, (d) $\lambda_{103}$, (e) $\lambda_{104}$, (f) $\lambda_{105}$, (g) $\lambda_{106}$, and (h) $\lambda_{107}$. In each plot of the probability density functions, the green curve represents the distribution by RMT.
Figure 8. Continued on the following page.
Figure 8. Eigenvectors (left column) and the probability density functions (right column) of eigenvectors $u_k$ of the IFIE matrix.

Eigenvectors associated with (a) $\lambda_{150}$, (b) $\lambda_{151}$, (c) $\lambda_{152}$, (d) $\lambda_{153}$, (e) $\lambda_{154}$, (f) $\lambda_{155}$, (g) $\lambda_{156}$, and (h) $\lambda_{157}$. In each plot of the probability density functions, the green curve represents the distribution by RMT.
Figure 9. Continued on the following page.
Figure 9. Eigenvectors (left column) and the probability density functions (right column) of eigenvectors $\mathbf{u}_k$ of the IFIE matrix. Eigenvectors associated with (a) $\lambda_{238}$, (b) $\lambda_{239}$, (c) $\lambda_{240}$, (d) $\lambda_{241}$, (e) $\lambda_{242}$, (f) $\lambda_{243}$, (g) $\lambda_{244}$, and (h) $\lambda_{245}$. In each plot of the probability density functions, the green curve represents the distribution by RMT.
Figure 10. Continued on the following page.
Figure 10. Eigenvectors (left column) and the probability density functions (right column) of eigenvectors \( \mathbf{u}_k \) of the IFIE matrix.

Eigenvectors associated with (a) \( \lambda_{235} \), (b) \( \lambda_{243} \), (c) \( \lambda_{232} \), (d) \( \lambda_{230} \), (e) \( \lambda_{237} \), (f) \( \lambda_{221} \), and (g) \( \lambda_{239} \). In each plot of the probability density functions, the green curve represents the distribution by RMT. The corresponding IPR values are \( I_{235} = 0.4536 \), \( I_{243} = 0.3149 \), \( I_{232} = 0.2854 \), \( I_{230} = 0.2814 \), \( I_{237} = 0.2433 \), \( I_{221} = 0.2420 \), and \( I_{239} = 0.2387 \).
Figure 11. 3D representation of the eigenvectors $u_1$ to $u_{12}$ from left to right and top to bottom.
The absolute value of each component is shown as the diameter of a sphere. The strength of the correlation is also expressed in order of red, orange, yellow, and green.
Figure 12. 3D representation of the eigenvectors with the largest IPRs, $u_{235}$, $u_{243}$, $u_{232}$, $u_{230}$, $u_{237}$, $u_{221}$, $u_{239}$, $u_{234}$, $u_{242}$, $u_{223}$, $u_{229}$, and $u_{244}$ from left to right and top to bottom in descending order of IPR. The absolute value of the components is shown as the diameter of a sphere. The strength of the correlation is also expressed in order of red, orange, yellow, and green.
Table 1. Correlation sectors obtained from $u_1$ to $u_4$, and from 10 eigenvectors in descending order of IPR.  
The sign in the $u_i$ column represents the sign of the respective components. The DNA fragments are the backbone. The residue and chain numbers follow the notation in PDB.

| fragment | chain | $u_1$ | $u_2$ | $u_3$ | $u_4$ | $u_{235}$ | $u_{243}$ | $u_{232}$ | $u_{230}$ | $u_{237}$ | $u_{221}$ | $u_{239}$ | $u_{234}$ | $u_{242}$ | $u_{223}$ |
|----------|-------|-------|-------|-------|-------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| A-2      | B     | +     |       |       |       |           |           |           |           |           |           |           |           |           |           |
| A-1      | B     | +     |       |       |       |           |           |           |           |           |           |           |           |           |           |
| A1       | B     | +     |       |       |       |           |           |           |           |           |           |           |           |           |           |
| A2       | B     | +     | +     | +     |       |           |           |           |           |           |           |           |           |           |           |
| A3       | B     | +     | +     | +     |       |           |           |           |           |           |           |           |           |           |           |
| T4       | B     | +     | +     | +     |       |           |           |           |           |           |           |           |           |           |           |
| G5       | B     | +     | +     |       |       |           |           |           |           |           |           |           |           |           |           |
| T6       | B     | +     |       |       |       |           |           |           |           |           |           |           |           |           |           |
| G7       | B     | +     |       |       |       |           |           |           |           |           |           |           |           |           |           |
| A8       | B     | +     |       |       |       |           |           |           |           |           |           |           |           |           |           |
| A9       | C     | +     | +     | +     |       |           |           |           |           |           |           |           |           |           |           |
| T8       | C     | +     |       |       |       |           |           |           |           |           |           |           |           |           |           |
| C7       | C     | +     |       |       |       |           |           |           |           |           |           |           |           |           |           |
| A6       | C     | +     |       |       |       |           |           |           |           |           |           |           |           |           |           |
| C5       | C     | +     |       |       |       |           |           |           |           |           |           |           |           |           |           |

Figure 13. $\delta$ dependence of the $z$ index. For $\delta = 0.25$, 0.2, 0.15, 0.1, and 0.05 from left to right and top to bottom.
A large $\delta$ indicates a cluster with a strong local correlation. The magnitude of the $z_i$ index is shown as the diameter of a sphere. The magnitude of $z_i$ is also expressed in order of red, orange, yellow, and green.
| Residue | Charge |
|---------|--------|
| A4      | C +    |
| T3      | C +    |
| T2      | C ++ + |
| T1      | C ++ + |
| T-1     | C ++   |
| CMP1    | A ++   |
| CMP2    | A ++   |
| ASP8    | A      |
| THR10   | A      |
| GLU12   | A +    |
| CYS18   | A +    |
| ILE20   | A      |
| HIS21   | A -    |
| LYS22   | A - +  |
| TYR23   | A -    |
| LYS26   | A - +  |
| THR28   | A +    |
| GLN32   | A      |
| GLU34   | A ++   |
| LYS35   | A - +  |
| GLU37   | A ++   |
| THR38   | A -    |
| LEU39   | A +    |
| TYR41   | A      |
| LYS44   | A -    |
| LYS52   | A -    |
| ASP53   | A ++   |
| GLU54   | A +    |
| GLU55   | A ++   |
| LYS57   | A - +  |
| GLU58   | A ++   |
| ASP68   | A ++   |
| ILE70   | A +    |
| GLU72   | A +    |
| GLY74   | A -    |
| PHE76   | A      |
| GLU77   | A +    |
| GLU78   | A + +  |
| GLU81   | A +    |
| ARG82   | A -    |
| SER83   | A +    |
| ARG87   | A - +  |
| LYS89   | A -    |
| GLU93   | A +    |
| GLU96   | A +    |
| LYS100  | A -    |
| LYS101  | A -    |
| ARG103  | A - +  |
In each plot of the probability density functions, the green curve represents the distribution by RMT (5). As shown in Fig. 5, the distribution of the eigenvector components with large eigenvalue deviates from the RMT. This deviation decreases such that the eigenvalue is located around the center of the RMT distribution (2), but a small difference remains (Figs. 8 and 9).
5. Discussion

5.1 The 1st eigenvector

In general, the sector seen from eigenvector $u_1$ is interpreted as the global trend of the whole system. In the present study, the distribution of the components is binary; this means that the fragments belonging to the corresponding sectors interact with the same weight. The fragments belonging to this sector are as follows: A-2;B, A-1;B, A1;B, A-3;B, T4;B, G5;B, T6;B, G7;B, A8;B, A9;C, T8;C, C7;C, A6;C, C5;C, A4;C, T3;C, T2;C, T1;C, T-1;C, CMP1, CMP2, GLU12, CYS18, THR28, GLU34, GLU37, LEU39, ASP53, GLU54, GLU55, GLU58, ASP68, ILE70, GLU72, GLU77, GLU78, GLU81, GLU96, ASP111, ILE112, MET120, GLU129, ASP138, ASP155, ASP161, GLN164, GLU171, GLU181, GLU191, ASP192, GLY207, HIS21, LYS22, TYR23, LYS26, LYS35, THR38, LYS44, LYS52, LYS57, GLY74, ARG82, ARG87, LYS89, LYS100, LYS101, ARG103, ARG115, ARG122, ARG123, LYS130, ARG142, LYS152, LYS166, ARG169, ARG180, ARG185, LYS188, LEU195, SER197, GLY200, LYS201; here, the letters B and C after the semicolon represent the chain numbers, and the fragments without a letter correspond to chain A and the former 53 fragments, i.e., up to GLY207, and the remaining 31 fragments, i.e., after HIS21, form two subsectors in the sense that their components have opposite signs in the eigenvector. These sectors consist of a DNA backbone (sugar and phosphate), CRP amino acids, and CMPs. This is extended over the entire complex, except for the bases of DNA. This mode is consistent with the interaction between a DNA backbone and CMP, as shown in the lower right panel of Fig.8 in Fukuzawa et al. [47].

5.2 The 2nd, 3rd, and 4th eigenvectors

The eigenvalues of the eigenvectors, $u_2$, $u_3$, and $u_4$, are outside the RMT boundary (2). The associated sectors are: A3;B, T4;B, G5;B, T2;C, T1;C, CMP1, CMP2, LYS22, GLU34, LYS35, GLU37, ASP68, GLU72, GLU77, GLU78, GLU81, ARG82, GLU93, GLU96, LYS100, LYS101, ARG103, ASP111, ARG115, ARG122, ARG123, ARG169, GLU181, ARG185, LYS188, LYS191, ASP192, and LYS201 extracted from $u_2$; A2;B, A3;B, A9;C, T8;C, T1;C, T-1;C LYS26, ASP53, GLU54, GLU55, LYS57, GLU58, ARG87, LYS89, ASP138, ARG142, GLU171, ARG180, GLU181, ARG185, LYS201, extracted from $u_3$, and A2;B, T4;B, G5;B, A9;C, T2;C, LYS26, LYS35, GLU37, LYS52, ASP53, GLU55, LYS57, GLU58, ARG87, LYS89, ARG103, ASP161, ARG169, GLU171, LYS188, GLU191, ASP192 extracted from $u_4$; where the criterion for their extraction is that the absolute value of the components of the normalized eigenvector must be 0.1 or more. These three sectors are subsets of $u_1$ and are interpreted as the dominant sectors in the decomposition of $u_1$. The sector associated with $u_2$ is interpreted as the interaction between the DNA backbone, CRP, and the CMPs. The sector associated with $u_3$ includes the fragments [47] related to Coulomb interactions between the DNA bases and CRP, such as ARG180, GLU181, and ALG185. However, this sector includes A2, A3, and A9 of the DNA backbone and does not include DNA bases such as G5 or G7. The sector associated with $u_4$ is interpreted as the result of the interaction between the DNA backbone and CRP.

5.3 Sectors with high IPRs

In RMT an eigenvector with an IPR that deviates from the RMT value describes a localized strong correlation sector. On the contrary, in PCA only some of the largest eigenvalues and eigenvectors are typically analyzed. In this IFIE matrix, there are eigenvectors whose IPRs are...
considerably larger than the RMT value. In particular, the sector associated with $u_{253}$ is composed of LYS22, TYR23, LYS44, and GLU93, and it is a subsector of $u_2$. The sector associated with $u_{243}$ is composed of LYS188, GLU191, and ASP192, and it is a subsector of $u_2$ and $u_4$. The sector associated with $u_{232}$ is composed of ASP138, VAL139, THR140, GLY141, ARG142, ILE143, GLN145, THR146, and GLY177, and it is a subsector of $u_3$. The sector associated with $u_{230}$ is composed of ASP8, THR10, GLU12, ASP111, ILE112, MET114, ARG115, and GLN119, and it is an isolated sector. The sector associated with $u_{237}$ is composed of GLU34, LYS35, GLU37, PHE76, GLU77, GLU78, and ARG103, and it is a subsector of $u_2$ and $u_4$. The sector associated with $u_{221}$ is composed of CMP1, CMP2, ILE20, HIS21, TYR23, TYR41, ASP68, ILE70, SER83, and GLU96, and it is a subsector of $u_2$. The sector associated with $u_{239}$ is composed of GLU34, LYS35, GLU37, GLU78, GLU81, and ARG103, and it is a subsector of $u_2$ and $u_4$. The sector associated with $u_{234}$ is composed of CMP1, CMP2, GLN32, GLU34, LEU39, ASP68, GLU72, GLU81, ARG82, SER83, and ARG122, and it is a subsector of $u_2$. The sector associated with $u_{223}$ is composed of ASP68, THR10, GLU12, ARG82, GLN119, ARG122, and ARG123, and it is a subsector of $u_2$. The sector associated with $u_{223}$ is composed of ASP68, THR10, GLU12, ARG82, GLN119, ARG122, and ARG123, and it is a subsector of $u_2$. The sector associated with $u_{234}$ is composed of CMP1, CMP2, GLN32, GLU34, LEU39, ASP68, GLU72, GLU81, ARG82, SER83, and ARG122, and it is a subsector of $u_2$. The sector associated with $u_{237}$ is composed of GLU34, LYS35, GLU37, GLU78, GLU81, and ARG103, and it is a subsector of $u_2$ and $u_4$. The sector associated with $u_{221}$ is composed of CMP1, CMP2, ILE20, HIS21, TYR23, TYR41, ASP68, ILE70, SER83, and GLU96, and it is an isolated sector. These findings are summarized in Table 1.

Next, we examine the $\delta$ dependence of the $z$ index defined by Eq. (7). The results are shown in Fig. 13 for $\delta = 0.25, 0.2, 0.15, 0.1,$ and $0.05$. Because $\delta$ is an IPR threshold, a large $\delta$ implies a cluster with strong local correlation. The results show that there are strong local correlations between the DNA-binding helix-turn-helix motif and the CRP on the opposite side of the DNA.

6. Summary

We studied the statistical properties of the IFIE matrix eigensystem of the FMO method using RMT and reported results in detail. The probability density of the eigenvalues is not consistent with Wigner's semicircle law. Conversely, the nearest and next-to-nearest neighbors unfolded eigenvalue statistics that are consistent with those of the Gaussian orthogonal and symplectic ensembles, respectively. The eigenvector $u_1$ is a unique feature. The distribution of the components is binary and the IPR agree with the RMT value, implying that only specific fragments perform interactions with a uniform strength and that those fragments are randomly scattered over the entire cyclic AMP receptor protein complex. Because all the correlation sectors obtained from $u_2$, $u_3$, and $u_4$ are subsets of $u_1$, it is understood that there exists a hierarchical structure, where specific fragments form subsectors with strong interactions between the fragments that belong to the $u_1$ sector. The IPRs of a large number of eigenvectors considerably deviate from the values of RMT, implying that there are many local sectors of fragments with strong correlation. Since most of these local sectors form a subsector of $u_2$, $u_3$, or $u_4$, we find that there is an additional hierarchical structure under the first hierarchy. It would be quite an interesting task to perform an in-depth analysis for other proteins, protein complexes, and complexes of proteins and nucleic acids including PIEDA.

7. Acknowledgements

I greatly appreciate the FMO tutorial course of the Chem-Bio Informatics (CBI) society annual meeting held in 2013, as the ideas underlying this report were inspired from this meeting. This study was performed using numerical data from the cpf file, trunc-DNA_CRP_mp2D_631G_FZC.cpf, from the Revolutionary Simulation Software for the 21st
century (RSS21) project [102]. The corresponding software is the BioStationViewer Version 16.00 produced by T. Nakano, K. Kitaura, N. Sasaki, Y. Mochizuki, and S. Amari and these results were achieved by using Innovative Simulation Software for an Industrial Science Project. Figs. 11, 12, and 13 were obtained using Rasmol [104].

References

[1] Wigner, E. P. On the statistical distribution of the widths and spacings of nuclear resonance levels. Proc. Cambridge Philos. Soc., 1951, 47, 790–798.
[2] Dyson, F. J. Statistical Theory of the Energy Levels of Complex Systems. I. J. Math. Phys. 1962, 3, 140–156.
[3] Dyson, F. J.; Mehta, M. L. Statistical Theory of the Energy Levels of Complex Systems. IV. J. Math. Phys. 1963, 4, 701–712.
[4] Dyson, F. J.; Mehta, M. L. Statistical Theory of the Energy Levels of Complex Systems. V. J. Math. Phys. 1963, 4, 713–719.
[5] Mehta, M. L. Random Matrices; Academic Press: New York, 1991.
[6] Polchinski, J. String Theory; Cambridge Monographs on Mathematical Physics Volume 2, Cambridge University Press: Cambridge, 1998.
[7] Pearson, K. On Lines and Planes of Closest Fit to Systems of Points in Space. Philos. Mag. 1901, 2, 559–572.
[8] Hotelling, H. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 1933, 24, 417–441.
[9] Hotelling, H. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 1933, 24, 498–520.
[10] Hotelling, H. Relations between two sets of variates. Biometrika, 1936, 28, 321–377.
[11] Laloux, L.; Cizeau, P.; Bouchaud, J. P.; Potters, M. Noise Dressing of Financial Correlation Matrices. Phys. Rev. Lett. 1999, 83, 1467–1470.
[12] Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L. A. N.; Stanley, H. E. Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series. Phys. Rev. Lett. 1999, 83, 1471–1474.
[13] Gopikrishnan, P.; Rosenow, B.; Plerou, V.; Stanley, H. E. Quantifying and Interpreting Collective Behavior in Financial Markets. Phys. Rev. E 2001, 64, 035106.
[14] Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L. A. N.; Guhr, T.; et al. Random matrix approach to cross correlations in financial data. Phys. Rev. E 2002, 65, 066126.
[15] Utsugi, A.; Ino, K.; Oshikawa, M. Random matrix theory analysis of cross correlations in financial markets. Phys. Rev. E 2004, 70, 026110.
[16] The Oxford Handbook of Random Matrix Theory; Akemann, G., Baik, J., Di Francesco, P., Eds.; Oxford University press: Oxford, 2011.
[17] Schaefer, J.; Yaris, R. Random Matrix Theory and Nuclear Magnetic Resonance Spectral Distributions J. Chem. Phys. 1969, 51, 4469–4474.
[18] Lacelle, S. Random matrix theory in biological nuclear magnetic resonance spectroscopy Biophys. J. 1984, 46, 181–186.
[19] Orland, H.; Zee, A. RNA folding and large $N$ matrix theory. Nucl. Phys. B 2002, 620, 456–476.
[20] Vernizzi, G.; Orland, H.; Zee, A. Enumeration of RNA Structures by Matrix Models. Phys. Rev. Lett. 2005, 94, 168103.
[21] Bon, M.; Vernizzi, G.; Orland, H.; Zee, A. Topological Classification of RNA Structures. J.
Mol. Bio. 2008, 379, 900–911.

[22] L’ener, M. K.; Schulten, K. General random matrix approach to account for the effect of static disorder on the spectral properties of light harvesting systems. Phys. Rev. E 2002, 65, 031916.

[23] Ciliberti, S.; De Los Rios, P.; Piazza, F. Glasslike Structure of Globular Proteins and the Boson Peak. Phys. Rev. Lett. 2006, 96, 198103.

[24] Luo, F.; Zhong, J.; Yang, Y.; Scheuermann, R. H.; Zhou, J. Application of random matrix theory to biological networks. Phys. Lett. A 2006, 357, 420–423.

[25] Bandyopadhyay, J. N.; Jalan, S. Universality in complex networks: Random matrix analysis. Phys. Rev. E 2007, 76, 026109.

[26] Potestio, R.; Caccioli, F.; Vivo, P. Random Matrix Approach to Collective Behavior and Bulk Universality in Protein Dynamics. Phys. Rev. Lett. 2009, 103, 268101.

[27] Halabi, N.; Rivoire, O.; Leibler, S.; Ranganathan R. Protein Sectors: Evolutionary Units of Three-Dimensional Structure. Cell 2009, 138, 774–786.

[28] Dahirel, V.; Shekhar, K.; Pereyra, F.; Miura, T.; Artyomov, M.; et al. Coordinate linkage of HIV evolution reveals regions of immunological vulnerability. Proc. Natl. Acad. of Sci. 2011, 108, 11530–11535.

[29] Yamanaka, M. Random Matrix Theory Analysis of Cross Correlations in Molecular Dynamics Simulations of Macro-Biomolecules. J. Phys. Soc. Jpn. 2013, 82, 083801.

[30] Palese, L. L. Random Matrix Theory in molecular dynamics analysis. Biophysical Chemistry 2015, 196, 1–9.

[31] Yamanaka, M. Random Matrix Theory of Rigidity in Soft Matter. J. Phys. Soc. Jpn. 2015, 84, 063801.

[32] Palese, L. L. Correlation Analysis of Trp-Cage Dynamics in Folded and Unfolded States J. Phys. Chem. B 2015, 119, 15568–15573.

[33] 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.

[34] Nakano, T.; Kaminuma, T.; Sato, T.; Akiyama, Y.; Uebayasi, M.; et al. Fragment molecular orbital method: application to polypeptides. Chem. Phys. Lett. 2000, 318, 614–618.

[35] Kitaura, K.; Sugiki, S. I.; Nakano, T.; Komeiji, Y.; Uebayasi, M. Fragment molecular orbital method: analytical energy gradients. Chem. Phys. Lett. 2001, 336, 163–170.

[36] Nakano, T.; Kaminuma, T.; Sato, T.; Fukuzawa, K.; Akiyama, Y.; et al. Fragment molecular orbital method: use of approximate electrostatic potential. Chem. Phys. Lett. 2002, 351, 475–480.

[37] Mochizuki, Y.; Fukuzawa, K.; Kato, A.; Tanaka, S.; Kitaura, K.; et al. A configuration analysis for fragment interaction. Chem. Phys. Lett. 2005, 410, 247–253.

[38] Fedorov, D. G.; Kitaura, K. Pair interaction energy decomposition analysis. J. Comp. Chem. 2007, 28, 222–237.

[39] Ishikawa, T.; Mochizuki, Y.; Amari, S.; Nakano, T.; Tokiwa, H.; et al. Fragment interaction analysis based on local MP2. Theor. Chem. Acc. 2007, 118, 937–945.

[40] Tanaka, S.; Watanabe, C.; Okiyama, Y.; Statistical correction to effective interactions in the fragment molecular orbital method. Chem. Phys. Lett. 2013, 556, 272–277.

[41] Kurauchi, R.; Watanabe, C.; Fukuzawa, K.; Tanaka, S. Novel type of virtual ligand screening on the basis of quantum-chemical calculations for protein-ligand complexes and extended clustering techniques. Comp. Theor. Chem. 2015, 1061, 12–22.

[42] Fukuzawa, K.; Kitaura, K.; Nakata, K.; Kaminuma, T.; Nakano, T. Fragment molecular orbital study of the binding energy of ligands to the estrogen receptor. Pure Appl. Chem. 2003, 75, 2405–2410.
[43] Fukuzawa, K.; Kitaura, K.; Uebayasi, M.; Nakata, K.; Kaminuma, T.; et al. Ab initio quantum mechanical study of the binding energies of human estrogen receptor with its ligands: An application of fragment molecular orbital method. *J. Comp. Chem.* 2005, 26, 1–10.

[44] Sugiki, S. -I.; Matsuoka, M.; Usuki, R.; Sengoku, Y.; Kurita, N.; et al. Density functional calculations on the interaction between catabolite activator protein and cyclic AMP using the fragment molecular orbital method. *J. Theor. Comp. Chem.* 2005, 4, 183–195.

[45] Nemoto, T.; Fedorov, D. G.; Uebayasi, M.; Kanazawa, K.; Kitaura, K.; et al. Ab initio fragment molecular orbital (FMO) method applied to analysis of the ligand-protein interaction in a pheromone-binding protein. *Comp. Biol. Chem.* 2005, 29, 434–439.

[46] Amari, S.; Aizawa, M.; Zhang, J.; Fukuzawa, K.; Mochizuki, Y.; et al. VISCANA: visualized cluster analysis of protein-ligand interaction based on the ab initio fragment molecular orbital method for virtual ligand screening. *J. Chem. Inf. Comp. Sci.* 2006, 46, 221–230.

[47] Fukuzawa, K.; Komeiji, Y.; Mochizuki, Y.; Kato, A.; Nakano, T.; et al. Intra- and intermolecular interactions between cyclic-AMP receptor protein and DNA: Ab initio fragment molecular orbital study. *J. Phys. Chem. B* 2006, 110, 16102–16110.

[48] Ito, M.; Fukuzawa, K.; Mochizuki, Y.; Nakano, T.; Tanaka, S. Ab initio fragment molecular orbital study of molecular interactions between liganded retinoid X receptor and its coactivator: Roles of helix 12 in the coactivator binding mechanism. *J. Phys. Chem. A* 2008, 112, 1986–1998.

[49] Iwata, T.; Fukuzawa, K.; Nakajima, K.; Aida-Hyugaji, S.; Mochizuki, Y.; et al. Theoretical analysis of binding specificity of influenza viral hemagglutinin to avian and human receptors based on the fragment molecular orbital method. *Comp. Biol. Chem.* 2008, 32, 198–211.

[50] Harada, T.; Yamagishi, K.; Nakano, T.; Tokiwa, H. Ab initio fragment molecular orbital study of ligand binding to human progesterone receptor ligand-binding domain. *Naunyn-Schmiedeberg's Arch. Pharmac.* 2008, 377, 607–615.

[51] Ito, M.; Fukuzawa, K.; Ishikawa, T.; Mochizuki, Y.; Nakano, T.; et al. Ab Initio Fragment Molecular Orbital Study of Molecular Interactions in Liganded Retinoid X Receptor: Specification of Residues Associated with Ligand Inducible Information Transmission. *J. Phys. Chem. B* 2008, 112, 12081–12094.
[58] Sawada, T.; Hashimoto, T.; Tokiwa, H.; Suzuki, T.; Nakano, H.; et al. Ab initio fragment molecular orbital studies of influenza virus hemagglutinin-sialosaccharide complexes toward chemical clarification about the virus host range determination. Glycoconj. J. 2008, 25, 805–815.

[59] Tada, M.; Nagasima, T.; Udagawa, T.; Tachikawa, M.; Sugawara H. Ab initio fragment molecular orbital (FMO) analysis of the structure of the phosphoinositide-binding peptide from gelsolin. J. Mol. Str. (THEOCHEM) 2009, 897, 149–153.

[60] Nakanishi, I.; Fedorov, D. G.; Kitaura, K. Detailed electronic structure studies revealing the nature of protein-ligand binding. in The fragment molecular orbital method: Practical applications to large molecular systems, Fedorov, D. G., Kitaura, K. Eds.; CRC Press: Boca Raton, FL, 2009; pp 171–192.

[61] Sawada, T.; Hashimoto, T.; Tokiwa, H.; Suzuki, T.; Nakano, H.; et al. How does FMO method help in studying viruses and their binding to receptors? in The fragment molecular orbital method: Practical applications to large molecular systems, Fedorov, D. G., Kitaura, K. Eds.; CRC Press: Boca Raton, FL, 2009; pp 193–216.

[62] Nagase, K.; Kobayashi, H.; Yoshikawa, E.; Kurita, N. Ab initio molecular orbital calculations on specific interactions between urokinase-type plasminogen activator and its receptor. J. Mol. Graph. Mod. 2009, 28, 46–53.

[63] Dedachi, K.; Khan, M. T. H.; Sylte, I.; Kurita, N. A combined simulation with ab initio MO and classical vibrational analysis on the specific interactions between thermolysin and dipeptide ligands. Chem. Phys. Lett. 2009, 479, 290–295.

[64] Yamagishi, K.; Yamamoto, K.; Mochizuki, Y.; Nakano, T.; Yamada, S.; et al. Flexible ligand recognition of peroxisome proliferator-activated receptor-gamma (PPARgamma). Bioorg. Med. Chem. Lett. 2010, 20, 3344–3347.

[65] Motoyoshi, S.; Yamagishi, K.; Yamada, S.; Tokiwa, H. Ligand-dependent conformation change reflects steric structure and interactions of a vitamin D receptor/ligand complex: a fragment molecular orbital study. J. Ster. Biochem. Mol. Biol. 2010, 121, 56–59.

[66] Yamagishi, K.; Tokiwa, H.; Makishima, M.; Yamada, S. Interactions between 1alpha,25(OH)2D3 and residues in the ligand-binding pocket of the vitamin D receptor: a correlated fragment molecular orbital study. J. Ster. Biochem. Mol. Biol. 2010, 121, 63–67.

[67] Yoshikawa, E.; Miyagi, S.; Dedachi, K.; Ishihara-Sugano, M.; Itoh, S.; et al. Specific interactions between aryl hydrocarbon receptor and dioxin congeners: Ab initio fragment molecular orbital calculations. J. Mol. Graph. Mod. 2010, 29, 197–205.

[68] Hitaoka, S.; Harada, M.; Yoshida, T.; Chuman, H. Correlation analyses on binding affinity of sialic acid analogues with influenza virus neuraminidase-1 using ab initio MO calculations on their complex structures. J. Chem. Inf. Model. 2010, 50, 1796–1805.

[69] Sawada, T.; Fedorov, D. G.; Kitaura, K. Binding of influenza A virus hemagglutinin to the sialoside receptor is not controlled by the homotropic allosteric effect. J. Phys. Chem. B 2010, 114, 15700–15705.

[70] Kurisaki, I.; Fukuzawa, K.; Nakano, T.; Mochizuki, Y.; Watanabe, H., et al. Fragment molecular orbital (FMO) study on stabilization mechanism of neuro-oncological ventral antigen (NOVA)-RNA complex system. J. Mol. Str. (THEOCHEM) 2010, 962, 45–55.

[71] Ohyama, T.; Hayakawa, M.; Nishikawa, S.; Kurita, N. Specific interactions between lactose repressor protein and DNA affected by ligand binding: Ab initio molecular orbital calculations. J. Comp. Chem. 2011, 32, 1661–1670.

[72] Tsuji, S.; Kasumi, T.; Nagase, K.; Yoshikawa, E.; Kobayashi, H.; et al. The effects of amino-acid mutations on specific interactions between urokinase-type plasminogen activator and its receptor: Ab initio molecular orbital calculations. J. Mol. Graph. Mod. 2011, 29,
975–984.

[73] Fukuzawa, K.; Omagari, K.; Nakajima, K.; Nobusawa, E.; Tanaka, S. Sialic Acid Recognition of the Pandemic Influenza 2009 H1N1 Virus: Binding Mechanism Between Human Receptor and Influenza Hemagglutinin. Prot. Pept. Lett. 2011, 18, 530–539.

[74] Ozawa, T.; Okazaki, K.; Kitaura. K. CH/π hydrogen bonds play a role in ligand recognition and equilibrium between active and inactive states of the β2 adrenergic receptor: An ab initio fragment molecular orbital (FMO) study. Bioorg. Med. Chem. 2011, 19, 5231–5237.

[75] Hitaoka, S.; Matoba, H.; Harada, M.; Yoshida, T.; Tsuji, D.; et al. Correlation Analyses on Binding Affinity of Sialic Acid Analogues and Anti-Influenza Drugs with Human Neuraminidase Using ab initio MO Calculations on Their Complex Structures LERE-QSAR Analysis (IV). J. Chem. Inf. Model. 2011, 51, 2706–2716.

[76] Yoshioka, A.; Takematsu, K.; Kurisaki, I.; Fukuzawa, K.; Mochizuki, Y.; et al. Antigen-antibody interactions of influenza virus hemagglutinin revealed by the fragment molecular orbital calculation. Theor. Chem. Acc. 2011, 130, 1197–1202.

[77] 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.

[78] Hirakawa, T.; Fujita, S.; Ohyama, T.; Dedachi, K.; Khan, M.T.H.; et al. Specific interactions and binding energies between thermolysin and potent inhibitors: Molecular simulations based on ab initio molecular orbital method. J. Mol. Graph. Mod. 2012, 33, 1–11

[79] Itoh, Y.; Sando, A.; Ikeda, K.; Suzuki, T.; Tokiwa, H. Origin of the inhibitory activity of 4-O-substituted sialic derivatives of human parainfluenza virus. Glycoconj. J. 2012, 29, 231–237.

[80] Ishikawa, T.; Burri, R. R.; Kamatari, Y. O.; Sakuraba, S.; Matubayasi, N.; et al. A theoretical study of the two binding modes between lysozyme and tri-NAG with an explicit solvent model based on the fragment molecular orbital method. Phys. Chem. Chem. Phys. 2013, 15, 3646–3654.

[81] Okiyama, Y.; Tsukamoto, T.; Watanabe, C.; Fukuzawa, K.; Tanaka, S.; et al. Modeling of peptide-silica interaction based on four-body corrected fragment molecular orbital (FMO4) calculations. Chem. Phys. Lett. 2013, 566, 25–31.

[82] Kasumi, T.; Araki, K.; Ohyama, T.; Tsuji, S.; Yoshikawa, E.; et al. The effects of vitronectin on specific interactions between urokinase-type plasminogen activator and its receptor: ab initio molecular orbital calculations. Mol. Sim. 2013, 39, 769–779.

[83] Ueno-Noto, A.; Ise, S.; Takano, K. Chemical description of the interaction between glycan ligand and Siglec-7 using ab initio FMO method and classical MD simulation. J. Theor. Comp. Chem. 2013, 12, 1350060.

[84] Watanabe, C.; Fukuzawa, K.; Tanaka, S.; Aida-Hyugaji, S. Charge clamps of lysines and hydrogen bonds play key roles in the mechanism to fix helix 12 in the agonist and antagonist positions of estrogen receptor alpha: intramolecular interactions studied by the ab initio fragment molecular orbital method. J. Phys. Chem. B 2014, 118, 4993–5008.

[85] Anzaki, S.; Watanabe, C.; Fukuzawa, K.; Mochizuki, Y.; Tanaka, S. Interaction energy analysis on specific binding of influenza virus hemagglutinin to avian and human siaosaccharide receptors: Importance of mutation-induced structural change. J. Mol. Graph. Model. 2014, 53, 48–58.

[86] Murakawa, T.; Matsushita, Y.; Suzuki, T.; Khan, M. T. H.; Kurita, N. Ab initio molecular simulations for proposing potent inhibitors to butyrylcholinesterases. J. Mol. Graph. Model. 2014, 54, 54–61.

[87] Prato, G.; Silvent, S.; Saka, S.; Lamberto, M.; Kosenkov, D. Thermodynamics of binding of di- and tetrasubstituted naphthalene diimide ligands to DNA G-quadruplex. J. Phys. Chem. B
2015, 119, 3335–3347.
[88] Ando, H.; Shigeta, Y.; Baba, T.; Watanabe, C.; Okiyama, Y.; et al. Hydration effects on enzyme-substrate complex of nylon oligomer hydrolase: inter-fragment interaction energy study by the fragment molecular orbital method. Mol. Phys. 2015, 113, 319–326.
[89] Miyagi, S.; Murata, K.; Sashino, K.; Sawamura, S.; Uruno, S.; et al. Binding affinity between AhR and exogenous/endogenous ligands: molecular simulations and biological experiment. Mol. Sim. 2015, 41, 555–563.
[90] Wigner, E. P. On the Distribution of the Roots of Certain Symmetric Matrices. Ann. Math. 1958, 67, 325–327.
[91] Brody, T. A.; Flores, J.; French, J. B.; Mello, P. A.; Pandey, A.; et al. Random-matrix physics: spectrum and strength fluctuations. Rev. Mod. Phys. 1981, 53, 385–479.
[92] Guhr, T.; Muller-Groeling, A.; Weidenmuller, H. A. Random-matrix theories in quantum physics: common concepts. Phys. Rep. 1998, 299, 189–425.
[93] Bell, R. J.; Dean, P. Atomic vibrations in vitreous silica. Discuss. Faraday Soc. 1970, 50, 55–61.
[94] Thouless, D. J. Electrons in disordered systems and the theory of localization. Phys. Rep. 1974, 13, 93–142.
[95] Wegner, F. Inverse participation ratio in $2+\varepsilon$ dimensions. Z. Phys. B 1980, 36, 209–214.
[96] Zirnbauer, M. R. Anderson localization and non-linear sigma model with graded symmetry. Nucl. Phys. B 1986, 265 [FS15], 375–408.
[97] Hikami, S. Localization Length and Inverse Participation Ratio of Two Dimensional Electron in the Quantized Hall Effect. Prog. Theor. Phys. 1986, 76, 1210–1221.
[98] Fyodorov, Y. V.; Mirlin, A. D. Analytical derivation of the scaling law for the inverse participation ratio in quasi-one-dimensional disordered systems. Phys. Rev. Lett. 1992, 69, 1093–1096.
[99] Fyodorov, Y. V.; Mirlin, A. D. Level-to-level fluctuations of the inverse participation ratio in finite quasi 1D disordered systems. Phys. Rev. Lett. 1992, 71, 412–415.
[100] Mirlin, A. D.; Fyodorov, Y. V. The statistics of eigenvector components of random band matrices: analytical results. J. Phys. A 1993, 26, L551–L558.
[101] Marcenko, V. A.; Pastur, L. A. Distribution of the eigenvalues for some sets of random matrices. Math. USSR Sb. 1967, 1, 457–483.
[102] URL: http://www.ciss.iis.u-tokyo.ac.jp/rss21/index.html
[103] Chen, S.; Vojtechovsky, J.; Parkinson, G. N.; Ebright, R. H.; Berman, H. M. Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: DNA binding specificity based on energetics of DNA kinking. J. of Mol. Bio. 2001, 314, 63–74.
[104] Sayle, R.; Milner-White, E. J. RASMOL: biomolecular graphics for all. Trends in Biochem. Sci. (TIBS), 1995, 20, 374–376.