Protein-ligand binding affinity prediction of cyclin-dependent kinase-2 inhibitors by dynamically averaged fragment molecular orbital-based interaction energy

Takaba, Kenichiro ; Watanabe, Chiduru ; Tokuhisa, Atsushi ; Akinaga, Yoshinobu ; Ma, Biao ; Kanada, Ryo ; Araki, Mitsugu ; Okuno, Yasushi ;…

(Citation)
Journal of Computational Chemistry,43(20):1362-1371

(Issue Date)
2022-07-30

(Resource Type)
journal article

(Version)
Accepted Manuscript

Protein–Ligand Binding Affinity Prediction of CDK2 Inhibitors by Dynamically Averaged FMO-based Interaction Energy

Kenichiro Takaba[1], Chiduru Watanabe[2], Atsushi Tokuhisa[3], Yoshinobu Akinaga[3,4] Biao Ma[3], Ryo Kanada[3], Mitsugu Araki[5], Yasushi Okuno[3,5], Yusuke Kawashima[6], Hirotomo Moriwaki[2], Norihito Kawashita[7], Teruki Honma[2], Kaori Fukuzawa [6,8, a)], and Shigenori Tanaka[9, a)]

[1] Pharmaceutical Research Center, Advanced Drug Discovery, Asahi Kasei Pharma Corporation, 632-1 Mifuku, Izunokuni, Shizuoka 410-2321, Japan

[2] RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan

[3] RIKEN Center for computational Science, 7-1-26 Minatojima-minami-machi, Chuo-ku Kobe, Hyogo 650-0047, Japan

[4] VINAS Co., Ltd., Keihan Dojima Bldg., Dojima 2 1 31, Kita-ku, Osaka 530-0003, Japan

[5] Graduate School of Medicine, Kyoto University, Shogoin-Kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan

[6] School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa, Tokyo 142-8501, Japan

[7] Faculty of Science and Engineering, Kindai University, 3-4-1 Kowakae, Higashiosaka, Osaka, 577-8502, Japan

[8] Department of Biomolecular Engineering, Graduate School of Engineering, Tohoku University, 6-6-11 Aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan

[9] Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan

a) Authors to whom correspondence should be addressed: tanaka2@kobe-u.ac.jp, k-fukuzawa@hoshi.ac.jp

ABSTRACT

Fragment molecular orbital (FMO) method is a powerful computational tool for structure-based drug design, in which protein–ligand interactions can be described by the inter-fragment interaction energy (IFIE) and its pair interaction energy decomposition analysis (PIEDA). Here, we introduced a dynamically averaged (DA) FMO-based approach in which molecular dynamics simulations were used to generate multiple protein–ligand complex structures for FMO calculations. To assess this approach, we examined the correlation between the experimental binding free energies and DA-IFIEs of six CDK2 inhibitors whose net charges are zero. The correlation between the experimental binding free energies and snapshot IFIEs for X-ray crystal structures was $R^2 = 0.75$. Using the DA-IFIEs, the correlation significantly improved to 0.99. When an additional CDK2 inhibitor with net charge of −1 was added, the DA FMO-based scheme with the dispersion energies still achieved $R^2 = 0.99$, whereas R^2 decreased to 0.32 employing all the energy terms of PIEDA.

KEYWORDS

Fragment molecular orbital method (FMO); Inter-fragment interaction energy (IFIE); Pair interaction energy decomposition analysis (PIEDA); Cyclin-dependent kinase-2 (CDK2) inhibitors; Dynamical average (DA)

I. INTRODUCTION

The fragment molecular orbital (FMO) method[1-4] is a powerful computational tool for structure-based drug design (SBDD).[5-7] Inter-fragment interaction energy (IFIE) analysis[3] examines protein–ligand interactions with quantum (electronic) effects at reasonable computational costs. Furthermore, by means of pair interaction energy decomposition analysis (PIEDA),[4] the interaction energy, ΔE, is decomposed into four energy terms: electrostatic (ES) exchange repulsion (EX), dispersion (DI), and charge transfer with high-order mixed terms (CT+mix), which offer detailed information regarding protein–ligand interactions. This information is valuable for rational drug design and constructive communications with medicinal chemists in drug discovery projects.

FMO-based interaction energies obtained by PIEDA indicate the stability of protein–ligand complexes; therefore, they can be used to predict protein–ligand binding affinities.[8-10] Although the FMO-based approach does not fully capture the complex phenomena of protein–ligand binding events and their associated energy terms, such as entropy, ligand deformation upon binding, and solvation/desolvation effects,[11] it is expected to provide reliable estimates of protein–ligand enthalpic binding contributions. Recently, many studies have shown that the FMO method can be used as an *in silico* approach to predict protein–ligand binding affinities. For example, Heifetz *et al.*[9] observed a moderate to high correlation range (R^2 = 0.58–0.76) between experimentally measured affinities and FMO-based interaction energies *in vacuo* for various G-protein coupled receptors by performing FMO calculations against computationally modeled structures. Watanabe *et al.*[10] utilized the FMO method with molecular-mechanics Poisson–Boltzmann surface area (FMO+MM-PBSA) to incorporate desolvation effects and studied the activity cliffs of serine/threonine kinase Pim1 inhibitors; comparing various properties of complex structures, they observed that quantum mechanics QM/MM-optimized structures provided the best correlation (R^2 = 0.85) between experimental pIC_{50} values and FMO+MM-PBSA binding energies.

The FMO-based interaction energy can be obtained from a single protein–ligand complex structure (snapshot), such as an X-ray crystal structure or a computationally modeled structure. However, previous studies[12-15] suggest that averaging the FMO-based interaction energies over multiple structures is more reliable than considering only a single structure. Ishikawa *et al.*[12] performed FMO calculations against 40 prion protein structures generated by molecular dynamics (MD) simulations and examined the influence of geometrical fluctuations on FMO calculations. Fedorov *et al.*[13] performed MD simulations with FMO calculations for two ligands bound to a Trp-cage mini-protein and systematically averaged the FMO-based interaction energies to evaluate the protein–ligand binding energies. Hatada *et al.*[14,15] also found the significance of structural fluctuations in the IFIE analysis on a complex system of SARS-CoV-2 main protease and N3 ligand. However, these studies have not addressed the ability of systematically averaged FMO-based interaction energies to improve protein–ligand binding affinity predictions using a set of multiple candidate compounds.

A limitation of FMO calculations performed *in vacuo* is the difficulty of predicting binding affinities for ligands of different net charges. This is because the total interaction energies are inclined to be dominated by strong electrostatic interactions between the protein and charged ligands. To overcome this issue, implicit solvent models such as the PBSA[16] and polarizable continuum model (PCM)[17] can be applied to consider the solvation and desolvation effects and fulfill the thermodynamic cycle. However, there is still a need for FMO calculations to be used to directly compare ligands with different net charges, particularly via dynamically averaged (DA) FMO calculations in solvent, where multiple FMO calculations are required to evaluate the affinity of a single ligand.

Here, we examined the potential for a DA-FMO-based approach in which FMO calculations are performed for multiple protein–ligand complex structures, generated by MD simulations for solvated systems, to improve the accuracy of protein–ligand binding affinity predictions for multiple compounds. The correlation between the experimental binding free energies and DA-FMO-based interaction energies of six cyclin-dependent kinase-2 (CDK2) inhibitors, whose net charges were expected to be zero under physiological conditions, was examined as a test case, which was previously studied by means of MD approach.[18] The correlation between the experimental binding free energies and FMO-based interaction energies *in vacuo* calculated with MM-optimized X-ray crystal structures was $R^2 = 0.75$, whereas the correlation was significantly improved to $R^2 = 0.99$, using the DA-FMO-based interaction energies. Additionally, we investigated the use of the DI energy term of PIEDA alone to predict the binding affinities of ligands with different net charges, without additional computational costs.

II. MATERIALS AND METHODS

A. Target system

CDK2 was used as the target system (Figure 1) to assess the DA-FMO-based approach. CDK2 belongs to the serine/threonine-protein kinase family, which controls the transition between phases during a cell cycle progression. CDK2 activity and regulation anomalies are critically associated with tumor growth, rendering CDK2 a candidate in the development of anticancer drugs.[19-21] Six CDK2 inhibitors with experimental binding free energies (ΔG_{exp}) ranging from −4.88 to −10.00 kcal/mol (Figure 1) were obtained from an earlier work by Araki *et al.*[18] The experimentally determined binding data (the dissociation constant, K_{d}, and inhibition constant, K_{i}), available in the Community Structure–Activity Resource (CSAR) database (http://www.csardock.org/),[22] were converted to binding free energy using

$\Delta G_{\text{exp}} = -RT \ln \left(1/K_{\text{d,i}} \right)$, where R is the gas constant and T is the absolute temperature. All inhibitors were expected to possess zero net charge under physiological conditions and were treated as electrically neutral ligands throughout the present study.

B. System preparation and MD simulation

The MD trajectories of the six CDK2 inhibitors were provided by Araki *et al.*[18] The X-ray crystal structure of each complex was obtained from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) (4FKL, 4FKI, 4FKQ, 4FKR, 4FKS, and 4FKW).[23] The system preparation procedures are summarized below, as described in a previous study.[18] Disordered loops and flexible side chains were modeled using the MODELER program,[24] and the structure preparation module was implemented using the Molecular Operating Module (MOE, Chemical Computing Group, Montreal, Canada). The dominant protonation state was assigned at pH of 7.0 for the titratable residues. Each system was solvated with a TIP3P[25] water model, and counter ions were added to electrically neutralize the system. The Amber ff99 force field[26] was used for the proteins and ions. Ligand parameters were determined by the general Amber force field (GAFF)[27] using the antechamber module of AmberTools12. The atomic partial charges of the ligand were obtained by the restrained electrostatic potential (RESP) model,[28] using the pre-optimized structures calculated by the General Atomic and Molecular Electronic Structure System (GAMESS) program[29] with the Hartree–Fock (HF)/6-31G* method. Afterward, the system was energetically minimized and equilibrated for 100 ps under constant temperature and volume (NVT), followed by 100 ps under constant pressure (NPT). Positional restraints were imposed on heavy protein and ligand atoms during the minimization and equilibration steps. Finally, five sets of 50 ns of production runs were performed under NPT conditions without positional restraints. The temperature and pressure were maintained at 298 K using the Nose–Hoover thermostat[30,31] and at 1 bar with the Berendsen barostat,[32] respectively. MD simulations were conducted under periodic boundary conditions, where the electrostatic interactions were calculated using the particle mesh Ewald (PME) method,[33] implemented in the Groningen Machine for Chemical Simulation (GROMACS) 4 program.[34] The trajectories obtained by the production run were used to calculate the DA-FMO-based interaction energies of the protein–ligand complex system. The MD snapshot structures were energetically optimized with the classical MM force field, prior to the FMO calculations.

C. Dynamical averaging of FMO interaction energy

MM-optimized MD snapshots were used to evaluate the DA-FMO-based interaction energies. First, the snapshots were extracted every 2 ns at and after 12 ns of the 50 ns production run, which yielded 100 MD snapshots (20 snapshots × 5 trials) for each protein–ligand complex system. All snapshots were energy-minimized using the MOE (ver. 2018.01) with an AMBER:10EHT force field to relieve the steric clashes, prior to the FMO calculations using Auto-FMO protocol,[35] which is an automated pretreatment protocol developed with the BIOVIA Pipeline Pilot.[36] The heavy protein atoms were fixed, and the heavy atoms of ligand and water molecules were positionally restrained with a force of 1.0 kcal/mol/Å, while all the hydrogen atoms were unrestrained. The restraint weight was defined such that the binding poses of the ligands did not deviate excessively far from their initial positions. All counter ions were removed, and water molecules within 6 Å of the protein–ligand complex system were included in the minimization step. Similar to the DNA hydration system we previously studied[37], water molecules within 6 Å of the complex structure were able to account for the solvation effect (see Table S1 and Figure S1). The MM-optimized structures of the MD snapshots were applied as input structures for the FMO calculations using the FMO2-MP2/6-31G* method[38,39]. FMO calculations were performed with the ABINIT-MP program of the Mizuho BioStation version 3.0 [40,41] using the K supercomputer[42] and other Japanese HPCI systems such as TSUBAME3.0 and Oakforest-PACS. The FMO-based inter-fragment interaction energy (IFIE), ΔE^{FMO}, at the MP2 level is given by[40, 41]

$$\Delta E_{IJ}^{\mathrm{FMO}} = \Delta E_{IJ}^{\mathrm{ES}} + \Delta E_{IJ}^{\mathrm{EX}} + \Delta E_{IJ}^{\mathrm{CT+mix}} + \Delta E_{IJ}^{\mathrm{DI}}, \qquad (1)$$

where $\Delta E_{IJ}^{\mathrm{FMO}}$ is the IFIE of the fragment pair *IJ*, while ΔE^{ES}, ΔE^{EX}, $\Delta E^{\mathrm{CT+mix}}$, and ΔE^{DI} represent the electrostatic, exchange repulsion, charge transfer with high-order mixed terms, and DI energy terms, obtained by PIEDA. The DA-FMO-based interaction energy of a ligand is evaluated by

$$\langle \Delta E^{\mathrm{FMO}} \rangle = \langle E^{\mathrm{FMO}} \rangle_{\mathrm{pl}} + \langle E^{\mathrm{FMO}} \rangle_{\mathrm{wl}} = \frac{1}{\mathrm{N}} \sum_{i=1}^{\mathrm{N}} \left(\Delta E_{i}^{\mathrm{FMO}} \right), \qquad (2)$$

where *i* and N denote the indices of the MD snapshots and the total number of MD snapshots, respectively. $\langle \Delta E^{\mathrm{FMO}} \rangle$ includes the IFIEs of protein–ligand $\langle E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ and water–ligand $\langle E^{\mathrm{FMO}} \rangle_{\mathrm{wl}}$ interaction pairs, whereas $\langle E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ only considers the IFIEs of the protein–ligand pairs. Here, the water-protein IFIEs are neglected because the protein (CDK2) is common for all the ligands and its interaction energy with water would be irrelevant for the binding of ligand in the present case. The FMO results obtained from a single structure can be distinguished from the DA-FMO results, where the FMO-based interaction energies in the former are expressed without brackets.

III. RESULTS AND DISCUSSION

A. Correlation between experimental binding free energies and DA-FMO-based interaction energies

The DA-FMO-based interaction energies were assessed in terms of the coefficient of determination (R^2) and root-mean-square error (RMSE) with respect to the experimental binding free energies,

ΔG_{exp}. The DA-FMO-based interaction energies of $\langle \Delta E^{\mathrm{FMO}} \rangle$ and $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ both showed an $R^2 >$ 0.9, whereas the result obtained using MM-optimized X-ray crystal structures, $\Delta E_{\mathrm{pl}}^{\mathrm{FMO}}$, was $R^2 = 0.75$ (Figure 2). This result suggests that averaging FMO-based interactions over multiple structures generated by MD simulations can offer improvement in predicting binding affinities, compared to those considering a single structure. It is also noted that the high correlations between the experimental binding free energy and the computational binding energy observed in this study may be due to the mutual cancellations of entropic, desolvation and ligand-distortion effects for the binding of similar ligands to a common protein.

Interestingly, $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ showed slightly better results than $\langle \Delta E^{\mathrm{FMO}} \rangle$. The R^2 and RMSE (in units of kcal/mol) were 0.99 and 0.16, respectively, for $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$; the corresponding values for $\langle \Delta E^{\mathrm{FMO}} \rangle$ were 0.91 and 0.52, respectively. $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ was successful in ranking all six inhibitors, according to their binding free energies, even though the ΔG_{exp} of the four most potent inhibitors differed only by approximately 1.5 kcal/mol. In contrast, $\langle \Delta E^{\mathrm{FMO}} \rangle$ ranked the four most potent inhibitors less accurately. While the sampling of phase space was limited by the simulation time, water molecules were expected to be highly mobile and more difficult to sample than the protein residues near the binding site. The exclusion of the water–ligand interaction pairs from the FMO-based interaction energy may have systematically reduced the noise associated with insufficient sampling. Preferably, the effect of water molecules could be excluded, except for those that are highly oriented and play a critical role in protein–ligand interactions, when averaging the FMO interaction energies.

On the other hand, the presence of water affects the interaction energy distribution. Figure S3 shows the distribution of IFIE/PIEDA energies of the complex structures sampled every 2 ns. Comparing $\langle E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ plus $\langle E^{\mathrm{FMO}} \rangle_{\mathrm{wl}}$ with water and $\langle E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ only without water, the total interaction energies were stabilized by about 40 kcal/mol (with the exception of CS4, whose ligand has a charge). For each component, the ES, CT+mix, and DI interactions tend to be more stable, and the steric repulsion represented by the EX term tends to be larger.

Next, we investigated the effects of the extraction numbers of MD snapshots and individual trials of MD simulations on the performance of $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$. The results of $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ were compared between those evaluated by MD snapshots that were extracted every 2 ns ($\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl,2ns}}$) and 4 ns ($\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl,4ns}}$). The R^2 and RMSE of $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl,4ns}}$ were 0.99 and 0.19, respectively, which is comparable with those of $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl,2ns}}$ ($R^2 = 0.99$ and RMSE = 0.16). The R^2 and RMSE fell in ranges of 0.95–0.99 and 0.16–0.38, respectively, depending on the MD simulation trials (Table 1). This suggests that averaging over multiple structures will probably be more affected by the distribution of structural conformations, rather than the number of snapshots.

B. Application of DI energy in PIEDA to comparison among ligands with different net charges

An additional study was performed to investigate the use of the DI energy term in PIEDA to predict the binding affinities of ligands with different net electric charges. DI energy is the electron correlation energy that describes the weak molecular interactions associated with the electronic polarization and van der Waals interactions, which play important roles in molecular recognition.[40, 41] An additional inhibitor, CS4 (Figure S2), with a net charge of -1 and $\Delta\mathrm{G}_{\mathrm{exp}} = -7.12$ kcal/mol, was added to the original dataset (Figure 1). The correlation between the $\Delta\mathrm{G}_{\mathrm{exp}}$ and DA-FMO-based interaction energy of the DI energy term, $\langle \Delta E_{\mathrm{DI}}^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$, was then investigated. The $\langle \Delta E_{\mathrm{DI}}^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ was evaluated using MD snapshots that were extracted every 2 ns from the production run.

The correlation plots for the $\Delta\mathrm{G}_{\mathrm{exp}}$ and the FMO-based interaction energies of $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ and $\langle \Delta E_{\mathrm{DI}}^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ are shown in Figure 3. The FMO-based interaction energy, $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$, of CS4 was lower by few tens of kcal/mol than those of the other inhibitors because of the strong electrostatic interactions in the protein–ligand complex and behaved as an outlier, as expected. The R^2 and RMSE of $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ were 0.32 and 1.37, respectively, indicating difficulty in directly comparing ligands with different net charges. In contrast, $\langle \Delta E_{\mathrm{DI}}^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ showed a significant correlation ($\mathrm{R}^2 = 0.99$), suggesting that the DI energy term of PIEDA can be used to predict the binding affinities of ligands with different net charges. The exclusion of the ES energy term from the FMO-based interaction energy may have implicitly accounted for the desolvation effect to some extent because of the offset relationship between the ES and desolvation energies.[41] The EX and CT+mix energy terms can be excluded, as these two energies are related to the short-range repulsive and attractive forces of the occupied orbitals, which tend to have offset relationships with each other.

In addition, we performed a dynamic averaging of the MM-PBSA energies with the same structures used in the FMO calculations. The MM-PBSA energies were calculated using mmpbsa.py,[44] implemented in AmberTools14, with an ionic strength of 150 mM. The Amber FF14SB force field[45] was used for the protein, and GAFF parameters were employed for the ligands with an empirical charge model, AM1-BCC, using the ANTECHAMBER AmberTools14 program. Furthermore, the FMO-based interaction energies of $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ were combined with the solvation energies obtained by PBSA calculations (FMO+MM-PBSA). The R^2 and RMSE evaluated by MM-PBSA were 0.60 and 1.05, respectively, while the corresponding values for FMO+MM-PBSA were 0.73 and 0.86, respectively. Although MM-PBSA and FMO+MM-PBSA showed moderate correlations with $\Delta\mathrm{G}_{\mathrm{exp}}$, the DA-FMO-based interaction energy of the DI energy term, $\langle \Delta E_{\mathrm{DI}}^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$, showed the best correlation ($\mathrm{R}^2 = 0.99$), among the methods examined in this study.

C. Interaction energy analysis between CDK2 and ligands

In order to understand the ligand binding properties of CDK2, we analyzed the interaction energies between the ligand and the surrounding residues in each complex. The results of PIEDA on the X-ray crystal structures (initial structures of MD calculation) are shown in Figures 4-5 (see also Figures S3 and S4). Figure 4a to Figure 4f are arranged in order of increasing ΔG_{exp}, showing that the number of interactions acquired increases as ΔG_{exp} is stabilized. A detailed analysis of CS262, which has the highest number of interactions and the largest absolute value of ΔG_{exp}, is shown in Figure 5. As shown in Figure 5a, the crystal structure suggests that CS262 forms hydrogen bonds with Lys33, Glu81, Leu83, Gln85, Asp86, and Lys89. The IFIE/PIEDA results (Fig. 5b-e) show that all of these residues have ES and CT+mix terms, indicating that they are hydrogen-bonded, and in particular, Asp86, followed by Lys89, Lys33, and Phe82 are strongly hydrogen-bonded. Note that the C=O group of Glu81 backbone belongs to the Phe82 fragment.[41] In addition, five residues, Ile10, Val18, Phe80, Leu134 and Asp145, whose main IFIE component is DI (-4.7~-5.6 kcal/mol), are located around the indole ring of CS262, indicating many CH/π interactions and dispersion interactions between hydrophobic groups. Structural analysis using the CHPI program[47] build into BioStation Viewer[48] showed that the indole ring of CS262 ligand is sandwiched between two CH bonds each with Leu134 and Ala31 and one CH bond each with Val18 and Ala144, all of which have CH/π interactions. The benzene ring of CS262 had CH/π interactions with Gln85 and Asp86, and their DI components were also substantial at -6.8 and 8.0 kcal/mol, respectively. Many of these DI energies vary with the induced fit of the ligand structure and are less affected by the charge distribution, which can be an effective indicator for predicting ligand binding affinity, as shown in Figure 3.

Comparing the correlation plots in Figs. 2 and 3 with the PIEDA results in Fig. 4, CS242 appeared to be over-stabilized as seen from PIEDA in Fig. 4c and was also an outlier in Fig. 2a, but was greatly improved in the DA-FMO results (Fig. 2b). Similarly, the relationship between the FMO energies of CS245 and CS246 was reversed in the crystal structure, but improved by DA-FMO. These facts suggest that the complex structure, which is over-stabilized by crystal packing and other effects, is relaxed by MD calculation, and that DA-FMO calculation is a more accurate evaluation method that reproduces the fluctuating structure in solution. The FMO calculation results of these X-ray crystal structures have already been registered in the FMO database (FMODB),[35,46] where simple analyses can be performed through the web interface and all calculation data files can be downloaded. The FMODB IDs of these FMO results are given in Figure 4. In addition, the FMO calculation results for MD snapshots have also been registered in the FMODB (see also Supplementary Material).

IV. CONCLUSION

The experimental binding free energies, ΔG_{exp}, of CDK2 inhibitors were compared with those from a DA-FMO-based approach in which MD simulations were used to generate multiple protein–ligand complex structures for FMO calculations. The DA-FMO-based approach was assessed using six CDK2 inhibitors whose net charges were expected to be zero under physiological conditions. The DA-FMO-based interaction energies of the protein–ligand complexes, $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$, showed higher correlations ($R^2 = 0.99$) with ΔG_{exp} than the correlations of those obtained with the MM-optimized X-ray crystal structure ($R^2 = 0.75$). This result suggests that the DA-FMO-based approach can further improve the accuracy of protein–ligand binding affinity predictions, compared to FMO-based methods that consider only a single structure. The DA-FMO-based approach showed a slightly higher R^2 value by excluding the water–ligand interaction pairs. The exclusion of the contribution of water molecules may have systematically reduced the noise because of insufficient sampling for water to an extent. The R^2 value was not significantly affected by the number of snapshots used to evaluate the interaction energies by the DA-FMO-based approach, indicating that averaging over multiple structures will probably be more affected by the distribution of the structural conformations, rather than the number of snapshots. Additionally, it was observed that the DI energy term of PIEDA can be used alone to better predict the binding affinities of ligands with different net charges. The DA-FMO-based interaction energies of the DI energy term, $\langle \Delta E_{\mathrm{DI}}^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$, achieved $R^2 = 0.99$, whereas R^2 decreased to 0.32, with $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$, which considered all four energy terms in PIEDA. The exclusion of the electrostatic and other energy terms of PIEDA may have implicitly accounted for the desolvation effect to an extent.

The interaction energy analysis of CDK2 with ligand also showed that the distorted interaction with surrounding residues in the X-ray crystal structure was improved by DA-FMO analysis. A large number of CH/π and DI interactions between CDK2 and the aromatic rings of the ligand, such as benzene and indole rings, were observed, which may be responsible for the high correlation between $\langle \Delta E_{\mathrm{DI}}^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ and ΔG_{exp}.

This is a case study in which the FMO-IFIE analysis has been carried out only for one protein (CDK2) and seven ligand compounds, thus demanding future investigations for more general cases. The DA-FMO-based approach can be computationally expensive if the FMO calculation is performed for numerous structures and ligands. To overcome this difficulty, a reinforcement learning-based approach called multiarmed bandits[49-51] can be used to allocate a limited amount of computational resources to more promising ligands. The ligands and FMO-based interaction energies are regarded as arms and rewards for the original bandit problems in this approach.

SUPPLEMENTARY MATERIAL

See the Supplementary Material for additional information on ligand binding energy, water shell, charged ligand and PIEDA.

DATA AVAILABILITY STATEMENTS

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

ACKNOWLEDGEMENTS

The authors thank Prof. Yuji Mochizuki at Rikkyo University, Dr. Tatsuya Nakano and Dr. Yoshio Okiyama at the National Institute of Health Sciences (NIHS) for general discussions related to FMO calculations, as well as Dr. Daisuke Takaya and Dr. Kikuko Kamisaka at RIKEN for supporting FMODB registration. In particular, we thank Prof. Mochizuki for providing the Oakforest-PACS version of ABINIT-MP and many advices on its use on the supercomputers. This research was done in activities of the FMO drug design consortium (FMODD). The results were obtained using the K computer (project ID: hp190119), TSUBAME3.0 (projectID: hp190133), and Oakforest-PACS (projectID: hp200101). This research was partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP21am0101113 as well as JSPS Kakenhi (JP17H06353, JP18K03825, and JP21K06098) and MEXT Quantum Leap Flagship Program (MEXT QLEAP) Grant Number JPMXS0120330644.

REFERENCES

1. K. Kitaura, E. Ikeo, T. Asada, T. Nakano, M. Uebayasi, Chem. Phys. Lett. 1999, 313, 701.
2. K. Kitaura, T. Sawai, T. Asada, T. Nakano, M. Uebayasi, Chem. Phys. Lett. 1999, 312, 319.
3. T. Nakano, T. Kaminuma, T. Sato, K. Fukuzawa, Y. Akiyama, M. Uebayasi, K. Kitaura, Chem. Phys. Lett. 2002, 351, 475.
4. D. G. Fedorov, K. Kitaura, J. Comput. Chem. 2007, 28, 222.
5. A. Wlodawer, J. Vondrasek, Annu. Rev. Biophys. Biomol. Struct. 1998, 27, 249.
6. S. K. Bagal, M. Andrews, B. M. Bechle, J. Bian, J. Bilsland, D. C. Blakemore, J. F. Braganza, P. J. Bungay, M. S. Corbett, C. N. Cronin, J. J. Cui, R. Dias, N. J. Flanagan, S. E. Greasley, R. Grimley, K. James, E. Johnson, L. Kitching, M. L. Kraus, I. McAlpine, A. Nagata, S. Ninkovic, K. Omoto, S. Scales, S. E. Skerratt, J. Sun, M. Tran-Dubé, G. J. Waldron, F. Wang, J. S. Warmus, J. Med. Chem. 2018, 61, 6779.
7. M. Atobe, T. Serizawa, N. Yamakawa, K. Takaba, Y. Nagano, T. Yamaura, E. Tanaka, A. Tazumi, S. Bito, M. Ishiguro, M. Kawanishi, J. Med. Chem. 2020, 63, 7143.
8. K. Fukuzawa, K. Kitaura, M. Uebayasi, K. Nakata, T. Kaminuma, T. Nakano, J. Comput. Chem. 2005, 26, 1.
9. A. Heifetz, E. I. Chudyk, L. Gleave, M. Aldeghi, V. Cherezov, D. G. Fedorov, P. C. Biggin, M. J. Bodkin, J. Chem. Inform. Model. 2016, 56, 159.
10. C. Watanabe, H. Watanabe, K. Fukuzawa, L. J. Parker, Y. Okiyama, H. Yuki, S. Yokoyama, H. Nakano, S. Tanaka, T. Honma, J. Chem. Inform. Model. 2017, 57, 2996.
11. C. Bissantz, B. Kuhn, M. Stahl, J. Med. Chem. 2010, 53, 5061.
12. T. Ishikawa, K. Kuwata, J. Chem. Theory Comput. 2010, 6, 538.
13. D. G. Fedorov, K. Kitaura, J. Phys. Chem. A 2018, 122, 1781.
14. R. Hatada, K. Okuwaki, K. Akisawa, Y. Mochizuki, Y. Handa, K. Fukuzawa, Y. Komeiji, Y. Okiyama, S. Tanaka, Appl. Phys. Express 2021, 14, 027003.
15. S. Tanaka, S. Tokutomi, R. Hatada, K. Okuwaki, K. Akisawa, K. Fukuzawa, Y. Komeiji, Y. Okiyama, Y. Mochizuki, J. Phys. Chem. B 2021, 125, 6501.
16. P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, T. E. Cheatham, Acc. Chem. Res. 2000, 33, 889.
17. J. Tomasi, B. Mennucci, R. Cammi, Chem. Rev. 2005, 105, 2999.
18. M. Araki, N. Kamiya, M. Sato, M. Nakatsui, T. Hirokawa, Y. Okuno, J. Chem. Inform. Model. 2016, 56, 2445.

19. J. Bai, Y. Li, G. Zhang, Cancer Biol. Med. 2017, 14, 348.

20. S. Tadesse, E. C. Caldon, W. Tilley, S. Wang, J. Med. Chem. 2019, 62, 4233.

21. M.P. Mazanetz, O. Ichihara, R.J. Law, M. Whittaker, J. Cheminf. 2011, 3, 2.

22. J. B. Dunbar, R. D. Smith, C. Y. Yang, P. M.-U. Ung, K. W. Lexa, N. A. Khazanov, J. A. Stuckey, S. Wang, H. A. Carlson, J. Chem. Inform. Model. 2011, 51, 2036.

23. S. K. Burley, C. Bhikadiya, C. Bi, S. Bittrich, L. Chen, G. V. Crichlow, C. H. Christie, K. Dalenberg, L. Di Costanzo, J. M. Duarte, S. Dutta, Z. Feng, S. Ganesan, D. S. Goodsell, S. Ghosh, R. K. Green, V. Guranović, D. Guzenko, B. P. Hudson, C. L. Lawson, Y. Liang, R. Lowe, H. Namkoong, E. Peisach, I. Persikova, C. Randle, A. Rose, Y. Rose, A. Sali, J. Segura, M. Sekharan, C. Shao, Y. P. Tao, M. Voigt, J. D. Westbrook, J. Y. Young, C. Zardecki, M. Zhuravleva, Nucl. Acids Res. 2021, 49, D437.

24. A. Šali, T. L. Blundell, J. Mol. Biol. 1993, 234, 779.

25. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, J. Chem. Phys. 1983, 79, 926.

26. J. Wang, P. Cieplak, P. A. Kollman, J. Comput. Chem. 2000, 21, 1049.

27. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, J. Comput. Chem. 2004, 25, 1157.

28. C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, J. Phys. Chem. 1993, 97, 10269.

29. M. W. Schmidt, K. K. Baldridge, J. A, Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery Jr., J. Computat. Chem. 1993, 14, 1347.

30. S. Nosé, Mol. Phys. 1984, 52, 255.

31. W. G. Hoover, Phys. Rev. A 1985, 31, 1695.

32. H. J. C. Berendsen, Postma, J. P. M.; W. F. van Gunsteren, A. DiNola, J. R. Haak, J. Chem. Phys. 1984, 81, 3684.

33. T. Darden, D. York, L. Pedersen, J. Chem. Phys. 1993, 98, 10089.

34. B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, J. Chem. Theory Comput. 2008, 4, 435.

35. C. Watanabe, H. Watanabe, Y. Okiyama, D. Takaya, K. Fukuzawa, S. Tanaka, T. Honma, Chem-Bio Inform. J. 2019, 19, 5.

36. D. S. Biovia, 2017 (BIOVIA Pipeline Pilot, 17.2.0.1361; Dassault Systèmes BIOVIA: 5005 Wateridge Vista Drive,San Diego, CA 92121 USA, 2017).

37. K. Fukuzawa, I. Kurisaki, C. Watanabe, Y. Okiyama, Y. Mochizuki, S. Tanaka, Y. Komeiji, Comput. Theor. Chem. 2015, 1054, 29.

38. Y. Mochizuki, T. Nakano, S. Koikegami, S. Tanimori, Y. Abe, U. Nagashima, K. Kitaura, Theor. Chem. Acc. 2004, 112, 442.

39. Y. Mochizuki, S. Koikegami, T. Nakano, S. Amari, K. Kitaura, Chem. Phys. Lett. 2004, 396, 473.

40 S. Tanaka, Y. Mochizuki, Y. Komeiji, Y. Okiyama, K. Fukuzawa, Phys. Chem. Chem. Phys. 2014, 16, 10310.

41 Y. Mochizuki, S. Tanaka, K. Fukuzawa, ed., Recent Advances of the Fragment Molecular Orbital Method, Springer, Singapore, 2021.

42. J. B. Brown, M. Nakatsui, Y. Okuno, Mol. Inform. 2014, 33, 732.

43. Y. Okiyama, K. Fukuzawa, Y. Komeiji, S. Tanaka, Taking Water into Account with the Fragment Molecular Orbital Method, in "Quantum Mechanics in Drug Discovery", edited by A. Heifetz, Methods in Molecular Biology vol. 2114 (Humana Press, New York, NY, 2020) Chap. 7, 105-122.

44. B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, A. E. Roitberg, J. Chem. Theory Comput. 2012, 8, 3314.

45. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, J. Chem. Theory Comput. 2015, 11, 3696.

46. D. Takaya, C. Watanabe, S. Nagase, K. Kamisaka, Y. Okiyama, H. Moriwaki, H. Yuki, T. Sato, N. Kurita, Y. Yagi, T. Takagi, N. Kawashita, T. Ozawa, M. Takimoto-Kamimura, S. Tanaka, K. Fukuzawa, T. Honma, J. Chem. Inf. Model. 2021, 61, 777.

47. Y. Umezawa, M. Nishio, Bioorg. Med. Chem. 1998, 6, 493.

48. BioStation Viewer; https://fmodd.jp/biostationviewer-dl/

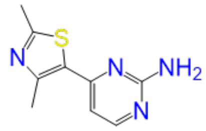
49. P. Auer, N. Cesa-Bianchi, P. Fischer, Machine Learning 2002, 47, 235.

50. J. -Y. Audibert, S. Bubeck, COLT - 23th Conference on Learning Theory -, 2010 Haifa, Israel, 2010-06-27-2010, p. 13.

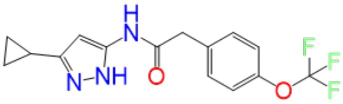
51. V. Gabillon, M. Ghavamzadeh, A. Lazaric, S. Bubeck, In Proceedings of the 24th International Conference on Neural Information Processing Systems; Curran Associates Inc.: Granada, Spain, 2011, 2222-2230.

Table 1. Correlation (R^2) and error (RMSE in units of kcal/mol) between experimental binding free energies and dynamically averaged (DA) FMO-based interactions, $\langle \Delta E \rangle_{\mathrm{pl}}$. Subscripts "2ns" and "4ns" represent the simulation time interval by which the molecular dynamics (MD) trajectories were sampled to calculate $\langle \Delta E \rangle_{\mathrm{pl}}$. Total IFIE energies (in units of kcal/mol) are also shown for comparison.

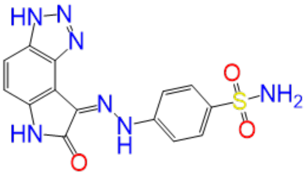
MD index	ALL	MD1	MD2	MD3	MD4	MD5
$\langle \Delta E \rangle_{\mathrm{pl,2ns}}$						
R^2	0.99	0.98	0.99	0.98	0.97	0.98
RMSE	0.16	0.23	0.21	0.25	0.31	0.23
Total IFIE	-604.71	-591.56	-622.09	-616.28	-598.38	-596.44
$\langle \Delta E \rangle_{\mathrm{pl,4ns}}$						
R^2	0.99	0.98	0.97	0.98	0.95	0.98
RMSE	0.19	0.24	0.32	0.24	0.38	0.27
Total IFIE	-603.54	-596.10	-616.99	-616.81	-604.04	-585.20



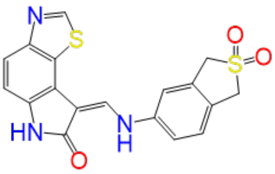
CS12 (PDB ID: 4FKL)
MW = 206.3
ΔG_{exp} = -4.88 [kcal/mol]



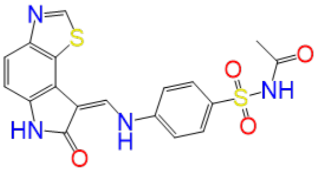
CS9 (PDB ID: 4FKI)
MW = 325.3
ΔG_{exp} = -6.50 [kcal/mol]



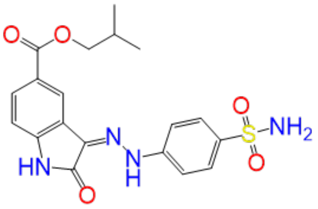
CS242 (PDB ID: 4FKQ)
MW = 357.4
ΔG_{exp} = -8.56 [kcal/mol]



CS245 (PDB ID: 4FKR)
MW = 383.4
ΔG_{exp} = -9.06 [kcal/mol]



CS246 (PDB ID: 4FKS)
MW = 414.5
ΔG_{exp} = -9.15 [kcal/mol]



CS262 (PDB ID: 4FKW)
MW = 416.5
ΔG_{exp} = -10.00 [kcal/mol]

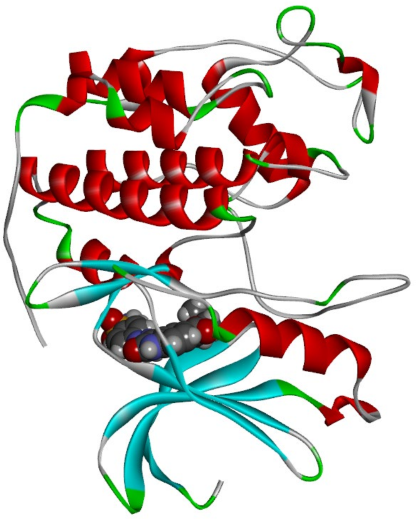


Figure 1. Ligand structures, experimental binding free energies (ΔG_{exp}), and the molecular weights (MWs) of six cyclin-dependent kinase-2 (CDK2) inhibitors. The protein data bank (PDB) codes are shown in parentheses. The complex structure of CDK2 and CS262 is also shown.

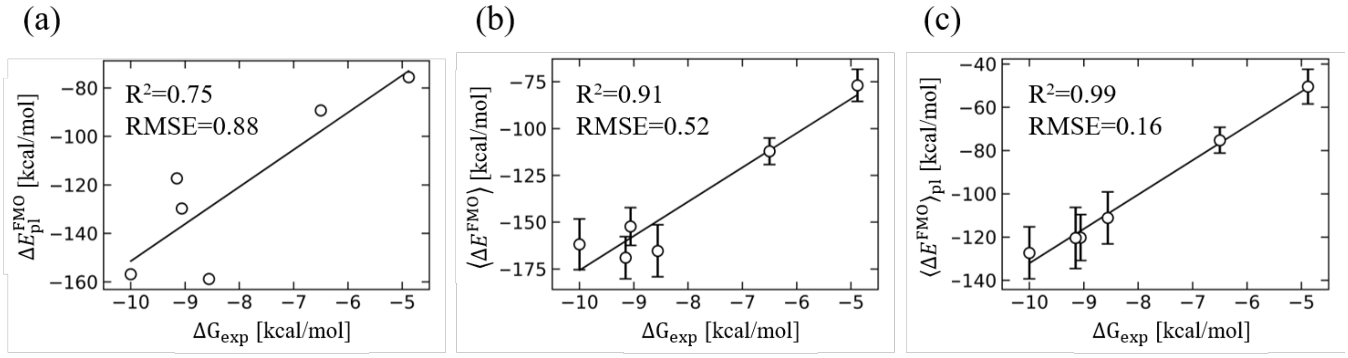


Figure 2. Plots of correlation between experimental binding free energies (ΔG_{exp}) and fragment molecular orbital (FMO)-based interaction energies, where R^2 denotes a squared correlation coefficient. RMSE is also shown in units of kcal/mol. (a) Correlation with FMO-based interaction energies using the molecular mechanics (MM)-optimized X-ray crystal structures, $\Delta E_{\mathrm{pl}}^{\mathrm{FMO}}$, obtained from protein–ligand (pl) interaction pairs. (b) Correlation with dynamically averaged (DA) FMO-based interactions, $\langle \Delta E^{\mathrm{FMO}} \rangle$, obtained from protein–ligand and water–ligand interaction pairs. (c) Correlation with FMO-based interactions, $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$, obtained from protein–ligand interaction pairs.

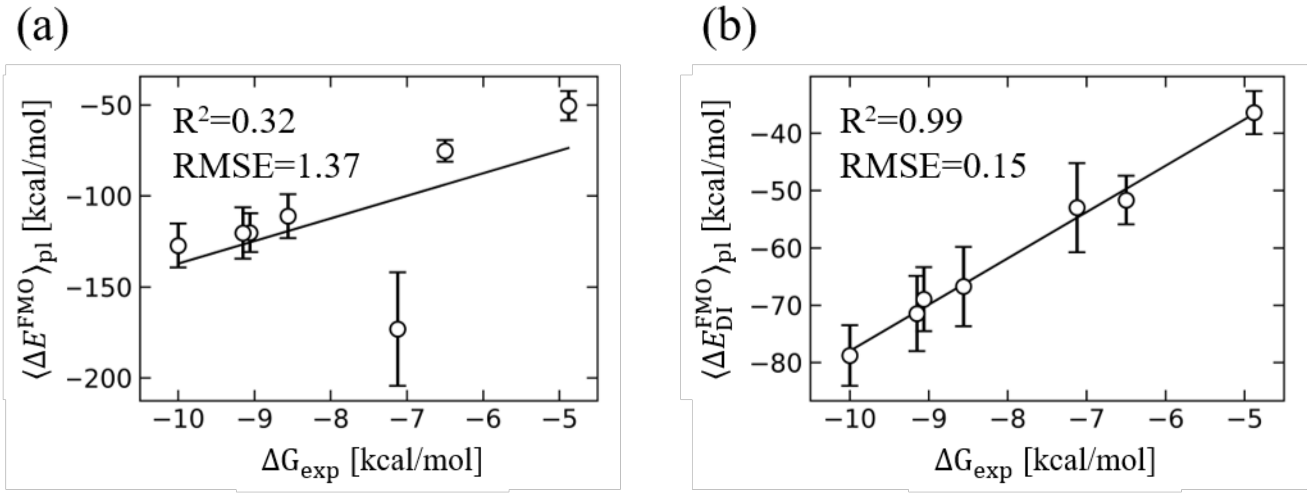


Figure 3. Plots of correlation between experimental binding free energies ($\Delta \mathrm{G}_{\mathrm{exp}}$) and fragment molecular orbital (FMO)-based interaction energies of (a) dynamically averaged (DA) FMO-based interactions of $\langle \Delta E^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ and (b) DA-FMO-based interactions of $\langle \Delta E_{\mathrm{DI}}^{\mathrm{FMO}} \rangle_{\mathrm{pl}}$ with dispersion (DI) energy, including a charged ligand (CS4). The meaning of R^2 and RMSE is the same as in Fig. 2.

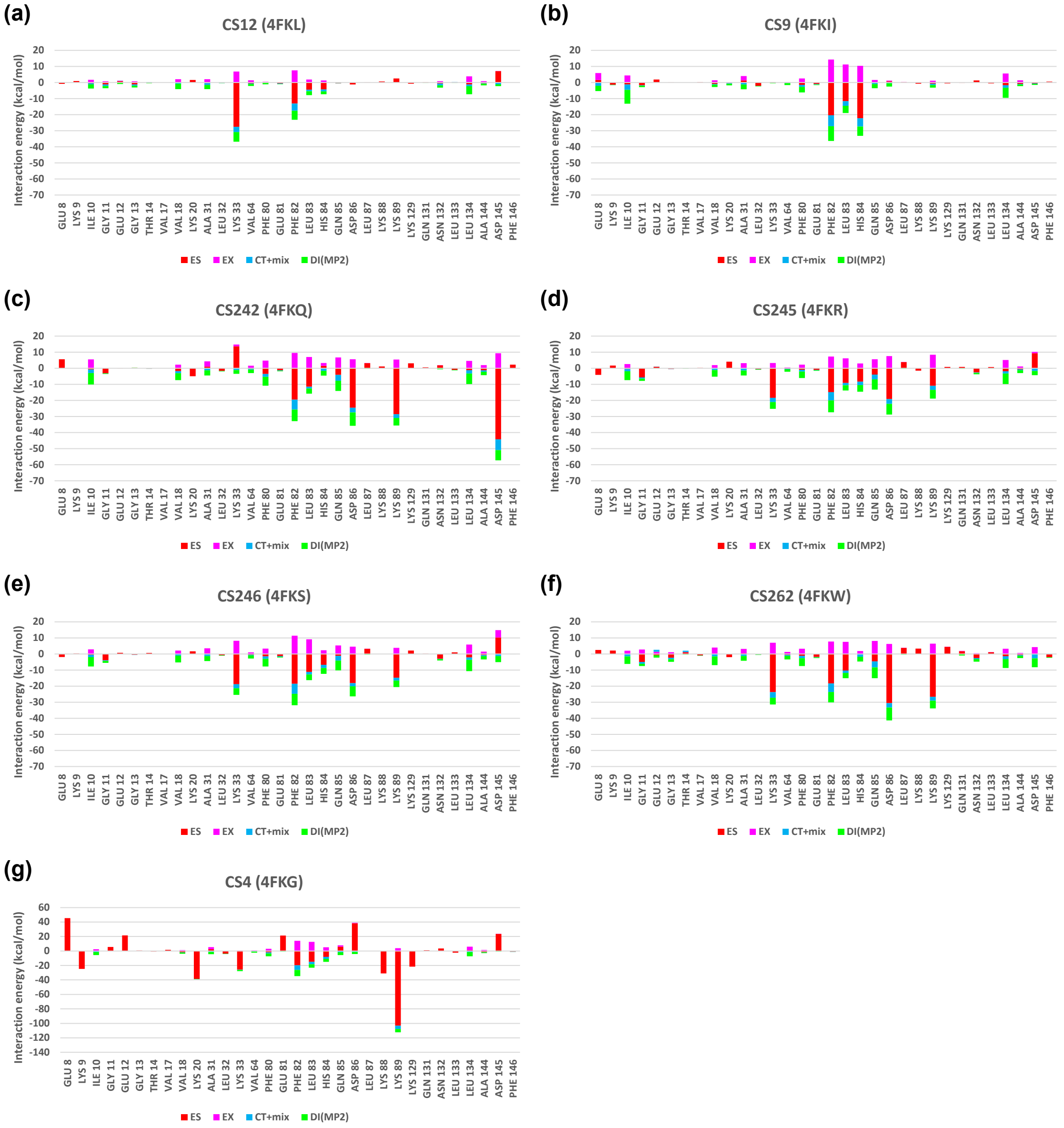


Figure 4. Pair interaction energy decomposition analysis (PIEDA) between ligand and each amino acid residue constituting the ligand-binding pocket for X-ray complex structure. The neutral ligands are (a) CS12 (PDB ID: 4FKL, FMODB ID: 666RZ), (b) CS9 (PDB ID: 4FKI, FMODB ID: R55K8), (c) CS242 (PDB ID: 4FKQ, FMODB ID: 7GGZK), (d) CS245 (PDB ID: 4FKR, FMODB ID: M332Z), (e) CS246 (PDB ID: 4FKS, FMODB ID: 9GG22), and (f) CS262 (PDB ID: 4FKW, FMODB ID: J33R9); and the charged ligand is (g) CS4 (PDB ID: 4FKG, FMODB ID: Q114Y). Here, FMODB means the FMO database.[35,46]

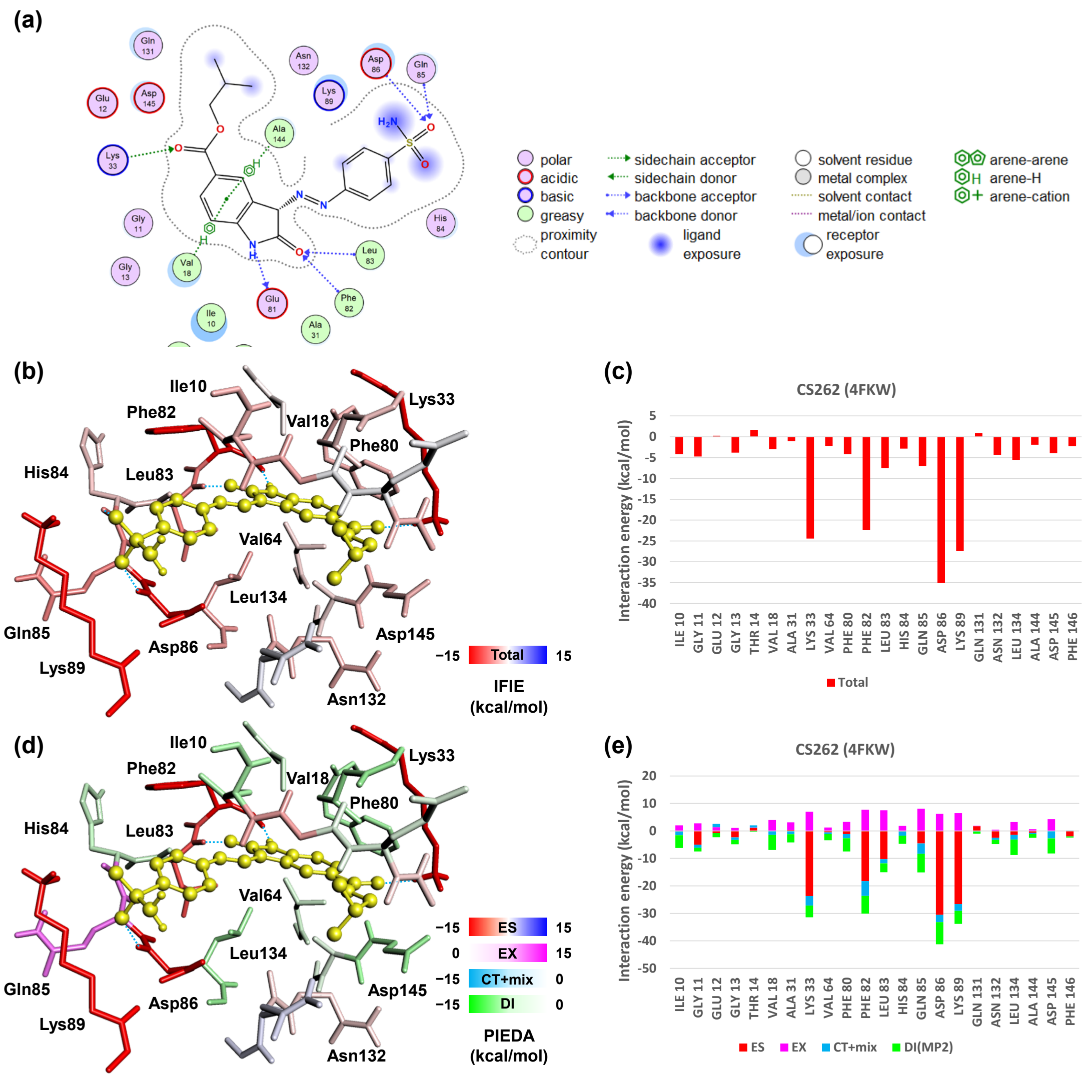
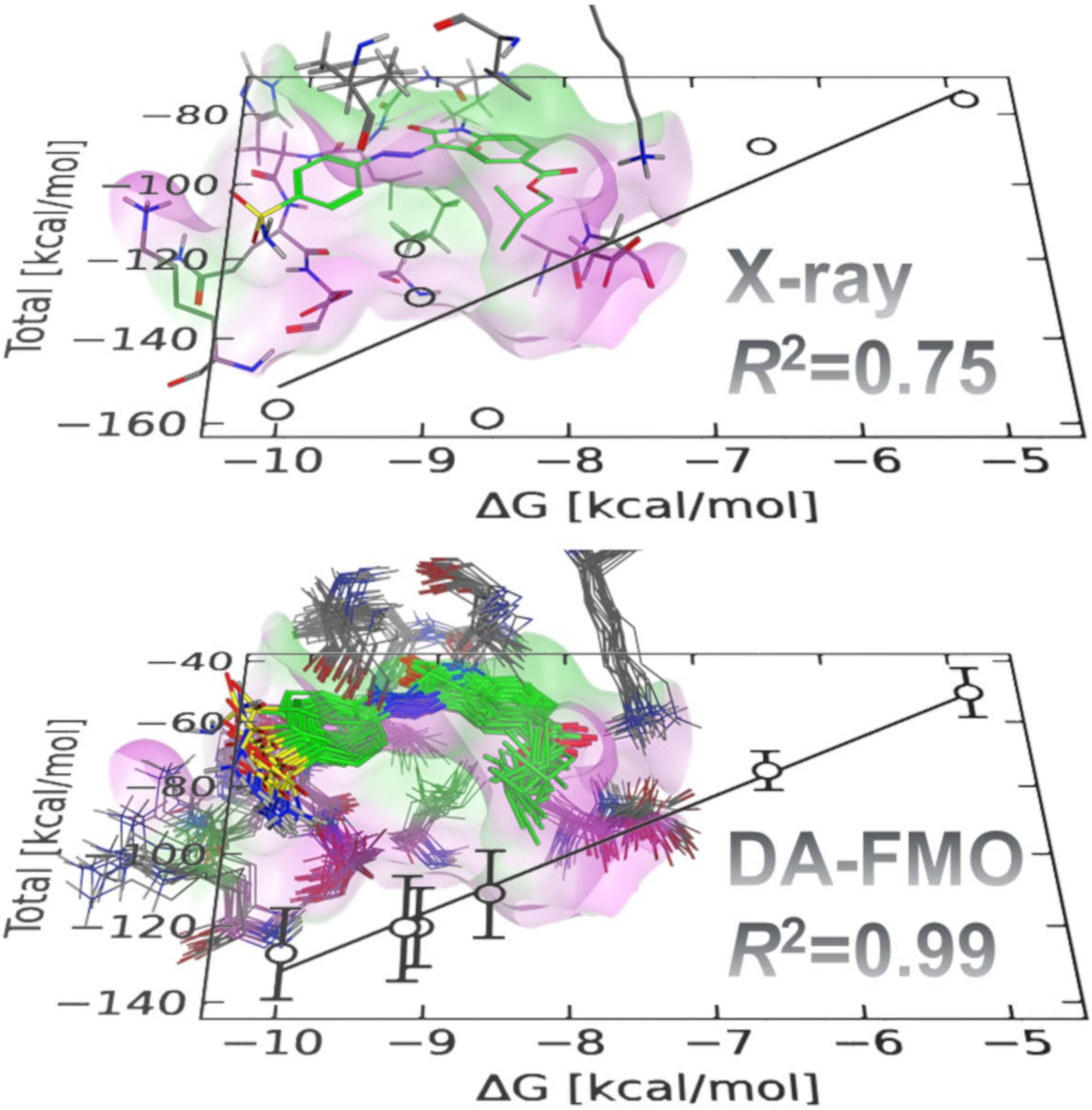


Figure 5. Intermolecular interaction energies between ligand and each amino acid residue constituting the ligand-binding pocket for X-ray complex structure with CS262 (PDB ID: 4FKW, FMODB ID: J33R9). A schematic diagram of the interaction around the ligand drawn by MOE is shown in (a), IFIE result is shown in (b) and (c), and PIEDA result is shown in (d) and (e). The ligand is represented by the yellow ball and stick model. The amino acid residues on CDK2 are represented by stick models that are colored with the gradation of the intensity of IFIE (b) and PIEDA main component (d).

GRAPHICAL TABLE OF CONTENTS



SUMMARY

Experimental binding free energies of CDK2 inhibitors were compared with the binding energies by DA-FMO-based approach in which MD simulations were used to generate multiple protein–ligand complex structures. For six CDK2 inhibitors with zero charges, the DA-FMO-based interaction energies of the protein–ligand complexes showed higher correlations with the experimental values than those calculated with the X-ray crystal structure. The dispersion term of DA-FMO-based interaction energies retained the high correlation even including a charged ligand.

Supplementary Material

Table S1. The $\Delta E_{\mathrm{pl}}^{\mathrm{FMO}}$ between CDK2 and CS4 ligand, number of fragments, and calculation time by water shell size. 256 nodes of the K supercomputer were used for calculations.

Shell size (Å)	3	4	5	6	7
$\Delta E_{\mathrm{pl}}^{\mathrm{FMO}}$(CS4) (kcal/mol)	−209.90	−208.40	−205.49	−204.78	−204.81
No. of fragments	814	1366	1849	2293	2806
Calculation time (h)	8.1	7.6	8.4	10.2	14.4

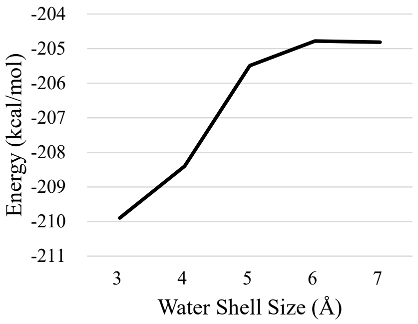
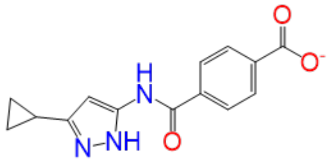
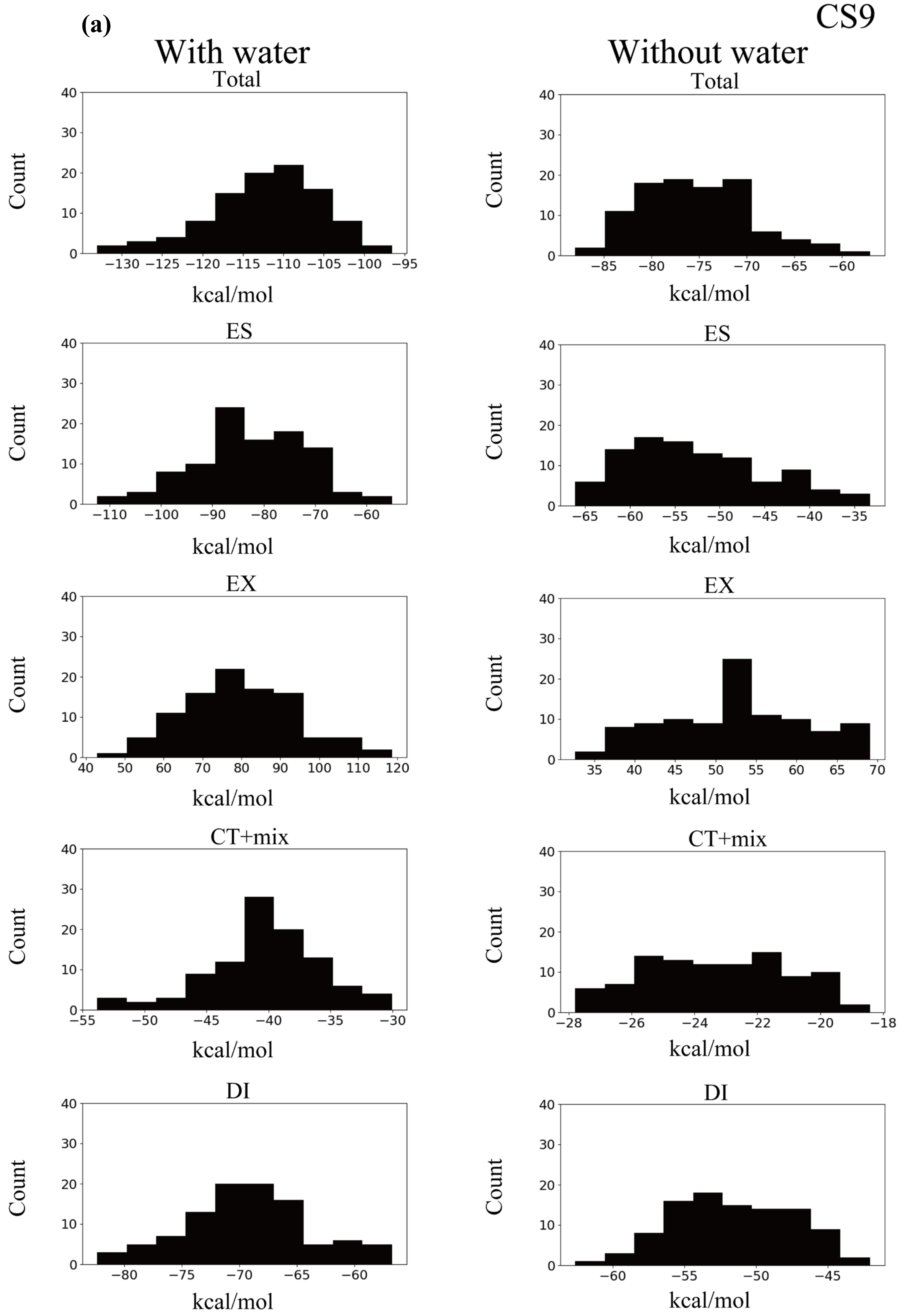


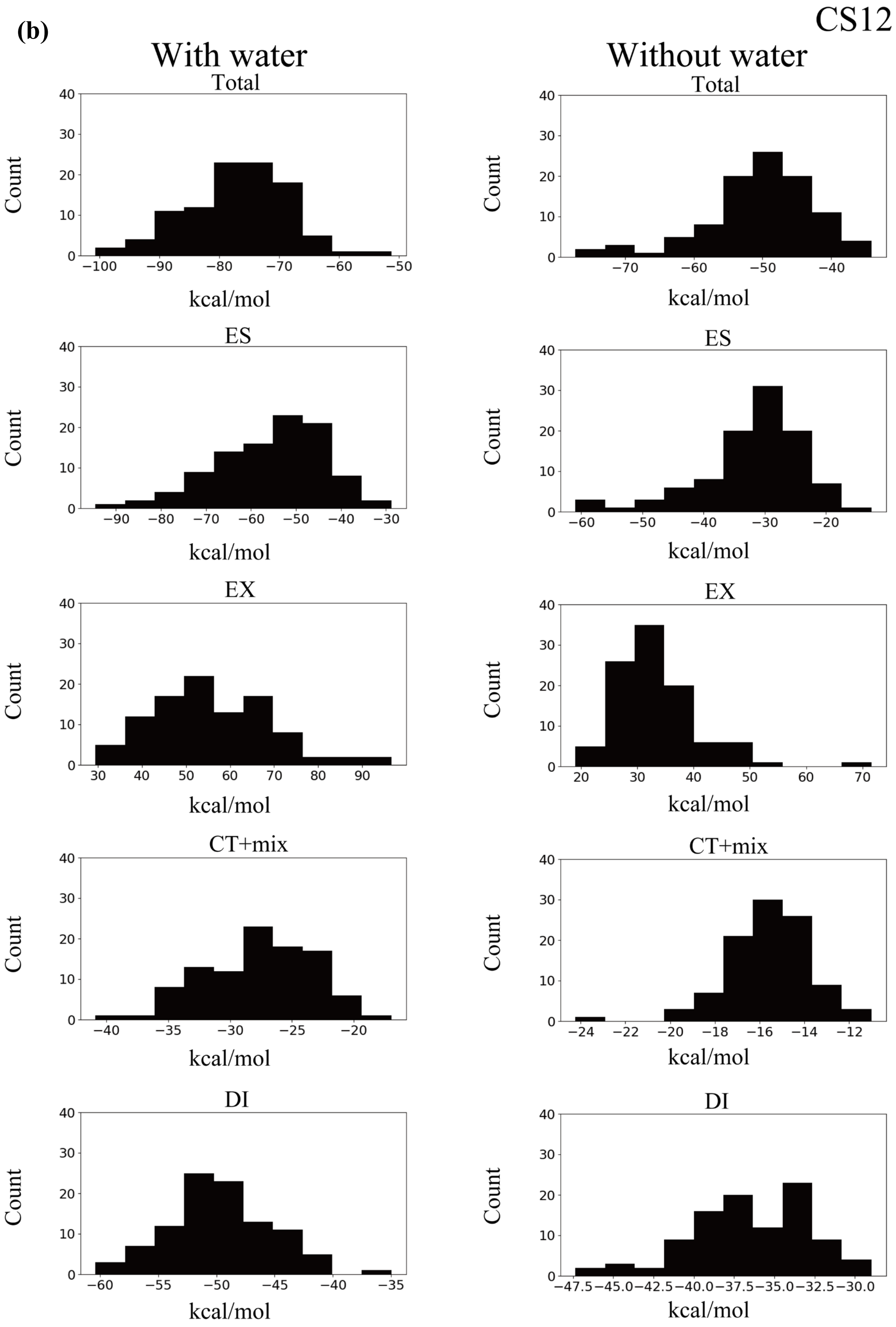
Figure S1. The relationship between water shell size and the $\Delta E_{\mathrm{pl}}^{\mathrm{FMO}}$ between CDK2 and CS4 ligand.

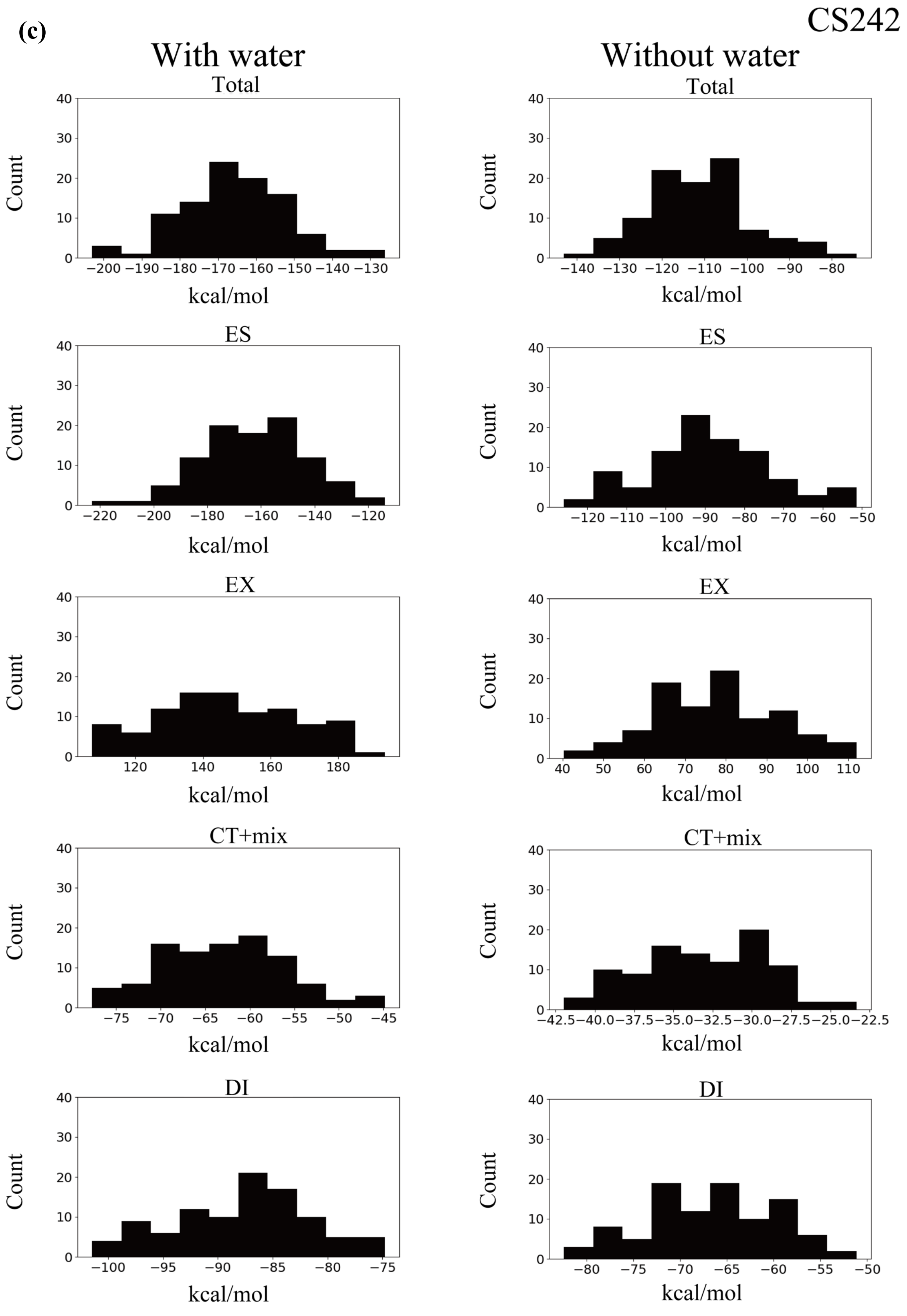


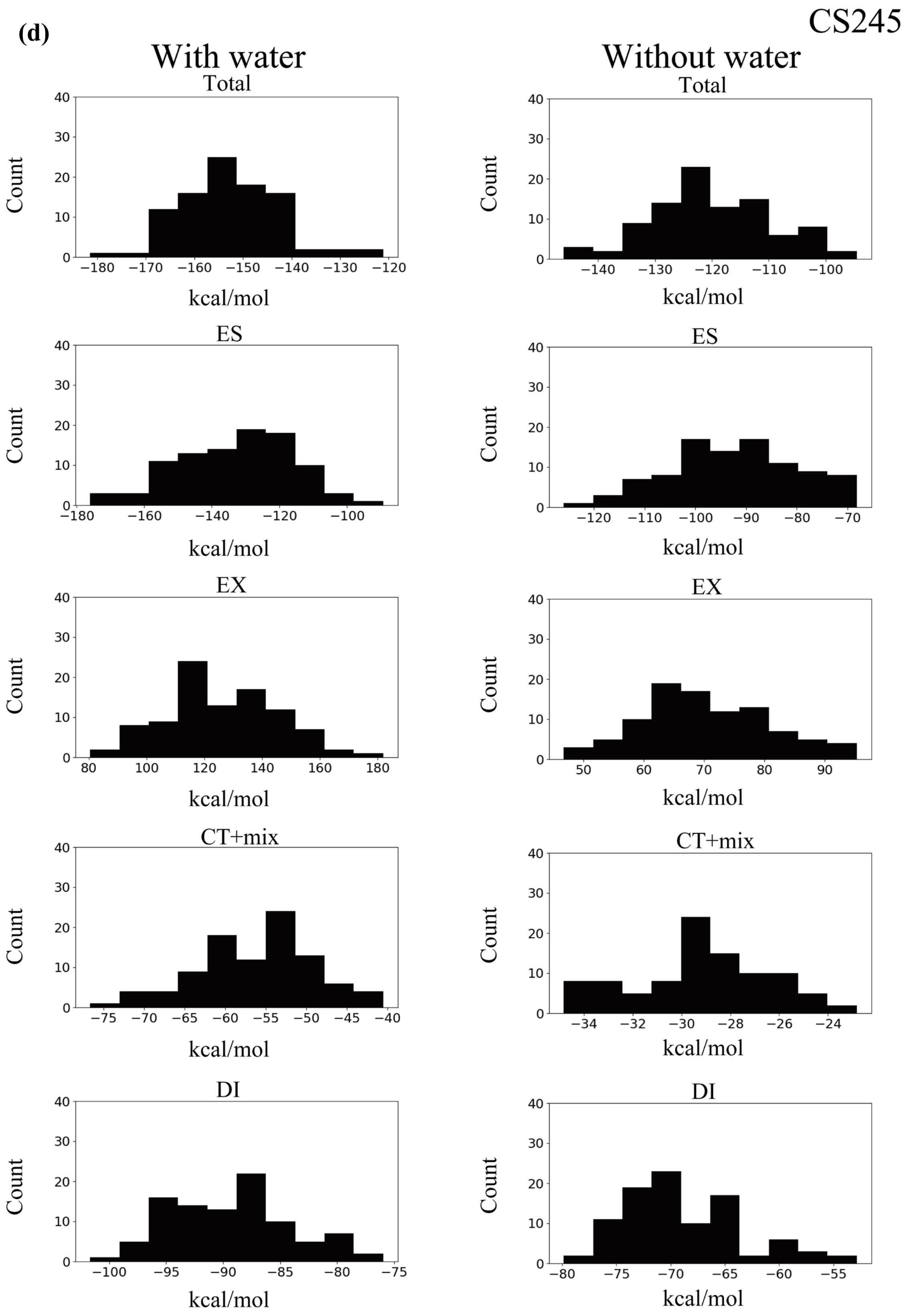
CS4 (PDB ID: 4FKG)
MW = 271.3
ΔG_{exp} = -7.12 [kcal/mol]

Figure S2. Additional ligand (CS4) structure with a net charge of −1. The Protein Data Bank (PDB) code is shown in parentheses.


(a)
CS9
With water
Without water
Total
ES
EX
CT+mix
DI
Count
kcal/mol
0
10
20
30
40
−130 −125 −120 −115 −110 −105 −100 −95
−85 −80 −75 −70 −65 −60
−110 −100 −90 −80 −70 −60
−65 −60 −55 −50 −45 −40 −35
40 50 60 70 80 90 100 110 120
35 40 45 50 55 60 65 70
−55 −50 −45 −40 −35 −30
−28 −26 −24 −22 −20 −18
−80 −75 −70 −65 −60
−60 −55 −50 −45

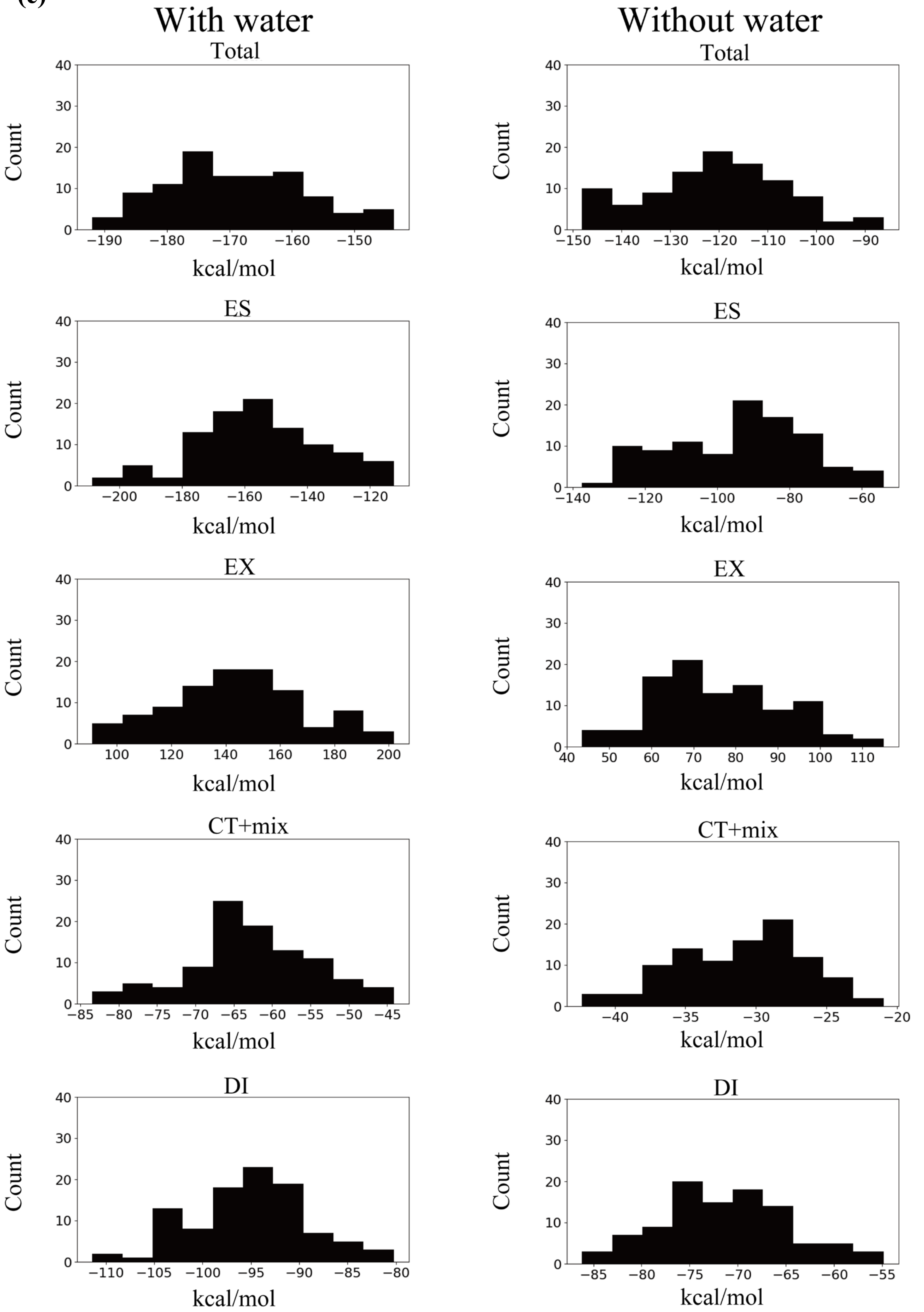

(b)
CS12
With water
Without water
Total
ES
EX
CT+mix
DI
Count
kcal/mol
0
10
20
30
40
−100
−90
−80
−70
−60
−50
−40
−30
30
40
50
60
70
80
90
−40
−35
−30
−25
−20
−60
−55
−50
−45
−35
20
−24
−22
−18
−16
−14
−12
−47.5
−45.0
−42.5
−40.0
−37.5
−35.0
−32.5
−30.0


(c)
CS242
With water
Without water
Total
ES
EX
CT+mix
DI
Count
kcal/mol
0
10
20
30
40
−200 −190 −180 −170 −160 −150 −140 −130
−140 −130 −120 −110 −100 −90 −80
−220 −200 −180 −160 −140 −120
−120 −110 −100 −90 −80 −70 −60 −50
120 140 160 180
40 50 60 70 80 90 100 110
−75 −70 −65 −60 −55 −50 −45
−42.5 −40.0 −37.5 −35.0 −32.5 −30.0 −27.5 −25.0 −22.5
−100 −95 −90 −85 −80 −75
−80 −75 −70 −65 −60 −55 −50


(d)
CS245
With water
Without water
Total
ES
EX
CT+mix
DI
Count
kcal/mol
0
10
20
30
40
−180
−170
−160
−150
−140
−130
−120
−110
−100
−90
−80
−75
−70
−65
−60
−55
−50
−45
−40
−34
−32
−30
−28
−26
−24
50
60
70
80
90
100
120
140
160
180

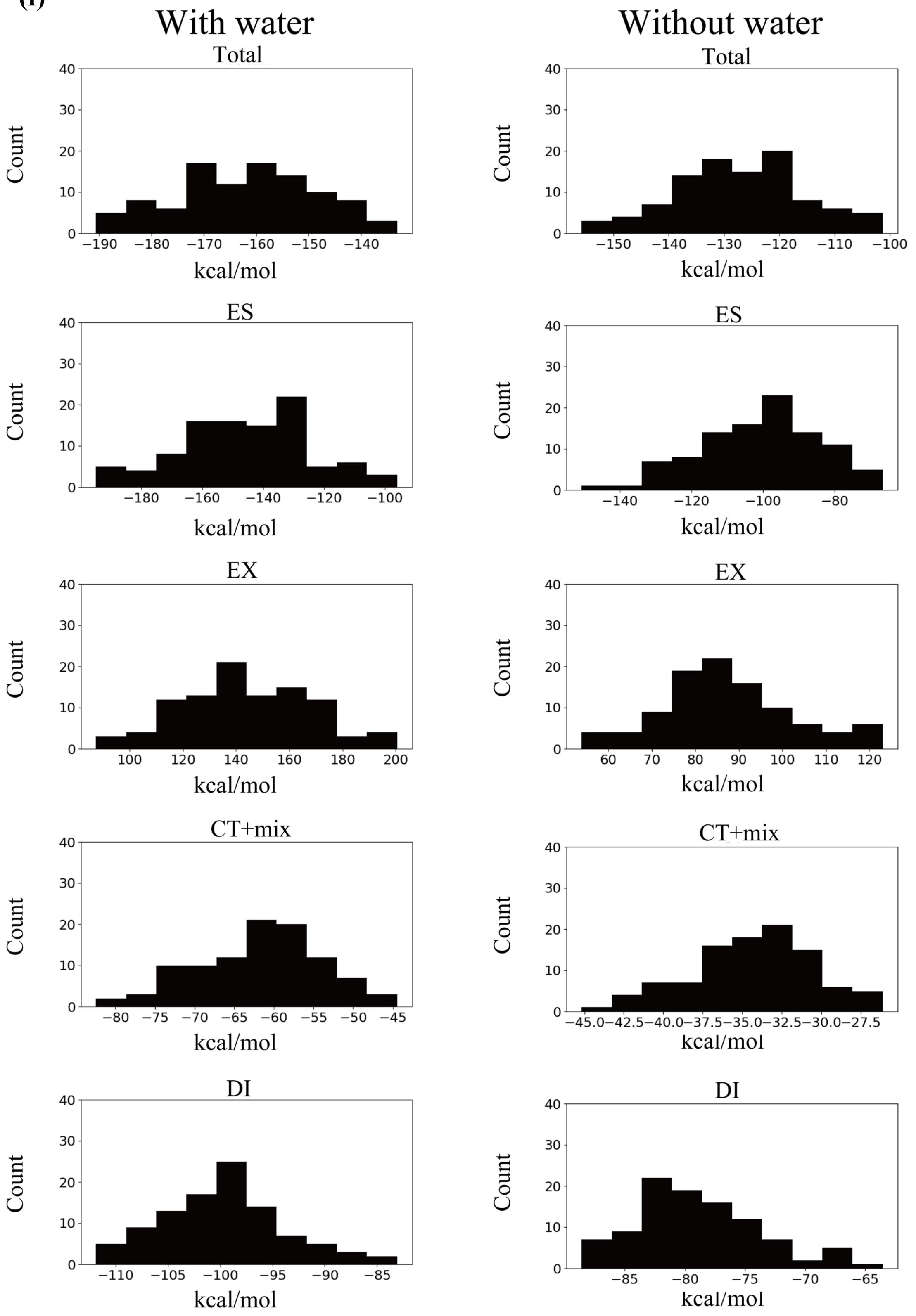
(e)

CS246


With water
Without water
Total
ES
EX
CT+mix
DI
Count
kcal/mol
0
10
20
30
40
−190
−180
−170
−160
−150
−150
−140
−130
−120
−110
−100
−90
−200
−180
−160
−140
−120
−140
−120
−100
−80
−60
100
120
140
160
180
200
40
50
60
70
80
90
100
110
−85
−80
−75
−70
−65
−60
−55
−50
−45
−40
−35
−30
−25
−20
−110
−105
−100
−95
−90
−85
−80
−85
−80
−75
−70
−65
−60
−55

(f)

CS262


With water
Without water
Total
ES
EX
CT+mix
DI
Count
kcal/mol
0
10
20
30
40
−190 −180 −170 −160 −150 −140
−150 −140 −130 −120 −110 −100
−180 −160 −140 −120 −100
−140 −120 −100 −80
100 120 140 160 180 200
60 70 80 90 100 110 120
−80 −75 −70 −65 −60 −55 −50 −45
−45.0 −42.5 −40.0 −37.5 −35.0 −32.5 −30.0 −27.5
−110 −105 −100 −95 −90 −85
−85 −80 −75 −70 −65

(g)

CS4

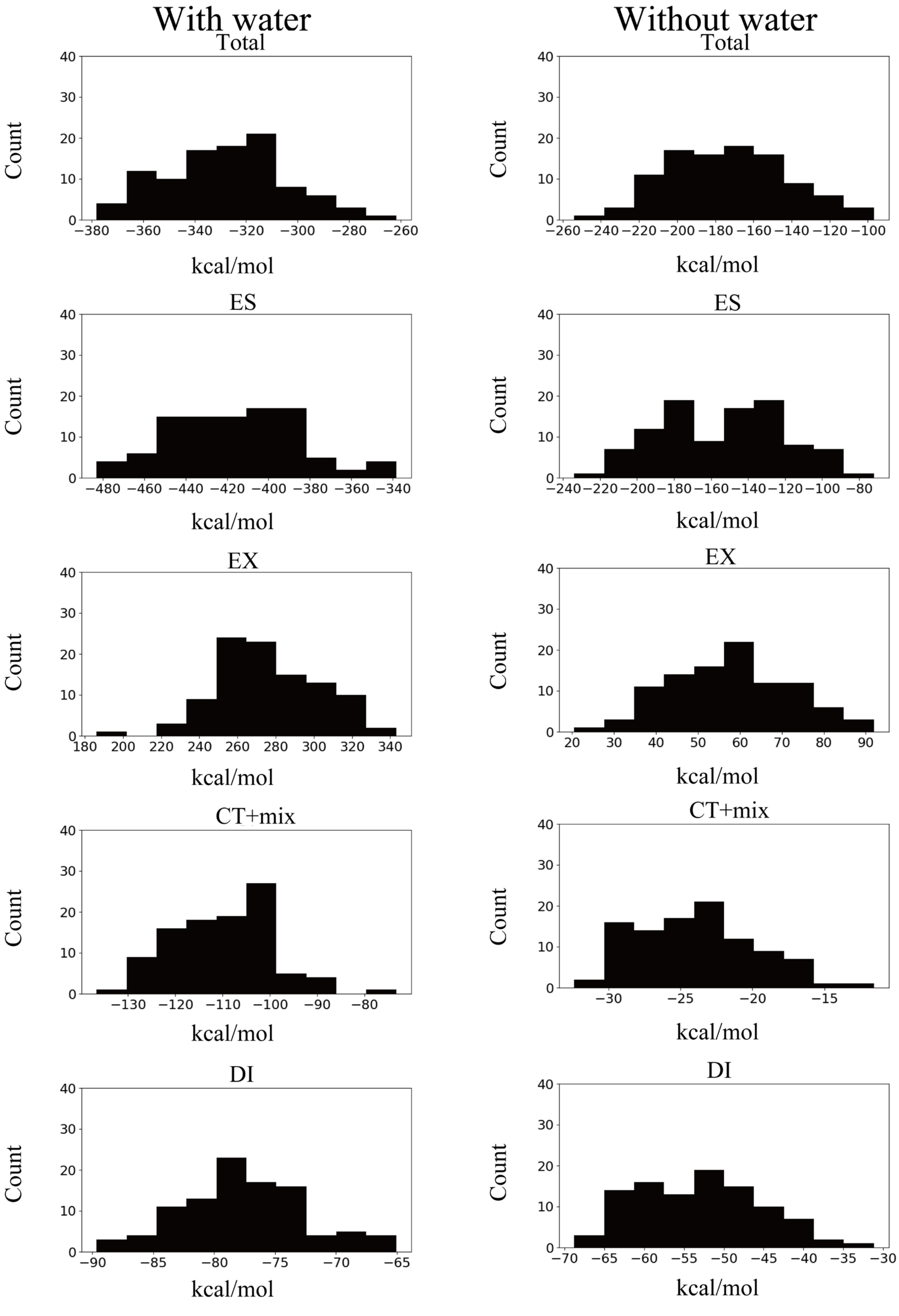

With water
Without water
Total
ES
EX
CT+mix
DI
Count
kcal/mol
0
10
20
30
40
−380 −360 −340 −320 −300 −280 −260
−260 −240 −220 −200 −180 −160 −140 −120 −100
−480 −460 −440 −420 −400 −380 −360 −340
−240 −220 −200 −180 −160 −140 −120 −100 −80
180 200 220 240 260 280 300 320 340
20 30 40 50 60 70 80 90
−130 −120 −110 −100 −90 −80
−30 −25 −20 −15
−90 −85 −80 −75 −70 −65
−70 −65 −60 −55 −50 −45 −40 −35 −30

Figure S3. The distribution of PIEDA energies. Total, ES, EX, CT+mix and DI, with and without water for the complex structures: (a) CS9, (b) CS12, (c) CS242, (d) CS245, (e) CS246, (f) CS262, and (g) CS4, sampled every 2 ns from MD simulations.

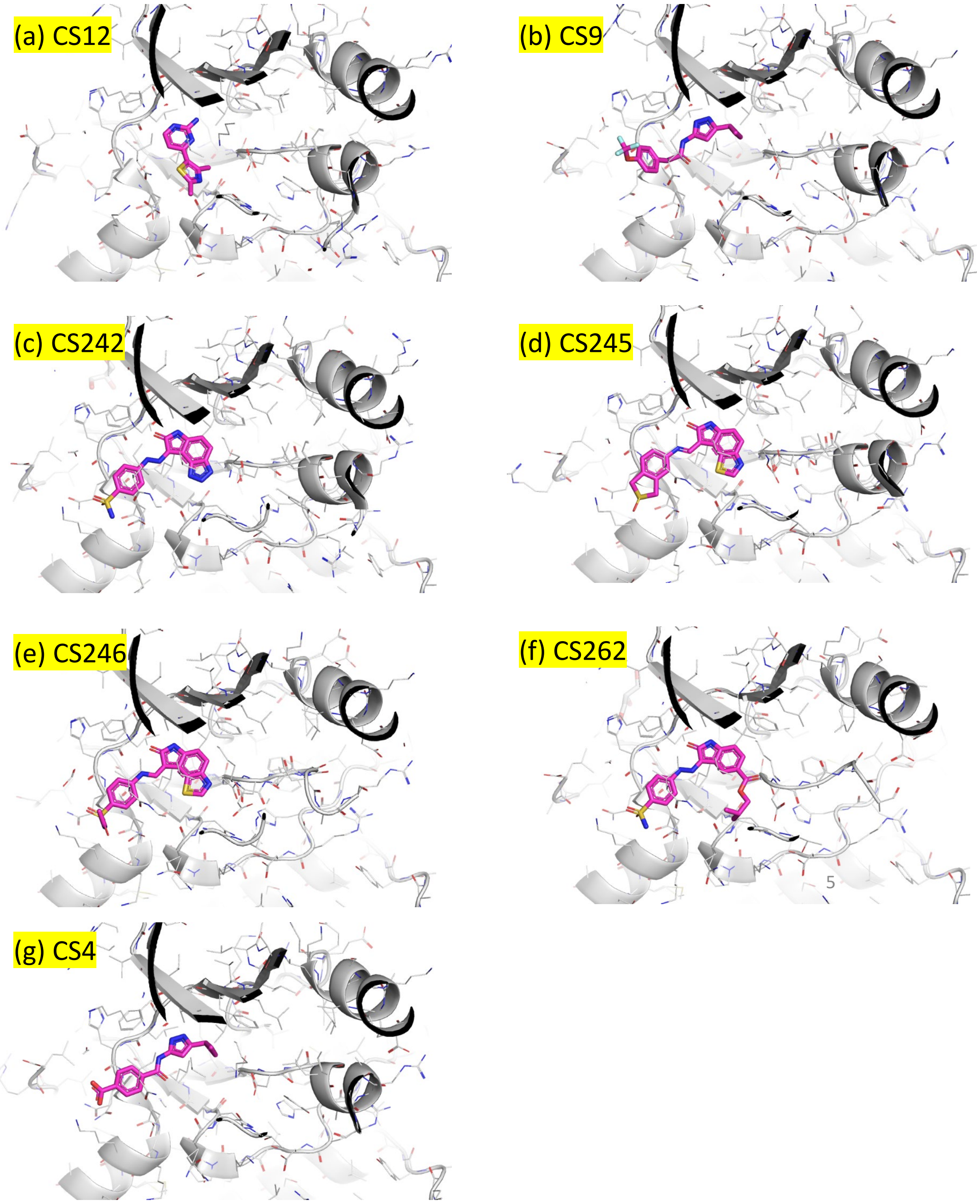


Figure S4. Ligand binding mode in the X-ray crystal structures. (a) CS12 (PDB ID: 4FK), (b) CS9 (PDB ID: 4FKI), (c) CS242 (PDB ID: 4FKQ), (d) CS245 (PDB ID: 4FKR), (e) CS246 (PDB ID: 4FKS), (f) CS262 (PDB ID: 4FKW), and the charged ligand, (g) CS4 (PDB ID: 4FKG).