Journal of Molecular Graphics and Modelling 29 (2010) 197–205

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Specific interactions between aryl hydrocarbon receptor and dioxin congeners: *Ab initio* fragment molecular orbital calculations

Eri Yoshikawa[a], Satoshi Miyagi[a], Kenichi Dedachi[a], Mitsuko Ishihara-Sugano[b], Satoshi Itoh[b], Noriyuki Kurita[a,*]

[a] *Department of Knowledge-based Information Engineering, Toyohashi University of Technology, Tempaku-cho, Toyohashi, Aichi 441-8580, Japan*
[b] *Corporate Research & Development Center, Toshiba Corporation, 1, Komukai-Toshiba-cho, Saiwai-ku, Kawasaki 212-8582, Japan*

ARTICLE INFO

Article history:
Received 2 April 2010
Received in revised form 16 June 2010
Accepted 17 June 2010
Available online 25 June 2010

Keywords:
Aryl hydrocarbon receptor
Dioxin
Molecular mechanics simulation
Fragment molecular orbital method
Ab initio molecular orbital calculation
Specific interaction between protein and ligand

ABSTRACT

Aryl hydrocarbon receptor (AhR) is a transcription factor and its function is activated by the binding of halogenated aromatic hydrocarbons such as 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and 1,2,4-trichlorodibenzo-p-dioxin (TrCDD). TCDD is highly toxic to rat, whereas its congener TrCDD shows only a weak effect on gene expression. In order to elucidate the reason of this remarkable difference in the effect of TCDD and TrCDD, we here obtained stable structures of the complexes with rat AhR (rAhR) and TCDD/TrCDD and investigated their electronic properties by using the *ab initio* fragment molecular orbital (FMO) method. The results indicate that TCDD binds more strongly to rAhR than TrCDD, which is consistent with the experimentally observed toxicity of TCDD and TrCDD. Furthermore, *ab initio* FMO calculations elucidate that His324 and Gln381 of rAhR are important for binding TCDD, while His324 and Ser334 are important for TrCDD binding.

1. Introduction

Aryl hydrocarbon receptor (AhR) is a ligand-activated transcription factor and contains a basic helix–loop–helix (bHLH) and Per-Arnt-Sim (PAS) structural motifs. The essential function of AhR is to regulate an expression of gene in a ligand-dependent manner [1]. In a ligand-free condition, AhR in cytoplasm forms a complex with HSP90 (a heat shock protein of 90 kDa), XAP2 (X-associated protein 2) and cochaperone p23. Ligand binding to AhR makes this complex transported into the nucleus, and the transported complex forms a heterodimer with the AhR nuclear translocator protein (ARNT) [2]. This heterodimer binds specifically to the promoter region of the target genes having the xenobiotic-responsive element (XRE) sequence, resulting in enhancing the expression of the target genes [3].

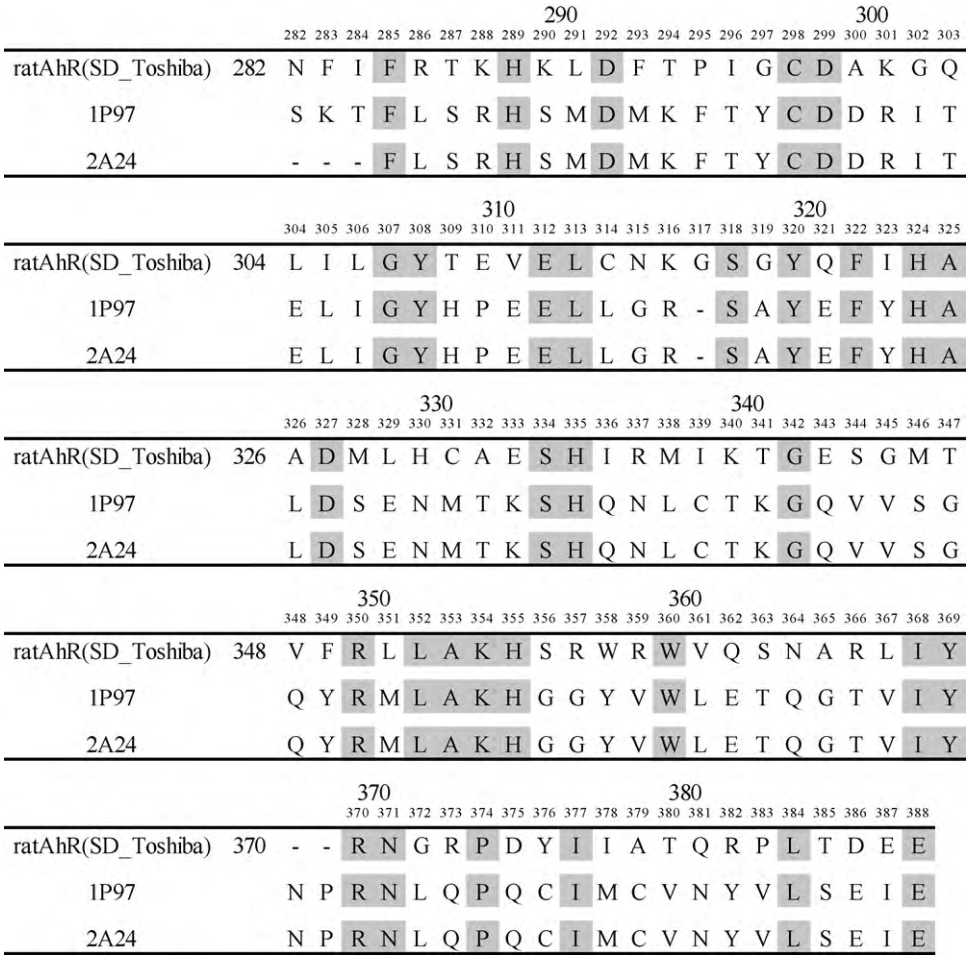
2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) is a prototype of environmental contaminants with high affinity to AhR [4]. Recent biomedical studies [5,6] revealed that the biological function of AhR is activated significantly by the TCDD binding to AhR, and that the degree of activation is largely dependent on the organismic species and strains. For example, LD_{50} (the dose of a hazardous chemical that produces death in 50% of the tested animals) of TCDD for the guinea pig is about 5000 times larger than that for the golden hamster [5]. This large difference in LD_{50} of TCDD between the animals was considered to come from the difference in structure of their AhRs. Furthermore, the Scatchard plot analysis [6] revealed that the dissociation constant of TCDD for the DBA/2J mouse AhR for TCDD is six times higher than that for the C57BL/6J mouse. The reason for this difference between the mouse strains was interpreted in terms of the difference in the 375th amino acid residue of PAS-B, which is one of the two structural repeats (PAS-A and PAS-B) within the ligand binding domain of PAS in AhR. This amino acid residue is alanine (C57BL/6J) and valine (DBA/2J), respectively, indicating that the size of the side chain of this amino acid residue may be important for the binding affinity of TCDD to these mouse AhRs.

In addition to TCDD, many types of planar ligands with the size below 14 Å × 12 Å × 5 Å are considered to bind into the ligand binding pocket of AhR [7]. However, the toxic potentials of TCDD congeners vary significantly according to their chemical structures. For example, TCDD is the most toxic among the congeners, whereas 1,2,4-trichlorodibenzo-p-dioxin (TrCDD) exhibits a weak effect on the gene expression regulated by AhR in the experiment using a reporter gene assay [8].

In order to reveal the difference in the effect of the TCDD/TrCDD binding on the biological functions of AhR, it is necessary to make

* Corresponding author. Tel.: +81 532 44 6875; fax: +81 532 44 6875.
E-mail addresses: kurita@tutkie.tut.ac.jp, kurita@cochem2.tutkie.tut.ac.jp (N. Kurita).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.06.008



		282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303
ratAhR(SD_Toshiba)	282	N	F	I	F	R	T	K	H	K	L	D	F	T	P	I	G	C	D	A	K	G	Q
1P97		S	K	T	F	L	S	R	H	S	M	D	M	K	F	T	Y	C	D	D	R	I	T
2A24		-	-	-	F	L	S	R	H	S	M	D	M	K	F	T	Y	C	D	D	R	I	T

		304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325
ratAhR(SD_Toshiba)	304	L	I	L	G	Y	T	E	V	E	L	C	N	K	G	S	G	Y	Q	F	I	H	A
1P97		E	L	I	G	Y	H	P	E	E	L	L	G	R	-	S	A	Y	E	F	Y	H	A
2A24		E	L	I	G	Y	H	P	E	E	L	L	G	R	-	S	A	Y	E	F	Y	H	A

		326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347
ratAhR(SD_Toshiba)	326	A	D	M	L	H	C	A	E	S	H	I	R	M	I	K	T	G	E	S	G	M	T
1P97		L	D	S	E	N	M	T	K	S	H	Q	N	L	C	T	K	G	Q	V	V	S	G
2A24		L	D	S	E	N	M	T	K	S	H	Q	N	L	C	T	K	G	Q	V	V	S	G

		348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369
ratAhR(SD_Toshiba)	348	V	F	R	L	L	A	K	H	S	R	W	R	W	V	Q	S	N	A	R	L	I	Y
1P97		Q	Y	R	M	L	A	K	H	G	G	Y	V	W	L	E	T	Q	G	T	V	I	Y
2A24		Q	Y	R	M	L	A	K	H	G	G	Y	V	W	L	E	T	Q	G	T	V	I	Y

		370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388
ratAhR(SD_Toshiba)	370	-	-	R	N	G	R	P	D	Y	I	I	A	T	Q	R	P	L	T	D	E	E
1P97		N	P	R	N	L	Q	P	Q	C	I	M	C	V	N	Y	V	L	S	E	I	E
2A24		N	P	R	N	L	Q	P	Q	C	I	M	C	V	N	Y	V	L	S	E	I	E

Fig. 1. Sequence of amino acid residues for the LBD of rAhR, HIF-2α (PDB code: 1P97) and the complex of ARNT + HIF-2α (PDB code: 2A24). The amino acid residues with hatching are common for the three sequences.

clear the structures of the AhR itself as well as the TCDD/TrCDD bounded AhR. However, there is no X-ray or NMR study elucidating the three-dimensional structure of AhR. On the other hand, some homology modeling studies [9–14] on the structure of the ligand binding domain (LBD) of AhR have been conducted. Recently, the site-directed mutagenesis experiments [12–14] followed by a homology modeling study were performed to identify the amino acid residues of AhR playing a key role in the ligand binding to AhR.

In the present study, to clarify the reason why the toxicities of TCDD and TrCDD are remarkably different each other, we obtain the stable structures for the complexes with the LBD of rat AhR (rAhR) and TCDD/TrCDD, and their electronic properties are investigated by *ab initio* molecular orbital (MO) calculations. The results obtained on the specific interactions between rAhR and TCDD/TrCDD can be related with the difference in the experimentally observed toxicity between TCDD and TrCDD. In addition, it is elucidated which amino acid residues of rAhR are important for the specific interactions between rAhR and TCDD/TrCDD at an electronic level.

2. Details of calculations

2.1. Construction and optimization of rAhR–TCDD and rAhR–TrCDD structures

Pandini et al. [12,13] constructed the three-dimensional structure of the LBD for the mouse AhR (mAhR) based on the PDB structures of HIF-2α and ARNT by homology modeling techniques. In their study, the amino acid residues from the 278th to the 384th of mAhR were considered for predicting the interactions between mAhR and TCDD. The results suggested some key features that are required for TCDD binding and activation of mAhR at a molecular level.

In the present study, we first constructed the three dimensional structure of the LBD for the Sprague–Dawley rat AhR (rAhR) in the similar way as Pandini et al. [12,13], by using the homology modeling program MODELLER [15]. In order to identify the template PDB structures for generating the model structure of rAhR–LBD, we searched PDB structures having a sequence of amino acids similar to rAhR–LBD by the amino-acids searching program PSI-BLAST [16]. The NMR structures of PAS-B domains for HIF-2α (PDB ID: 1P97) and ARNT (PDB ID: 1X0O) were detected. In addition, we found the NMR structure (PDB ID: 2A24) for a docked complex with HIF-2α and ARNT constructed by the HADDOCK (high ambiguity driven biomolecular docking) [17] program. The PSI-BLAST scores for the similarity of amino-acid sequence are 187 (HIF-2α), 63 (ARNT) and 185 (docked structure of HIF-2α and ARNT), respectively. In Fig. 1, the sequence of amino acid residues from the 282nd to the 388th for rAhR–LBD are compared with those for HIF-2α and the complex of ARNT + HIF-2α. The sequence identities of rAhR–LBD with the template structures are about 30%. We thus employed the NMR structures of HIF-2α and the ARNT + HIF-2α complex as template structures in homology modeling, because these structures have large PSI-BLAST scores. This choice of the template structures is different from that in the modeling procedure by Pandini et al. [12,13].

Table 1
Molpdf scores obtained by MODELLER for 9 model structures of rat AhR.

1P97	2A24	molpdf
7	4	3713.0
7	8	3757.4
7	18	3726.7
13	4	3657.8
13	8	3710.6
13	18	3643.0
6	4	3689.3
6	8	3695.4
6	18	3730.0

There are 20 different NMR structures in each PDB file (PDB ID:1P97 and 2A24). Pandini et al. selected the most representative one among the 20 structures by using the NMRCLUST program [18]. On the other hand, we used Ramachandran plots for comparing the backbone structures of these 20 structures with those of normal proteins registered in PDB and obtained three most similar structures among the 20 structures in each PDB file. By choosing one structure from the both PDB files, we built up 9 (3×3) types of combinations for the template structures in homology modeling and generated 10 candidate structures for each combination by using MODELLER program [15].

In order to select an appropriate model structure from the 90 candidate structures for the rAhR–LBD modeled by MODELLER, we evaluated Molpdf scores [19] for these model structures. Molpdf scores can be used to rank the model structures obtained from the same alignment. In fact, the model structure with smaller Molpdf score is more similar to the template PDB structures. The smallest Molpdf score for each combination of the template PDB structures is listed in Table 1, which indicates that the model structure obtained from the combination of the 13th structure in 1P97 PDB and the 18th one in 2A24 PDB has the smallest Molpdf score. Therefore, the model structure obtained from this combination was employed as the initial structure of the rAhR–LBD, in the present study.

In order to obtain the candidate structures for the complexes with rAhR and TCDD/TrCDD, TCDD/TrCDD was docked into the ligand binding pocket of rAhR by using the automated protein–ligand docking program AutoDock 4.4 [20]. In the docking procedure, the grid box was set as the 18.75 Å × 18.75 Å × 18.75 Å cubic centered at the middle of the ligand binding pocket of rAhR. The spacing between the nearest neighboring grid points was set to be 0.375 Å, which is the default value of AutoDock 4.4. Because TCDD and TrCDD have no rotatable bond, these structures were fixed during the docking. 256 conformations of both TCDD and TrCDD docked in the ligand binding pocket of rAhR were created by AutoDock 4.4, and they were classified into some clusters according to the root mean square distance (RMSD) values (0.5 Å) between the conformations. We adopted the most stable structures in each cluster as the candidate structures for the rAhR–TCDD and rAhR–TrCDD complexes.

Subsequently, hydrogen atoms were added to the candidate structures, and the positions of hydrogen atoms were optimized to minimize the strain energy around the hydrogen atoms, by the classical molecular mechanics (MM) and molecular dynamics (MD) simulations program AMBER9 [21]. Finally, we added solvating water molecules around the rAhR–TCDD/TrCDD complexes with a 9 Å layer and optimized the whole structure by AMBER9 to obtain stable structures of the solvated rAhR–TCDD/TrCDD complexes. In these optimizations, the PARM99 [22] and TIP3P [23] force fields were used for rAhR–TCDD/TrCDD and water molecules, respectively. The atomic charges of TCDD and TrCDD were determined from the RESP (restrained electrostatic potential) analysis [24] based on the MP2/6-31G(d,p) calculations in the *ab initio* MO program package Gaussian03 [25]. The threshold value of the energy-gradient for the convergence of the AMBER9 optimization was set as 0.0001 kcal/mol/Å. The most stable structure among the optimized ones was adopted as the solvated structure of the complex.There are two types of isomers for the uncharged His residues, depending on the position of hydrogen atom. Since there is no experimental information about the protonated state of His residues in rAhR, we considered the both isomers of His. One His has hydrogen atom at the δ site of the imidazole ring of His, which is indicated as HID, while the other (HIE) has hydrogen atom at the ε site. We considered the two types of rAhR models containing with HID or HIE and optimized the both structures in water. From the comparison of their total energies obtained by the *ab initio* MO calculations, the most stable structures for the rAhR–TCDD and rAhR–TrCDD complexes were determined. As for the other ionisable amino acid residues contained in rAhR, the ionized state was adopted.

2.2. Molecular orbital calculations for rAhR–TCDD and rAhR–TrCDD complexes

The electronic properties of these optimized structures for the solvated rAhR–TCDD and rAhR–TrCDD complexes were investigated by the *ab initio* MO calculations based on fragment MO (FMO) method [26] to elucidate which amino acid residues in rAhR are important for the specific interactions between rAhR and TCDD/TrCDD. We used the FMO program ABINIT-MP Ver. 4.1 [27,28], in which the target molecule is divided into units called fragment, and the electronic properties of the target molecule are estimated from the electronic properties of the monomers and dimers of the fragments. Since the electronic properties of the dimers are calculated in FMO, we can obtain the interaction energies between specific fragments considering with the effect from the other fragment.

In the present study, we investigated electronic properties of the solvated structures of rAhR–TCDD and rAhR–TrCDD with considering water molecules existing within a 3 Å distance from the surface of the complex, because the FMO calculation for the whole structure of the solvated complex is not practical because of the limitation of our computer facilities. Each amino acid residue of rAhR, TCDD, TrCDD and each water molecule were assigned as a fragment, respectively. This fragmentation enables us to evaluate the interaction energy between the amino acid residue of rAhR and TCDD/TrCDD. From the comparison of the obtained interaction energies, it was elucidated which amino acid residues of rAhR have large contribution to the specific interactions between rAhR and TCDD/TrCDD. In the present FMO calculations, *ab initio* MP2/6-31G method was employed to investigate the π–π stacking and CH–π interactions between the amino acid residues of rAhR and TCDD/TrCDD accurately.

In order to elucidate binding affinity between rAhR and TCDD/TrCDD quantitatively, binding free energy should be evaluated. However, it is not practical to calculate free energies for the solvated rAhR–TCDD/TrCDD complexes by *ab initio* MO method. In the present study, we thus investigated the binding energies between rAhR and TCDD/TrCDD to estimate binding affinity qualitatively, on the assumption that the entropic effects on rAhR–TCDD and rAhR–TrCDD are almost the same.

3. Results and discussion

3.1. Optimized structures of rAhR–TCDD and rAhR–TrCDD complexes

By using MODELLER program [15], 90 candidate structures for rAhR were predicted from the template structures of HIF-2α (PDB

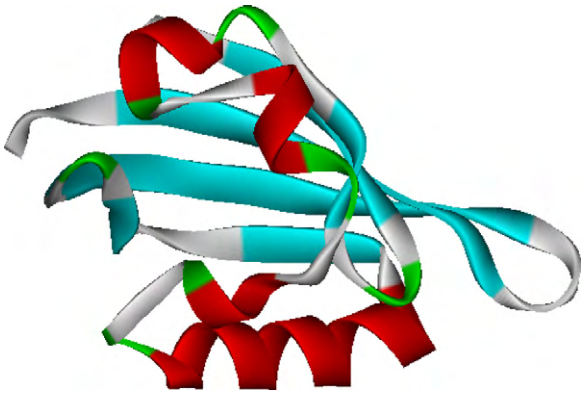


Fig. 2. Structure of ligand binding domain of rAhR obtained by MODELLER program [15].

ID: 1P97) [29] and the complex of ARNT and HIF-2α (PDB ID: 2A24) [30]. Among these 90 structures, the structure (Fig. 2) with the smallest score of Molpdf in MODELLER was employed as the model structure of rAhR. As shown in Fig. 2, the ligand binding pocket of rAhR is surrounded by one α-helix and some β-sheets and has enough space for the TCDD and TrCDD binding.

By using the protein–ligand docking program AutoDock 4.4 [20], 256 candidate structures were obtained for both the rAhR–TCDD and rAhR–TrCDD complexes. The structures of rAhR–TCDD were classified into four clusters based on their structural similarity, while those of rAhR–TrCDD were classified into six clusters. We adopted the most stable structures in each cluster as the candidate structures for these complexes and optimized their solvated structures by AMBER9 [21]. The optimizations elucidated that there is no water molecule in the ligand-binding pocket of the rAhR–TCDD/TrCDD complexes. This result seems to come from the hydrophobic nature of TCDD/TrCDD as well as the amino acid residues existing in the pocket. The optimized structures for all clusters are compared in Figs. 3 and 4, where the solvating water molecules are not shown for clarifying the conformations of TCDD and TrCDD in the complexes.

As shown in Fig. 3a and b, TCDD has a similar conformation in the two clusters of the rAhR(HID)–TCDD complex. In a similar way, in the two rAhR(HIE)–TCDD clusters, the conformation of TCDD is similar to each other as shown in Fig. 3c and d. On the other hand, TrCDD has a variety of conformations in the ligand-binding pocket of both the rAhR(HID) and rAhR(HIE). As shown in Fig. 4a–c, the directions of TrCDD in the clusters 2 and 3 of rAhR(HID)–TrCDD are similar to each other, while that in the cluster 1 is completely different. In the rAhR(HIE)–TrCDD complex, TrCDD has many conformations as shown in Fig. 4d–f. In comparison with the high symmetry TCDD, TrCDD has low symmetry, so that TrCDD can take many conformations in the ligand binding pocket of rAhR. Moreover, the comparison between the optimized structures shown in Figs. 3 and 4 elucidates that the protonation state of His residues in rAhR affects remarkably the conformations of TCDD and TrCDD in the ligand binding pocket of rAhR.

To determine the most stable structure among these optimized structures, we investigated their electronic properties by the *ab initio* FMO method (MP2/6-31G). Table 2 lists the obtained total energies. The energy difference between the most stable structure and the other ones is from 8.4 to 44.5 kcal/mol. In particular, the difference is large for the rAhR(HID)-TrCDD complexes, indicating that TrCDD prefers to exist at a specific position of the ligand-binding pocket of rAhR(HID).

The most stable structures for rAhR(HID)–TCDD and rAhR(HIE)–TCDD are shown in Fig. 3a and d, respectively. In these complexes, TCDD is stabilized at almost the center of the ligand binding pocket of rAhR. On the other hand, as shown in Fig. 4c and d, in the most stable structures of rAhR(HID)–TrCDD

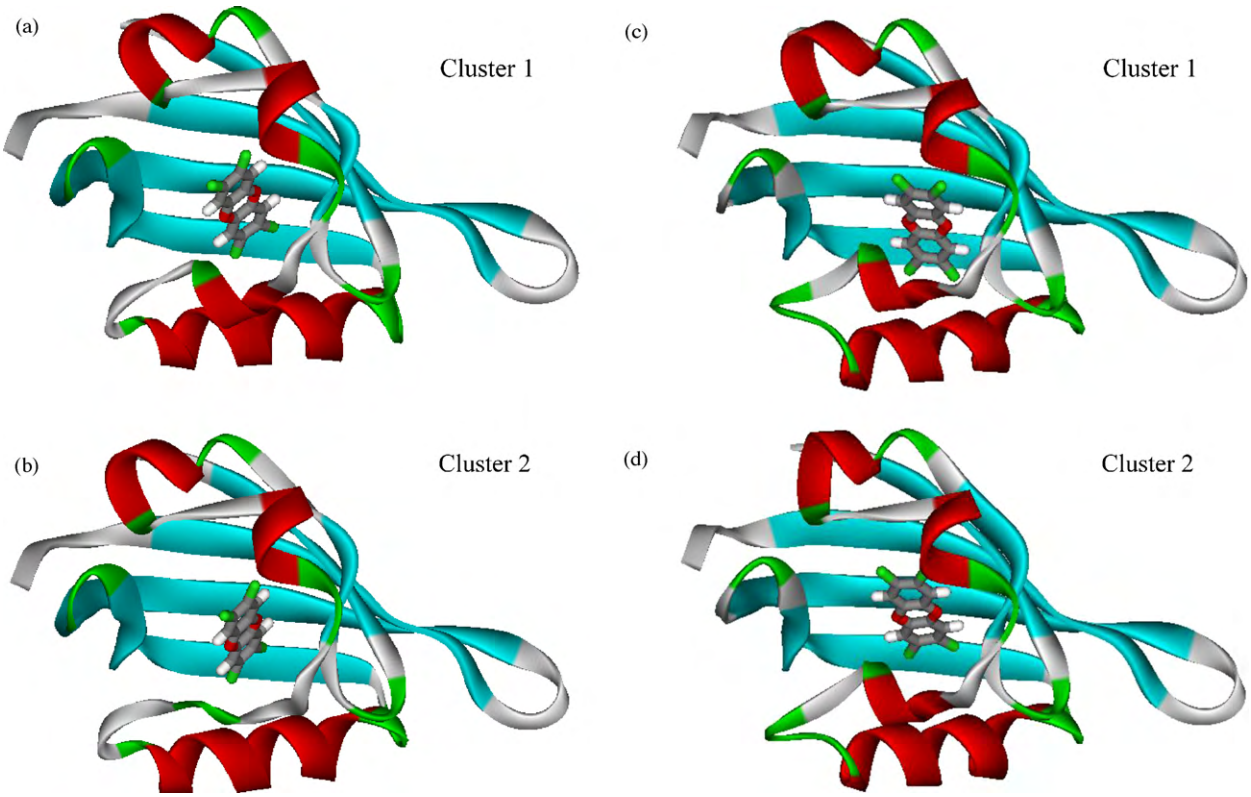


Fig. 3. Structures of solvated rAhR–TCDD complex optimized by AMBER9: (a) and (b) rAhR(HID)–TCDD, (c) and (d) rAhR(HIE)–TCDD.

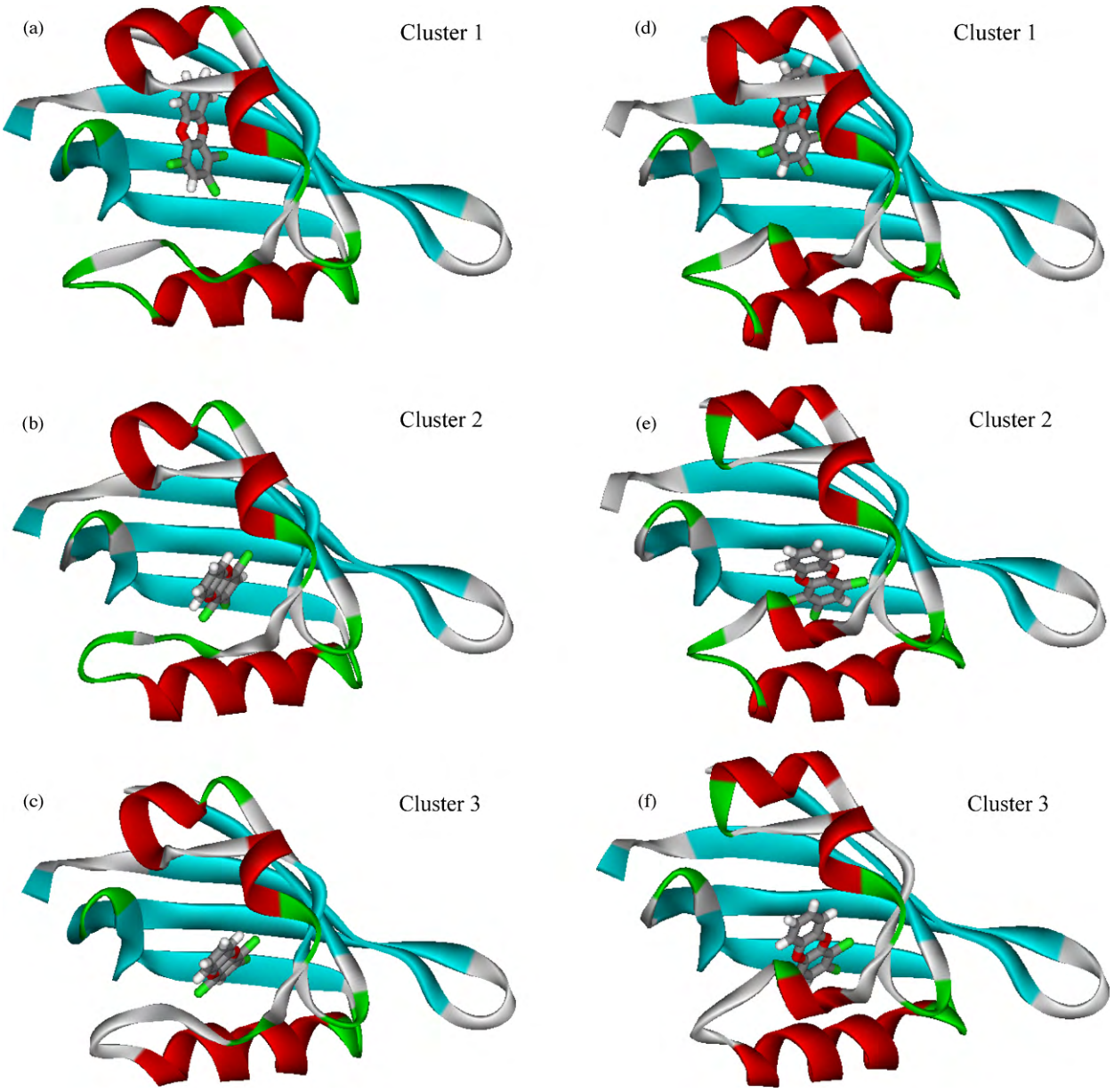


Fig. 4. Structures of solvated rAhR–TrCDD complex optimized by AMBER9: (a)–(c) rAhR(HID)–TrCDD; (d)–(f) rAhR(HIE)–TrCDD.

and rAhR(HIE)–TrCDD, TrCDD is stabilized near the side wall of the ligand binding pocket of rAhR. It seems that this result comes mainly from the large electric dipole moment of TrCDD compared with that of high symmetry TCDD. In this way, the binding conformations of TCDD and TrCDD to the ligand-binding pocket of rAhR are remarkably different from each other, resulting in the difference in the specific interaction between rAhR and TCDD/TrCDD, as will be shown below.

Table 2
Total energies of each cluster for rAhR–TCDD/TrCDD obtained by FMO calculations.

Complex	Cluster no.	Total energy (Hartree)	Relative (kcal/mol)
rAhR(HID)–TCDD	1	−46,509.858	0.0
	2	−46,509.839	11.6
rAhR(HIE)–TCDD	1	−46,509.801	8.4
	2	−46,509.814	0.0
rAhR(HID)–TrCDD	1	−46,050.908	32.8
	2	−46,050.889	44.5
	3	−46,050.960	0.0
rAhR(HIE)–TrCDD	1	−46,050.951	0.0
	2	−46,050.925	16.7
	3	−46,050.935	9.9

3.2. *Electronic properties of rAhR–TCDD and rAhR–TrCDD complexes*

The total energies listed in Table 2 indicate that the cluster 1 of rAhR(HID)-TCDD (Fig. 3a) is the most stable among the four structures of rAhR–TCDD complex, while the cluster 3 of rAhR(HID)-TrCDD (Fig. 4c) is the most stable. In order to compare the binding affinities for the rAhR–TCDD and rAhR–TrCDD complexes, the binding energies between rAhR and TCDD/TrCDD were evaluated by the FMO calculations. The binding energy was estimated as the difference between the total energy of rAhR–TCDD complex and the sum of the total energies of rAhR and TCDD. As shown in Table 3, the binding energies are −34.05 (rAhR–TCDD)

Table 3
Total energies of rAhR–TCDD/TrCDD, rAhR, TCDD and TrCDD, and binding energies between rAhR and TCDD/TrCDD.

	Total energy (Hartree)			Binding energy (kcal/mol)
	rAhR + ligand	rAhR	Ligand	
rAhR–TCDD	−46,509.8576	−44,064.3426	−2445.4607	−34.05
rAhR–TrCDD	−46,050.9600	−44,064.3726	−1986.5456	−26.18

and −26.18 (rAhR–TrCDD) kcal/mol, respectively, indicating that both TCDD and TrCDD bind strongly to rAhR in the ligand binding pocket. In addition, it was clarified that TCDD binds more strongly to rAhR than TrCDD. This result is qualitatively consistent with the experimental result [8] showing the larger toxicity of TCDD than TrCDD. In the experiment, transactivation ability of rAhR was evaluated by means of a reporter gene assay, in which TCDD/TrCDD molecule binds to rAhR and promotes the rAhR binding to specific sequence of DNA, resulting in an induction of gene expression. The experimental results elucidated that the transactivation ability of rAhR is increased about four times by the addition of 1000 pg/ml TCDD to the assay, while the TrCDD addition has little effect on the rAhR-mediated gene expression. Although the difference in the transactivation abilities of rAhR activated by TCDD/TrCDD is not directly related with the binding affinity between rAhR and these dioxin congeners, the present computed results indicate that the change in transactivation abilities of rAhR induced by dioxin congeners can be related with the binding affinity between rAhR and dioxin congeners.

To make clear the reason why TCDD binds more strongly than TrCDD, we investigated the electronic properties of the optimized rAhR–TCDD/TrCDD complexes by the FMO calculations and elucidated which amino acid residues in rAhR or solvating water molecules are important for the specific interactions between rAhR and TCDD/TrCDD. Table 4 lists the amino acid residues of rAhR and their interaction energies to TCDD/TrCDD in a decreasing order of magnitude of attractive interaction energy. Some residues with large repulsive interaction are also listed. Many amino acid residues of rAhR contribute to the binding between rAhR and TCDD/TrCDD, while solvating water molecules do not have significant contribution to the specific interaction between rAhR and TCDD/TrCDD.

Table 4
Interaction energies (kcal/mol) between amino acid residues of rAhR and (a) TCDD or (b) TrCDD. The amino acids are listed in a decreasing order of the magnitude of attractive interaction energy, and some amino acids with large repulsive interaction are also listed.

(a) rAhR–TCDD		(b) rAhR–TrCDD	
Amino acids	Energy	Amino acids	Energy
His324	−8.0	His324	−5.7
Gln381	−5.2	Ser334	−5.7
Thr287	−2.8	Phe293	−3.1
Ile323	−2.7	Arg366	−2.9
Phe293	−2.5	Gln381	−2.8
Phe349	−2.2	Phe349	−2.6
Ser363	−2.0	His289	−2.4
His289	−2.0	Ile323	−2.1
Asp327	−1.9	Phe322	−2.0
Glu388	−1.3	Arg382	−1.7
Leu351	−1.1	Asn364	−1.6
Cys331	−1.0	Thr380	−1.3
Arg337	−1.0	Thr287	−1.3
Ser334	−1.0	Glu388	−1.3
Pro295	−0.9	Met346	−1.1
Met346	−0.8	Leu351	−1.0
Arg366	−0.8	Asp327	−1.0
...	...	...	...
Glu343	0.7	Gly345	1.0
Arg359	0.9	Ala365	1.0
Leu352	0.9	His335	1.1
Arg350	1.0	Lys316	1.2
Glu310	1.0	Glu343	1.6
Ala332	2.6	Ala332	3.7

As shown in Table 4a, His324, Gln381, Thr287, Ile323, Phe293, Phe349, Ser363 and His289 have attractive interaction energy larger than −2 kcal/mol in the rAhR–TCDD complex. In particular, His324 has large (−8.0 kcal/mol) interaction energies. On the other hand, in the rAhR–TrCDD complex, His324, Ser334, Phe293, Arg366, Gln381, Phe349, His289, Ile323 and Phe322 have large attractive interaction energy as shown in Table 4b. Among them, His324 and Ser334 have rather large (−5.7 kcal/mol) attractive interaction energy. Therefore, it was elucidated that His324 of rAhR has large contribution to the binding of both TCDD and TrCDD in the ligand binding pocket of rAhR.

In addition, we investigated the interaction energies for all amino acid residues of rAhR with TCDD and TrCDD to make clear the difference in specific interactions between the rAhR–TCDD and rAhR–TrCDD complexes. In Fig. 5a and b, the amino acid residues with large interaction energy are clearly indicated. As shown in Fig. 5a, Ala332 has large (2.6 kcal/mol) repulsive interaction to TCDD. In rAhR–TrCDD, Ala332 has rather large (3.7 kcal/mol) repulsive interaction to TrCDD, as shown in Fig. 5b. It was also elucidated from the comparison of Fig. 5a and b that more amino acid residues of rAhR have repulsive interactions to TrCDD compared with those to TCDD. This fact is expected to cause the smaller binding energy between rAhR and TrCDD than that between rAhR and TCDD. The difference in interaction energies between rAhR–TCDD and rAhR–TrCDD is shown in Fig. 5c, indicating that Ser334, Gln381, His324 and Arg366 have remarkable difference (larger than 2 kcal/mol) in interaction energy between TCDD and TrCDD. These amino acid residues are expected to play an important role in distinguishing between the TCDD and TrCDD bindings to rAhR.

To elucidate the reason for this large difference between the rAhR–TCDD and rAhR–TrCDD complexes, the interacting structures and distances between these amino acid residues of rAhR and TCDD/TrCDD were analyzed in detail. As for His324 having the largest attractive interaction to both TCDD and TrCDD, its interacting structures to TCDD and TrCDD are compared in Fig. 6a. The imidazole ring of His324 binds to the oxygen and hydrogen atoms of TCDD with the distance of 2.46 and 2.62 Å, respectively. On the other hand, His324 binds to TrCDD with the distance of 3.19 and 2.62 Å, respectively. Because of the chlorine atom existing near the oxygen atom in TrCDD, the hydrogen atom of His324 cannot get closer to the oxygen atom of TrCDD, so that the distance between His324 and the oxygen atom of TrCDD is much longer than that in the rAhR–TCDD complex as shown in Fig. 6a. As a result, the interaction energy between His324 and TrCDD is 2.3 kcal/mol smaller than that between His324 and TCDD. Consequently, it is concluded that the difference in the position of chlorine atoms between TCDD and TrCDD causes the large difference in the interaction between His324 and the oxygen atom of TCDD/TrCDD.

The interacting structures between Gln381 and TCDD/TrCDD are compared in Fig. 6b. The amino group of Gln381 interacts to TCDD/TrCDD by a similar way as a CH–π interaction. The distances between the hydrogen atom of the amino group and the center of hexagonal ring of TCDD/TrCDD are 2.39 and 2.70 Å, respectively, so that the interaction energy between Gln381 and TCDD is about 2.4 kcal/mol larger than that between Gln381 and TrCDD as shown in Table 4. The difference in interaction energies between His324/Gln381 and TCDD/TrCDD is expected to be a main factor for the difference in binding energy between rAhR and TCDD/TrCDD.

Table 4 also indicates that Ser334 and Arg366 have large attractive interaction energy only in the rAhR–TrCDD complex. Ser334 has the largest difference in interaction energy between TCDD and TrCDD as shown in Fig. 5c. To elucidate the reason of this large difference, we compared the structures around Ser334 in

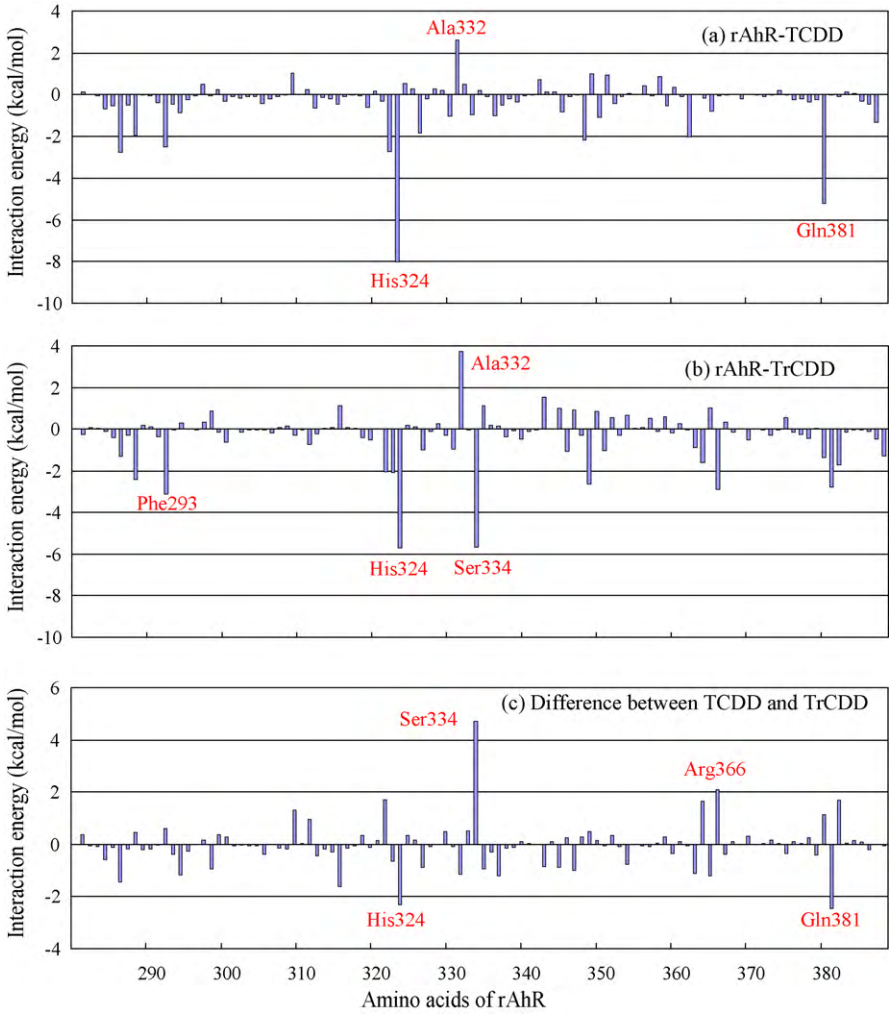


Fig. 5. Interaction energies between amino acid residues of rAhR and (a) TCDD, (b) TrCDD and (c) the difference between them.

the rAhR–TCDD and rAhR–TrCDD complexes. As shown in Fig. 6c, Ser334 interacts with TrCDD at a 2.49 Å distance, while the nearest distance between Ser334 and TCDD is 3.01 Å. This large difference in interacting structure causes the large (−4.7 kcal/mol) difference in interaction energy between Ser334 and TCDD/TrCDD. It is noted that the hydrogen atom of TrCDD contributes to this interaction, while the chlorine atom of TCDD dose. Therefore, it is expected that the difference in the position of chlorine atoms between TCDD and TrCDD has a remarkable influence in the interaction between Ser334 and TCDD/TrCDD.

In addition, Arg366 has different interaction energies to TCDD and TrCDD as shown in Fig. 5c. Fig. 6d indicates that there are two weak interactions between the hydrogen atom of the backbone NH group of Arg366 and the two chlorine atoms of TrCDD, while there is only one interaction in rAhR–TCDD. As a result, the interaction energy between Arg366 and TrCDD is 2.1 kcal/mol larger compared with that between Arg366 and TCDD.

The binding energy between rAhR and TCDD is about 8 kcal/mol larger than that between rAhR and TrCDD as shown in Table 3. From the above mentioned analysis of interaction energies between each amino acid residue of rAhR and TCDD/TrCDD shown in Fig. 5c, it was clarified that His324 and Gln381 have positive contribution to stabilizing the rAhR–TCDD complex compared with rAhR–TrCDD, while Ser334 and Arg366 have negative contribution. In addition, Fig. 5b indicates that there are many amino acid residues of rAhR having the repulsive interaction with TrCDD, resulting in the smaller binding energy between rAhR and TrCDD.

In the experimental study [14] on the specific interactions between the mouse AhR (mAhR) and TCDD, it was concluded that Phe318, Ile319 and His320 are important for the TCDD binding. In addition, the experimental study for the avian AhR [31] indicated that Ile319 plays an essential role in the binding between TCDD and AhR. These amino acid residues correspond to the Phe322, Ile323 and His324 of rAhR. The present calculated result for rAhR–TCDD shown in Fig. 5a indicates that His324 has the largest interaction energy to TCDD and that Ile323 of rAhR also has large interaction energy to TCDD, while Phe322 of rAhR has small interaction energy. Consequently, it can be concluded that the present result for Ile323 and His324 of rAhR–TCDD shown in Fig. 5a is consistent with the experimental results [14,31]. On the other hand, the present result for Phe322 of rat AhR shown in Fig. 5a is not comparable to the previous experimental result for mouse AhR. This difference may come from the difference in organism species.

In the previous homology modeling and mutagenesis experimental studies for mouse AhR [12,13], it was predicted that Cys327, His285 and Ala375 are important for the binding between mouse AhR-LBD and TCDD. These amino acid residues correspond to the Cys331, His289 and Ala379 of rAhR, respectively. As shown in Table 4, Cys331 and His289 have −1.0 and −2.0 kcal/mol attractive interaction energy to TCDD, respectively, in the present *ab initio* FMO study for the rAhR–TCDD complex. The interaction energy

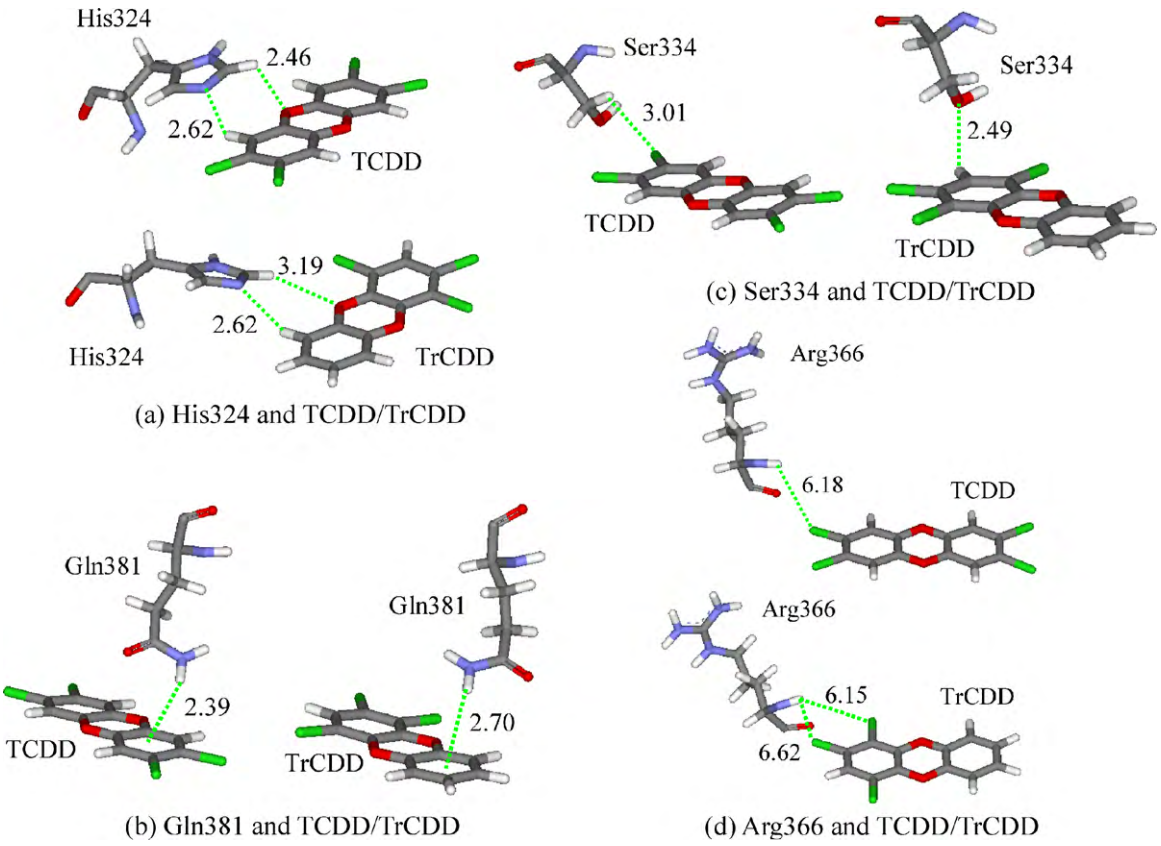


Fig. 6. Interacting structures and distances (Å) between amino acid residues of rAhR and TCDD/TrCDD: (a) His324, (b) Gln381, (c) Ser334, and (d) Arg366.

between Ala379 and TCDD is −0.4 kcal/mol, so that the present result on Ala379 of rat AhR cannot be compared with the previous result for mouse AhR [12,13]. *Ab initio* FMO investigations for the mouse AhR–TCDD complex are underway to make clear the difference in TCDD binding properties between mouse and rat AhRs.

4. Conclusions

In order to elucidate the reason for the large difference in toxicity between TCDD and TrCDD, we investigated the specific interactions between the model structure for the LBD of rat AhR (rAhR) and TCDD/TrCDD by the molecular simulations combined with the homology modeling, protein–ligand docking, classical MM and *ab initio* FMO methods. The binding energy between rAhR and TCDD is larger than that between rAhR and TrCDD, which is consistent with the trend of toxicity for TCDD and TrCDD. Furthermore, the interaction energies between each amino acid residue of rAhR and TCDD/TrCDD were investigated to find that His324 and Gln381 of rAhR are important for the binding between rAhR and TCDD, while His324, Ser334 are important for the interaction between rAhR and TrCDD. The analyses for the structures of these amino acid residues existing around TCDD/TrCDD elucidate that the difference in the position of chlorine atoms between TCDD and TrCDD causes a remarkable difference in specific interactions between His324/Ser334 and TCDD/TrCDD.

Acknowledgments

This work was supported in part by the research grants from Murata Science Foundation, Iketani Science and Technology Foundation, Tatematsu Foundation and CASIO Science Promotion Foundation.

References

[1] H. Reyes, S. Reisz-Porszansz, O. Hankinson, Identification of the Ah receptor nuclear translocator protein (Arnt) as a component of the DNA binding form of the Ah receptor, Science 256 (1992) 1193–1195.
[2] J.R. Petrulis, G.H. Perdew, The role of chaperone proteins in the aryl hydrocarbon receptor core complex, Chem. Biol. Interact. 141 (2002) 25–40.
[3] C.L. Wilson, S. Safe, Mechanisms of ligand-induced aryl hydrocarbon receptor-mediated biochemical and toxic responses, Toxicol. Pathol. 26 (1998) 657–671.
[4] S. Safe, Polychlorinated biphenyls (PCBs), dibenzo-pdioxins (PCDDs), dibenzofurans (PCDFs), and related compounds: environmental and mechanistic considerations which support the development of toxic equivalency factors (TEFs), Crit. Rev. Toxicol. 21 (1990) 51–88.
[5] O. Hankinson, The aryl hydrocarbon receptor complex, Annu. Rev. Pharmacol. Toxicol. 35 (1995) 307–340.
[6] M. Ema, N. Ohe, M. Suzuki, J. Mimura, K. Sogawa, S. Ikawa, Y. Fujii-Kuriyama, Dioxin binding activities of polymorphic forms of mouse and human aryl hydrocarbon receptors, J. Biol. Chem. 269 (1994) 27337–27343.
[7] C.L. Waller, J.D. McKinney, Three-dimensional quantitative structure-activity relationships of dioxins and dioxin-like compounds: model validation and Ah receptor characterization, Chem. Res. Toxicol. 8 (1995) 847–858.
[8] S. Uruno, S. Yoshimura, E. Akahoshi, M. Ishihara-Sugano, Applicability of tyrosine hydroxylase promoter activation assay (TH assay) for screening of dioxin-like compounds, in: 29th International Symposium on Halogenated Persistent Organic Pollutants (DIOXIN 2009), 2009, pp. 10–13.
[9] M. Procopio, A. Lahm, A. Tramontano, L. Bonati, D. Pitea, A model for recognition of polychlorinated dibenzo-p-dioxins by the aryl hydrocarbon receptor, Eur. J. Biochem. 269 (2002) 13–18.
[10] M.S. Denison, A. Pandini, S.R. Nagy, E.P. Baldwin, L. Bonati, Ligand binding and activation of the Ah receptor, Chem. Biol. Interact. 141 (2002) 3–24.
[11] M.N. Jacobs, M. Dickins, D.F. Lewis, Homology modeling of the nuclear receptors: human estrogen receptorβ (hERβ), the human pregnane-X-receptor (PXR), the Ah receptor (AhR) and the constitutive androstane receptor (CAR) ligand binding domains from the human estrogen receptor α (hERα) crystal structure, and the human peroxisome proliferator activated receptor α (PPARα) ligand binding domain from the human PPARγ crystal structure, J. Steroid. Biochem. Mol. Biol. 84 (2003) 117–132.
[12] A. Pandini, M.S. Denison, Y. Song, A.A. Soshilov, L. Bonati, Structural and functional characterization of the aryl hydrocarbon receptor ligand binding domain by homology modeling and mutational analysis, Biochemistry 46 (2007) 696–708.

[13] A. Pandini, A.A. Soshilov, Y. Song, J. Zhao, L. Bonati, M.S. Denison, Detection of the TCDD binding-fingerprint within the Ah receptor ligand binding domain by structurally driven mutagenesis and functional analysis, Biochemistry 48 (2009) 5972–5983.
[14] K. Goryo, A. Suzuki, C.A. Del Carpio, K. Siizaki, E. Kuriyama, Y. Mikami, K. Kinoshita, K. Yasumoto, A. Rannug, A. Miyamoto, Y. Fujii-Kuriyama, K. Sogawa, Identification of amino acid residues in the Ah receptor involved in ligand binding, Biochem. Biophys. Res. Commun. 354 (2007) 396–402.
[15] A. Sali, T.L. Blundell, Comparative protein modeling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[16] S.F. Altschul, F. Stephen, L. Thomas, A. Alejandro, J. Zhang, J. Zhang, W. Miller, J. Lipman, Gapped blast and PSI-BLAST—a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[17] A.D.J. Van Dijik, S.J. De Vries, C. Dominguez, H. Chen, H.X. Zhou, A.M.J.J. Bonvin, Data-driven docking: HADDOCK's adventures in CAPRI, Protein 60 (2005) 232–238.
[18] L.A. Kelley, S.P. Gardner, M.J. Sutcliffe, An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies, Protein Eng. 9 (1996) 1063–1065.
[19] M.A. Marti-Renom, A.C. Stuart, A. Fiser, R. Sánchez, F. Melo, A. Sali, Comparative protein structure modeling of genes and genomes, Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 291–325.
[20] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, J. Comput. Chem. 19 (1998) 1639–1662.
[21] D.A. Case, T.E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K.M. Merz Jr., A. Onufriev, C. Simmerling, B. Wang, R. Woods, The amber biomolecular simulation programs, J. Comput. Chem. 26 (2005) 1668–1688.
[22] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, J. Am. Chem. Soc. 117 (1995) 5179–5197.
[23] W.L. Jorgensen, J. Chandrasekhar, J. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[24] B.H. Besler, K.M. Merz Jr., P.A. Kollman, Atomic charges derived from semiempirical methods, J. Comput. Chem. 11 (1990) 431–439.
[25] M.J. Frisch, et al., Gaussian 03 Revision D.01, Gaussian, Inc, Pittsburgh, PA, 2003.
[26] K. Kitaura, S. Sugiki, T. Nakano, Y. Komeiji, M. Uebayasi, Fragment molecular orbital method: analytical energy gradients, Chem. Phys. Lett. 336 (2001) 163–170.
[27] T. Nakano, et al., ABINIT-MP Ver 4.1, 2007, http://moldb.nihs.go.jp/abinitmp/.
[28] Y. Mochizuki, K. Fukuzawa, A. Kato, S. Tanaka, K. Kitaura, T. Nakano, A configuration analysis for fragment interaction, Chem. Phys. Lett. 410 (2005) 247–253.
[29] P.J. Erbel, P.B. Card, O. Karakuzu, R.K. Bruick, K.H. Gardner, Structural basis for PAS domain heterodimerization in the basic helix–loop–helix–PAS transcription factor hypoxia-inducible factor, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 15504–15509.
[30] P.B. Card, P.J. Erbel, K.H. Gardner, Structural basis of ARNT PAS-B dimerization: use of a common beta-sheet interface for hetero- and homodimerization, J. Mol. Biol. 353 (2005) 664–678.
[31] S.I. Karchner, D.G. Franks, S.W. Kennedy, M.E. Hahn, The molecular basis for differential dioxin sensitivity in birds: Role of the aryl hydrocarbon receptor, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 6252–6257.